

UNIVERSIDAD AUTÓNOMA AGRARIA ANTONIO NARRO  
SUBDIRECCIÓN DE POSTGRADO



ASOCIACIÓN DE REGIONES GENÓMICAS CON LAS CONDICIONES DE  
ARIDEZ Y SU RELACIÓN CON EL HÁBITO EN TRIGO HARINERO (*Triticum  
aestivum* L.).

Tesis

Que presenta ANA LUISA GÓMEZ ESPEJO

Como requisito parcial para obtener el Grado de

DOCTOR EN CIENCIAS EN RECURSOS FITOGENÉTICOS PARA ZONAS  
ÁRIDAS

Saltillo, Coahuila

Octubre 2022

UNIVERSIDAD AUTÓNOMA AGRARIA ANTONIO NARRO  
SUBDIRECCIÓN DE POSTGRADO



ASOCIACIÓN DE REGIONES GENÓMICAS CON LAS CONDICIONES DE  
ARIDEZ Y SU RELACIÓN CON EL HÁBITO EN TRIGO HARINERO (*Triticum  
aestivum* L.).

Tesis

Que presenta ANA LUISA GÓMEZ ESPEJO

Como requisito parcial para obtener el Grado de

DOCTOR EN CIENCIAS EN RECURSOS FITOGENÉTICOS PARA ZONAS  
ÁRIDAS

Dr. Manuel Humberto Reyes Valdés.  
Director interno UAAAN

Dra. Carolina Paola Sansaloni.  
Co-director externo CIMMYT

Dr. Marcelino Cabrera de la Fuente.  
Subdirector de Postgrado  
UAAAN

Saltillo, Coahuila

Octubre 2022

ASOCIACIÓN DE REGIONES GENÓMICAS CON LAS CONDICIONES DE  
ARIDEZ Y SU RELACIÓN CON EL HÁBITO EN TRIGO HARINERO (*Triticum  
aestivum* L.).

Tesis

Elaborada por ANA LUISA GÓMEZ ESPEJO como requisito parcial para obtener el  
grado de Doctor en Ciencias en Recursos Fitogenéticos para Zonas Áridas con la  
supervisión y aprobación del Comité de Asesoría

Dr. Manuel Humberto Reyes Valdés.  
Asesor principal

Dra. Carolina Paola Sansaloni.  
Asesor

Fernando H. Ribeiro Barrozo Toledo.  
Asesor

Dra. Hermila Trinidad García Osuna.  
Asesor

Dr. Adalberto Benavides Mendoza.  
Asesor

Dr. Marcelino Cabrera de la Fuente.  
Subdirector de Postgrado  
UAAAN

## AGRADECIMIENTOS

A todas las personas e instituciones que participaron de alguna manera en la presente investigación, mis más sinceros agradecimientos:

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por otorgar la beca para la realización de mis estudios de doctorado.

A la Universidad Autónoma Agraria Antonio Narro y al Departamento de Control Escolar de Postgrado, por el apoyo y facilidades otorgadas durante el ingreso y desarrollo del programa de doctorado.

Al Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT) a través de la iniciativa Seeds of Discovery (SeeD) por el apoyo en la obtención y acceso a los datos genotípicos utilizados en la presente investigación.

Al Comité particular de asesoría, por todas sus contribuciones y aportaciones para el desarrollo de este trabajo de investigación.

Particularmente al Dr. Humberto Reyes por su orientación en la dirección de la investigación, por sus atinadas correcciones en los manuscritos y por el conocimiento compartido tanto en sus clases, como en los distintos foros de apoyo.

A los doctores adscritos al CIMMYT: Carolina Sansaloni, Juan Burgueño y Fernando Toledo, por las facilidades otorgadas para el acceso a la información del CIMMYT y por las contribuciones en la corrección de los manuscritos.

A los doctores: Francisca Rivero, Adalberto Benavides y Hermila García por su amabilidad, accesibilidad, por los conocimientos compartidos en mi formación doctoral, y por su participación en la revisión de la presente investigación.

A mis compañeros del doctorado, por sus consejos, amistad y disponibilidad de apoyo, les deseo mucho éxito en sus proyectos profesionales.

A las amables y gentiles colaborados del programa REFIZA: Dulce Mendoza y Norma González, por su amistad y apoyo.

## Carta de artículo aceptado y publicado



ISSN: 2007-9028 impreso

ISSN: 2007-901X electrónico

Año 8 Núm. 2 Mayo - Agosto 2021

Fecha de publicación: 2021-06-27

Citación:

Gómez-Espejo, A. L., Sansaloni, C. P., Burgueño, J., Toledo, F. H., & Reyes-Valdés, M. H. (2021). Asociación de genoma completo para el hábito de crecimiento en trigo harinero (*Triticum aestivum* L.). *Ecosistemas Y Recursos Agropecuarios*, 8(2). <https://doi.org/10.19136/era.a8n2.2912>

<https://era.ujat.mx/index.php/rera/article/view/2912>

## Carta de artículo enviado

9/9/22, 08:15

Gmail - [Plants] Manuscript ID: plants-1871600 - Submission Received



Ana Gómez <gomez.ana.mvz@gmail.com>

### [Plants] Manuscript ID: plants-1871600 - Submission Received

Editorial Office <plants@mdpi.com>

1 de agosto de 2022, 13:08

Responder a: plants@mdpi.com

Para: "M. Humberto Reyes-Valdés" <mathgenome@gmail.com>

CC: "Ana L. Gómez-Espejo" <gomez.ana.mvz@gmail.com>, Carolina Paola Sansaloni <c.sansaloni@cgiar.org>, Juan Burgueño <j.burgueno@cgiar.org>, "Fernando H. Toledo" <f.toledo@cgiar.org>, Adalberto Benavides-Mendoza <abenmen@gmail.com>

Dear Dr. Reyes-Valdés,

Thank you very much for uploading the following manuscript to the MDPI submission system. One of our editors will be in touch with you soon.

Journal name: Plants

Manuscript ID: plants-1871600

Type of manuscript: Article

Title: Worldwide selection footprints for drought and heat in bread wheat

(*Triticum aestivum* L.)

Authors: Ana L. Gómez-Espejo, Carolina Paola Sansaloni, Juan Burgueño, Fernando H. Toledo, Adalberto Benavides-Mendoza, M. Humberto Reyes-Valdés \*

Received: 1 August 2022

E-mails: gomez.ana.mvz@gmail.com, c.sansaloni@cgiar.org, j.burgueno@cgiar.org, f.toledo@cgiar.org, abenmen@gmail.com, mathgenome@gmail.com

Submitted to section: Plant Genetics, Genomics and Biotechnology,

[https://www.mdpi.com/journal/plants/sections/Plant\\_Genetics](https://www.mdpi.com/journal/plants/sections/Plant_Genetics)

From Genome to Phenome: A Bridge to Accelerate Adaptation of Crops to Climate Change

[https://www.mdpi.com/journal/plants/special\\_issues/genome\\_phenome\\_crops](https://www.mdpi.com/journal/plants/special_issues/genome_phenome_crops)

You can follow progress of your manuscript at the following link (login

required):

[https://susy.mdpi.com/user/manuscripts/review\\_info/786ee6530a7190400290a4c32ff28d6](https://susy.mdpi.com/user/manuscripts/review_info/786ee6530a7190400290a4c32ff28d6)

8/9/22, 13:57

Gmail - [Plants] Manuscript ID: plants-1871600 - Accepted for Publication



Ana Gómez <gomez.ana.mvz@gmail.com>

### [Plants] Manuscript ID: plants-1871600 - Accepted for Publication

Plants Editorial Office <plants@mdpi.com>

29 de agosto de 2022, 01:12

Responder a: Bianca Wu <bianca.wu@mdpi.com>, Plants Editorial Office <plants@mdpi.com>

Para: "M. Humberto Reyes-Valdés" <mathgenome@gmail.com>

CC: "Ana L. Gómez-Espejo" <gomez.ana.mvz@gmail.com>, Carolina Paola Sansaloni <c.sansaloni@cgiar.org>, Juan Burgueño <j.burgueno@cgiar.org>, "Fernando H. Toledo" <f.toledo@cgiar.org>, Adalberto Benavides-Mendoza <abenmen@gmail.com>, Plants Editorial Office <plants@mdpi.com>, Bianca Wu <bianca.wu@mdpi.com>

Dear Dr. Reyes-Valdés,

Congratulations on the acceptance of your manuscript, and thank you for submitting your work to Plants:

Manuscript ID: plants-1871600

Type of manuscript: Article

Title: Worldwide selection footprints for drought and heat in bread wheat

(*Triticum aestivum* L.)

Authors: Ana L. Gómez-Espejo, Carolina Paola Sansaloni, Juan Burgueño, Fernando H. Toledo, Adalberto Benavides-Mendoza, M. Humberto Reyes-Valdés \*

Received: 1 August 2022

E-mails: gomez.ana.mvz@gmail.com, c.sansaloni@cgiar.org, j.burgueno@cgiar.org, f.toledo@cgiar.org, abenmen@gmail.com, mathgenome@gmail.com

Submitted to section: Plant Genetics, Genomics and Biotechnology,

[https://www.mdpi.com/journal/plants/sections/Plant\\_Genetics](https://www.mdpi.com/journal/plants/sections/Plant_Genetics)

From Genome to Phenome: A Bridge to Accelerate Adaptation of Crops to Climate Change

[https://www.mdpi.com/journal/plants/special\\_issues/genome\\_phenome\\_crops](https://www.mdpi.com/journal/plants/special_issues/genome_phenome_crops)

[https://susy.mdpi.com/user/manuscripts/review\\_info/786ee6530a7190400290a4c32ff28d6](https://susy.mdpi.com/user/manuscripts/review_info/786ee6530a7190400290a4c32ff28d6)

Gómez-Espejo, A.L.; Sansaloni, C.P.; Burgueño, J.; Toledo, F.H.; Benavides-Mendoza, A.; Reyes-Valdés, M.H. Worldwide Selection Footprints for Drought and Heat in Bread Wheat (*Triticum aestivum* L.). *Plants* 2022, 11, 2289. <https://doi.org/10.3390/plants11172289>

<https://www.mdpi.com/2223-7747/11/17/2289/htm>

## INTRODUCCIÓN

El trigo harinero (*Triticum aestivum* L.) se encuentra entre los tres cultivos alimentarios más extendidos en el mundo, ya que es cultivado en más de 218 millones de hectáreas (Giraldo *et al.*, 2019; Rasheed & Xia, 2019). Lo anterior, debido a que posee una alta capacidad de adaptación causada por la diversidad de sus genes fenológicos que garantizan su éxito reproductivo en una amplia gama de ambientes (Nazim Ud Dowla *et al.*, 2018).

En los últimos años, el cambio climático, caracterizado por altas temperaturas e imprevisibilidad en los patrones de precipitación, se ha convertido en la principal limitante para su producción (Daloz *et al.*, 2021; Valizadeh *et al.*, 2014). Asimismo, con el aumento constante de la población mundial, que se estima será de más de nueve mil millones para el año 2050 (United Nations *et al.*, 2019), se requerirá de una mejora significativa en la producción de este cereal.

Entre los rasgos fundamentales de adaptación a diferentes entornos, se encuentra su fenología, la arquitectura de las plantas, y la tolerancia al estrés climático, que son caracteres cuantitativos controlados por múltiples genes, cuyo estudio es sumamente complejo (Lamaoui *et al.*, 2018). Por otro lado, dado que el cambio climático puede influir potencialmente en la variación de estos rasgos (Xiao *et al.*, 2018), es importante comprender los genes que subyacen a esta variación y cómo interactúan entre sí, con otros rasgos y con el entorno del cultivo (Hyles *et al.*, 2020).

Los genes fenológicos regulan los procesos fisiológicos del trigo (Barrett *et al.*, 2002), entre sus principales determinantes se consideran la respuesta a la temperatura de vernalización y la duración del día, los cuales están regulados por la variación alélica en los genes de vernalización “*Vrn*” y fotoperíodo “*Ppd*” (Whittal *et al.*, 2018). Estos genes en conjunto con los genes de precocidad per se “*Eps*” controlan la sincronía de la floración bajo condiciones estacionales óptimas (Khumalo *et al.*, 2017).

Pese a que se han estudiado y manipulado diversidad de genes que subyacen a la expresión de muchos rasgos fenológicos (Bendix *et al.*, 2015; Kippes *et al.*, 2015; Ochagavía *et al.*, 2019; Trevaskis *et al.*, 2007; Yan *et al.*, 2004, 2006; Yoshida *et al.*, 2010), aún se desconocen muchos aspectos relacionados con la respuesta ante condiciones climáticas

limitantes que podrían alterar las fases de desarrollo y el rendimiento del cultivo (Lopes *et al.*, 2015), por lo que es importante cuantificar el impacto en la productividad por parámetros climáticos como cambios de temperatura y precipitación (Daloz *et al.*, 2021).

Para Nazim Ud Dowla *et al.* (2018) la optimización de la fenología del trigo mediante una mejor utilización de los recursos genéticos disponibles contribuirá a mantener y maximizar el potencial de rendimiento en ambientes con limitaciones. En este sentido, las variedades locales representan una fuente potencial de diversidad genética, ya que pueden proporcionar los recursos genéticos útiles para hacer frente a los desafíos climáticos en el mejoramiento del cultivo (Tehseen *et al.*, 2021).

Las variedades locales han evolucionado a través ciclos continuos de selección natural y humana para subsistir a entornos desafiantes (Corrado & Rao, 2017; Sansaloni *et al.*, 2020; Vikram *et al.*, 2016). Por lo que albergan en su genoma las firmas de adaptación a las diferentes condiciones agroecológicas de sus centros de origen (Contreras-Moreira *et al.*, 2019; Singh *et al.*, 2018).

Hoy en día, con la disponibilidad de nuevas herramientas bioinformáticas y recursos genómicos, como las tecnologías de genotipado de polimorfismos de un solo nucleótido (SNP) de alta densidad y rendimiento, puede ser posible diseccionar los mecanismos genéticos de rasgos complejos que subyacen al proceso adaptativo (Pavlidis *et al.*, 2010). Mejor aún, existen los estudios de asociación del genoma completo (GWAS) que representan un enfoque estadístico muy eficiente, en relación con el tiempo, el costo y la precisión, para la investigación de variaciones naturales asociadas a diversos rasgos usando grandes poblaciones (Nakano & Kobayashi, 2020; Tang *et al.*, 2016).

GWAS puede detectar la región genómica involucrada en un rasgo objetivo basado en el análisis de asociación entre la variación fenotípica y los SNP de todo el genoma (Rodríguez *et al.*, 2020). La detección de asociaciones se basa principalmente en la cobertura de marcadores genéticos, el número de individuos estudiados y el desequilibrio de ligamiento (LD) entre los polimorfismos causales y vinculados (Turuspekov *et al.*, 2017).



Existen varios enfoques GWAS para descubrir signos de la selección natural, así como detectar genes y factores ambientales implicados en la adaptación local (Collevatti *et al.*, 2019; De La Torre *et al.*, 2019). Sin embargo, un enfoque que integra con éxito el entorno, que es una de las principales fuerzas impulsoras de la selección natural son las asociaciones genoma-ambiente (GEA), que representan una mejor alternativa para detectar loci adaptativos (Rellstab *et al.*, 2015).

El enfoque GEA se ha empleado con éxito en la detección de genes específicos involucrados en los procesos de adaptación de poblaciones locales (Pluess *et al.*, 2016), y consiste en asociar SNP distribuidos a lo largo del genoma con predictores ambientales individuales, determinados en los sitios de muestreo de las accesiones evaluadas (López-Hernández & Cortés, 2019). Además, tiene el potencial de predecir los factores involucrados en la adaptación al estrés abiótico en las plantas (Elias *et al.*, 2021).

A través de la iniciativa Seeds of Discovery (SeeD; <http://seedsofdiscovery.org/>) del Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT) se han reportado progresos importantes en los recursos de secuenciación de todo el genoma de trigo, con la caracterización de más de 80,000 accesiones de trigo (Sansaloni *et al.*, 2020), las cuales forman parte de una vasta colección resguardada en los bancos de germoplasma de la institución y del Centro Internacional para la Investigación Agrícola en Zonas Áridas (ICARDA).

Dado el supuesto de que la variación genética necesaria para adaptar los cultivos al cambio climático puede ser proporcionada por las variedades locales y esta puede ser identificada empleando los diversos enfoques GWAS (Gates *et al.*, 2019), existe la oportunidad de explorar la información genotípica disponible de trigo en el estudio de rasgos adaptativos que conduzcan al descubrimiento de nuevas fuentes de variabilidad genética para las Zonas áridas. Por ello, esta investigación tiene por objetivo la identificación de regiones genómicas relacionadas con el proceso de adaptación a climas áridos y con el hábito de crecimiento en *Triticum aestivum* mediante estudios de asociación del genoma completo.

## Antecedentes

La iniciativa SeeD, es una iniciativa conjunta del CIMMYT y la Secretaría de Agricultura de México (SAGARPA) creada con el objetivo de facilitar el uso efectivo de los recursos genéticos de maíz y trigo para contribuir en el desarrollo de variedades mejoradas acordes a los desafíos del cambio climático y a las necesidades cambiantes de los agricultores y consumidores (Pixley *et al.*, 2018). Entre sus servicios, cuenta con un laboratorio de genotipado rentable, denominado Servicio de Análisis Genético para la Agricultura (SAGA) en la sede del CIMMYT, en el que se utiliza tecnología de secuenciación de última generación DArTSeq para la caracterización genotípica de diversas accesiones (<https://www.diversityarrays.com/>).

Desde mediados de 2016, SAGA genotipificó dos de los principales cultivos alimentarios de la humanidad, aproximadamente 100,000 muestras de trigo y 44,500 de maíz (Pixley *et al.*, 2018), produciendo datos de varios cientos de miles de marcadores, tanto de polimorfismo de un solo nucleótido (SNP), como de presencia, ausencia y varianza (PAV), en cada genotipo secuenciado (Sansaloni *et al.*, 2020). Esta información se encuentra disponible públicamente en el sitio web de SeeD, a través de las plataformas *Germinate* (Shaw *et al.*, 2017) y *Dataverse* (<https://data.cimmyt.org/>).

A la fecha, la información producida por SeeD se ha empleado con éxito en más de 28 publicaciones científicas, entre las que destacan la construcción de un mapa genético de trigo (H. Li *et al.*, 2015), múltiples estudios de diversidad (Lopes *et al.*, 2015; Robbana *et al.*, 2019; Romero-Navarro *et al.*, 2017; Sansaloni *et al.*, 2020; Sehgal *et al.*, 2015; Vikram *et al.*, 2016) y en la identificación de regiones genómicas asociadas con rasgos de interés (J. Chen *et al.*, 2016; Crossa *et al.*, 2016; Ledesma-Ramírez *et al.*, 2019; Saint Pierre *et al.*, 2016; Singh *et al.*, 2018).

## REVISIÓN DE LITERATURA

### **Biología del trigo y su interacción con el ambiente.**

El trigo común hexaploide (*Triticum aestivum* L.) fue uno de los primeros cultivos en ser domesticado hace más de 10,000 años en el Medio Oriente (Charmet, 2011). Posteriormente, se extendió a un amplio rango de latitudes, desde los 67° norte en Rusia hasta los 50° sur en Chile, también se cultiva en diferentes épocas del año y en casi todas las regiones climáticas (Guzmán *et al.*, 2019). Lo anterior, debido a su alta “plasticidad genómica”, conferida por su naturaleza alopoliploide (Feldman & Levy, 2012).

De acuerdo con Haider, (2013) la especie *Triticum aestivum* se originó a través de uno o más eventos raros de hibridación entre *Aegilops tauschii* (DD) y el tetraploide *Triticum turgidum* (AABB). Es decir, su genoma se deriva de los tres genomas homólogos, A, B y D, albergados por las especies diploides *T. urartu*, *A. speltoides* (o una especie estrechamente relacionada) y *A. tauschii*, cada uno de los cuales aporta siete pares de cromosomas ( $2n = 6x = 42$ ) al genoma total (L. Li *et al.*, 2015).

Después de la domesticación, los procesos evolutivos del trigo involucraron numerosos cambios morfológicos, fisiológicos y genéticos complejos denominados "síndrome de domesticación" (Peleg *et al.*, 2011). Los cambios fenotípicos más comunes se centraron en el modo de dispersión de la semilla, la pérdida de la rotura de las semillas, la minimización de la latencia de las semillas y el aumento tanto del tamaño como del número de semillas (Rahman *et al.*, 2020; Spor *et al.*, 2020).

El estudio de la domesticación permitió estructurar una visión histórica de la selección y mejora realizada por nuestros antepasados durante varios miles de años en su transición de cazadores/recolectores a pastores/agricultores (Sakuma *et al.*, 2011). Además, se descubrió que su base genética consistía en unos pocos grupos de genes de grandes efectos, a lo que se debía la mayor parte de la variación asociada entre las formas domesticadas y sus progenitores salvajes (Peng *et al.*, 2003).

Como resultado, los efectos combinados de la domesticación, la selección natural y la humana llevaron a la arquitectura de genotipos con mucha variación fenológica, ideales

para el desarrollo de cultivares autóctonos (razas locales) adaptados a las condiciones específicas de sus hábitats (Jaradat, 2013).

Las razas locales se definen como poblaciones complejas, variables, genéticamente dinámicas y diversas, poseedoras de altos niveles de tolerancia a los estreses abióticos y bióticos (Lodhi *et al.*, 2020). Con el paso del tiempo, se convirtieron en la base genética de los cultivares modernos de trigo obtenidos a través de programas mejoramiento (Royo *et al.*, 2020).

Fisiológicamente se suelen distinguir cuatro etapas de desarrollo en trigo: germinación a emergencia (E); etapa de crecimiento 1 (GS1) desde la emergencia hasta la doble cresta; etapa de crecimiento 2 (GS2) de doble cresta a antesis; y etapa de crecimiento 3 (GS3), que incluye el período de llenado del grano, desde la antesis hasta la madurez (Acevedo *et al.*, 2002).

Un punto fundamental en el ciclo de vida del trigo son los factores ambientales y regionales, como la temperatura, la humedad disponible, la latitud, la salinidad y la duración del día, que tienen una influencia directa sobre la duración de las fases de desarrollo (Trethowan *et al.*, 2006).

Para Hyles *et al.* (2020) esta dependencia con respecto al entorno de crecimiento representa una fuerte interacción entre el genotipo y el entorno, y actúa para sincronizar el ciclo de vida con las condiciones externas, lo que da como resultado la producción de biomasa y el rendimiento. En este sentido, la interacción entre el medio ambiente con los genes de fenología y desarrollo de las plantas determinan el éxito adaptativo del cultivo, por lo que es vital comprender los genes que subyacen a la variación del desarrollo.

Los genes que controlan la fenología y desarrollo de las plantas se han estudiado en múltiples sistemas de plantas, mediante los cuales se ha proporcionado conocimientos profundos sobre las vías genéticas que controlan los patrones de crecimiento de las plantas y que son fundamentalmente importantes para la biología de los cultivos y la agricultura (Trevaskis, 2018).

### **Floración en trigo.**

Diversos comportamientos de floración estacional impulsan la adaptación global del trigo harinero (Trevaskis, 2015). Por lo tanto, el espigado y la floración son etapas críticas en el desarrollo del cultivo, ya que tienen un papel importante en la adaptación, el potencial de rendimiento y la calidad del grano (Maccaferri *et al.*, 2008).

Se han identificado tres grupos de genes fenológicos relacionados con el control del tiempo de floración en trigo, que incluyen el requisito de vernalización, la sensibilidad al fotoperíodo y la “precocidad *per se*” (Gupta *et al.*, 2020), cuyo conocimiento ha permitido la selección de cultivares de trigo con ciclos de vida adaptados a diferentes condiciones estacionales.

La vernalización promueve el cambio de la fase vegetativa de la planta a la fase reproductiva por exposición al frío (Trevaskis *et al.*, 2007; Zhang *et al.*, 2008). Según los requisitos de vernalización, el trigo se clasifica como de hábito de crecimiento invernal o primaveral (Dreisigacker *et al.*, 2021). Las variedades de trigo de invierno requieren un cierto período de frío para la expresión de activadores florales, mientras que las variedades de primavera florecen sin vernalización (Kamran *et al.*, 2014).

Cuatro genes mayores denominados *Vrn1*, *Vrn2*, *Vrn3* y *Vrn4* han sido relacionados con la respuesta a la vernalización en diferentes cultivares de trigo (Danyluk *et al.*, 2003; Kippes *et al.*, 2015; Nazim Ud Dowla *et al.*, 2018; Trevaskis, 2010; Yan *et al.*, 2004, 2006). Particularmente, el locus Vernalization1 (*Vrn1*) ubicado en el brazo largo del cromosoma 5 (A, B y D), es señalado como el responsable de la supresión del requisito de vernalización (Y. Li *et al.*, 2020).

El trigo generalmente se clasifica como una planta de día largo, porque la mayoría de las variedades florecen antes cuando se exponen a días más largos (Dubcovsky *et al.*, 2006). Sin embargo, existen algunas variedades actuales que poseen diferentes niveles de respuesta al fotoperíodo, incluida la insensibilidad.

Los genes del fotoperíodo son los responsables de la insensibilidad en el trigo y juegan un papel importante en el desarrollo de las espigas y el tiempo de floración (L. Chen *et al.*, 2018). La sensibilidad al fotoperíodo está controlada por la presencia del alelo insensible (a) o del alelo sensible (b) en los genes *Ppd-A1*, *Ppd-B1* y *Ppd-D1* del cromosoma 2 (Yang *et al.*, 2009).

Las variedades insensibles al fotoperíodo son de floración temprana en condiciones de día corto ( $\leq 10$  hrs de luz); mientras que en las sensibles, tienen el efecto contrario, la fase vegetativa es prolongada y se retrasa el desarrollo reproductivo (S. Kumar *et al.*, 2012). En estas, los días largos ( $\geq 10$  hrs de luz) reducen el tiempo de iniciación del desarrollo floral (Beales *et al.*, 2007). Por su parte, los cultivos de trigo de invierno, luego de cubrir el requisito de vernalización, requieren exposición a día largo durante varias o muchas semanas para reducir el tiempo de floración (Trevaskis, 2010).

La precocidad *per se* es la diferencia en el tiempo de floración de variedades cuyos requisitos de vernalización y fotoperíodo han sido cumplidos, y es un rasgo heredado cuantitativamente controlado por varios genes menores (Royo *et al.*, 2020). Para Zikhali & Griffiths, (2015) el papel de los genes *Ppd* y *Vrn* representan el ajuste del control de floración grueso al influir en la adaptación a mega-ambientes, mientras que los genes *Eps* adaptan la floración dentro de los mega-ambientes, lo que permite una manipulación y ajuste fino del tiempo de floración.

Zikhali *et al.* (2016) propuso a un ortólogo regulador del reloj circadiano EARLY FLOWERING 3 (*ELF3*) como un gen candidato dentro del locus *Eps-D1* para la precocidad *per se* en *Triticum aestivum*, ya que una delección que contiene *ELF3* se asocia con una floración avanzada. Por otro lado, Ochagavía *et al.*, (2019) reporta algunos genes subyacentes y polimorfismos causales de los genes *Eps* en trigo hexaploide.

### **Adaptación a sequía y calor.**

Las plantas han adoptado diferentes mecanismos de respuesta para contrarrestar la amplia gama de estreses bióticos y abióticos que enfrentan (Gull *et al.*, 2019), para el cual activan varias vías integradas en respuesta al estrés externo (Verma *et al.*, 2016).

Entre las vías conservadas evolutivamente, las proteínas quinasas activadas por mitógenos (MAPK) representan una de las vías de señalización más importantes en la transmisión de estímulos relacionados con el estrés (Jalmi & Sinha, 2015). Las MAPK desempeñan un papel importante en la transducción de los estímulos extracelulares reconocidos por los receptores de la célula a un gran número de moléculas blanco para integrar las respuestas intracelulares altamente específicas al estímulo inicial (Danquah *et al.*, 2014).

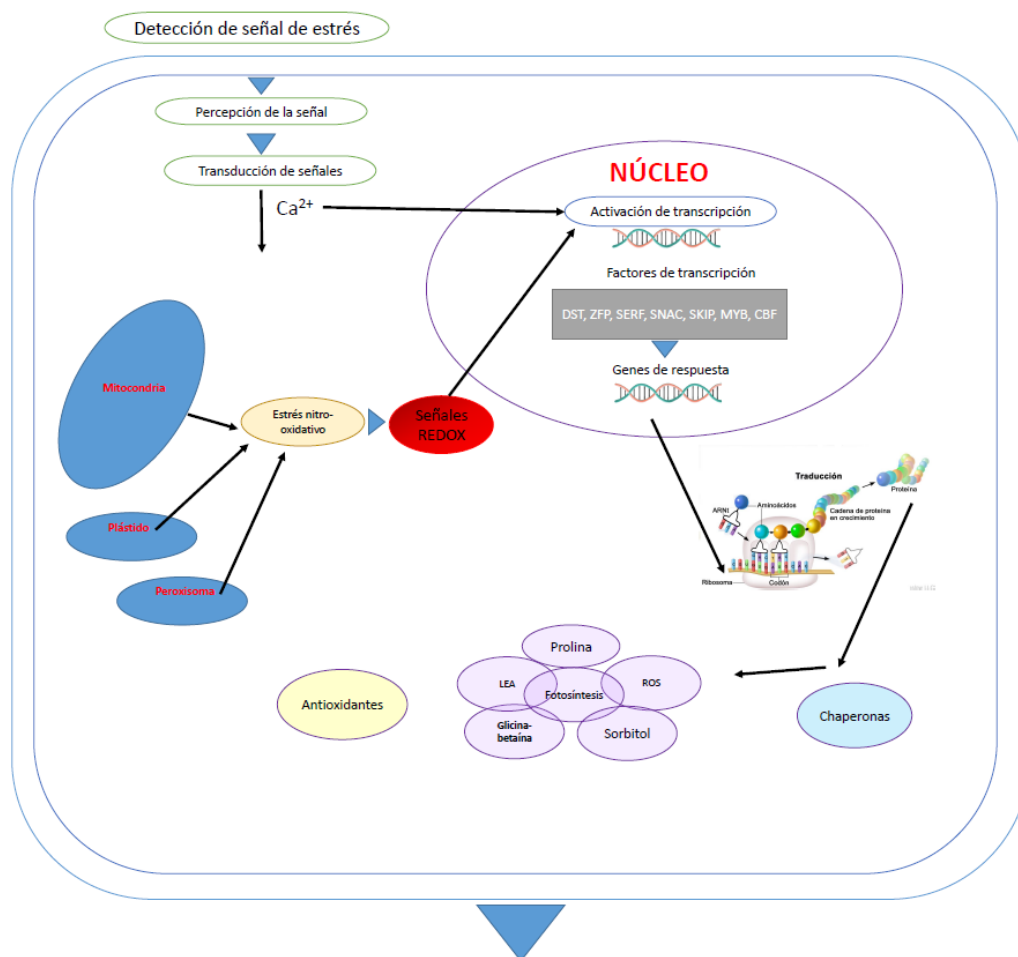


Figura 1. Mecanismos de respuesta a estrés abiótico en las células vegetales.

Por otro lado, las vías de señalización mediada por las hormonas vegetales, también son consideradas como importantes reguladores centrales de la respuesta al estrés abiótico en

las plantas, ya que desencadenan cambios importantes en la expresión génica y las respuestas fisiológicas adaptativas (EL Sabagh *et al.*, 2022).

Ante un escenario de cambio climático, el estrés producido por la escasez hídrica y las altas temperaturas destacan entre los principales factores ambientales que afectan el crecimiento, desarrollo, producción y calidad del trigo (Akter & Islam, 2017).

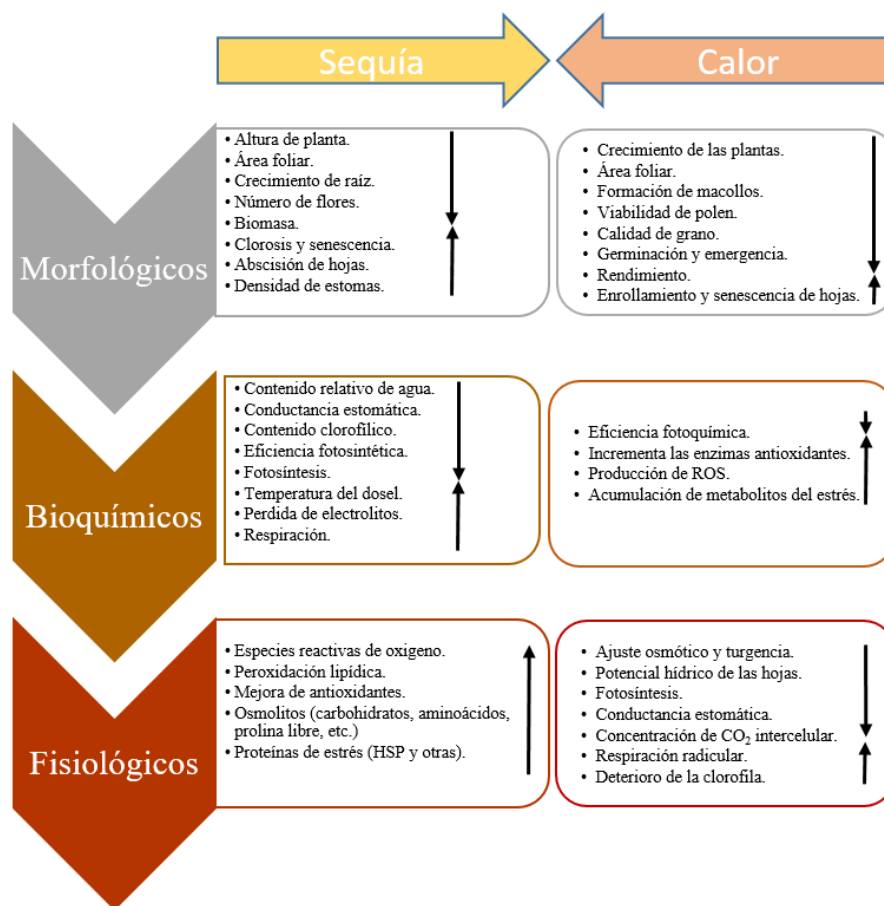


Figura 2. Principales efectos del estrés por sequía y calor en las plantas.

Ambos estreses tienen un efecto sinérgico, alterando el metabolismo y la expresión génica de formas distintas a las inducidas de forma independiente (Qaseem *et al.*, 2018). Además, pueden afectar las características agronómicas en cada etapa de desarrollo y provocar impactos negativos en varios procesos fisiológicos, celulares y moleculares en las células vegetales (López-Hernández & Cortés, 2019).



Por su parte, el estrés por calor produce afectaciones considerables en las etapas de prefloración y floración en trigo, pues causa reducciones en el número de granos por espiga y el rendimiento (Shenoda *et al.*, 2021). Específicamente, afecta la estructura celular y las vías metabólicas relacionadas con la termoestabilidad de la membrana, la fotosíntesis y la síntesis de almidón (Ni *et al.*, 2018).

En este sentido, se han descubierto muchos genes clave que gobiernan los rasgos morfofisiológicos de respuesta al estrés producido por sequía y calor (Kulkarni *et al.*, 2017). Sin embargo, la mayoría de los rasgos relacionados con tolerancia al estrés son poligénicos y muy complejos de entender en el nivel fisiológico y molecular (Sallam *et al.*, 2019).

Entre los genes mejor estudiados como reguladores de la respuesta a la sequía destacan los asociados con la arquitectura de la raíz y el desarrollo de los estomas desempeñan funciones importantes en la extracción de la humedad del suelo y su retención (Kulkarni *et al.*, 2017). También se han identificado algunos grupos de enzimas y proteínas responsables de tolerancia al estrés en condiciones de sequía, como las proteínas abundantes en la embriogénesis tardía (LEA), RuBisCO, y helicasas (Nezhadahmadi *et al.*, 2013).

En cuanto a los genes asociados con la tolerancia al estrés por calor, se reconocen a las proteínas de choque térmico (HSP) como las candidatas idóneas para el desarrollo de cultivos tolerantes al calor (Xu *et al.*, 2011). Estas proteínas funcionan como chaperonas moleculares en el mantenimiento de la homeostasis del plegamiento de proteínas y están relacionadas con la adquisición de termotolerancia (Kumar *et al.*, 2020).

### **Estudios de asociación del genoma completo.**

Los estudios de asociación del genoma completo (GWAS) tienen sus inicios con la era de la secuenciación del genoma humano para estudiar enfermedades de importancia (Ikegawa, 2012). GWAS es utilizado para identificar variantes genéticas (SNP) en muchos genomas que están asociadas estadísticamente con un riesgo de una enfermedad o un rasgo particular (Scherer & Christensen, 2016).

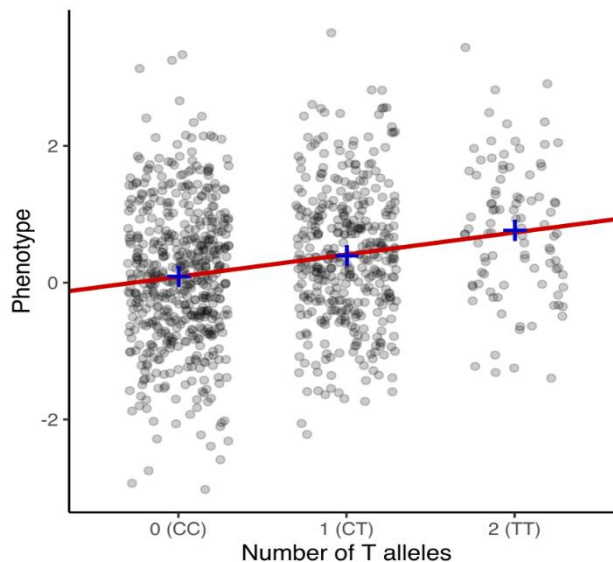


Figura 3. Ilustración de un genotipo simulado por regresión de fenotipo para un solo SNP, donde cada punto representa a un individuo. El GWAS de un rasgo continuo consiste esencialmente en repetir este análisis en cada SNP.

By D Wells - Own work, CC BY-SA 4.0,

<https://commons.wikimedia.org/w/index.php?curid=110760724>

Más recientemente, la tecnología GWAS ha experimentado un rápido crecimiento en muchas especies de animales, plantas y organismos modelo, pero sobre todo en el campo de las ciencias agronómicas enfocados en mejorar los programas de mejoramiento y la producción de alimentos (Gurgul *et al.*, 2014).

En este campo, los GWAS generalmente enfatizan las asociaciones entre los SNP y los rasgos de importancia, por ejemplo, el ADN de las variedades de plantas se compara con diferentes fenotipos para un rasgo particular (Challa & Neelapu, 2018). Por lo que han sido ampliamente utilizados en la identificación de la base genética de rasgos cuantitativos responsables de características como el rendimiento, la calidad, la tolerancia o resistencia a diversos tipos de estrés (Huang *et al.*, 2010; H. Li *et al.*, 2013; Pasam *et al.*, 2012; Sukumaran *et al.*, 2015).

La implementación de GWAS se basa en diseños, tecnologías de genotipado y conceptos estadísticos para el análisis, replicación, interpretación y seguimiento de los resultados de asociación (Bush & Moore, 2012). En este sentido, el poder estadístico para detectar

asociaciones genéticas verdaderas está determinado por factores como la variación fenotípica, el número de individuos, la estructura de la población, la frecuencia alélica y el desequilibrio de ligamiento (Alqudah *et al.*, 2020).

Un aspecto fundamental para tenerse en cuenta es el filtrado de los datos a utilizar en cuanto a su calidad. El pre-procesamiento de los datos genotípicos implica la eliminación o “filtrado” de los SNP que no se incluirán en los análisis (Reed *et al.*, 2015). Para ello, existen excelentes protocolos teóricos y prácticos para el control de la calidad general de los datos (Anderson *et al.*, 2010; Reed *et al.*, 2015; Weale, 2010).

Por lo general, en la selección de marcadores se considera el contenido de información y la selección de variantes más frecuentes, ya que el poder estadístico para inferir una relación significativa entre el SNP y el rasgo bajo estudio es más poderoso en estos casos (Guo & Elston, 1999; Zhao *et al.*, 2005). El umbral para considerar la rareza de un SNP dependerá del tamaño del estudio y del efecto esperado, en poblaciones pequeñas se recomienda un límite más estricto (1 al 5 %), por el contrario con tamaños de muestra muy grandes, puede ser beneficioso evitar la eliminación de los SNP raros (Laurie *et al.*, 2010).

La estratificación de la población ocurre cuando los individuos estudiados difieren tanto en su ascendencia genética como en el fenotipo bajo investigación (Hellwege *et al.*, 2017). Por lo que se debe tener el máximo conocimiento de la estructura de los datos, ya que las diferencias de ascendencia causan la aparición de asociaciones espúreas. Una técnica útil para evaluar este aspecto son los análisis de componentes principales (PCA), ya que realizan una correcta clasificación de individuos en grupos de ascendencia según su composición genética (Turkmen *et al.*, 2019).

**PRIMER ARTÍCULO**

**Asociación de genoma completo para el hábito de crecimiento en trigo harinero**  
*(Triticum aestivum L.)*

## Asociación de genoma completo para el hábito de crecimiento en trigo harinero (*Triticum aestivum* L.)

### Genome-wide association for growth habit in bread wheat (*Triticum aestivum* L.)

Ana L. Gómez-Espejo<sup>1</sup>  
 Carolina P. Sansaloni<sup>2</sup>  
 Juan Burgueño<sup>2</sup>  
 Fernando H. Toledo<sup>2</sup>  
 M. Humberto Reyes-Valdés<sup>1\*</sup>

<sup>1</sup>Programa de Doctorado en Recursos Fitogenéticos para Zonas Áridas-Universidad Autónoma Agraria Antonio Narro. Calzada Antonio Narro 1923, Buenavista, CP. 25315. Saltillo, Coahuila, México.

<sup>2</sup>Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT), CP. 56237. Texcoco, México.

\*Autor de correspondencia: mathgenome@gmail.com

#### Artículo científico

Recibido: 02 de marzo 2021

Aceptado: 10 de mayo 2021

**Como citar:** Gómez-Espejo AL, Sansaloni CP, Burgueño J, Toledo FH, Reyes-Valdés MH (2021) Asociación de genoma completo para el hábito de crecimiento en trigo harinero (*Triticum aestivum* L.). Ecosistemas y Recursos Agropecuarios 8(2): e2912. DOI: 10.19136/era.a8n2.2912

**RESUMEN.** El hábito de crecimiento de los trigos harineros (primavera e invierno) está determinado primordialmente por el proceso de vernalización, regulado por los genes *Vrn*. Algunos estudios sugieren la implicación de otros genes en la diferenciación de dicha característica. Por ello, se tiene por objetivo la identificación genómica de regiones con una posible relación funcional al hábito de crecimiento. Se realizó un estudio de asociación del genoma completo (GWAS) con dos enfoques de análisis: a) FarmCPU (Unificación de Probabilidad Circulante de Modelo fijo y aleatorio) y b) BLINK (Información Bayesiana y Desequilibrio de ligamiento Anidado Iterativamente) implementados en el paquete GAPIT de R. Se utilizaron datos de 25 681 loci de polimorfismo de un solo nucleótido (SNP's) de 7 757 colectas de *T. aestivum* del banco de germoplasma del CIMMYT clasificadas en cuanto a su hábito de crecimiento. Se identificaron 593 marcadores SNP's con una alta asociación al hábito de crecimiento ( $p < 0.01$ ). Pero solo 70 marcadores formaban bloques de asociación (BA) en 11 cromosomas. Como se esperaba, se identificó un BA cercano al gen *Vrn1* en el cromosoma 5A. Adicionalmente se identificaron 20 regiones genómicas (BA) con asociación al hábito de crecimiento, en las que se detectó la presencia de proteínas de respuesta al estrés implicadas en procesos de aclimatación y des-aclimatación. Los resultados sugieren la participación de nuevos genes en la determinación del hábito de crecimiento en *Triticum aestivum* y su potencial predictivo en colectas no clasificadas.

**Palabras clave:** Estrés, GWAS, hábito de crecimiento, SNP's, trigo harinero.

**ABSTRACT.** The growth habit in bread wheat (spring and winter) is mainly determined by the process of vernalization, regulated by the *Vrn* genes. Several studies suggest the involvement of other genes in this differential trait. Therefore, the aim of this research was the genomic identification of regions with a possible functional relationship to growth habit. A genome wide association study (GWAS) was performed by using two different approaches: a) FarmCPU (Fixed and random model Circulating Probability Unification) and b) BLINK (Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway), both implemented in the GAPIT package of R. The source was a data set of 25 681 single nucleotide polymorphism loci (SNP's) from 7 757 accessions of *T. aestivum* stored in the CIMMYT germplasm bank and classified according to growth habit. A total of 593 SNP's markers strongly associated to their growth habit was identified ( $p < 0.01$ ). But only 70 markers formed association blocks (BA) within 11 chromosomes. As expected, at least one block was located close to the *Vrn1* gene of the 5A chromosome. Additionally, 20 genomic regions (BA) were identified with putative functional relationship to growth habit, with response stress proteins involved in acclimation and de-acclimation processes. The findings suggest of new genes implicated in the growth habit for *Triticum aestivum* and its predictive potential for unclassified genetic accessions.

**Key words:** Stress, GWAS, growth habit, SNP's, bread wheat.

## INTRODUCCIÓN

El trigo harinero, es el cereal más cultivado en el mundo, por su adaptación exitosa a diferentes condiciones agroecológicas (Li *et al.* 2020). Se clasifica en dos hábitos de crecimiento, primavera o invierno, según sus patrones de floración bajo diversos rangos de temperatura ambiental (Trevaskis *et al.* 2007, Zhang *et al.* 2008). Dicha característica fenológica asociada con el proceso de vernalización, se encuentra determinada por los genes *Vrn*, los cuales en interacción con genes del fotoperíodo (*Ppd*) y genes que controlan la tasa de desarrollo (*Eps*) regulan la floración oportuna bajo condiciones estacionales óptimas (Whittal *et al.* 2018, Khumalo *et al.* 2017).

En algunas especies de plantas de clima templado el requisito de vernalización es indispensable para la expresión de activadores florales después de un periodo prolongado de temperaturas frías (Dixon *et al.* 2019, Li *et al.* 2020). En este sentido, cuatro genes mayores denominados *Vrn1*, *Vrn2*, *Vrn3* y *Vrn4* han sido relacionados con la respuesta a la vernalización en diferentes cultivares de trigo (Yan *et al.* 2004, 2006, Trevaskis *et al.* 2003, Danyluk *et al.* 2003, Kippes *et al.* 2015, Nazim *et al.* 2018). Particularmente, se ha demostrado que variaciones alélicas en el gen *Vrn1*, cuyas copias homólogas (*Vrn-A1*, *Vrn-B1* y *Vrn-D1*) se ubican en los cromosomas 5A, 5B y 5D, respectivamente, son las responsables de la supresión del requisito de vernalización, ya que la presencia de alelos dominantes en tan solo un loci es suficiente para conferir el hábito de crecimiento primaveral (Li *et al.* 2020).

La caracterización de los genes *Vrn* ha permitido un avance significativo al entendimiento de la regulación genética del proceso de vernalización de los cereales (Kippes *et al.* 2015). Sin embargo, algunos estudios realizados en líneas endogámicas de cebada (Saisho *et al.* 2011) y mutantes de trigo tetraploide (Chen y Dubcovsky 2012) proporcionan evidencia sobre la contribución de genes desconocidos en la diferenciación de los dos hábitos de crecimiento y en el proceso de floración (Trevaskis 2015). Considerando lo anterior

y dada la importancia de la vernalización en la adaptación a entornos específicos de los cereales (Goncharov 2003), esta investigación tuvo como objetivo estudiar la asociación entre marcadores SNP's y el hábito de crecimiento en colectas de *T. aestivum* a través de un estudio de asociación de genoma completo (GWAS), con la finalidad de localizar regiones que tengan una posible relación funcional con dichos patrones.

## MATERIALES Y MÉTODOS

### Obtención de datos de hábito de crecimiento

Se utilizaron los datos de pasaporte de 174 553 colectas del banco de germoplasma de CIMMYT, cada entrada de los datos de pasaporte contiene una URL pública que permite acceso a información adicional de cada colecta. A través del desarrollo de código escrito en el lenguaje R y con ayuda del paquete "rvest" v0.3.6 (Wickham 2020), se programó el acceso automático a cada URL para la búsqueda de patrones que indicaran la presencia de información de hábito de crecimiento, y su categoría como cultivo de primavera o de invierno y que fueran *Triticum aestivum*. La búsqueda arrojó un total de 20 380 colectas con información sobre hábito de crecimiento. El proceso automático se validó a través de una búsqueda manual de una muestra de 100 colectas dentro de dicho conjunto, y se verificó una concordancia del 100% en la determinación de las categorías de primavera o invierno.

### Obtención y limpieza de datos genotípicos

La información molecular fue obtenida a partir de una tabla de datos en formato HapMap con información de 45 871 colectas del banco de germoplasma de trigo del Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT), las cuales fueron genotipificadas con tecnología DArTseq con 86 683 loci SNP en el CIMMYT. Los datos se filtraron para incluir solo colectas de *Triticum aestivum* caracterizadas por su hábito de crecimiento. Se seleccionaron solo marcadores informativos (polimórficos para las accesiones seleccionadas) con una tasa de datos perdidos del 10%. Se detectó la ubicación física de cada marcador con datos obtenidos por BLAST, y se mantuvieron solo aquellos

que podían ser localizados de una manera no ambigua en el genoma de trigo. Bajo este filtrado de datos, se obtuvo un conjunto con 7 956 colectas caracterizadas con 25 681 loci de SNP, en un formato HapMap, los cuales dieron cobertura amplia a los 21 cromosomas de *T. aestivum*. Las colectas seleccionadas tienen diverso origen geográfico y son una mezcla con diferente nivel genético, ya que se consideraron desde variedades criollas hasta líneas mejoradas.

### **Análisis exploratorio**

Los datos se exploraron mediante pruebas exactas de Fisher para todos los loci, en cuanto a su asociación con el hábito de crecimiento. Se detectó un número considerable de asociaciones bajo el criterio de Bonferroni para un valor  $p = 0.01$ , por lo que se procedió a realizar las pruebas adecuadas para el GWAS.

### **Análisis de asociación**

Considerando las dimensiones del conjunto de datos, además del tiempo y demanda computacional que implica realizar un número muy grande de pruebas estadísticas, una por cada SNP's genotificado, los GWAS se ejecutaron con un código de lectura de SNP's fragmentado para múltiples archivos. Mientras que los cálculos de parentesco y componentes principales (CP) se estimaron bajo un esquema de muestreo aleatorio. Para ello, la tabla de datos genotípicos HapMap fue dividida en 21 archivos, uno para cada cromosoma. Los datos fenotípicos de las accesiones fueron incluidos como un archivo independiente con el hábito de crecimiento clasificado en 0 y 1, donde 0 son las accesiones de hábito invernal y 1 las de primavera.

Se definieron dos enfoques de estudio para determinar la asociación: 1) El de Unificación de Probabilidad Circulante de Modelo fijo y aleatorio (FarmCPU) (Liu *et al.* 2016) y 2) BLINK (Huang *et al.* 2018), el cual está basado en información bayesiana. El procesamiento con estos dos enfoques se realizó con el paquete GAPIT (Herramienta Integrada de Predicción y Asociación del Genoma) versión 3.0 de R (Lipka *et al.* 2012). Estos enfoques se caracterizan por usar

iteraciones para seleccionar un conjunto de marcadores asociados con un rasgo de interés. Los marcadores asociados se ajustan como covariables para probar los marcadores en todo el genoma, una a la vez.

La estructura poblacional se determinó automáticamente aplicando un análisis de componentes principales (ACP) al ejecutar cada GWAS. Con las gráficas de componentes principales se identificaron colectas fuera de tipo, con posibles errores de clasificación, las cuales fueron depuradas del conjunto de datos genotípicos. Los gráficos de ACP se generaron con el paquete "ggplot2" v3.3.2 de R (Wickham 2009).

Los análisis de asociación bajo los modelos mencionados se ejecutaron nuevamente con los datos depurados (7 757 colectas y 25 681 marcadores SNP's). Para este efecto, la estructura poblacional se ajustó mediante los primeros cinco componentes principales. Los marcadores significativos en cada modelo fueron seleccionados considerando su valor  $p$  ajustado por la tasa de falso descubrimiento (FDR) a un alfa de 0.01.

### **Mapeo de marcadores**

Mediante el uso de la herramienta BLAST y las secuencias de nucleótidos de cada SNP significativo se corroboró su ubicación física con base en el genoma de referencia de trigo IWGSC\_refseqv1.0 (IWGSC, 2018) publicado en la base de datos de Ensembl Plants (2021). Las posiciones físicas fueron seleccionadas con base en los alineamientos con más altos porcentajes de identidad (> 90%). Con la información obtenida se elaboró una representación esquemática del mapa físico del trigo harinero mediante el paquete "LinkageMapView" v2.1.2 de R (Ouellette *et al.* 2018). En el mapa se incluyeron genes de vernalización, fotoperíodo y precocidad *perse*, reportados en bases de datos. Una vez elaborado el mapa físico se identificaron las regiones genómicas conformadas en bloques de asociación (BA).

### **Análisis bioinformático**

Las regiones genómicas conformadas en BA fueron sometidas a un análisis bioinformático para conocer las

propiedades funcionales de las mismas. Para ello, se exploró a detalle cada región genómica en la base de datos de Ensembl Plants con la finalidad de identificar posibles genes candidatos. La identificación de proteínas funcionales codificadas por genes candidatos se consultó en la base de datos UniProt (2021).

## RESULTADOS

### Análisis de asociación del genoma completo

Considerando los valores  $p$  ajustados por la tasa de falso descubrimiento bajo dos enfoques de análisis de asociación (FarmCPU y BLINK) se identificó un conjunto de 593 marcadores SNP's con una asociación altamente significativa (0.01) al hábito de crecimiento. Del conjunto total de marcadores elegidos, el modelo FarmCPU fue el más abundante con respecto al número de marcadores asociados, con un aporte de 518 marcadores SNP's (Figura 1). La mayor presencia de estos marcadores se observa en los cromosomas 2A, 2B, 2D, 3A, 3B y 5A con más de 30 SNP's en cada uno. Por el contrario, el modelo BLINK aportó únicamente 144 marcadores, con una mayor incidencia en los cromosomas 2B y 2D, con 16 SNP's cada uno (Figura 2). Algunos marcadores SNP's fueron coincidentes en los resultados de ambos modelos. En la Tabla 1 se presentan los resultados GWAS para 70 marcadores SNP's conformados en bloques de asociación.

La estructura de la población estimada a través de cinco componentes principales reveló la estrecha base genética de los trigos de tipo invernal con 938 colectas, en comparación con las 6 819 de primavera (Figura 3). El patrón de dispersión puede ser explicado en gran medida por la variabilidad geográfica y genética a la que pertenecen las accesiones bajo estudio. Sin embargo, se puede diferenciar con claridad como las colectas de hábito invernal pertenecen a un conjunto compacto, el cual fue conformado en su mayoría con cultivares y líneas mejoradas.

### Mapeo de marcadores

El mapa físico de los marcadores SNP's asociados con el hábito de crecimiento reveló que 70 de estos formaban BA en 11 cromosomas (Figura 4 y 5). De los bloques

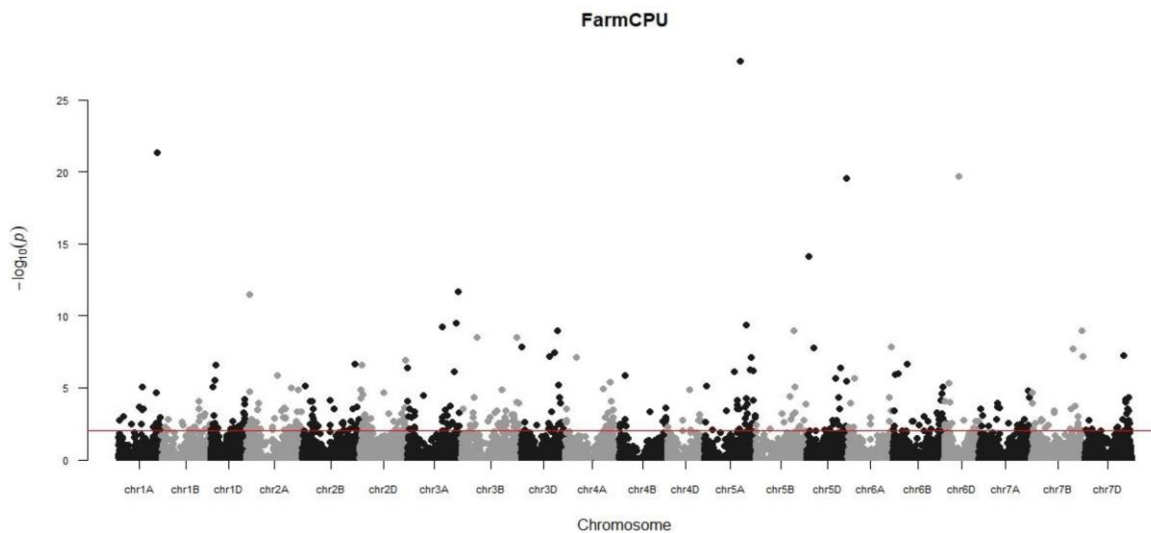
mencionados, dos comparten cercanía genómica con un gen de vernalización (Vrn-A1) ubicado en el cromosoma 5A y otro con un gen de fotoperíodo (Ppd-D1) del cromosoma 2D. También se observaron 3 marcadores SNP's (109087597|F|0-9:A>T-9:A>T - 2A, 09271053|F|0-12:A>G-12:A>G - 5D y 109533531|F|0-29:A>G-29:A>G - 7B) cercanos a la regiones genómicas de los genes Ppd-A1 (fotoperíodo), Vrn-D1 y Vrn-B3 (vernalización).

Los BA con mayor número de marcadores se ubicaron en regiones genómicas presentes en los brazos largos de los cromosomas 1D (468309273-470535989 pb), 2B (748982899-750002913 pb), 2D (615469335-620270385 pb), 4A (658343324-661569447 pb), 4B (649577977-653001877 pb) y 5A (586669360-589287661 pb). El SNP 109372697|F|0-26:C>G-26:C>G (745524854 pb) ubicado en el cromosoma 2A es coincidente con regiones seleccionadas en los BA de los cromosomas 2B y 2D.

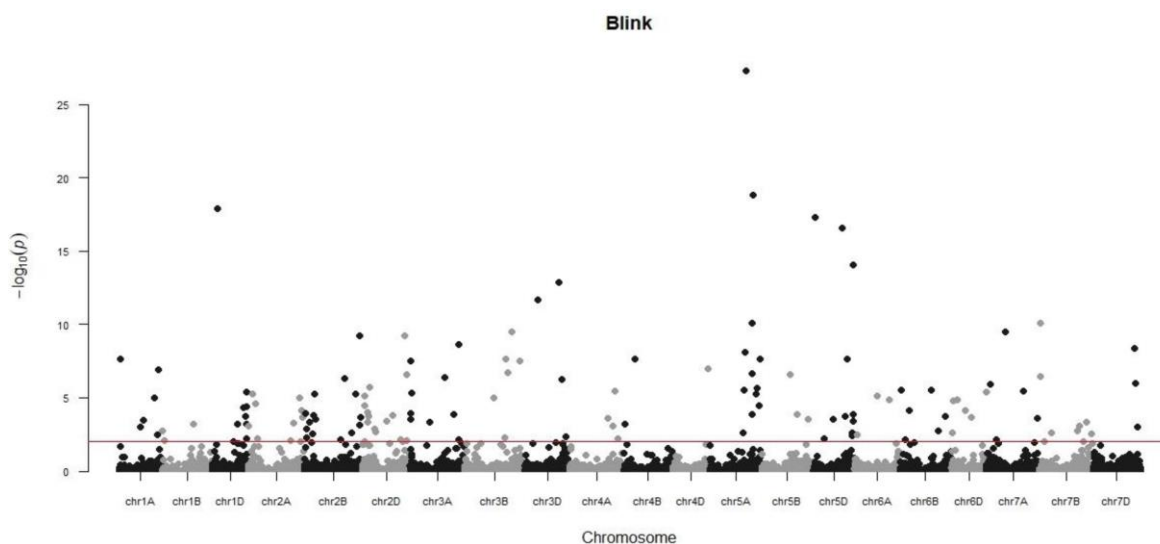
### Análisis bioinformático

Las regiones genómicas conformadas por los BA en los distintos cromosomas fueron exploradas a detalle en las bases de datos de Ensembl Plants y UniProt. La información bioinformática de cada BA muestra 20 regiones genómicas con información relevante de genes y proteínas con posible funcionalidad en la determinación del hábito de crecimiento (Tabla 2). Estos resultados revelan la abundancia de proteínas de dominio quinasas en muchos BA, como Quinasas de serina y treonina (RSTK), Receptores de serina y treonina (RLK), Quinasas dependiente de ciclina N (RSTK), Diacilglicerol quinasa (DGK) y Tiamina pirofosfoquinasa (TPK). También se notaron algunas proteínas de dominio de caja, como Homeobox, F-box, U-box y MADS-box, esta última vinculada con el gen de vernalización Vrn1 del cromosoma 5A. Se reportan dos genes ortólogos de Arabidopsis thaliana (SEP2 y EDA15) coincidentes con BA ubicados en los cromosomas 2B y 5A. En algunos BA se identificaron los genes TthV (1D), DREB2 (2B), GRF4-D1 (2D) y D1-DWARF1 (2D), reportados en bases de datos para *T. aestivum*.





**Figura 1.** Gráfico de Manhattan del modelo FarmCPU que muestra la asociación de 25 681 marcadores SNP's en 7 757 colectas de *T.aestivum* con el hábito de crecimiento.



**Figura 2.** Gráfico de Manhattan del modelo BLINK que muestra la asociación de 25 681 marcadores SNP's en 7 757 colectas de *T.aestivum* con el hábito de crecimiento.

## DISCUSIÓN

Algunas de las asociaciones entre marcadores SNP's y el hábito de crecimiento detectadas por el GWAS implican polimorfismos dentro de genes de vernalización documentados en *T. aestivum*, como es el caso del gen *Vrn1*, cuyas copias homologas se localizan en los cromosomas 5A, 5B y 5D (Li *et al.* 2020).

Este gen posee particular importancia en la diferenciación del hábito de crecimiento, ya que la presencia de variaciones alélicas dominantes causadas por mutaciones en regiones reguladoras confieren el hábito primaveral (Konopatskaia *et al.* 2016). En este sentido, existe congruencia en los resultados obtenidos, ya que se observó un conjunto de marcadores SNP's formando un marcado BA con el gen *Vrn1* en el cromosoma 5A.

**Tabla 1.** Resultados GWAS de marcadores significativos pertenecientes a bloques de asociación con el hábito de crecimiento en 7 757 colectas de *T. aestivum*.

Crom	SNP	Valor <i>P</i>	MAF	Posición Física (pb)	BA
1D	109374796 F 0-9:C>T-9:C>T	1.25E-03	1.04E-02	468309273	1
	108513969 F 0-11:G>A-11:G>A	5.74E-03	1.68E-02	469542771	
	108713509 F 0-41:T>C-41:T>C	6.33E-05	4.16E-02	470220690	
	108147322 F 0-16:A>G-16:A>G	7.07E-04	3.73E-02	470523298	
	109349553 F 0-34:C>G-34:C>G	4.27E-06	3.40E-02	470523363	
109349552 F 0-19:C>A-19:C>A	1.38E-04	3.69E-02	470535921	2	
108256603 F 0-47:T>G-47:T>G	4.09E-03	4.50E-01	58386894		
109340887 F 0-22:C>T-22:C>T	2.01E-03	4.42E-01	59120190		
108620226 F 0-51:G>A-51:G>A	5.18E-06	2.13E-02	59334130		
109372697 F 0-26:C>G-26:C>G	5.55E-03	2.13E-01	745524854		
109118300 F 0-28:T>C-28:T>C	5.79E-03	4.91E-01	150710793	4	
108926834 F 0-68:T>G-68:T>G	1.37E-03	4.91E-01	150985370		
109302969 F 0-60:C>T-60:C>T	4.85E-03	4.70E-01	151895514		
108203258 F 0-9:A>G-9:A>G	5.46E-03	2.28E-01	152503627		
108144780 F 0-66:T>C-66:T>C	3.42E-03	1.60E-01	169861838		
107873878 F 0-33:A>C-33:A>C	2.73E-03	2.53E-01	170086991	5	
108303724 F 0-60:G>C-60:G>C	6.49E-03	2.44E-01	170781731		
108939717 F 0-34:G>T-34:G>T	7.21E-04	4.84E-01	748982898		
109318254 F 0-49:A>G-49:A>G	8.43E-03	1.98E-01	749044179		
109486206 F 0-29:G>C-29:G>C	3.29E-03	1.90E-01	749922381		
109355990 F 0-35:T>C-35:T>C	2.18E-04	1.92E-01	750002860	6	
109136240 F 0-18:C>T-18:C>T	2.91E-05	1.20E-01	30934992		
108547082 F 0-63:A>G-63:A>G	2.95E-03	1.26E-01	31502743		
109067136 F 0-25:T>C-25:T>C	9.70E-03	1.33E-01	31502808		
109232275 F 0-40:A>G-40:A>G	9.75E-03	4.38E-03	545940420		
109545052 F 0-31:C>G-31:C>G	1.99E-03	3.37E-02	547286664	8	
109545052 F 0-28:G>T-28:G>T	9.99E-03	3.71E-02	547559641		
109235294 F 0-10:G>C-10:G>C	2.90E-03	1.93E-01	615469335		
109486206 F 0-24:G>C-24:G>C	3.26E-03	1.87E-01	615503378		
108800261 F 0-24:C>G-24:C>G	2.27E-03	2.24E-01	615943480		
109098245 F 0-58:A>G-58:A>G	6.03E-03	1.97E-01	616526082	9	
108144213 F 0-30:G>A-30:G>A	1.13E-03	2.07E-01	618071961		
109137337 F 0-32:C>A-32:C>A	5.22E-03	1.97E-01	618826852		
109153060 F 0-47:C>A-47:C>A	6.94E-03	2.13E-01	618828092		
109068785 F 0-19:C>A-19:C>A	2.51E-04	2.85E-02	620160840		
108145366 F 0-36:G>A-36:G>A	6.41E-03	8.27E-02	620270317	10	
107697013 F 0-24:A>G-24:A>G	3.60E-03	4.25E-01	123356856		
109241628 F 0-16:A>C-16:A>C	9.99E-03	4.45E-01	124354979		
108085036 F 0-61:T>G-61:T>G	7.68E-03	4.79E-01	124817091		
109256730 F 0-63:C>G-63:C>G	3.77E-03	2.67E-01	795329079		
108713853 F 0-12:G>C-12:G>C	3.18E-03	2.85E-01	795429435	11	
107877816 F 0-46:A>C-46:A>C	1.20E-04	2.87E-01	796008897		
108994529 F 0-6:T>C-6:T>C	1.92E-03	9.39E-02	615181053		
108994530 F 0-37:T>A-37:T>A	3.89E-03	9.08E-02	615181053		
108476242 F 0-11:G>C-11:G>C	8.39E-03	1.30E-02	616934964		
109364118 F 0-8:C>T-8:C>T	4.86E-03	3.68E-01	658343324	13	
109193592 F 0-14:A>T-14:A>T	3.22E-04	4.72E-01	659678760		
108434554 F 0-15:T>G-15:T>G	5.74E-03	4.77E-01	660455556		
107881764 F 0-40:T>G-40:T>G	1.12E-03	3.26E-01	661569379		

Tabla 6. Continuación.

Crom	SNP	Valor P	MAF	Posición Física (pb)	BA
4B	108998096 F 0-43:C>A-43:C>A	1.67E-03	1.41E-01	649577977	14
	109229236 F 0-17:C>T-17:C>T	2.33E-04	5.52E-02	650880267	
	109229237 F 0-49:T>C-49:T>C	1.19E-03	4.15E-02	650883267	
	109222860 F 0-60:C>T-60:C>T	4.10E-03	3.87E-02	653001809	
5A	109187296 F 0-15:T>C-15:T>C	1.61E-04	1.81E-01	461231162	15
	109208559 F 0-9:T>C-9:T>C	2.99E-04	3.47E-01	461899837	
	108587064 F 0-68:C>A-68:C>A	1.25E-04	4.24E-01	581488583	16
	108957420 F 0-65:G>A-65:G>A	5.72E-05	4.87E-01	582636444	
6B	108142772 F 0-28:G>T-28:G>T	4.59E-10	3.91E-01	586669360	17
	108440456 F 0-50:C>G-50:C>G	1.60E-19	2.92E-01	588401400	
	109266488 F 0-41:G>A-41:G>A	6.53E-04	1.56E-01	588761898	
	107702923 F 0-64:G>C-64:G>C	1.40E-04	4.64E-01	589287593	
7B	109077677 F 0-39:G>A-39:G>A	3.42E-03	4.13E-03	644135224	18
	109441349 F 0-64:G>A-64:G>A	3.45E-03	8.12E-03	645483651	
	109137306 F 0-50:T>G-50:T>G	1.55E-03	3.74E-03	645537586	
7D	109078487 F 0-20:G>A-20:G>A	9.70E-04	4.90E-01	640866076	19
	109127303 F 0-32:A>C-32:A>C	2.90E-03	4.75E-01	640905664	
7D	109568141 F 0-23:A>G-23:A>G	3.33E-03	1.31E-01	641806854	20
	108550833 F 0-50:G>T-50:G>T	4.56E-05	7.59E-02	606323492	
	108884325 F 0-11:C>G-11:C>G	3.73E-04	6.83E-02	607228637	
	108201673 F 0-12:G>A-12:G>A	1.89E-03	7.33E-02	607297673	

Crom = Cromosoma; P = valor de p al 0.01 ajustado por tasa de falso descubrimiento (FDR); MAF = frecuencia mínima de alelos. El símbolo E en los valores de P y MAF indica x10n.

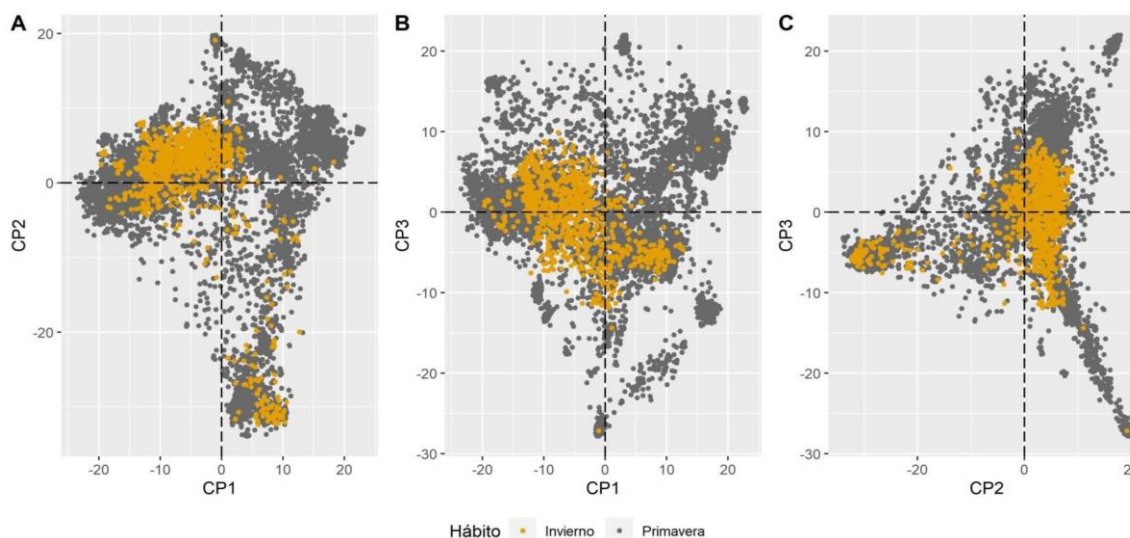


Figura 3. Grafico ACP de 7 757 colectas de *T. aestivum* clasificadas en hábito de crecimiento. A) CP1 vs CP2, B) CP1 vs CP3 y C) CP2 vs CP3.

Los BA revelaron la presencia de proteínas con diversidad de dominios (Tabla 2) entre las que destacan las relacionadas con cambios estructurales y metabólicos de

respuesta al estrés de las plantas. En este sentido Kosová *et al.* (2011) argumentan que los cambios en la acumulación de las proteínas durante un periodo de estrés están estrechamente relacionados con el proceso

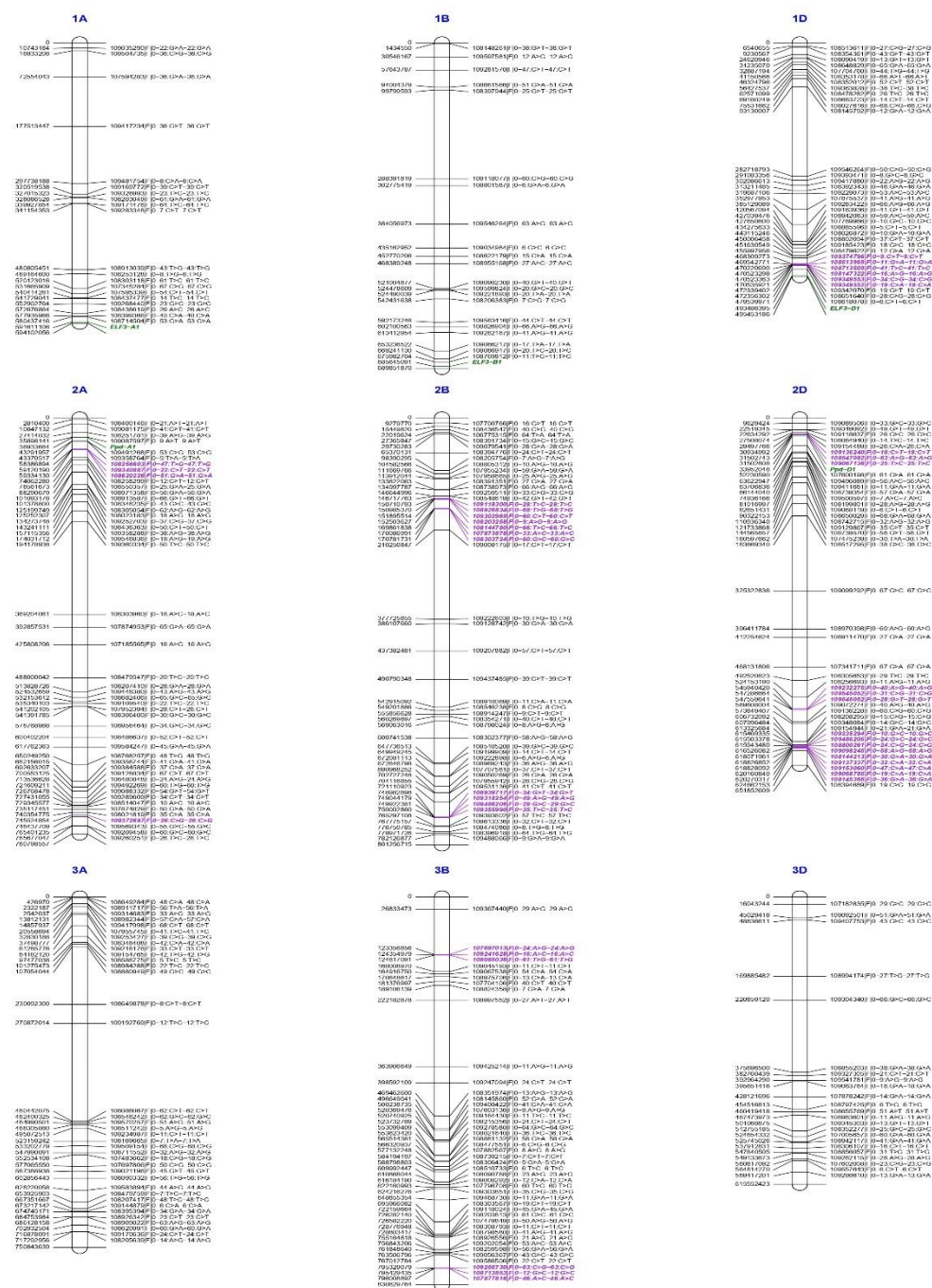


Figura 4. Mapa físico de la ubicación de 593 marcadores SNP's con asociación altamente significativa (0.01) al hábito decrecimiento en 7 757 colectas de *T. aestivum* (cromosomas 1A al 3D).



Figura 5. Mapa físico de la ubicación de 593 marcadores SNPs con asociación altamente significativa (0.01) al hábito de crecimiento en 7 757 colectas de *T. aestivum* (cromosomas 4A al 7D).

Tabla 7. Información bioinformática de bloques de asociación al hábito de crecimiento.

BA	Ensembl Plants Genes Asociados	ID	Dominio	UniProt-Proteína Función Biológica
1	TraesCS1D02G404300	Q56TP4	XTH	Pared celular de tejidos en crecimiento.
	TraesCS1D02G405600	Q05806	TTHV	Respuesta de defensa.
	TraesCS1D02G407600	A0A1D5SRX7	CytB <sub>561</sub>	Respuesta al déficit hídrico y exceso de luz.
	TraesCS1D02G404400	A0A1D6RG09	ARF	Crecimiento y desarrollo celular.
	TraesCS1D02G405100	A0A341PSZ5	F-box	Respuesta de defensa hormonal y estrés abiótico.
2	TraesCS1D02G406900	A0A1D5SWV6	Homeobox	Procesos de desarrollo y respuesta hormonal.
	TraesCS2A02G107000	A0A341PZ27	RSTK	Respuesta a estrés biótico y abiótico.
	TraesCS2A02G106700	A0A1D5TN11	POD	Mecanismos de defensa.
3	TraesCS2A02G106800	A0A341PUV9	TPK	Resistencia a estrés biótico y abiótico.
	TraesCS2A02G525500	W5ASP5	LOX	Crecimiento y desarrollo, estrés, senescencia, etc.
4	TraesCS2B02G176000	A0A1D5UCP9	DIR	Respuesta a sequía, salinidad y estrés oxidativo.
	TraesCS2B02G175800	A0A1D5UFD5	RSTK	Respuesta a estrés biótico y abiótico.
	TraesCS2B02G176200	A0A080YU40	UGT	Desarrollo y crecimiento, resistencia a enfermedades.
	TraesCS2B02G177100	A0A341PYM4	NA	Respuesta al frío, déficit de agua, estrés salino, etc.
5	TraesCS2B02G193500	W5BPC0	RING-E3	Respuesta a estrés abiótico.
	TraesCS2B02G193600	W5BPN4	ALT1,2	Respuesta hipoxia y anoxia.
	TraesCS2B02G193700	A0A1D5U885	F-box	Respuesta de defensa hormonal y estrés abiótico.
6	TraesCS2B02G553600	A0A1D5TZL6	SEP2	Respuesta al estrés lumínico (K-box y MADS-box).
	TraesCS2B02G554500	G3FIPO	DREB2	Respuesta a la deshidratación.
	TraesCS2B02G554600	A0A341QXE7	Hidrotropismo	Reacción al estímulo de agua o humedad.
	TraesCS2B02G554700	A0A341R834	RLK	Respuesta de defensa.
7	TraesCS2D02G073600	A0A341R4Y5	GLP	Respuestas al estrés abiótico, exposición al calor, la sal y el aluminio.
	TraesCS2D02G435200	A0A1D5UNL9	GRF4-D1	Regulación del crecimiento.
8	TraesCS2D02G435300	A0A2X0SA68	D1-DWARF1	Respuesta celular al frío y ácido abscísico
	TraesCS2D02G528400	A0A1D5UR29	LOX	Crecimiento y desarrollo, estrés, senescencia, etc.
9	TraesCS2D02G534400	A0A1D5UH93	RSTK	Respuesta a estrés biótico y abiótico.
	TraesCS2D02G534600	A0A341RJ78	DGK	Respuesta al estrés por congelación.
	TraesCS3B02G137800	A0A341SBA6	Tudor	Tolerancia al estrés salino.
10	TraesCS3B02G138400	A0A1D5VQ05	Homeobox	Procesos de desarrollo y respuesta hormonal.
	TraesCS3B02G562700	A0A077S8R7	β-1,3-glucanos	División celular, resistencia a estrés abiótico.
12	TraesCS4A02G330400	A0A341TN72	EF-Tu	Aclimatación al calor.
	TraesCS4A02G331000	W5DR31	F-box	Respuesta de defensa hormonal y estrés abiótico.
	TraesCS4A02G381100	A0A341TWL3	RSTK	Respuesta a estrés biótico y abiótico.
13	TraesCS4A02G382200	W5DQV3	Homeobox	Procesos de desarrollo y respuesta hormonal.
	TraesCS4A02G382600	A0A341TWM7	F-box	Tolerancia a estrés biótico y abiótico.
	TraesCS4A02G382700	A0A341TUP2	E3-U	Respuesta a estrés abiótico.
	TraesCS4B02G361500	A0A1D5XNM3	NA	Respuesta de defensa.
14	TraesCS4B02G364500	A0A1D5XKM0	UGT	Desarrollo y crecimiento, resistencia a enfermedades.
	TraesCS4B02G363400	A0A1D5XHF3	NA	Ciclo celular (Meiosis), microsporogénesis, etc.
	TraesCS5A02G247000	A0A1D5YFJ0	RSTK	Ciclo celular (Mitosis), división celular.
15	TraesCS5A02G383800	A0A1D5YRN0	HSF	Respuesta celular al calor.
	TraesCS5A02G384500	A0A1D5YDWO	ARN Helicasa	Respuesta estrés salino, déficit de agua, etc.
17	TraesCS5A02G391700	Q5ETV1	MADS-box WM6	Vernalización (Vrn-A1)
18	TraesCS6B02G370800	A0A1D6ANZ6	U-box	Respuesta estrés abiótico.
	TraesCS6B02G368300	A0A1D6ATD5	XTH	Biogénesis de la pared celular.
19	TraesCS7B02G375600	A0A1D6C5B8	POD	Respuesta estrés oxidativo.
	TraesCS7B02G376400	A0A1D6C2E6	Kinase	Respuesta a estrés biótico y abiótico.
	TraesCS7D02G500300	A0A341Z6N1	RSTK	Respuesta a estrés biótico y abiótico.

Crom = Cromosoma; BA = Bloque de asociación; XTH = Xiloglucano endoTransglucosilasa-Hidrolasa; TTHV = Tionina tipo 5; CytB<sub>561</sub> = CitocromoB<sub>561</sub>; ARF = Factores de respuesta a auxinas; RSTK = Quinasas de serina/treonina; POD = Peroxidasas; TPK = Tiamina pirofosfoquinasa; LOX = Lipogénasas; DIR = Proteína dirigente; UGT = UDP-glicosiltransferasa; RING-E3 = Ligasas RING tipo E3; ALT1,2 = Alanina aminotransferasa 1 y 2; SEP2 = ortólogo del gen SEPALLATA 2 de *A. thaliana*; DREB2 = Proteína de unión a elementos sensibles a la deshidratación; RLK = Receptor de serina/treonina; GLP = Germina; GRF4-D1 = Factor de regulación de crecimiento D1; DGK = Diacilglicerol quinasa; β-1,3-glucanos = Glucano endo-1,3-β-D-glucosidasa; EF-Tu = Factores de elongación Tu; E3-U = Ligasas de ubiquitina E3; HSF = Factor de transcripción de choque térmico; ARN Helicasa = ortólogo del gen de EDA15 de *A. thaliana*.

de aclimatación. Estos resultados son similares a los encontrados por Kosová *et al.* (2013) quienes observaron la expresión de proteínas de regulación de respuesta al estrés y del desarrollo, además de proteínas de defensa y proteínas involucradas en la restauración de la división celular durante el proceso de aclimatación al frío en trigos de primavera e invierno. Otro estudio realizado por Wójcik-Jagła *et al.* (2021) también mostró la expresión de proteínas de respuesta al estrés, además de proteínas de choque térmico relacionadas con la tolerancia a la congelación y proteínas implicadas en funciones estructurales al estudiar la respuesta a la desaclimatación en cebada. Particularmente, notaron una sobreexpresión de enzimas Peroxidasas en variedades susceptibles, lo que muestra la importancia de dichas proteínas en los mecanismos de respuesta a la desaclimatación al frío.

En este sentido, las proteínas quinasas en sus distintas formas han sido reportadas como elementos clave en la traducción de señales que responden a diversos tipos de estrés en las plantas (Calliste *et al.* 2008, Sinha *et al.* 2011). Un ejemplo de ello, son las quinasas RLK implicadas en la señalización durante el reconocimiento de patógenos y la posterior activación de los mecanismos de defensa (Afzal *et al.* 2008). Otras quinasas denominadas RSTK intervienen en la regulación del desarrollo y las vías de señalización de autoincompatibilidad, así como en procesos de resistencia a enfermedades (Goring y Walker 2004).

En trigo, el gen WPK4 que codifica una proteína RSTK es reconocido por desempeñar funciones de control en el metabolismo del nitrógeno y la asimilación del carbono durante períodos de baja temperatura (Ikeda *et al.* 1999). Por otro lado, el estudio de genes homólogos de WPK4 en maíz y arroz sugieren la expresión diferencial del gen en las distintas especies en respuesta a estímulos ambientales, sobre todo en la exposición a bajas temperaturas (Ohba *et al.* 2000). De igual manera, las proteínas de caja F desempeñan funciones importantes en las respuestas a tensiones bióticas a través de la vía de la ubiquitina (Yee y Goring 2009, Zhou *et al.* 2015). También se ha notado su

expresión en procesos cruciales en las plantas, como la embriogénesis, las respuestas hormonales, el desarrollo de las plántulas, la organogénesis floral y la senescencia (Xu *et al.* 2009). En trigo, algunos genes de caja se asocian con la tolerancia a estrés de tipo oxidativo, hídrico y salino (Zhou *et al.* 2014, Wang *et al.* 2020). Por su parte, los genes homeobox son señalados por intervenir en el desarrollo reproductivo y en la señalización durante estrés abiótico (Tan e Irish 2006, Jain *et al.* 2008). Mientras tanto, a las enzimas Glicosiltransferasas (UGT) se les atribuye procesos de regulación hormonal y de biosíntesis de compuestos secundarios implicados en el crecimiento, desarrollo y resistencia a enfermedades en algodón (He *et al.* 2018).

Algunos marcadores significativos situados en los brazos largos de los cromosomas 2A, 2B y 2D evidencian regiones genómicas coincidentes. Particularmente en los cromosomas 2B y 2D donde se pueden observar BA conformados por varios marcadores. Dichas regiones evidencian la presencia de proteínas Lipogenasas (LOX), las cuales desempeñan un papel fundamental en procesos fisiológicos de las plantas, como la germinación, crecimiento, desarrollo y respuestas al estrés por calor y salinidad (Shaban *et al.* 2018, Viswanath *et al.* 2020). Por otro lado, el BA situado en el cromosoma 2B expresó algunos genes documentados en bases de datos, el primero es un ortólogo del gen SEP2 reportado en *Arabidopsis thaliana*, el cual es un gen perteneciente a la familia de reguladores transcripcionales de caja MADS de tipo II involucrado en la diferenciación y regulación floral (Moore *et al.* 2005). El otro gen es DREB2, cuya expresión es inducida por el frío y la deshidratación (Akbudak *et al.* 2018).

Dado que la mayoría de las proteínas expresadas en los genes pertenecientes a los BA formados por los marcadores significativos apuntan a la expresión de factores de respuesta a diversos tipos de estrés, los cuales dependen en gran medida de factores exógenos y endógenos como la temperatura del ambiente, la disponibilidad de agua, el fotoperíodo, entre otros, se puede denotar su inferencia en la determinación del hábito de crecimiento, que como bien se menciona ha sido atribuida

principalmente a mutaciones genéticas del gen de vernalización *Vrn1*. Lo anterior, reafirma las teorías de algunos autores (Saisho *et al.* 2011, Chen y Dubcovsky 2012, Shcherban y Salina 2017) sobre la posible acción de genes desconocidos en la expresión de *Vrn1* y en el complejo genético regulador de la floración.

### CONCLUSIONES

Se proporciona información adicional para comprender las funciones potenciales de nuevos genes en la determinación del hábito de crecimiento en *Triticum aestivum*. Los resultados identifican 20 regiones genómicas ubicadas en 11 cromosomas con posibles genes candidatos, cuyo estudio y caracterización puede ser base en la predicción genómica del hábito de crecimiento en colectas no clasificadas.

La presencia de un abundante bloque de asociación cercano al gen de vernalización *Vrn1*, que es sumamente importante en la determinación del hábito de crecimiento, es un indicador de la confiabilidad de los métodos utilizados en este estudio.

### AGRADECIMIENTOS

A la Universidad Autónoma Agraria Antonio Narro por el apoyo financiero para el desarrollo de este proyecto. Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por su apoyo a través de una beca de doctorado. Agradecemos al CIMMYT a través de su iniciativa *Seeds of Discovery* por generar y compartir los datos genotípicos de trigo utilizados en la presente investigación.

### LITERATURA CITADA

- Afzal AJ, Wood AJ, Lightfoot DA (2008) Plant receptor-like serine threonine kinases: Roles in signaling and plant defense. *Molecular Plant Microbe Interactions* 21: 507-517.
- Akbudak MA, Filiz E, Kontbay K (2018) DREB2 (dehydration-responsive element-binding protein 2) type transcription factor in sorghum (*Sorghum bicolor*): genome-wide identification, characterization and expression profiles under cadmium and salt stresses. *3 Biotech* 8: Article number 426. DOI: 10.1007/s13205-018-1454-1.
- Calliste DJ, Popova OV, Dietz KJ, Gollack D (2008) The SNF1-type serine-threonine protein kinase SAPK4 regulates stress-responsive gene expression in rice. *BMC Plant Biology* 8: Article number 49. DOI: 10.1186/1471-2229-8-49.
- Chen A, Dubcovsky J (2012) Wheat TILLING mutants show that the vernalization gene VRN1 down-regulates the flowering repressor VRN2 in leaves but is not essential for flowering. *PLoS Genet* 8: e1003134. DOI: 10.1371/journal.pgen.1003134.
- Danyluk J, Kane NA, Breton G, Limin AE, Fowler DB, Sarhan F (2003) TaVRT-1, a putative transcription factor associated with vegetative to reproductive transition in cereals. *Plant Physiology* 132: 1849-1860.
- Dixon LE, Karsai I, Kiss T, Adamski NM, Liu Z, Ding Y, Allard V, Boden SA, Griffiths S (2019) VERNALIZATION1 controls developmental responses of winter wheat under high ambient temperatures. *Development* 146: dev.172684. DOI: 10.1242/dev.172684.
- Goncharov NP (2003) Genetics of growth habit (spring vs winter) in common wheat: Confirmation of the existence of dominant gene *Vrn4*. *Theoretical and Applied Genetics* 107: 768-772.



- Goring DR, Walker JC (2004) Self-rejection - A new kinase connection. *Science* 303: 1474-1475.  
Ensembl Plants (2021) [http://plants.ensembl.org/Triticum\\_aestivum](http://plants.ensembl.org/Triticum_aestivum). Data consulted: January 16, 2021.
- Ensembl Plants (2021) [http://plants.ensembl.org/Triticum\\_aestivum](http://plants.ensembl.org/Triticum_aestivum). Data consulted: January 16, 2021.
- He Y, Ahmad D, Zhang X, Zhang Y, Wu L, Jiang P, Ma H (2018) Genome-wide analysis of family-1 UDP glyco- syltransferases (UGT) and identification of UGT genes for FHB resistance in wheat (*Triticum aestivum* L.). *BMC Plant Biology* 18: Article number 67. DOI: 10.1186/s12870-018-1286-5
- Huang M, Liu X, Zhou Y, Summers RM, Zhang Z (2018) BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *GigaScience* 8: giy154. DOI: 10.1093/gigascience/giy154
- Ikeda Y, Koizumi N, Kusano T, Sano H (1999) Sucrose and cytokinin modulation of WPK4, a gene encoding a SNF1-related protein kinase from wheat. *Plant Physiology* 121: 813-820.
- IWGSC (2018) International Wheat Genome Sequencing Consortium. *Science* 361: eaar7191. DOI: 10.1126/ science.aar7191
- Jain M, Tyagi AK, Khurana JP (2008) Genome-wide identification, classification, evolutionary expansion and expression analyses of homeobox genes in rice. *FEBS Journal* 275: 2845-2861.
- Khumalo TP, Barnard A, Maphosa L, Tsilo TJ (2017) Impact of growth habit and architecture genes on adaptation and performance of bread wheat. In: Wanyera R (Ed) *Wheat Improvement, Management and Utilization*. IntechOpen. United Kingdom. pp: 77-95.
- Kippes N, Debernardi JM, Vasquez-Gross HA, Akpinar BA, Budak H, Kato K, Chao S, Akhunov E, Dubcovsky J (2015) Identification of the VERNALIZATION 4 gene reveals the origin of spring growth habit in ancient wheats from South Asia. *Proceedings of the National Academy of Sciences of the United States of America* 112: E5401-E5410. DOI: 10.1073/pnas.1514883112.
- Konopatskaia I, Vavilova V, Kondratenko EY, Blinov A, Goncharov NP (2016) VRN1 genes variability in tetraploid wheat species with a spring growth habit. *BMC Plant Biology* 16: Article number 244. DOI: 10.1186/s12870- 016-0924-z
- Kosová K, Vítámvás P., Prášil IT, Renaut J (2011) Plant proteome changes under abiotic stress - Contribution of proteomics studies to understanding plant stress response. *Journal of Proteomics* 74: 1301-1322.
- Kosová K, Vítámvás P, Planchon S, Renaut J, Vanková R, Prášil IT (2013) Proteome analysis of cold response in spring and winter wheat (*Triticum aestivum*) crowns reveals similarities in stress adaptation and differences in regulatory processes between the growth habits. *Journal of Proteome Research* 12: 4830-4845.
- Li Y, Xiong H, Guo H, Zhou C, Xie Y, Zhao L, Gu J, Zhao S, Ding Y, Liu L (2020) Identification of the vernalization gene VRN-B1 responsible for heading date variation by QTL mapping using a RIL population in wheat. *BMC Plant Biology* 20: Article number 331. DOI: 10.1186/s12870-020-02539-5.

- Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z (2012) GAPIT: Genome association and prediction integrated tool. *Bioinformatics* 28: 2397-2399.
- Liu X, Huang M, Fan B, Buckler ES, Zhang Z (2016) Iterative usage of fixed and random effect models for powerful and efficient genome - wide association studies. *PLoS Genetics* 12: e1005767. DOI: 10.1186/1471-2156-13-100.
- Moore RC, Grant SR, Purugganan MD (2005) Molecular population genetics of redundant floral-regulatory genes in *Arabidopsis thaliana*. *Molecular Biology and Evolution* 22: 91-103.
- Nazim Ud Dowla MAN, Edwards I, O'Hara G, Islam S, Ma W (2018) Developing wheat for improved yield and adaptation under a changing climate: Optimization of a few key genes. *Engineering* 4: 514-522.
- Ohba H, Steward N, Kawasaki S, Berberich T, Ikeda Y, Koizumi N, Kusano T, Sano H (2000) Diverse response of rice and maize genes encoding homologs of WPK4, an SNF1-related protein kinase from wheat, to light, nutrients, low temperature and cytokinins. *Molecular and General Genetics* 263: 359-366.
- Ouellette LA, Reid RW, Blanchard SG, Brouwer CR (2018) LinkageMapView-rendering high-resolution linkage and QTL maps. *Bioinformatics* 34: 306-307.
- Saisho D, Ishii M, Hori K, Sato K (2011) Natural variation of barley vernalization requirements: Implication of quantitative variation of winter growth habit as an adaptive trait in East Asia. *Plant and Cell Physiology* 52: 775-784.
- Shaban M, Ahmed MM, Sun H, Ullah A, Zhu L (2018) Genome-wide identification of lipoxygenase gene family in cotton and functional characterization in response to abiotic stresses. *BMC Genomics* 19: Article number 599. DOI: 10.1186/s12864-018-4985-2.
- Shcherban AB, Salina EA (2017) Evolution of VRN-1 homoeologous loci in allopolyploids of Triticum and their diploid precursors. *BMC Plant Biology* 17: Article number 188. DOI: 10.1186/s12870-017-1129-9.
- Sinha AK, Jaggi M, Raghuram B, Tuteja N (2011) Mitogen-activated protein kinase signaling in plants under abiotic stress. *Plant Signaling and Behavior* 6: 196-203.
- Tan QKG, Irish VF (2006) The *Arabidopsis* zinc finger-homeodomain genes encode proteins with unique bio- chemical properties that are coordinately expressed during floral development. *Plant Physiology* 140: 1095- 1108.
- Trevaskis B, Bagnall DJ, Ellis MH, Peacock WJ, Dennis ES (2003) MADS box genes control vernalization-induced flowering in cereals. *Proceedings of the National Academy of Sciences of the United States of America* 100: 13099-13104.
- Trevaskis B, Hemming MN, Dennis ES, Peacock WJ (2007) The molecular basis of vernalization-induced flowering in cereals. *Trends in Plant Science* 12: 352-357.
- Trevaskis B (2015) Wheat gene for all seasons. *Proceedings of the National Academy of Sciences of the United States of America* 112: 11991-11992.

- UniProt (2021) The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. D480-D489 *Nucleic Acids Research*. 49. <https://www.uniprot.org>. Data consulted: January 22, 2021. DOI: 10.1093/nar/gkaa1100.
- Viswanath KK, Varakumar P, Pamuru RR, Basha SJ, Mehta S, Rao AD (2020) Plant lipoxygenases and their role in plant physiology. *Journal of Plant Biology* 63: 83-95.
- Wang W, Wang W, Wu Y, Li Q, Zhang G, Shi R, Yang J, Wang Y, Wang W (2020) The involvement of wheat U-box E3 ubiquitin ligase TaPUB1 in salt stress tolerance. *Journal of Integrative Plant Biology* 62: 631-651.
- Whittal A, Kaviani M, Graf R, Humphreys G, Navabi A (2018) Allelic variation of vernalization and photoperiod response genes in a diverse set of North American high latitude winter wheat genotypes. *PLoS ONE* 13: e0203068. DOI: 10.1371/journal.pone.0203068.
- Wickham H (2009) *ggplot2: Elegant graphics for data analysis*. 2th Edition. Vol. 35. Springer. New York, USA. 217p.
- Wickham H (2020) *rvest: Easily Harvest (Scrape) Web Pages*. R package version 0.3.6. <https://CRAN.R-project.org/package=rvest>. Data consulted: January 22, 2021.
- Wójcik-Jagła M, Daszkowska-Golec A, Fiust A, Kopec P, Rapacz M (2021) Identification of the genetic basis of response to de-acclimation in winter barley. *International Journal of Molecular Sciences* 22: Article number 1057. DOI: 10.3390/ijms22031057.
- Xu G, Ma H, Nei M, Kong H (2009) Evolution of F-box genes in plants: Different modes of sequence divergence and their relationships with functional diversification. *Proceedings of the National Academy of Sciences of the United States of America* 106: 835-840.
- Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J (2004) The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. *Science* 303: 1640-1644.
- Yan L, Fu D, Li C, Blechl A, Tranquilli G, Bonafede M, Sanchez A, Valarik M, Yasuda S, Dubcovsky J (2006) The wheat and barley vernalization gene VRN3 is an orthologue of FT. *Proceedings of the National Academy of Sciences of the United States of America* 103: 19581-19586.
- Yee D, Goring DR (2009) The diversity of plant U-box E3 ubiquitin ligases: From upstream activators to down-stream target substrates. *Journal of Experimental Botany* 60: 1109-1121.
- Zhang XK, Xiao YG, Zhang Y, Xia XC, Dubcovsky J, He ZH (2008) Allelic variation at the vernalization genes Vrn-A1, Vrn-B1, Vrn-D1, and Vrn-B3 in Chinese wheat cultivars and their association with growth habit. *Crop Science* 48: 458-470.
- Zhou SM, Sun XD, Yin S, Kong X, Zhou S, Xu Y, Luo Y, Wang W (2014) The role of the F-box gene TaFBA1 from wheat (*Triticum aestivum* L.) in drought tolerance. *Plant Physiology and Biochemistry* 84: 213-223.
- Zhou SM, Kong XZ, Kang HH, Sun XD, Wang W (2015) The involvement of wheat F-box protein gene TaFBA1 in the oxidative stress tolerance of plants. *PLoS ONE* 10: e0122117. DOI: 10.1371/journal.pone.0122117

## SEGUNDO ARTÍCULO

**Worldwide selection footprints for drought and heat in bread wheat (*Triticum aestivum* L.)**

# Worldwide selection footprints for drought and heat in bread wheat (*Triticum aestivum* L.)

Ana L. Gómez-Espejo<sup>1</sup>, Carolina P. Sansaloni<sup>2</sup>, Juan Burgueño<sup>2</sup>, Fernando H. Toledo<sup>2</sup>, Adalberto Benavides-Mendoza<sup>1</sup>, M. Humberto Reyes-Valdés<sup>1\*</sup>

<sup>1</sup> Programa de Doctorado en Recursos Fitogenéticos para Zonas Áridas, Universidad Autónoma Agraria Antonio Narro (UAAAN), Saltillo 25315, Coahuila, México; [algomez@innovakglobal.com](mailto:algomez@innovakglobal.com) / [gomez.ana.mvz@gmail.com](mailto:gomez.ana.mvz@gmail.com) (A.L.G.E.); [adalberto.benavides@uaaan.edu.mx](mailto:adalberto.benavides@uaaan.edu.mx) (A.B.-M.)

<sup>2</sup> International Maize and Wheat Improvement Center (CIMMYT), Texcoco56237, México; [C.Sansaloni@cgiar.org](mailto:C.Sansaloni@cgiar.org) (C.P.S); [J.Burgueno@cgiar.org](mailto:J.Burgueno@cgiar.org) (J.B.); [F.TOLEDO@cgiar.org](mailto:F.TOLEDO@cgiar.org) (F.H.T.)

\* Correspondence: [manuel.reyes@uaaan.edu.mx](mailto:manuel.reyes@uaaan.edu.mx) (M.H.R.-V).

**Abstract:** Genome-environment associations (GEA) or environmental genome-wide association scans (EnvGWAS) have been poorly applied for studying the genomics of adaptive traits in bread wheat landraces (*Triticum aestivum* L.). We analyzed 990 landraces and seven climatic variables (mean temperature, maximum temperature, precipitation, precipitation seasonality, heat index of mean temperature, heat index of maximum temperature, and drought index) in GEA using the FarmCPU approach with GAPIT. Historical temperature and precipitation values were obtained as monthly averages from 1970 to 2000. Based on 26,064 high-quality SNP loci, landraces were classified into ten subpopulations exhibiting high genetic differentiation. The GEA identified 59 SNPs, and near 89 protein-encoding genes involved in response processes to abiotic stress. Genes related to biosynthesis and signaling are mainly mediated by *auxins*, *abscisic acid* (ABA), *ethylene* (ET), *salicylic acid* (SA), and *jasmonates* (JA), which are known to operate together in modulation responses to heat stress and drought in plants. In addition, we identified some proteins associated with the response and tolerance to stress by high temperatures, water deficit, and cell wall functions. The results provide candidate regions for selection aimed to improve drought and heat tolerance in bread wheat and provide insights into the genetic mechanisms involved in adaptation to extreme environments.

**Keywords:** *Triticum aestivum* L.; landraces; adaptation; drought stress; heat stress; Genome-environment Associations (GEA).

## 1. Introduction

Since its domestication more than 10,000 years ago, common wheat has experienced a series of selective events caused by humans and the environment, contributing to the increase in its genetic diversification [1]. Climate change has severely reduced wheat production in recent years due to extreme temperature episodes and unpredictable precipitation patterns [2,3]. Simulation models predict losses of more than 20% in agricultural production by 2,050 [4].

There is an urgent need to discover new sources of adaptation to drought and heat that contribute to maintaining crop productivity [2]. To address this scenario, landraces are valuable because they have developed survival mechanisms for challenging environments through natural and human selection [5,6]. This is why they preserve loci of adaptation to climate change in their places of origin [7].

Recent advances in sequencing technologies have allowed the exploration of entire genomes in various species with increasingly dense *single nucleotide polymorphism* (SNP) data that identify selective events [8]. Likewise, novel bioinformatic analysis approaches, such as genome-wide

association studies (GWAS), are very efficient in time, cost, and precision for identifying genes that control important agricultural traits [9].

Genome-Environment Associations (GEA) or Environmental Genome-Wide Association scans (EnvGWAS) have been used successfully for studying adaptive traits in local populations. They consist of associating SNPs distributed throughout the genome with environmental variables of the accession sampling sites [10].

Associations between the genome and environment of origin were initially documented in wild populations, with successful identification of adaptive loci and prediction of phenotypic variation [11-14]. However, its application to crops is recent, with research carried out on sorghum through bioclimatic and soil gradients to predict adaptive traits [15]. Subsequent applications have been made in crops such as corn [16], beans [10,17], barley [7,18], soybean [19], tomato [20], chickpea [21], peach [22] and wheat [23].

In the last decade, significant progress has been reported in the characterization of wheat genomes through high-throughput genotyping with DArT-seq technology. More than 100,000 accessions belonging to the germplasm bank of the International Maize and Wheat Improvement Center (CIMMYT) have been characterized through the Seeds of Discovery initiative [24,25].

Considering the effectiveness of GWAS for the identification of genomic regions associated with traits of agronomic importance, as well as the functional genetic variation to adapt crops to climate change, this research aims to identify genomic regions related to the adaptation process to arid climates through genome-environment association studies in the *Triticum aestivum* collection maintained in the CIMMYT germplasm bank.

## 2. Results

### 2.1. Exploratory analysis

As expected, the highly significantly ( $p \leq 0.001$ ) and correlated variables share with each other temperature in their definition (Table 1). For instance (Table 1a), AMT was consistently significantly and positively correlated with the variables MaxTWM ( $r = 0.79$ ), MeanTDQ ( $r = 0.71$ ), and MeanTWQ ( $r = 0.77$ ), similar to MaxTWM with MeanTWQ ( $r = 0.94$ ). For those based on precipitation, only PDM with PDQ ( $r = 0.99$ ) exhibited a significant correlation. Similarly, Table 1b shows positive correlations between the variables associated with temperature, such as the case of AMT with the variables MaxT ( $r = 0.98$ ), HITmead ( $r = 0.92$ ), and HITmax ( $r = 0.93$ ), MaxT with HITmead ( $r = 0.92$ ) and HITmax ( $r = 0.96$ ), and HITmead with HITmax ( $r = 0.97$ ). On the other hand, negative correlations were observed only for MeanTDQ with PWQ ( $r = -0.84$ ) and AP with DI ( $r = -0.83$ ), which reflects contrasting trends between temperature and precipitation.

In the PCA, we observed that the evaluated variables had contrasting contributions to the total variation of each component, especially for the second set of variables (Figure S1 and S2). In the biplot graphs of the PCA (Figure 1), we identified that the accessions sites are well differentiated with respect to their Köppen-Geiger climate, revealing a greater representation of collections from temperate (C) and cold (D) groups. For the first set (Figure 1a), those related to temperature (AMT, MaxTWM, MeanTDQ, and MeanTWQ) are oriented along PC1, while the precipitation-derived variables (PDM, PS, and PDQ) predominate in PC2. Similarly, for the second set (Figure 1b), PC1 includes the temperature-related variables (AMT, HITmead, MaxT, and HITmax), while PC2 includes AP and DI. The PCA has helpfully identified some variables with discriminating potential of accessions according to the Köppen-Geiger climate groups (Figure 1b). For example, the high DI sites are related to dry climates (B), whereas temperature-related variables (HITmax, HITmead, AMT, and MaxT) define gradients from tropical (A) to temperate (C) groups. On the other hand, we identify the variables of most significant importance concerning the monitoring, follow-up, and informativeness of drought and heat stress events and their contribution to the total variation.

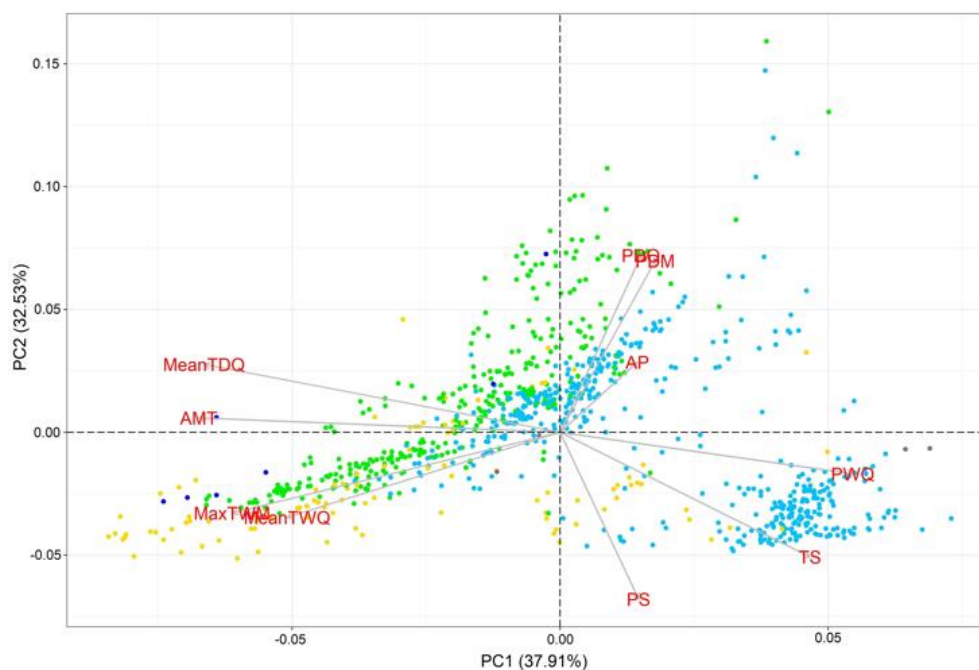
**Table 1.** Above the diagonal, correlations between bioclimatic variables, indices, and elevation. Below the diagonals,  $p$ -values for the significance test.

**(a) Correlation table for bioclimatic variables related to aridity.**

Variable	AMT	TS	MaxTWM	MeanTDQ	MeanTWQ	AP	PDM	PS	PDQ	PWQ
AMT		-0.65	0.79	0.71	0.77	0.07	-0.03	-0.08	0.01	-0.4
TS	0.000		-0.14	-0.76	-0.03	-0.13	-0.3	0.53	-0.35	0.56
MaxTWM	0.000	0.000		0.5	0.94	-0.17	-0.38	0.21	-0.38	-0.34
MeanTDQ	0.000	0.000	0.000		0.32	-0.08	-0.01	-0.5	0.03	-0.84
MeanTWQ	0.000	0.374	0.000	0.000		-0.02	-0.29	0.33	-0.29	-0.09
AP	0.023	0.000	0.000	0.009	0.552		0.41	0.01	0.44	0.38
PDM	0.408	0.000	0.000	0.645	0.000	0.000		-0.59	0.99	0.19
PS	0.012	0.000	0.000	0.000	0.000	0.814	0.000		-0.62	0.49
PDQ	0.784	0.000	0.000	0.357	0.000	0.000	0.000	0.000		0.16
PWQ	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.000	

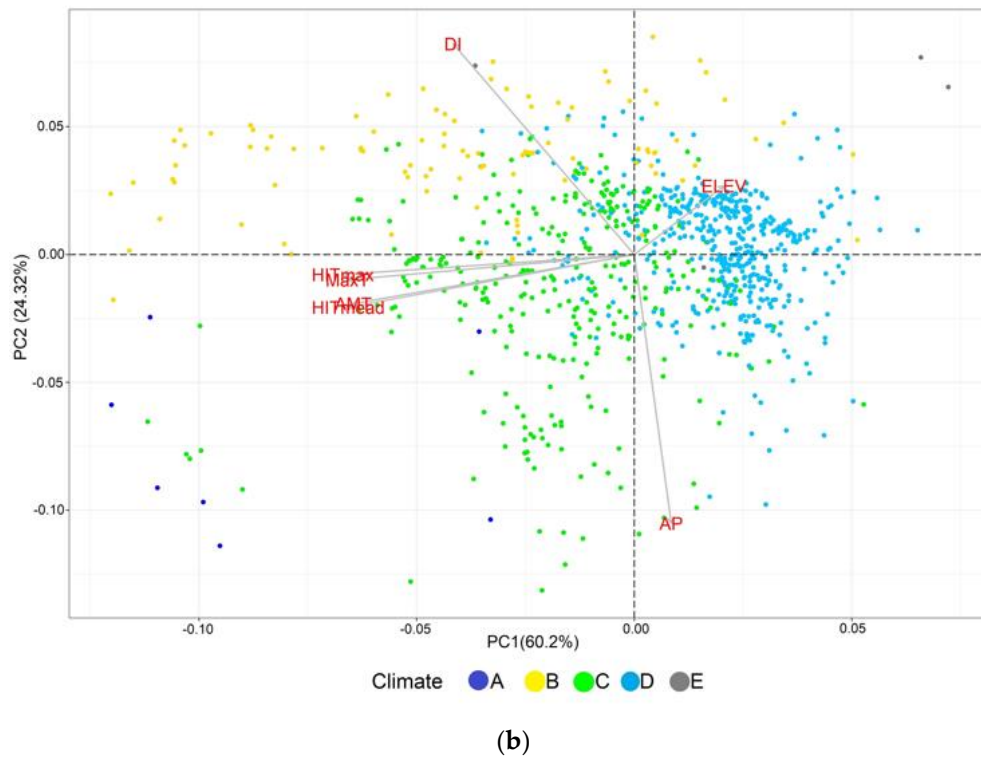
**(b) Correlation table for precipitation, temperature-related variables, calculated indices, and elevation.**

Variable	AMT	MaxT	HITmead	HITmax	AP	DI	ELEV
AMT		0.98	0.92	0.93	0.07	0.42	-0.16
MaxT	0.000		0.92	0.96	-0.02	0.49	-0.14
HITmead	0.000	0.000		0.97	0.06	0.45	-0.38
HITmax	0.000	0.000	0.000		-0.05	0.53	-0.27
AP	0.023	0.602	0.042	0.117		-0.83	-0.13
DI	0.000	0.000	0.000	0.000	0.000		-0.11
ELEV	0.000	0.000	0.000	0.000	0.000	0.000	



**(a)**

Figure 1. Cont.

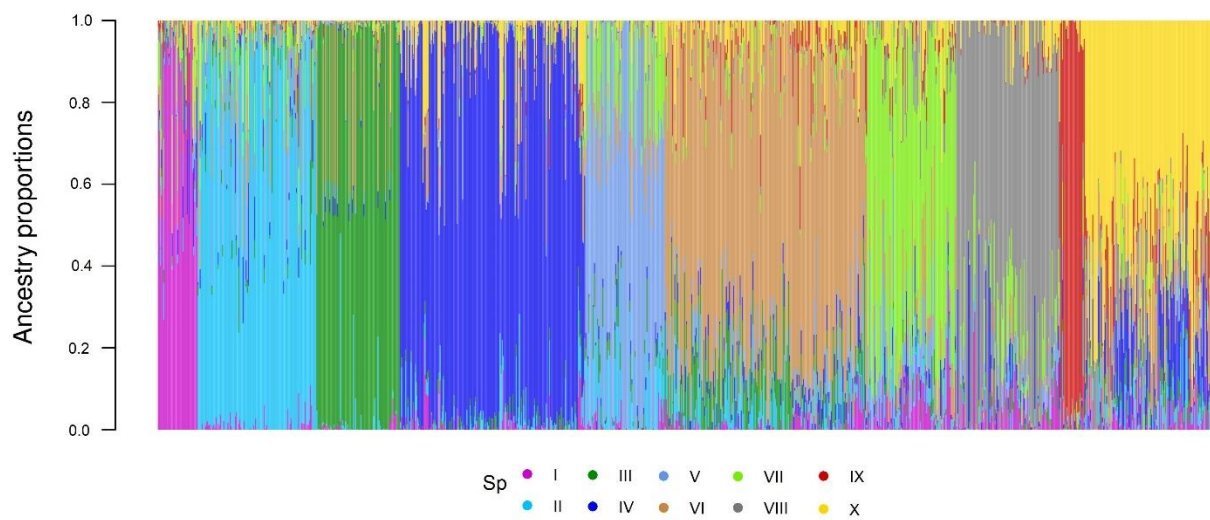


**Figure 1.** Principal component analysis for bioclimatic variables and elevation: (a) Bioclimatic variables related to aridity; (b) Temperature related variables, calculated indices, AP, and ELEV. A = Tropical; B = Dry; C = Temperate; D = Continental and E = Polar.

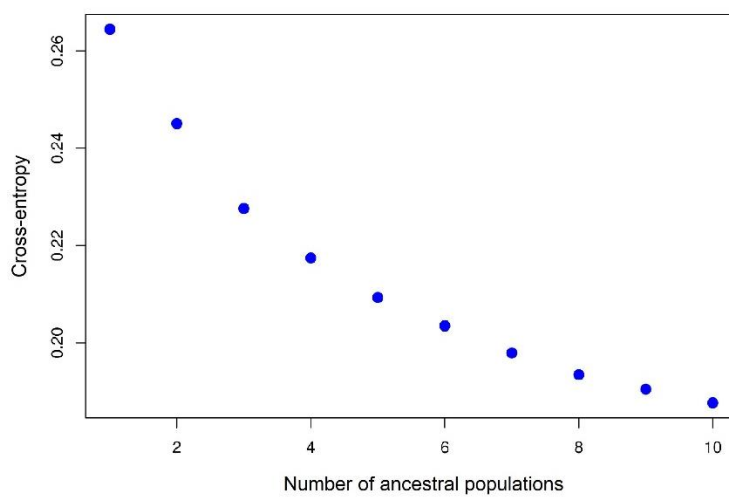
## 2.2. Population structure

The cross-entropy validation implemented in the LEA package, based on SNP markers, suggested an optimal number of 10 subpopulations (Figure 2). This subdivision reflects diversity according to the climate in the subpopulations since there is no explicit representation between these and their geographic or regional origin. The main contribution of accessions comes from Turkey and China, with 49.5 % and 27.8 %, respectively. It is followed by Afghanistan, Tajikistan, and Iran, with 5.2 %, 4.2 %, and 3.4 %, respectively. The rest of the countries contribute with less than 2% of accessions. Subpopulations II, IV, VI, and X concentrate 60.4% of the total population under study. On the other hand, the dispersion pattern observed in the molecular PCA biplot revealed that the subpopulations were well differentiated, reflecting the high genetic diversity of the analyzed accessions (Figure 3a). Correspondence analysis (Figure 3b) between subpopulations and Köppen-Geiger climate groups shows that dry weather accessions (B) are better associated with subpopulation VII. This subpopulation comprises 83 accessions from eight countries (Afghanistan, Armenia, Azerbaijan, China, Iran, Iraq, Tajikistan, and Turkey) cataloged principally within the region of Southern Asia (SAS) with records of aridity index (DI) greater than 5, whose values indicate water deficiency.



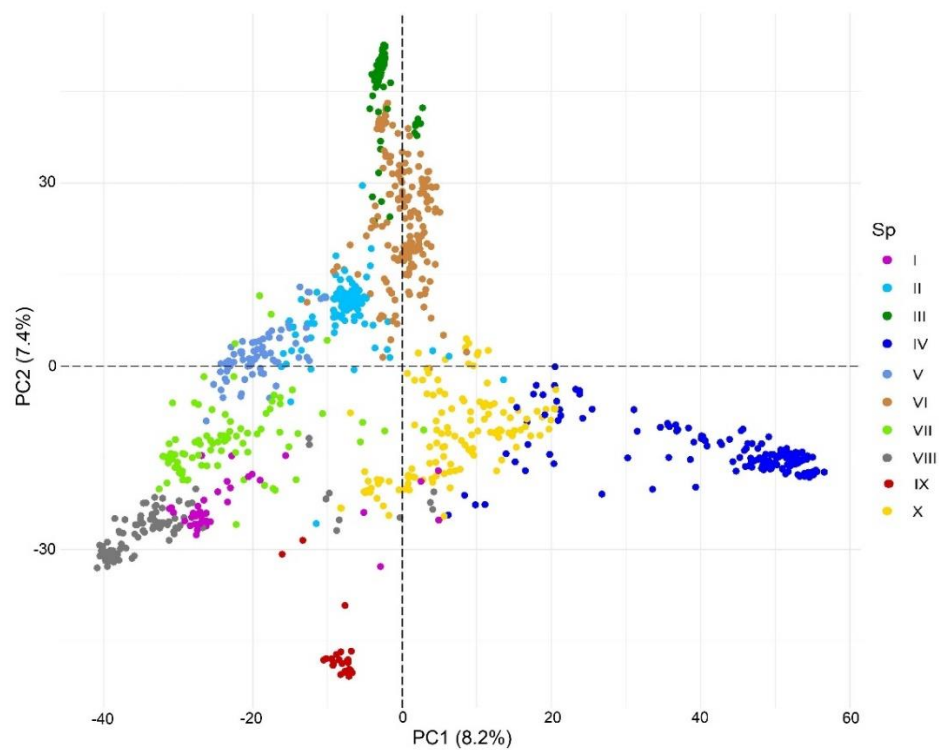


(a)

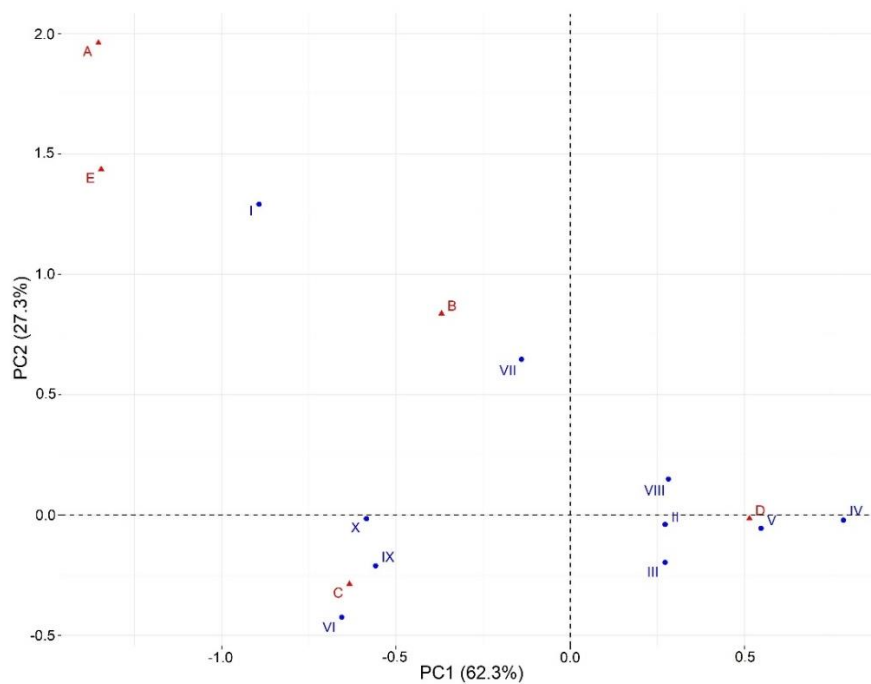


(b)

**Figure 2.** (a) Grouping of subpopulations ( $K=10$ ), each collection is represented by a thin vertical line, divided into colored segments representing the estimated probabilities of belonging ( $Q$ ) to each subpopulation; (b) Ancestry coefficients estimated by LEA.



(a)



(b)

**Figure 3.** Population structure: (a) Biplot for molecular PCA; (b) Biplot of correspondence analysis, where A = Tropical; B = Dry; C = Temperate; D = Continental and E = Polar.

### 2.3. Genome-Wide Association Studies

We identified 59 SNP markers associated with the climatic variables evaluated in all 21 bread wheat chromosomes (Table 2). The chromosomes with the highest number of associated markers were 2B (seven SNPs), 7A (six SNPs), 3B (five SNPs), and 5B (five SNPs). The chromosomes with a single associated marker were 1B, 5A, 6A, 6D, and 7D.

**Table 2.** Number of genes detected from Genome-Environment Associations (GEA) for seven climatic variables with 26,064 SNPs in 990 landraces of bread wheat.

Chr	Physical position (bp)	SNP	Variables	Genes
1A	277825486	108825112 F 0-19:T>A-19:T>A	AMT, MaxT y HITmax	1
1A	388440316	108891114 F 0-33:A>G-33:A>G	AMT, MaxT, HITmead y HITmax	1
1A	588647447	108256081 F 0-24:C>G-24:C>G	AP	1
1B	201700273	108853213 F 0-26:C>T-26:C>T	HITmead	1
1D	2113561	108739422 F 0-31:T>A-31:T>A	AMT	1
1D	413939387	107880401 F 0-66:C>T-66:C>T	HITmead	2
1D	485539732	107874524 F 0-37:G>C-37:G>C	PS	3
2A	640221452	109058266 F 0-7:C>T-7:C>T	AP	2
2A	711971015	108020469 F 0-43:T>C-43:T>C	PS	2
2A	729345577	108514047 F 0-10:A>C-10:A>C	PS	2
2B	29129965	107488994 F 0-17:T>C-17:T>C	HITmead	3
2B	38321752	108024073 F 0-47:C>T-47:C>T	HITmax	2
2B	76022529	107794074 F 0-23:T>C-23:T>C	PS	2
2B	91596926	107797386 F 0-11:T>G-11:T>G	AP	1
2B	584134131	108980638 F 0-33:T>C-33:T>C	DI	1
2B	768567114	107593284 F 0-50:T>A-50:T>A	AP	1
2B	795754781	109021888 F 0-34:T>C-34:T>C	PS	1
2D	79989524	108968402 F 0-23:C>T-23:C>T	AMT	2
2D	302776397	107489027 F 0-34:C>T-34:C>T	DI	1
3A	502971897	106510612 F 0-30:C>T-30:C>T	MaxT	2
3A	507114154	108476623 F 0-32:T>A-32:T>A	MaxT y HITmax	2
3A	574516154	108028914 F 0-20:A>G-20:A>G	MaxT y HITmax	2
3B	50526285	108953425 F 0-21:A>G-21:A>G	PS	1
3B	535641207	109122135 F 0-40:A>G-40:A>G	HITmax	1
3B	741467423	107601308 F 0-23:G>A-23:G>A	PS	0
3B	758205945	108516380 F 0-24:T>C-24:T>C	PS	1
3B	764282419	108146856 F 0-65:T>A-65:T>A	HITmead	1
3D	97153088	107698139 F 0-48:C>T-48:C>T	AMT y HITmead	1
3D	546672913	108308207 F 0-18:G>A-18:G>A	AMT y HITmax	2
3D	575639014	109610030 F 0-20:C>T-20:C>T	HITmax	2
4A	598521397	109242168 F 0-13:A>G-13:A>G	HITmead y HITmax	2
4A	714179675	108145069 F 0-52:C>A-52:C>A	AP	2
4B	39759168	106772473 F 0-11:T>G-11:T>G	PS	1
4B	56276689	108145974 F 0-24:T>C-24:T>C	HITmead	1
4D	62303691	109363420 F 0-34:G>A-34:G>A	HITmead	2
4D	97959582	108773963 F 0-37:C>T-37:C>T	HITmead	2
4D	370691683	109431634 F 0-54:G>A-54:G>A	AMT, MaxT, HITmead y HITmax	3
5A	619468262	107001591 F 0-37:T>C-37:T>C	PS	2

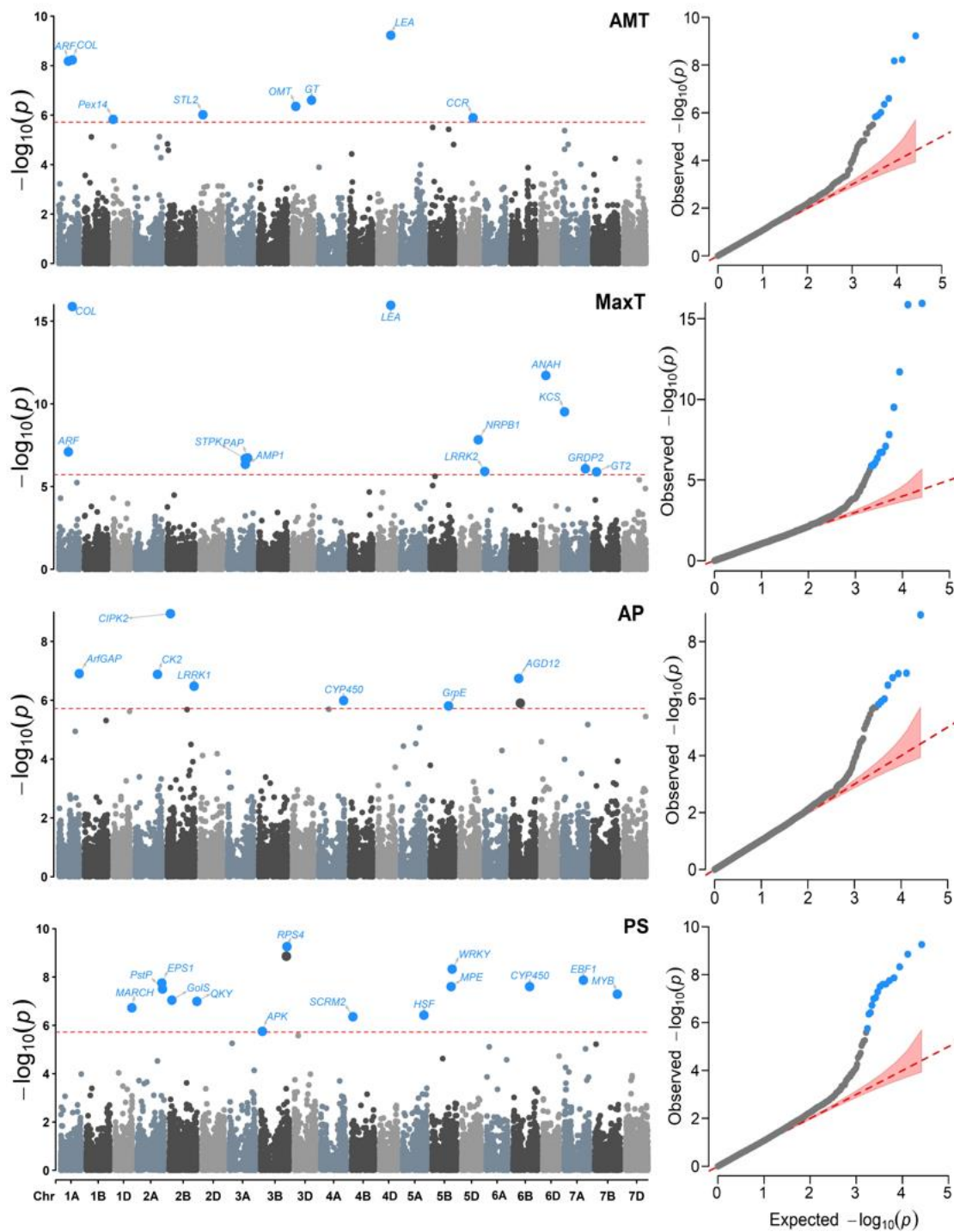
Table 2. Cont.

Chr	Physical position (bp)	SNP	Variables	Genes
5B	68925700	109240982 F 0-38:G>C-38:G>C	HITmead	1
5B	526249013	109305833 F 0-34:C>G-34:C>G	AP y DI	1
5B	530916378	108773279 F 0-51:C>T-51:C>T	HITmead	2
5B	548120559	108348543 F 0-58:A>G-58:A>G	PS	1
5B	577227418	109119258 F 0-68:T>C-68:T>C	PS	2
5D	379130055	108305241 F 0-68:A>G-68:A>G	AMT	1
5D	528660566	108652995 F 0-58:G>C-58:G>C	MaxT	1
6A	5854616	109126792 F 0-49:C>T-49:C>T	MaxT	2
6B	223441723	109177937 F 0-11:A>G-11:A>G	AP	1
6B	269818400	109523315 F 0-24:C>T-24:C>T	AP	0
6B	485290761	109354013 F 0-58:A>G-58:A>G	PS	2
6D	143959936	108582786 F 0-49:T>G-49:T>G	MaxT y HITmax	2
7A	64789408	108830300 F 0-9:T>G-9:T>G	MaxT	2
7A	67617920	108206445 F 0-49:C>A-49:C>A	DI	1
7A	498861613	107952026 F 0-63:G>C-63:G>C	DI	1
7A	552606647	108619258 F 0-17:C>T-17:C>T	PS	1
7A	628906923	107178221 F 0-24:T>G-24:T>G	HITmead	2
7A	662017143	109126469 F 0-35:G>A-35:G>A	MaxT	2
7B	107522176	107878167 F 0-16:G>A-16:G>A	MaxT	2
7B	650581291	108981313 F 0-8:C>T-8:C>T	PS	1
7D	88318125	109035950 F 0-29:C>T-29:C>T	DI	1
			<b>Total</b>	<b>89</b>

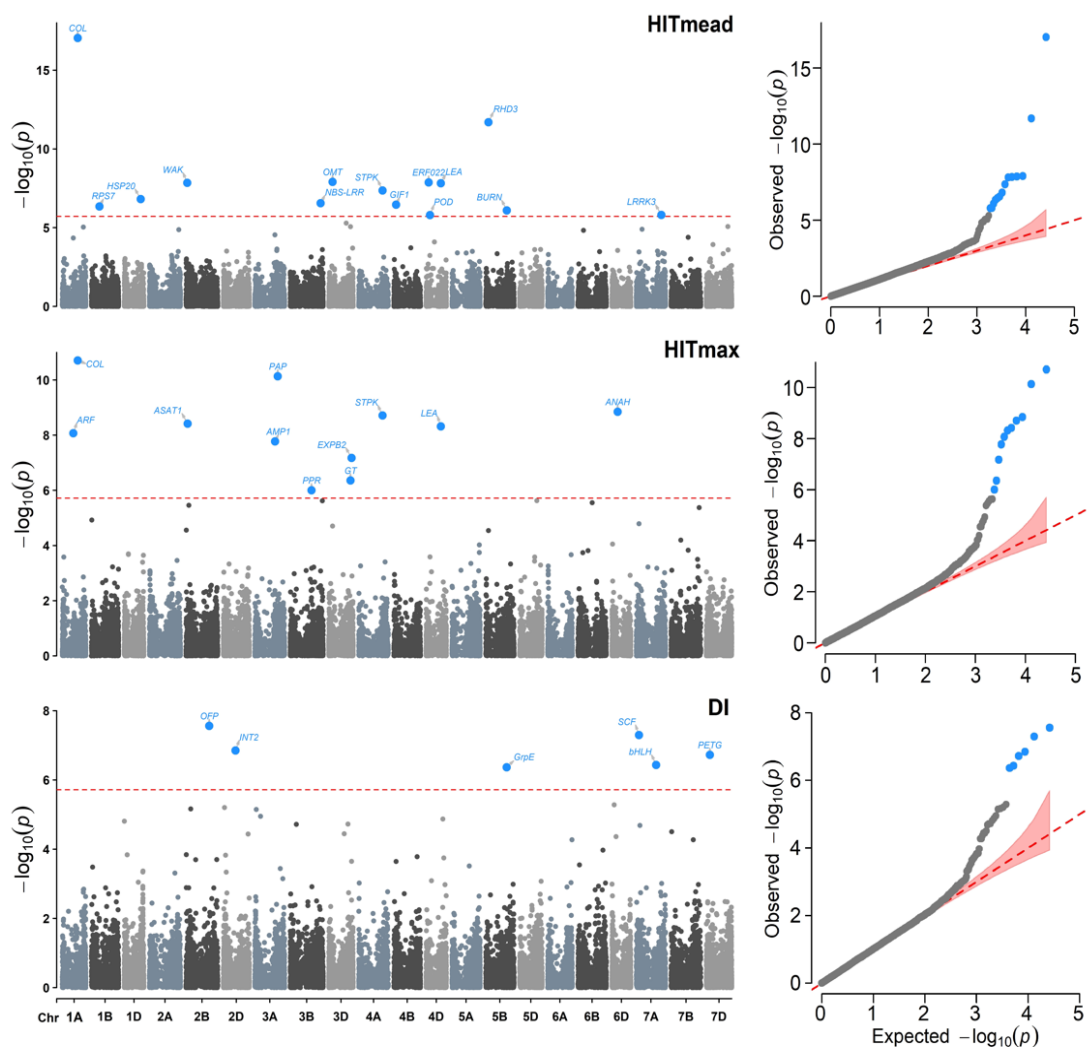
For all variables, the QQ plots (Figures 4 and 5) show a good adjustment, with most  $-\log_{10}(p\text{-values})$  for the null hypothesis of no association, being close to the diagonal. In contrast, some points at the top of each plot may be in *linkage disequilibrium* (LD) with a causal polymorphism, indicating that the model has reasonable control for both false positives and negatives.

Regarding the individual detection of association for each variable, there were different levels of associated SNPs (Table S1). These differences in detecting different SNPs in highly correlated variables may be due to the possible presence of atypical data in the collection sites. Likewise, it is known that the effects produced by drought and heat are differential in certain phases of reproductive development, during which plants are more susceptible. We observed a lower association with variables DI, AMT, and AP, with less than eight markers for each one. In contrast, the variables with more associated loci were PS, HITmead, MaxT, and HITmax, with 15, 14, 12, and 11 SNPs, respectively.

At least 10 SNP located in seven chromosomes (1A, 3A, 3D, 4A, 4D, 5B, and 6D) were detected more than once for variables AMT, MaxT, HITmead, and HITmax. The most frequent SNPs were 108825112|F|0-19:T>A-19:T>A and 108891114|F|0-33:A>G-33:A>G on chromosome 1A, and allele 109431634|F|0-54:G>A-54:G>A on chromosome 4D. These SNPs are related to genes *Auxin response factor* (ARF), *CONSTANS-like* (COL), and *proteins abundant in late embryogenesis* (LEA).



**Figure 4.** Manhattan (left) and QQ (right) plots for Genome-Environment Associations (GEA) analyses with the FarmCPU model and the variables: Annual mean temperature (AMT), maximum temperature (MaxT), annual precipitation (AP), and precipitation seasonality (PS) in 990 landraces of bread wheat and 26,064 SNPs markers. The red dashed horizontal line marks the  $-\log_{10}(p)$  threshold after Bonferroni correction for multiple comparisons. The proteins associated to the significant SNPs are labeled in each graph.



**Figure 5.** Manhattan (left) and QQ (right) plots for Genome-Environment Associations (GEA) analyses with the FarmCPU model and the variables: heat index of mean temperature (HITmead), heat index of maximum temperature (HITmax), and drought index (DI) in 990 landraces of bread wheat and 26,064 SNPs markers. The red dashed horizontal line marks the  $-\log_{10}(p)$  threshold after Bonferroni correction for multiple comparisons. The proteins associated to the significant SNPs are labeled in each graph.

#### 2.4. Identification of genes related to adaptation to abiotic stress in plants

For the associated regions, we found 89 candidate genes encoding proteins related to various biological processes in plants (Table 2). Among these, we identified the significant presence of 26 proteins involved in the signaling network (Table 3), 15 cell wall structural proteins (Table 4), 21 response proteins to various types of abiotic stress (Table 5), and seven proteins related to morphological changes (Table S2).

Given that response to stress begins with the perception and signal transduction of environmental stimuli, it is not surprising that we found an abundance of signaling proteins with a well-documented role in plant response to drought and heat stress. This is the case for protein kinases serine-threonine (*Pstp* - 2A, *STPK* - 2B, 3A, 4A, 4D, *CIPK2* - 2B) and some proteins activated by stress-related plant hormones, such as ethylene (ET), abscisic acid (ABA), jasmonic acid (JA), salicylic acid (SA), and auxins, reported mainly by the BIO15, HITmax, and MaxT variables and chromosomes 7A, 4D, 2B, and 1D (Table 3).

**Table 3.** Signaling genes and proteins identified for Genome-Environment Associations (GEA) with seven climatic variables.

Chr	SNP	Gene	Variable	Protein	Function
1A	108825112 F 0-19:T>A-19:T>A	TraesCS1A02G156600	AMT, MaxT and HITmax	Auxin response factor ( <i>ARF</i> )	<i>Auxin</i> -activated signaling.
1D	108739422 F 0-31:T>A-31:T>A	TraesCS1D02G003900	AMT	Peroxisome membrane anchor ( <i>PEX14</i> ).	Transduction of stress signals by <i>ROS</i> .
1D	107880401 F 0-66:C>T-66:C>T	TraesCS1D02G319600	HITmead	S-adenosyl-L-methionine-dependent methyltransferases ( <i>SAME</i> ).	Methylation of DNA and proteins, <i>ET</i> biosynthesis, phenylpropanoid biosynthesis.
1D	107874524 F 0-37:G>C-37:G>C	TraesCS1D02G438700	PS	Swi-Independent 3 (SIN3)-Like 1 ( <i>SNL1</i> ).	<i>ABA</i> and <i>ET</i> -activated signaling.
2A	108020469 F 0-43:T>C-43:T>C	TraesCS2A02G467400	PS	Protein enhanced pseudomonas susceptibility 1 ( <i>EPS1</i> ).	<i>SA</i> biosynthesis and response to <i>JA</i> .
2A	108514047 F 0-10:A>C-10:A>C	TraesCS2A02G500200	PS	Ser/Thr phosphatase ( <i>PstP</i> ).	Signaling in response to <i>ABA</i> .
2B	108024073 F 0-47:C>T-47:C>T	TraesCS2B02G071900	HITmax	Ser/Thr kinase ( <i>STPK</i> ).	Signaling cascades.
2B	107794074 F 0-23:T>C-23:T>C	TraesCS2B02G112600	PS	MYB108 TF ( <i>MYB</i> ).	Response to signaling by <i>ABA</i> and <i>JA</i> .
2B	107797386 F 0-11:T>G-11:T>G	TraesCS2B02G123900	PS	Nonspecific Ser/Thr kinase ( <i>CIPK2</i> ).	Signaling cascades.
2D	107489027 F 0-34:C>T-34:C>T	TraesCS2D02G252400	DI	Sugar/inositol transporter 2 ( <i>INT2</i> ).	Transduction of hormonal signals.
3A	106510612 F 0-30:C>T-30:C>T	TraesCS3A02G274000	MaxT	Ser/Thr kinase ( <i>STPK</i> ).	Signaling cascades.
3B	109122135 F 0-40:A>G-40:A>G	TraesCS3B02G331800	HITmax	Pentatricopeptide repeat ( <i>PPR</i> ).	Signaling in response to <i>ABA</i> .
3B	108516380 F 0-24:T>C-24:T>C	TraesCS3B02G516800	PS	Mitochondrial ribosomal S4 ( <i>RPS4</i> ).	Proteins encoded in the mitochondrial genome exported to the cytoplasm.
3D	108308207 F 0-18:G>A-18:G>A	TraesCS3D02G433200	AMT and HITmax	Similar to helix-loop-helix DNA ( <i>bHLH</i> )	Signaling in response to <i>auxin</i> and cell wall modification.
4A	109242168 F 0-13:A>G-13:A>G	TraesCS4A02G301600	HITmead and HITmax	2-methyl-6-phytyl-1,4-hydroquinone methyltransferase ( <i>VTE3</i> ).	Vitamin <i>E</i> biosynthesis and stress signaling.
		TraesCS4A02G302000		Ser/Thr kinase ( <i>STPK</i> ).	Signaling cascades.

Table 3. Cont.

Chr	SNP	Gene	Variable	Protein	Function
4D	109363420 F 0-34:G>A-34:G>A	TraesCS4D02G087000	HITmead	Ethylene-responsive TF ( <i>ERF022</i> ).	<i>ET</i> -activated signaling.
4D	109431634 F 0-54:G>A-54:G>A	TraesCS4D02G216300	AMT, MaxT, HITmead and	Ethylene-responsive TF ( <i>ERF014</i> ).	<i>ET</i> -activated signaling.
		TraesCS4D02G216600	HITmax	Ser/Thr kinase ( <i>STPK</i> ).	Signaling cascades.
5B	108348543 F 0-58:A>G-58:A>G	TraesCS5B02G369300	PS	Metallophone domain ( <i>MPE</i> ).	GPI biosynthesis the cell membrane.
7A	107952026 F 0-63:G>C-63:G>C	TraesCS7A02G340300	DI	Basic helix-loop-helix ( <i>bHLH</i> ).	Signaling in response to <i>auxin</i> and cell wall modification.
7A	108619258 F 0-17:C>T-17:C>T	TraesCS7A02G377500	PS	EIN3-binding F-box 1 ( <i>EBF1</i> ).	<i>ET</i> -activated signaling.
7A	107178221 F 0-24:T>G-24:T>G	TraesCS7A02G435700	HITmead	IAA-amino acid hydrolase ( <i>ILL</i> ).	<i>Auxin</i> metabolic process.
		TraesCS7A02G465400		Kinase (Kinase).	Signaling cascades.
7A	109126469 F 0-35:G>A-35:G>A	TraesCS7A02G465500	MaxT	Glycine-rich domain 2 ( <i>GRDP2</i> ).	<i>Auxin</i> -activated signaling.
7B	108981313 F 0-8:C>T-8:C>T	TraesCS7B02G385700	PS	HTH myb-type domain ( <i>MYB</i> ).	Response to signaling by <i>ABA</i> and <i>JA</i> .



**Table 4.** Cell wall genes and proteins identified for Genome-Environment Associations (GEA) with seven climatic variables.

Chr	SNP	Gene	Variable	Protein	Function
1A	108256081 F 0-24:C>G-24:C>G	TraesCS1A02G439300	AP	ArfGAP domain 2G ( <i>ArfGAP</i> ).	Membrane trafficking.
1D	107874524 F 0-37:G>C-37:G>C	TraesCS1D02G439900	PS	RING-CH-type domain/E3 ubiquitin ligase ( <i>MARCH</i> ).	Protein degradation by the ubiquitin pathway with abnormalities.
2B	107488994 F 0-17:T>C-17:T>C	TraesCS2B02G059400	HITmead	Wall-associated kinase ( <i>WAK</i> ).	Regulation of wall functions and signaling of extracellular environment.
2B	108024073 F 0-47:C>T-47:C>T	TraesCS2B02G071600	HITmax	Acyl-CoA--sterol O-acyltransferase ( <i>ASAT1</i> ).	Synthesis of long-chain esters (waxes).
2B	109021888 F 0-34:T>C-34:T>C	TraesCS2B02G621600	PS	C2 calcium/lipid-phosphoribosyltransferase ( <i>QKY</i> ).	Signal transduction or calcium-dependent membrane trafficking.
2D	108968402 F 0-23:C>T-23:C>T	TraesCS2D02G136300	AMT	Glycosyltransferase STELLO2 ( <i>STL2</i> ).	Cell wall cellulose biosynthesis.
3A	106510612 F 0-30:C>T-30:C>T	TraesCS3A02G274200	MaxT	Mannan endo-1,4-beta-mannosidase 2 ( <i>MAN2</i> ).	Lignocellulose component in primary cell walls.
3A	108476623 F 0-32:T>A-32:T>A	TraesCS3A02G277100	MaxT and HITmax	Glucan endo-1,3-beta-glucosidase 13 ( <i>1,3-β-glucanasa</i> ).	Defense against pathogens, cell wall biogenesis and reorganization.
3D	107698139 F 0-48:C>T-48:C>T	TraesCS3D02G138700	AMT and HITmead	O-methyltransferase ( <i>OMT</i> ).	Lignin biosynthesis.
3D	108308207 F 0-18:G>A-18:G>A	TraesCS3D02G433400	AMT and HITmax	Glycosyltransferase ( <i>GT</i> ).	Biosynthesis polysaccharides of cell walls: cellulose, hemicellulose, and pectin.
3D	109610030 F 0-20:C>T-20:C>T	TraesCS3D02G474800	HITmax	Putative expansin-B2 ( <i>EXPB2</i> ).	Loosening of plant cell walls.
5B	109240982 F 0-38:G>C-38:G>C	TraesCS5B02G061500	HITmead	Root hair defective 3 ( <i>RDH2</i> ).	Biogenesis of the cell wall and organization of the cytoskeleton.
5D	108305241 F 0-68:A>G-68:A>G	TraesCS5D02G276400	AMT	Cinnamoyl-CoA reductase 4 ( <i>CCR</i> ).	Primary alcohols and leaf cuticular wax synthesis.
7A	108830300 F 0-9:T>G-9:T>G	TraesCS7A02G107500	MaxT	3-ketoacyl-CoA synthase ( <i>KCS</i> ).	Synthesis of long-chain esters (waxes).
7B	107878167 F 0-16:G>A-16:G>A	TraesCS7B02G093900	MaxT	Glyco_trans_2-like ( <i>GT2</i> ).	Cell wall organization.

Within the genes associated with the cell wall (Table 4), proteins involved in the biosynthesis of cuticle and cell wall components stand out, such as cuticular wax (*ASAT1*, *CCR*, and *KCS*), polysaccharides such *cellulose*, *hemicellulose*, and *pectin* (*STL2* and *MAN2*), *lignin* (*OMT*), and *glycoproteins* (*GT*, *1,3- $\beta$ -glucanase*) on chromosomes 2B, 2D, 3A, 3D, 5B, 5D, and 7A. We also found some structural proteins (*WAK90*, *EXPB2*, *RD22*, and *GT2*), a membrane trafficking protein (*ArfGAP*), and two signal transduction proteins (*MARCH* and *QKY*) on chromosomes 1A, 1D, 2B, 3D, 5B, and 7B.

Many morphological changes are induced when the plant is subjected to long periods of environmental stress (Table S2). For this reason, we also observed some genes regulating multiple development processes (*FBX* - 2B and 4D), organ morphology (*OFP* - 2B), stomatal differentiation (*SCRM2* - 4B), cell expansion (*GIF1* - 4B), proteins that affect gravitropism (*AGD12* - 6B) and sheet curling (*ROC5* - 6B).

### 2.5. Drought and heat adaptation genes

Through genomic association with the variables AP, PS, MaxT, HITmead, and DI, we identified 12 genes related to the response and tolerance of plants to drought and heat stress (Table 5 and Figure 6) on chromosomes 1D (one gene), 2B (two genes), 4D (one gene), 5A (one gene), 5B (three genes), 5D (one gene), 6A (one gene), and 7A (two genes). Proteins from different domains represent biological processes related to water deficit: enzymes with *galactinol synthase* activity (*Gols* - 2B), *leucine-rich repeat kinase proteins* (*LRRK* - 2B, 6A, and 7A), *late embryogenesis abundant proteins* (*LEA* - 4D), *BURP* domain proteins (*BURP* - 5B), *WRKY* transcription factors (*WRKY*- 5B), and component box proteins of the *SKP-Cullin-Fbox E3 complex* (*SCF* - 7A). Likewise, the response to heat stress by some *heat shock proteins* (*HSP20* - 1D, *HSFA2E* - 5A), a *chronic heat stress thermotolerance protein* (*GrpE* - 5B), and a *DNA-dependent RNA polymerase* (*NRPB1* - 5D).

### 2.6. Other genes

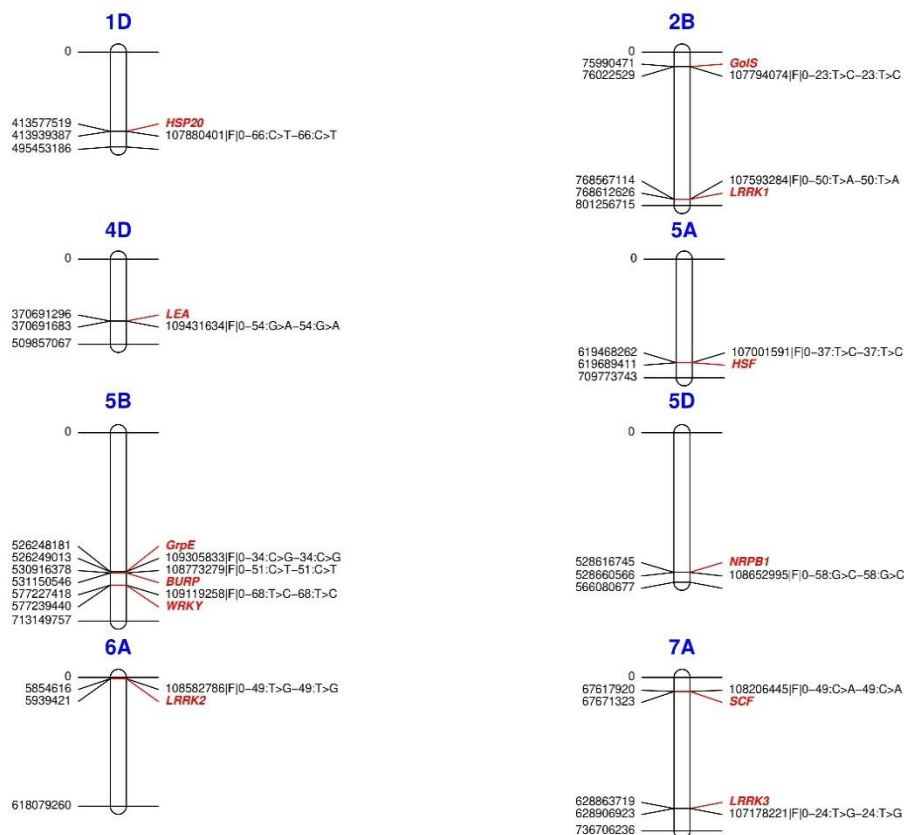
Additionally, we observed seven genes involved in the response to biotic stress (*PELO*, *RIP*, *NBS-LRR*, *LFG4*, *FMO*, and *CNGC2*), six genes involved in photosynthesis (*GLO1*, *CHLH*, *PSB33*, *psbL*,  $\gamma$ *CA1*, and *PETG*), four genes involved in flowering (*COL*, *CK2*, *FPF1*, and *APK*), and three genes involved in nutrient assimilation (*YSL1*, *PAP*, and *wSs2a-3*) on chromosomes 1A, 2A, 2B, 2D, 3A, 3B, 3D, 4A, 4D, 5B, 5D, 7A, 7B and 7D (Table S2).

**Table 5.** Abiotic stress genes and proteins identified for Genome–Environment Associations (GEA) with seven climatic variables.

Chr	SNP	Gene	Variable	Protein	Function
1B	108853213 F 0-26:C>T-26:C>T	TraesCS1B02G146100	HITmead	40S ribosomal S7 ( <i>RPS7</i> ).	Response to environmental signals.
1D	107880401 F 0-66:C>T-66:C>T	TraesCS1D02G319400	HITmead	Heat shock protein class VI ( <i>HSP20</i> ).	Heat and salt tolerance.
1D	107874524 F 0-37:G>C-37:G>C	TraesCS1D02G439800	PS	Trimethylguanosine synthase ( <i>TGS</i> ).	Cold tolerance.
2B	107794074 F 0-23:T>C-23:T>C	TraesCS2B02G112800	PS	Galactinol synthase 7 ( <i>GolS</i> ).	Tolerance to drought, salinity, and cold.
2B	107593284 F 0-50:T>A-50:T>A	TraesCS2B02G581100	AP	Leucine-Rich Repeat Kinase ( <i>LRRK1</i> ).	Tolerance to drought.
3A	108476623 F 0-32:T>A-32:T>A	TraesCS3A02G276800	MaxT and HITmax	Glutamate carboxypeptidase ( <i>AMP1</i> ).	Responses to oxidative stress.
4A	108145069 F 0-52:C>A-52:C>A	TraesCS4A02G446900	AP	Cytochrome P450 709B3 ( <i>CYP450</i> ).	Biosynthesis of secondary metabolites and phytohormones in response to stress.
4D	108773963 F 0-37:C>T-37:C>T	TraesCS4D02G117200	HITmead	Peroxidase ( <i>POD</i> ).	Response to oxidative stress.
4D	109431634 F 0-54:G>A-54:G>A	TraesCS4D02G216700	AMT, MaxT, HITmead and HITmax	Late embryogenesis abundant ( <i>LEA</i> ).	Osmotic stress.
5A	107001591 F 0-37:T>C-37:T>C	TraesCS5A02G437900	PS	Heat shock factor ( <i>HSF</i> ).	Heat shock proteins.
5B	109305833 F 0-34:C>G-34:C>G	TraesCS5B02G341100	AP and DI	GrpE protein homolog ( <i>GrpE</i> ).	Thermotolerance to chronic heat stress.
5B	108773279 F 0-51:C>T-51:C>T	TraesCS5B02G350000	HITmead	BURP domain ( <i>BURP</i> ).	Responses to drought stress by <i>ABA</i> .
5B	109119258 F 0-68:T>C-68:T>C	TraesCS5B02G399900	PS	Probable WRKY TF 57 ( <i>WRKY</i> ).	Response to osmotic stress, salt, and drought.
5D	108652995 F 0-58:G>C-58:G>C	TraesCS5D02G498900	MaxT	DNA-directed RNA polymerase subunit ( <i>NRPB1</i> ).	Response to heat stress.

Table 5. Cont.

Chr	SNP	Gene	Variable	Protein	Function
6A	108582786 F 0-49:T>G-49:T>G	TraesCS6A02G012100	MaxT	Cytochrome P450 709B1 ( <i>CYP450</i> ).	Biosynthesis of secondary metabolites and phytohormones in response to stress.
		TraesCS6A02G013100		Leucine-rich repeat receptor-like kinase ( <i>LRRK2</i> ).	Tolerance to drought.
6B	109354013 F 0-58:A>G-58:A>G	TraesCS6B02G269500	PS	Cytochrome P450 ( <i>CYP450</i> ).	Biosynthesis of secondary metabolites and phytohormones in response to stress.
6D	108582786 F 0-49:T>G-49:T>G	TraesCS6D02G164900	MaxT and HITmax	Adenine nucleotide alpha hydrolases ( <i>ANAH</i> ).	Response to salt stress.
		TraesCS6D02G165100		Cytochrome P450 ( <i>CYP450</i> ).	Biosynthesis of secondary metabolites and phytohormones in response to stress.
7A	108206445 F 0-49:C>A-49:C>A	TraesCS7A02G110500	DI	Fbox component of the SKP-Cullin-Fbox E3 ( <i>SCF</i> ).	Water deficit.
7A	107178221 F 0-24:T>G-24:T>G	TraesCS7A02G435300	HITmead	Leucine-rich repeat receptor-like kinase ( <i>LRRK3</i> ).	Tolerance to drought.



**Figure 6.** Physical map of the location of SNPs associated with heat and drought tolerance related genes in *Triticum aestivum*. For each locus the position is written on the left side, whereas SNPs are coded on the right side. Gene names are colored. HSP20 = Heat shock protein class VI; GolS = Galactinol synthase 7; LEA = Late embryogenesis abundant; HSF = Heat shock factor; GrpE = GrpE protein homolog; BURP = BURP domain; WRKY = Probable WRKY TF 57; NRPB1 = DNA-directed RNA polymerase subunit, SCF = Fbox component of the SKP-Cullin-Fbox E3; LRRK = Leucine-rich repeat receptor-like kinase.

### 3. Discussion

#### 3.1. Environmental variables involved in the detection of adaptive loci by GEA

Climate change affects diverse geographic areas throughout the world. However, its effects in arid and semiarid climatic zones have devastating impacts [26]. These selective environmental effects play an essential role in the local adaptation, genetic diversity, and population structure of wild accessions [17]. Therefore, using climatic variables to represent selective environmental pressure can be valuable to capture important components of the mechanisms of resistance and tolerance to abiotic stress. We observed local adaptation footprints in multiple genomic regions along the 21 wheat chromosomes, which each climatic variable having different numbers of genomic associations of biological importance.

It should be remarked that climate data come from records spanning 1970-2000, which naturally does not cover the adaptation period of the accessions. However, these 30-year records, are good indicators of the prevailing climate type in the collection sites. On the other hand, the accessions were collected from 1983 to 2011, with the main bulk of collection occurring in 1984 and 2011. One must be aware of the noise arising from the migration dynamics of those materials, especially for the most recent collection efforts. Thus, the herein reported results are subject to validation by different approaches.

Exceptionally, the seasonality of precipitation (PS) had the largest number of significant SNP markers along different chromosomes, placed close to unique genes of resistance and tolerance to abiotic stress. This makes sense because it is considered an important variable in influencing the distribution of species through water availability [27]. On the other hand, measuring the variation in precipitation [28] at sites of origin of the collections over three decades (1970 to 2000) faithfully represents the alterations in the uniformity and distribution of precipitation.

The research of Cortés et al. [29] reaffirms the above; they found a strong influence of rainfall patterns on population structure and the ecological diversity to tolerance drought in wild beans. Usually, prolonged periods of drought cause the expression of genomic regions associated with the activation of plant survival mechanisms [17].

The associations with maximum temperature (MaxT), heat index (HITmead), annual precipitation (AP), and drought index (DI) identified several adaptation genes to drought and heat stress. This is primarily explained by the nature and importance of these variables in the monitoring of conditions of meteorological drought. Both AP and DI are valid descriptors for measuring drought intensity [30]. DI is calculated through a combination of climatic and meteorological variables, among which precipitation is the most important [31]. In addition, the estimate values DI presented an excellent discriminating potential of accessions from arid climates (B), which gives reliability to its use. On the other hand, according to López-Hernández & Cortés [10] the maximum temperature and the heat index are better estimators of natural adaptation to high temperatures and identify successfully associated genetic factors markers.

Different variables shared associations with some loci, suggesting that their selective pressures can shape the same genomic regions [22] and therefore remain stable in the landraces of *Triticum aestivum*. The most frequent loci on chromosomes 1A and 4D are related to two genes: *CONSTANS-like* and *proteins abundant in late embryogenesis (LEA)*. The first is involved in various biological processes of plants, such as the control of flowering time, regulation of growth and development, and responses to abiotic stress [32-34]. *LEA* proteins are recognized during adaptation to abiotic stress, which includes dehydration, salinity, high temperature, and cold [35-37].

On the other hand, the only matching locus between AP and DI flanked an F-box domain gene, which is a homolog of the *GrpE* protein. In *Arabidopsis*, this gene acts as a nucleotide exchange factor of the 70 kD *heat shock protein complex (HSP70)*, which specializes in thermotolerance to heat stress [38,39].

### 3.2. Adaptation to drought and heat stress

Despite their coexistence in a climate change scenario, the combined effects of drought and heat stress have been poorly studied [40]. They have a synergistic effect, altering metabolism and gene expression in ways other than those induced independently [41]. These combined effects affect several physiological, cellular, and molecular processes in plant cells [10].

The stress response mechanism in plants is very complex and requires several integrated pathways to be activated in response to external stress [42]. Plant hormones, such as *auxins*, *abscisic acid (ABA)*, *ethylene (ET)*, *salicylic acid (SA)*, and *jasmonates (JA)* operate together in the modulation of the plants' heat and drought stress responses [43,44].

Typically, *auxin* and the *auxin* pathway regulate thermomorphogenesis in plants, coordinating growth and defense against heat stress [45]. While, *ABA* and *ET* interact positively to activate or repress the expression of numerous stress response genes, such as *LEA* proteins and *dehydrins* [46,47]. Likewise, *SA* is related to synthesizing of proteins chaperones, heat shock proteins, protective membrane proteins, antioxidants, and secondary metabolites [46,48,49]. *Jasmonates* significantly improve tolerance to heat and drought stress through various *TFs*, which induce

responsive gene expression and organic osmoprotectant activation, osmotic adjustment, and antioxidant activity [50-52].

Many biochemical and physiological impacts affect the growth and development of plants [40], as they affect the photosynthetic system, gas exchange, and water relations [45]. Consequently, a series of physiological and molecular responses are produced, which include root increase, reduction in the number and conductance of stomata, decrease in leaf area, and morphological changes in leaves [53]. On the other hand, among molecular responses, one should consider the production of antioxidants and osmolytes for osmotic adjustment and the expression of various proteins, such as *HSP*, *WRKY*, *MYB*, *LEA*, and *GrpE* [40].

The genes reported in this work are involved in most of the mentioned biological processes, including genes with signaling roles and genes associated with the cell walls and membranes, photosynthesis, flowering, and, of course, proteins involved in response to heat and drought stress on various chromosomes. Our findings are consistent with Y. Li et al. [22] affirmation: plant genomes have been shaped by natural selection during local adaptation to different environmental conditions, so there is a close relationship between species survival and response to climate change.

## 4. Materials and Methods

### 4.1. Geographical data

Through code written in the R language v.3.4.4 [54], the passport data of 174,553 accessions from the CIMMYT Wheat Germplasm Bank were filtered to select 1,151 landraces of *Triticum aestivum* landraces with unique and sensible geographic coordinates. Subsequently, the location mapping was carried out through the geographic information system QGIS version 2.18 [55]. This filtering process yielded 1,151 landraces with validated geographic data. This filtering process yielded 990 landraces with validated geographic data. The accessions come from 33 countries distributed in 13 geographic regions (Table 6) [56].

**Table 6.** Landraces countries of origin and geographic regions.

Country code	Country	Region	Region code	Total landraces
AFG	Afghanistan	Southern Asia	SAS	51
ARM	Armenia	Western Asia	WAS	7
AUS	Australia	Australia and New Zealand	AUS	1
AUT	Austria	Western Europe	WEU	1
AZE	Azerbaijan	Western Asia	WAS	5
CAN	Canada	Northern America	NAM	1
CHN	China	Eastern Asia	EAS	275
DEU	Germany	Western Europe	WEU	1
DZA	Algeria	Northern Africa	NAF	1
ESP	Spain	Southern Europe	SEU	3
ETH	Ethiopia	Eastern Africa	EAF	4
GEO	Georgia	Western Asia	WAS	10
GRC	Greece	Southern Europe	SEU	1
IND	India	Southern Asia	SAS	17
IRN	Iran	Southern Asia	SAS	34
IRQ	Iraq	Western Asia	WAS	8
ITA	Italy	Southern Europe	SEU	1
JPN	Japan	Eastern Asia	EAS	1
LBN	Lebanon	Western Asia	WAS	1
MEX	Mexico	Central America	CAM	9
PAK	Pakistan	Southern Asia	SAS	1
PER	Peru	South America	SAM	1
POL	Poland	Eastern Europe	EEU	1

**Table 6.** *Cont.*

Country code	Country	Region	Region code	Total landraces
PRT	Portugal	Southern Europe	SEU	8
RUS	Russia	Eastern Europe	EEU	2
SRB	Serbia	Southern Europe	SEU	1
SYR	Syria	Western Asia	WAS	3
TJK	Tajikistan	Central Asia	CAS	42
TUN	Tunisia	Northern Africa	NAF	2
TUR	Turkey	Western Asia	WAS	490
USA	United States	Northern America	NAM	4
UZB	Uzbekistan	Central Asia	CAS	2
VEN	Venezuela	South America	SAM	1

#### 4.2. Genotypic data

Germplasm genotyping was carried out through DArT-seq technology in CIMMYT under the Seeds of Discovery initiative for 45,871 accessions belonging to the wheat germplasm bank. The information was integrated using R v.3.4.4 [54] in a data table with the HapMap format containing the information of 86,683 SNP loci. The marker physical locations were obtained by reference genome sequence provided by Diversity Arrays Technology (wheat\_ChineseSpring04), and only markers unambiguously located in the wheat genome were retained. Subsequently, they were filtered for quality control through a selection of the cleanest and most informative SNPs with a maximum missing data rate of 20%, Shannon entropy greater than zero, and variants with a minor allele frequency (MAF)  $\geq 2\%$ . The filtered table contained 26,064 SNP loci in 990 landraces with geographic data.

#### 4.3. Climatic data

For each collection site, we extracted the values of altitude, temperature, precipitation and eight bioclimatic variables related to drought and heat stress (Table 7) at a spatial resolution of 2.5 min (4.5 km) from the WorldClim platform (<https://www.worldclim.org/>) using the *getData* function of the R raster package version 3.5-15 [57]. Bioclimatic variables are derived from the monthly temperature and rainfall historical climate data from 1970 to 2000 in order to generate more biologically meaningful variables [58].

**Table 7.** List of geographic, climatic and bioclimatic variables downloaded from the WorldClim platform.

Abbreviation	Variable description and unit
ELEV	Altitude, meters.
MaxT	Maximum temperature, °C x 10.
AMT	Annual mean temperature, °C x 10.
TS	Temperature seasonality, standard deviation x 100.
MaxTWM	Maximum temperature of warmest month, °C x 10.
MeanTDQ	Mean temperature of driest quarter, °C x 10.
MeanTWQ	Mean temperature of warmest quarter, °C x 10.
AP	Annual precipitation, mm.
PDM	Precipitation of driest month, mm.
PS	Precipitation seasonality, mm.
PDQ	Precipitation of driest quarter, mm.
PWQ	Precipitation of warmest quarter, mm.

Additionally, we determined the Köppen-Geiger main climate groups (Table 8) through the R *kgc* package version 1.0.0.2 [59]. Köppen classification was constructed based on five vegetation groups that distinguish between plants of the equatorial zone (A), the arid zone (B), the warm



temperate zone (C), the snow zone (D), and the polar zone (E); in the sub classification, the second letter consider the precipitation, while the third letter the air temperature [60].

**Table 8.** Climate types under the Köppen-Geiger climate classification.

Group	Climates
A - Tropical	Tropical rainforest ( <i>Af</i> ), Tropical monsoon ( <i>Am</i> ) and Tropical savanna ( <i>Aw</i> , <i>As</i> ).
B - Dry	Desert ( <i>BWh</i> , <i>BWk</i> ) and Semi-arid ( <i>BSh</i> , <i>BSk</i> ).
C - Temperate	Humid subtropical ( <i>Cfa</i> , <i>Cwa</i> ), Oceanic ( <i>Cfb</i> , <i>Cwb</i> , <i>Cfc</i> , <i>Cwc</i> ) and Mediterranean ( <i>Csa</i> , <i>Csb</i> , <i>Csc</i> ).
D - Continental	Humid continental ( <i>Dfa</i> , <i>Dwa</i> , <i>Dfb</i> , <i>Dwb</i> , <i>Dsa</i> , <i>Dsb</i> ) and Subarctic ( <i>Dfc</i> , <i>Dwc</i> , <i>Dfd</i> , <i>Dwd</i> , <i>Dsc</i> , <i>Dsd</i> ).
E - Polar	Tundra ( <i>ET</i> ), Ice cap ( <i>EF</i> ) and Alpine ( <i>ET</i> , <i>EF</i> ).

#### 4.4. Index estimation

Two heat indices (HIT) were estimated by the Thornthwaite model [61] using values of mean and maximum temperature from 1970-2000. We denominated  $HIT_{mean}$  and  $HIT_{max}$ , respectively:

$$HIT_{mean} = \sum_{i=1}^{12} (tmean_i/5)^{1.514} \quad (1)$$

$$HIT_{max} = \sum_{i=1}^{12} (tmax_i/5)^{1.514}, \quad (2)$$

for  $tmean_i$ ,  $tmax_i > 0$ , where  $tmean_i$  is the average monthly temperature, and  $tmax_i$  is the maximum monthly, respectively, for the  $i$ -th month.

Furthermore, a drought index was calculated [29]. This index is based on the relationship between the potential evapotranspiration and the annual precipitation of each collection site:

$$DI = 100 * [(PET - AP) / PET], \quad (3)$$

where  $DI$  is the drought index,  $PET$  is the potential evapotranspiration and  $AP$  is the annual precipitation. In this index, negative values indicate excessive precipitation, while positive values indicate water deficiency. Calculation of  $PET$  was done with the *thornthwaite* function of the R SPEI package version 1.7 [62], through the values of monthly average temperature from 1970-2000 ( $tmed_i > 0$ ) with the estimated solar radiation being based on the latitude of each collection site.

#### 4.5. Exploratory analysis of climatic variables

To look for patterns of relationships among climate variables, we used the Pearson correlation coefficient ( $r$ ) and principal component analysis (PCA). The standardized variables are used in PCA to estimate the correlation matrix and determine the principal components (PC). Bioclimatic variables were grouped into two sets, with the first one containing bioclimatic variables related to aridity (AMT, TS, MaxTWM, MeanTDQ, MeanTWQ, AP, PS, PDM, PDQ, and PWQ). The second set included AMT, MaxT, AP, constructed indices ( $HIT_{mean}$ ,  $HIT_{max}$  and  $DI$ ), and ELEV. AMT and AP were included in both sets because they are considered the main variables for indices related to aridity. The biplot graphs were constructed with the *factoextra* R package version 1.0.7 [63]. In the latter, the vector's length and the angle's cosine were used to group the variables into different groups. The Köppen-Geiger climate groups of each collection were also included.

#### 4.6. Population structure

The stratification of the collection was explored by two methods. The first one was based on the Landscape and Ecological Associations studies (LEA) package [64], with SNPs coded in numerical form (0, 1, 2). The *smmf* function was used to estimate the ancestry coefficients (K) with cross-entropy. This algorithm was executed with 10 replications and a K value from 1 to 10. The optimal number of K was defined to assign genotypes to subpopulations according to estimates individual admixture coefficients from the genotypic matrix. Subsequently, we visualized principal components through the GAPIT R package version 3.0 [65]. We compared the results of the population stratification in subpopulations with the Köppen-Geiger climate groups using correspondence analysis (CA), which reveals the close relationships between and within two groups of categorical variables, based on data provided in a contingency table.

#### 4.7. Association analysis

We used 990 landraces of *Triticum aestivum* genotyped with 26,064 SNP loci and seven variables (AMT, MaxT, AP, PS, HITmead, HITmax, and DI) to run Genome-Environment Association (GEA) studies with the Fixed and random model Circulating Probability Unification multiple-locus model multiple-locus model “FarmCPU” [66] implemented in the R GAPIT (Genome Association and Prediction Integrated Tool) package version 3.0 [65]. FarmCPU is characterized by iteratively using two models, a linear mixed model (MLM) and a fixed-effects model, to select a set of markers associated with a trait of interest.

The significant SNPs were determined according to the Bonferroni threshold to an alpha of 0.05, with a threshold value of  $-\log_{10}(0.05/26064) = 5.72$ , coupled by the visual interpretation of the Q-Q plots. The Manhattan and Q-Q plots were built with the R CMplot package version 4.0.0 [67].

#### 4.8. Candidate genes and their annotation

The sequence of the significant SNP markers was blasted in the wheat reference genome IWGSC\_refseqv1.0 [68] published in the Ensembl Plants database ([www.plants.ensembl.org/](http://www.plants.ensembl.org/) accessed on 31 July 2022) to identify candidate genes. For this, the genes found in the overlapping region and within one Mb upstream and downstream of the matched regions were selected as candidate genes and their molecular functions were determined. We identified the proteins in the UniProt (<https://www.uniprot.org/> accessed on 31 July 2022) and InterPro (<https://www.ebi.ac.uk/interpro> accessed on 31 July 2022) databases encoded by the candidate genes, the functionality of the sequences, their domain, and classification. Finally, we elaborated a schematic representation of the physical map of bread wheat with the significant SNPs associated with response proteins to water stress and heat stress with the R LinkageMapView package version 2.1.2 [69].

### 5. Conclusions and Practical Implications

The results suggest that local adaptations have footprints along the 21 wheat chromosomes in multiple genomic regions. We found vital genes that include several critical points on abiotic stress response mechanisms—highlighting a considerable number of signaling genes mediated by plant hormones, regulatory processes of the cell wall, morphophysiological changes, photosynthesis, flowering, and some response mechanisms to abiotic stress.

We showed that the climatic variables estimated with historical data help capturing the environmental variability that occurred in the collection sites of the landraces. The variables of maximum temperature, annual precipitation, precipitation seasonality, and heat and drought indices had relevant participation in identifying genes related to the response to water deficit and

high temperatures. This confirms its representativeness in determining aridity in some climatic regions.

This study points to 89 genes involved in the adaptation of bread wheat to its native habitats by association with seven specific climatic variables. The results are consistent with the idea that environmental pressure has modeled, through natural selection, the structure of genomic regions in local wheat populations over time.

Our findings constitute a new resource to select accessions carrying alleles linked to specific climatic responses, which can be exploited through genomic prediction tools to select germplasms with genetic potential for adaptation to climate change.

**Supplementary Materials:** The following are available online at: <https://www.mdpi.com/article/10.3390/plants11172289/s1>: Figure S1: Contributions to the total variation of each component for the evaluated variables in the PCA; Figure S2: Contributions to component principals 1 and 2 for the evaluated variables in the PCA; Table S1: Detection of associated SNPs for seven climatic variables; Table S2: Other genes and proteins identified by Genome-Environment Association with seven climatic variables.

**Author Contributions:** Conceptualization, data curation, formal analysis, investigation and writing—original draft, A.L.G.-E.; Conceptualization, methodology, project administration, supervision, writing—review & editing, M.H.R.-V.; Data curation, funding acquisition, supervision, C.P.S.; Funding acquisition, investigation, methodology, supervision, writing-review & editing, J.B.; Methodology, supervision, writing-review & editing, F.H.T.; and Investigation, writing-review & editing, A.B.-M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was implemented by CIMMYT as part of the Seeds of Discovery (SeeD) Initiative in collaboration with Universidad Autónoma Agraria Antonio Narro (UAAAN), made possible by the generous support of Sustainable Modernization of Traditional Agriculture (MasAgro) Project funded by the Ministry of Agriculture, and Rural Development (SADER) of the Government of Mexico. Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of SADER.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** We are grateful to the Consejo Nacional de Ciencia y Tecnología (CONACYT, México) for the scholarship provided to the first author for her doctoral studies.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

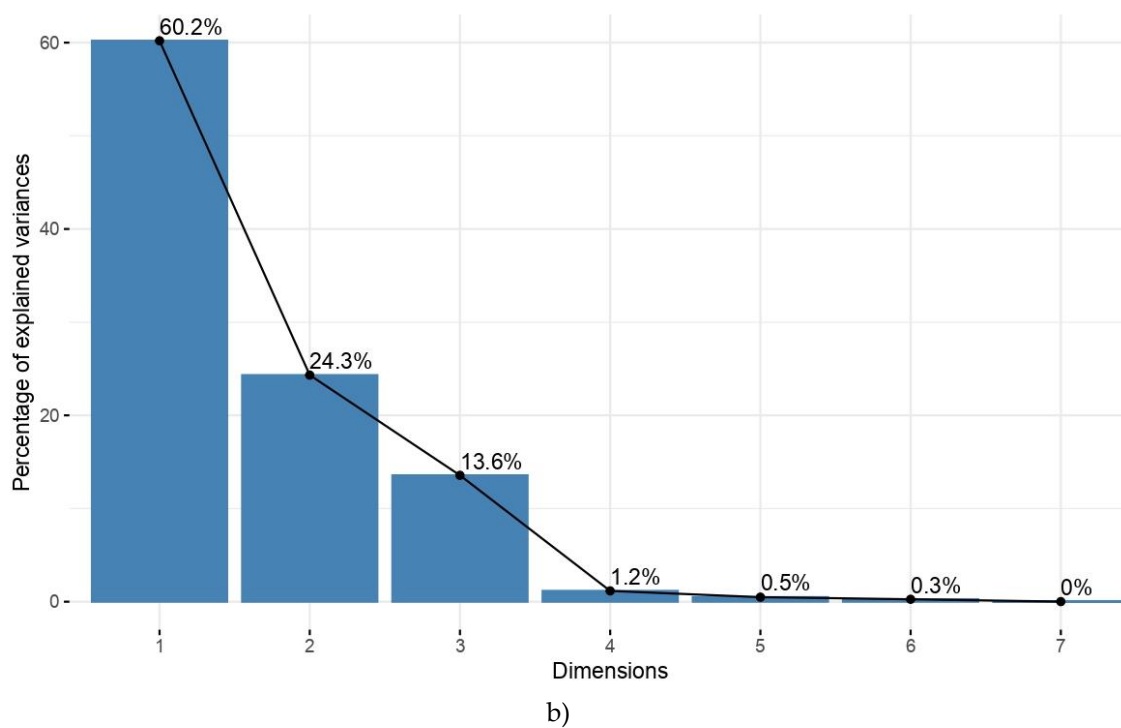
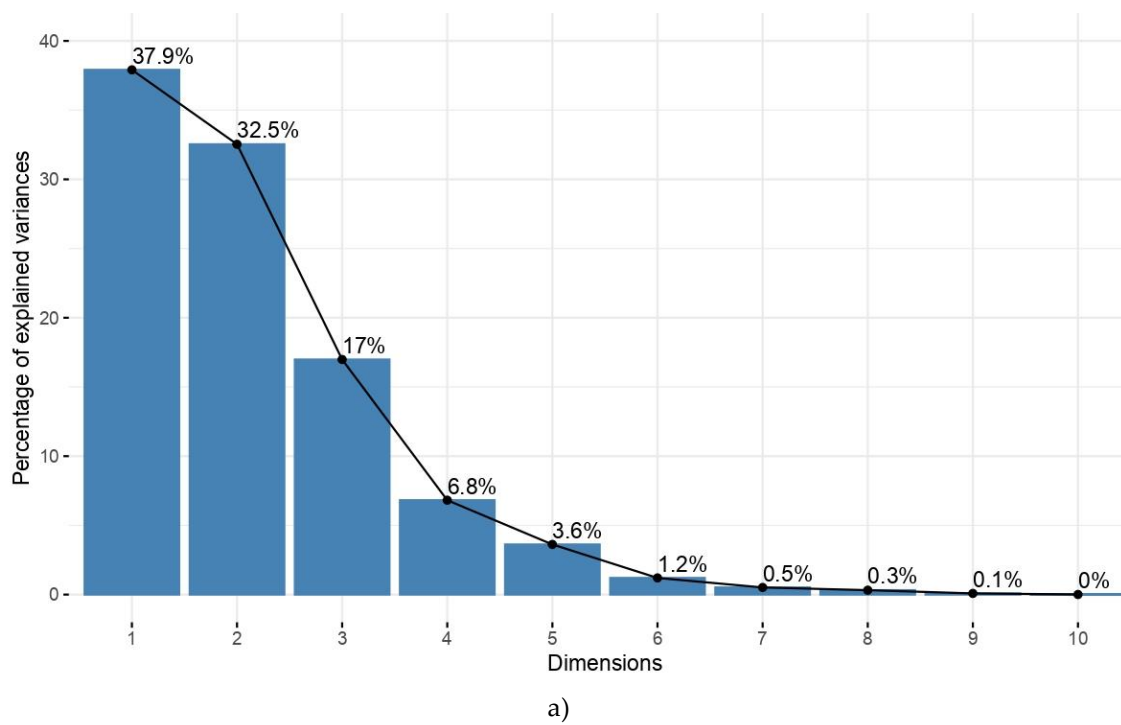
1. Chao, S.; Dubcovsky, J.; Dvorak, J.; Luo, M.-C.; Baenziger, S.P.; Matnyazov, R.; Clark, D.R.; Talbert, L.E.; Anderson, J.A.; Dreisigacker, S.; et al. Population- and Genome-Specific Patterns of Linkage Disequilibrium and SNP Variation in Spring and Winter Wheat (*Triticum Aestivum* L.). *BMC Genomics* **2010**, *11*, 727, doi:10.1186/1471-2164-11-727.
2. Lopes, M.S.; El-Basyoni, I.; Baenziger, P.S.; Singh, S.; Royo, C.; Ozbek, K.; Aktas, H.; Ozer, E.; Ozdemir, F.; Manickavelu, A.; et al. Exploiting Genetic Diversity from Landraces in Wheat Breeding for Adaptation to Climate Change. *J. Exp. Bot.* **2015**, *66*, 3477–3486, doi:10.1093/jxb/erv122.
3. Nazim Ud Dowla, M.A.N.; Edwards, I.; O'Hara, G.; Islam, S.; Ma, W. Developing Wheat for Improved Yield and Adaptation under a Changing Climate: Optimization of a Few Key Genes. *Engineering* **2018**, *4*, 514–522, doi:10.1016/j.eng.2018.06.005.
4. Challinor, A.J.; Watson, J.; Lobell, D.B.; Howden, S.M.; Smith, D.R.; Chhetri, N. A Meta-Analysis of Crop Yield under Climate Change and Adaptation. *Nat. Clim. Change* **2014**, *4*, 287–291, doi:10.1038/nclimate2153.
5. Corrado, G.; Rao, R. Towards the Genomic Basis of Local Adaptation in Landraces. *Diversity* **2017**, *9*, 51, doi:10.3390/d9040051.
6. Sansaloni, C.; Franco, J.; Santos, B.; Percival-Alwyn, L.; Singh, S.; Petrolí, C.; Campos, J.; Dreher, K.; Payne, T.; Marshall, D.; et al. Diversity Analysis of 80,000 Wheat Accessions Reveals Consequences

- and Opportunities of Selection Footprints. *Nat. Commun.* **2020**, *11*, 4572, doi:10.1038/s41467-020-18404-w.
7. Contreras-Moreira, B.; Serrano-Notivoli, R.; Mohammed, N.E.; Cantalapiedra, C.P.; Beguería, S.; Casas, A.M.; Igartua, E. Genetic Association with High-resolution Climate Data Reveals Selection Footprints in the Genomes of Barley Landraces across the Iberian Peninsula. *Mol. Ecol.* **2019**, *28*, 1994–2012, doi:10.1111/mec.15009.
  8. Pavlidis, P.; Jensen, J.D.; Stephan, W. Searching for Footprints of Positive Selection in Whole-Genome SNP Data from Nonequilibrium Populations. *Genetics* **2010**, *185*, 907–922, doi:10.1534/genetics.110.116459.
  9. Tang, Y.; Liu, X.; Wang, J.; Li, M.; Wang, Q.; Tian, F.; Su, Z.; Pan, Y.; Liu, D.; Lipka, A.E.; et al. GAPIT Version 2: An Enhanced Integrated Tool for Genomic Association and Prediction. *Plant Genome* **2016**, *9*, doi:10.3835/plantgenome2015.11.0120.
  10. López-Hernández, F.; Cortés, A.J. Last-Generation Genome–Environment Associations Reveal the Genetic Basis of Heat Tolerance in Common Bean (*Phaseolus Vulgaris* L.). *Front. Genet.* **2019**, *10*, 954, doi:10.3389/fgene.2019.00954.
  11. Hancock, A.M.; Brachi, B.; Faure, N.; Horton, M.W.; Jarymowycz, L.B.; Sperone, F.G.; Toomajian, C.; Roux, F.; Bergelson, J. Adaptation to Climate across the *Arabidopsis Thaliana* Genome. *Science* **2011**, *334*, 83–86, doi:10.1126/science.1209244.
  12. Jones, F.C.; Grabherr, M.G.; Chan, Y.F.; Russell, P.; Mauceli, E.; Johnson, J.; Swofford, R.; Pirun, M.; Zody, M.C.; White, S.; et al. The Genomic Basis of Adaptive Evolution in Threespine Sticklebacks. *Nature* **2012**, *484*, 55–61, doi:10.1038/nature10944.
  13. Turner, T.L.; Bourne, E.C.; Von Wettberg, E.J.; Hu, T.T.; Nuzhdin, S.V. Population Resequencing Reveals Local Adaptation of *Arabidopsis Lyrata* to Serpentine Soils. *Nat. Genet.* **2010**, *42*, 260–263, doi:10.1038/ng.515.
  14. Yoder, J.B.; Stanton-Geddes, J.; Zhou, P.; Briskine, R.; Young, N.D.; Tiffin, P. Genomic Signature of Adaptation to Climate in *Medicago Truncatula*. *Genetics* **2014**, *196*, 1263–1275, doi:10.1534/genetics.113.159319.
  15. Lasky, J.R.; Upadhyaya, H.D.; Ramu, P.; Deshpande, S.; Hash, C.T.; Bonnette, J.; Juenger, T.E.; Hyma, K.; Acharya, C.; Mitchell, S.E.; et al. Genome–Environment Associations in Sorghum Landraces Predict Adaptive Traits. *Sci. Adv.* **2015**, *1*, e1400218, doi:10.1126/sciadv.1400218.
  16. Romero-Navarro, J.A.; Willcox, M.; Burgueño, J.; Romay, C.; Swarts, K.; Trachsel, S.; Preciado, E.; Terron, A.; Delgado, H.V.; Vidal, V.; et al. A Study of Allelic Diversity Underlying Flowering-Time Adaptation in Maize Landraces. *Nat. Genet.* **2017**, *49*, 476–480, doi:10.1038/ng.3784.
  17. Cortés, A.J.; Blair, M.W. Genotyping by Sequencing and Genome–Environment Associations in Wild Common Bean Predict Widespread Divergent Adaptation to Drought. *Front. Plant Sci.* **2018**, *9*, 128, doi:10.3389/fpls.2018.00128.
  18. Lei, L.; Poets, A.M.; Liu, C.; Wyant, S.R.; Hoffman, P.J.; Carter, C.K.; Shaw, B.G.; Li, X.; Muehlbauer, G.J.; Katagiri, F.; et al. Environmental Association Identifies Candidates for Tolerance to Low Temperature and Drought. *G3 GenesGenomesGenetics.* **2019**, *9*, 3423–3438, doi:10.1534/g3.119.400401.
  19. Li, Y.; Li, D.; Jiao, Y.; Schnable, J.C.; Li, Y.; Li, H.; Chen, H.; Hong, H.; Zhang, T.; Liu, B.; et al. Identification of Loci Controlling Adaptation in Chinese Soya Bean Landraces via a Combination of Conventional and Bioclimatic GWAS. *Plant Biotechnol. J.* **2020**, *18*, 389–401, doi:10.1111/pbi.13206.
  20. Rodriguez, M.; Scintu, A.; Posadinu, C.M.; Xu, Y.; Nguyen, C.V.; Sun, H.; Bitocchi, E.; Bellucci, E.; Papa, R.; Fei, Z.; et al. GWAS Based on RNA-Seq SNPs and High-Throughput Phenotyping Combined with Climatic Data Highlights the Reservoir of Valuable Genetic Diversity in Regional Tomato Landraces. *Genes.* **2020**, *11*, 1387, doi:10.3390/genes11111387.
  21. Sokolkova, A.; Bulyntsev, S.V.; Chang, P.L.; Carrasquilla-Garcia, N.; Igolkina, A.A.; Noujdina, N.V.; von Wettberg, E.; Vishnyakova, M.A.; Cook, D.R.; Nuzhdin, S.V.; et al. Genomic Analysis of Vavilov’s Historic Chickpea Landraces Reveals Footprints of Environmental and Human Selection. *Int. J. Mol. Sci.* **2020**, *21*, 3952, doi:10.3390/ijms21113952.
  22. Li, Y.; Cao, K.; Li, N.; Zhu, G.; Fang, W.; Chen, C.; Wang, X.; Guo, J.; Wang, Q.; Ding, T.; et al. Genomic Analyses Provide Insights into Peach Local Adaptation and Responses to Climate Change. *Genome Res.* **2021**, *31*, 592–606, doi:10.1101/gr.261032.120.
  23. Hanif, U.; Alipour, H.; Gul, A.; Jing, L.; Darvishzadeh, R.; Amir, R.; Munir, F.; Ilyas, M.K.; Ghafoor, A.; Siddiqui, S.U.; et al. Characterization of the Genetic Basis of Local Adaptation of Wheat Landraces from Iran and Pakistan Using Genome-wide Association Study. *Plant Genome.* **2021**, *14*, doi:10.1002/tpg2.20096.

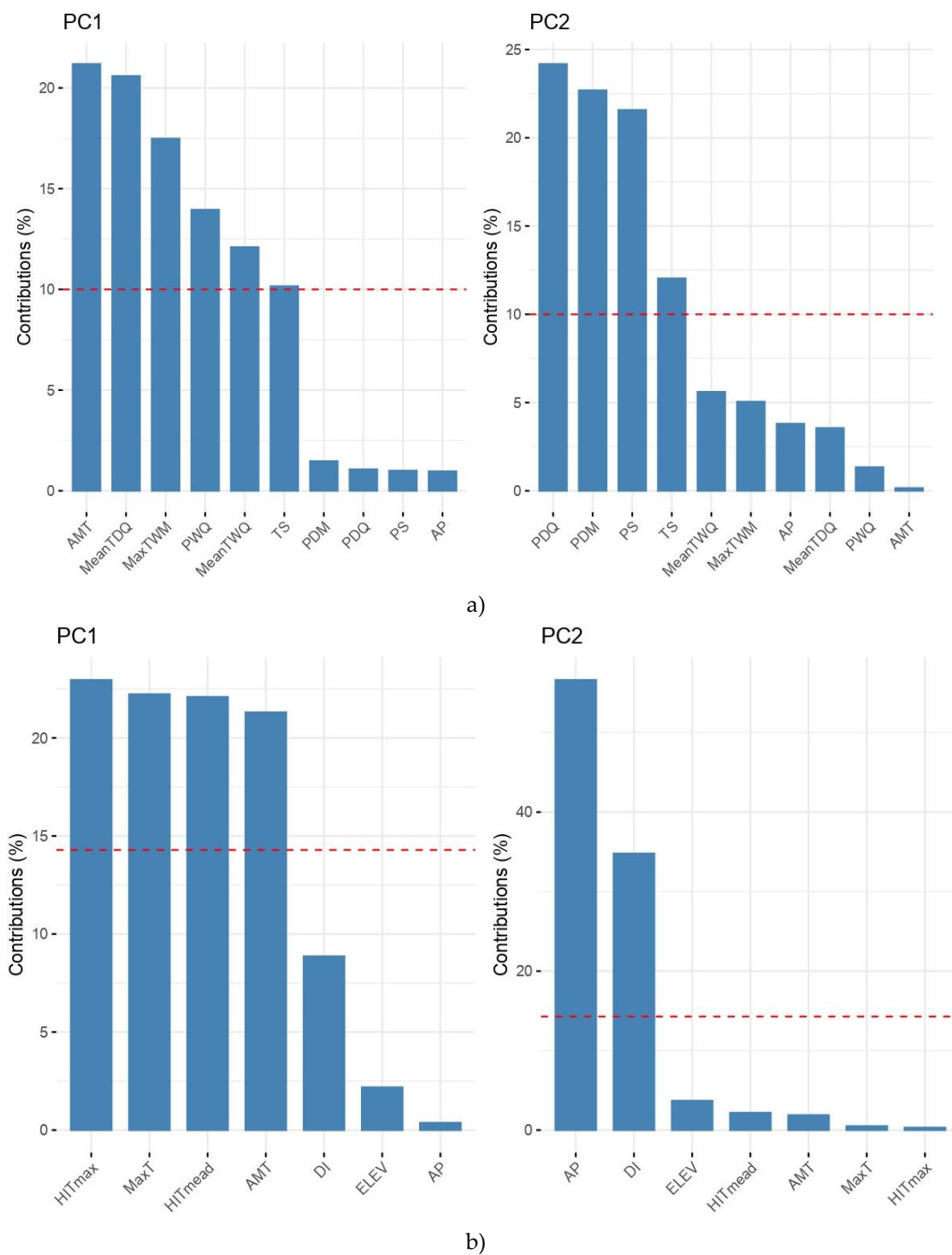
24. Pixley, K.V.; Salinas-Garcia, G.E.; Hall, A.; Kropff, M.; Ortiz, C.; Bouvet, L.C.; Suhalia, A.; Vikram, P.; Singh, S. CIMMYT's Seeds of Discovery Initiative: Harnessing Biodiversity for Food Security and Sustainable Development. *Indian J. Plant Genet. Resour.* 2018, 31, 1, doi:10.5958/0976-1926.2018.00001.3.
25. Vikram, P.; Franco, J.; Burgueño-Ferreira, J.; Li, H.; Sehgal, D.; Saint Pierre, C.; Ortiz, C.; Sneller, C.; Tattaris, M.; Guzman, C.; et al. Unlocking the Genetic Diversity of Creole Wheats. *Sci. Rep.* 2016, 6, 23092, doi:10.1038/srep23092.
26. Vangelis, H.; Tigkas, D.; Tsakiris, G. The Effect of PET Method on Reconnaissance Drought Index (RDI) Calculation. *J. Arid Environ.* 2013, 88, 130–140, doi:10.1016/j.jaridenv.2012.07.020.
27. O'Donnell, M.S.; Ignizio, D.A. *Bioclimatic Predictors for Supporting Ecological Applications in the Conterminous United States*; Data Series; U.S. Geological Survey, 2012.
28. Noce, S.; Caporaso, L.; Santini, M. A New Global Dataset of Bioclimatic Indicators. *Sci. Data.* 2020, 7, 398, doi:10.1038/s41597-020-00726-5.
29. Cortés, A.J.; Monserrate, F.A.; Ramírez-Villegas, J.; Madriñán, S.; Blair, M.W. Drought Tolerance in Wild Plant Populations: The Case of Common Beans (*Phaseolus Vulgaris* L.). *PLoS ONE* 2013, 8, e62898, doi:10.1371/journal.pone.0062898.
30. van der Schrier, G.; Jones, P.D.; Briffa, K.R. The Sensitivity of the PDSI to the Thornthwaite and Penman-Monteith Parameterizations for Potential Evapotranspiration. *J. Geophys. Res.* 2011, 116, D03106, doi:10.1029/2010JD015001.
31. Salehnia, N.; Alizadeh, A.; Sanaeinejad, H.; Bannayan, M.; Zarrin, A.; Hoogenboom, G. Estimation of Meteorological Drought Indices Based on AgMERRA Precipitation Data and Station-Observed Precipitation Data. *J. Arid Land* 2017, 9, 797–809, doi:10.1007/s40333-017-0070-y.
32. Khatun, K.; Debnath, S.; Robin, A.H.K.; Wai, A.H.; Nath, U.K.; Lee, D.-J.; Kim, C.-K.; Chung, M.-Y. Genome-Wide Identification, Genomic Organization, and Expression Profiling of the CONSTANS-like (COL) Gene Family in *Petunia* under Multiple Stresses. *BMC Genomics.* 2021, 22, 727, doi:10.1186/s12864-021-08019-w.
33. Qin, W.; Yu, Y.; Jin, Y.; Wang, X.; Liu, J.; Xi, J.; Li, Z.; Li, H.; Zhao, G.; Hu, W.; et al. Genome-Wide Analysis Elucidates the Role of CONSTANS-like Genes in Stress Responses of Cotton. *Int. J. Mol. Sci.* 2018, 19, 2658, doi:10.3390/ijms19092658.
34. Wong, A.C.S.; Hecht, V.F.G.; Picard, K.; Diwadkar, P.; Laurie, R.E.; Wen, J.; Mysore, K.; Macknight, R.C.; Weller, J.L. Isolation and Functional Analysis of *CONSTANS-LIKE* Genes Suggests That a Central Role for CONSTANS in Flowering Time Control Is Not Evolutionarily Conserved in *Medicago Truncatula*. *Front. Plant Sci.* 2014, 5, doi:10.3389/fpls.2014.00486.
35. Battaglia, M.; Covarrubias, A.A. Late Embryogenesis Abundant (LEA) Proteins in Legumes. *Front. Plant Sci.* 2013, 4, doi:10.3389/fpls.2013.00190.
36. Liu, H.; Xing, M.; Yang, W.; Mu, X.; Wang, X.; Lu, F.; Wang, Y.; Zhang, L. Genome-Wide Identification of and Functional Insights into the Late Embryogenesis Abundant (LEA) Gene Family in Bread Wheat (*Triticum Aestivum*). *Sci. Rep.* 2019, 9, 13375, doi:10.1038/s41598-019-49759-w.
37. Mertens, J.; Aliyu, H.; Cowan, D.A. LEA Proteins and the Evolution of the WHy Domain. *Appl. Environ. Microbiol.* 2018, 84, e00539-18, doi:10.1128/AEM.00539-18.
38. Hu, C.; Lin, S.; Chi, W.; Charng, Y. Recent Gene Duplication and Subfunctionalization Produced a Mitochondrial GrpE, the Nucleotide Exchange Factor of the Hsp70 Complex, Specialized in Thermotolerance to Chronic Heat Stress in *Arabidopsis*. *Plant Physiol.* 2012, 158, 747–758, doi:10.1104/pp.111.187674.
39. Su, P.-H.; Lin, H.-Y.; Lai, Y.-H. Two *Arabidopsis* Chloroplast GrpE Homologues Exhibit Distinct Biological Activities and Can Form Homo- and Hetero-Oligomers. *Front. Plant Sci.* 2020, 10, 1719, doi:10.3389/fpls.2019.01719.
40. El Habti, A.; Fleury, D.; Jewell, N.; Garnett, T.; Tricker, P.J. Tolerance of Combined Drought and Heat Stress Is Associated with Transpiration Maintenance and Water Soluble Carbohydrates in Wheat Grains. *Front. Plant Sci.* 2020, 11, 568693, doi:10.3389/fpls.2020.568693.
41. Qaseem, M.F.; Qureshi, R.; Muqaddasi, Q.H.; Shaheen, H.; Kousar, R.; Röder, M.S. Genome-Wide Association Mapping in Bread Wheat Subjected to Independent and Combined High Temperature and Drought Stress. *PLOS ONE* 2018, 13, e0199121, doi:10.1371/journal.pone.0199121.
42. Verma, V.; Ravindran, P.; Kumar, P.P. Plant Hormone-Mediated Regulation of Stress Responses. *BMC Plant Biol.* 2016, 16, 86, doi:10.1186/s12870-016-0771-y.
43. Miura, K.; Tada, Y. Regulation of Water, Salinity, and Cold Stress Responses by Salicylic Acid. *Front. Plant Sci.* 2014, 5, doi:10.3389/fpls.2014.00004.

44. Sabagh, A.E.; Mbarki, S.; Hossain, A.; Iqbal, M.A.; Islam, M.S.; Raza, A.; Llanes, A.; Reginato, M.; Rahman, M.A.; Mahboob, W.; et al. Potential Role of Plant Growth Regulators in Administering Crucial Processes against Abiotic Stresses. *Front. Agron.* 2021, 3, 648694, doi:10.3389/fagro.2021.648694.
45. Li, N.; Euring, D.; Cha, J.Y.; Lin, Z.; Lu, M.; Huang, L.-J.; Kim, W.Y. Plant Hormone-Mediated Regulation of Heat Tolerance in Response to Global Climate Change. *Front. Plant Sci.* 2021, 11, 627969, doi:10.3389/fpls.2020.627969.
46. Binder, B.M. Ethylene Signaling in Plants. *J. Biol. Chem.* 2020, 295, 7710–7725, doi:10.1074/jbc.REV120.010854.
47. Müller, M. Foes or Friends: ABA and Ethylene Interaction under Abiotic Stress. *Plants* 2021, 10, 448, doi:10.3390/plants10030448.
48. Khan, M.I.R.; Fatma, M.; Per, T.S.; Anjum, N.A.; Khan, N.A. Salicylic Acid-Induced Abiotic Stress Tolerance and Underlying Mechanisms in Plants. *Front. Plant Sci.* 2015, 6, doi:10.3389/fpls.2015.00462.
49. Wani, S.H.; Kumar, V.; Shriram, V.; Sah, S.K. Phytohormones and Their Metabolic Engineering for Abiotic Stress Tolerance in Crop Plants. *Crop J.* 2016, 4, 162–176, doi:10.1016/j.cj.2016.01.010.
50. Qiu, Z.; Guo, J.; Zhu, A.; Zhang, L.; Zhang, M. Exogenous Jasmonic Acid Can Enhance Tolerance of Wheat Seedlings to Salt Stress. *Ecotoxicol. Environ. Saf.* 2014, 104, 202–208, doi:10.1016/j.ecoenv.2014.03.014.
51. Su, Y.; Huang, Y.; Dong, X.; Wang, R.; Tang, M.; Cai, J.; Chen, J.; Zhang, X.; Nie, G. Exogenous Methyl Jasmonate Improves Heat Tolerance of Perennial Ryegrass through Alteration of Osmotic Adjustment, Antioxidant Defense, and Expression of Jasmonic Acid-Responsive Genes. *Front. Plant Sci.* 2021, 12, 664519, doi:10.3389/fpls.2021.664519.
52. Sharma, M.; Laxmi, A. Jasmonates: Emerging Players in Controlling Temperature Stress Tolerance. *Front. Plant Sci.* 2016, 6, doi:10.3389/fpls.2015.01129.
53. Lamaoui, M.; Jemo, M.; Datla, R.; Bekkaoui, F. Heat and Drought Stresses in Crops and Approaches for Their Mitigation. *Front. Chem.* 2018, 6, 26, doi:10.3389/fchem.2018.00026.
54. R Core Team R: *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2017; ISBN <https://www.r-project.org/>.
55. QGIS Development Team. *QGIS Las Palmas. Open Source Geospatial Foundation Project*; QGIS: Austin, TX, USA, 2017.
56. Balfourier, F.; Bouchet, S.; Robert, S.; De Oliveira, R.; Rimbart, H.; Kitt, J.; Choulet, F.; IWGSC; BreedWheat, C.; Paux, E. Worldwide Phylogeography and History of Wheat Genetic Diversity. *Sci. Adv.* 2019, 11.
57. Hijmans, R.J. *Raster: Geographic Data Analysis and Modeling*; R Package Version: Vienna, Austria, 2022;
58. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-km Spatial Resolution Climate Surfaces for Global Land Areas. *Int. J. Climatol.* 2017, 37, 4302–4315, doi:10.1002/joc.5086.
59. Chelsey, B.; Wheeler, N.R.; Rubel, F.; French, R.H. *Kgc: Koeppen-Geiger Climatic Zones*; R Package Version: Vienna, Austria, 2017.
60. Kottek, M.; Grieser, J.; Beck, C.; Rudolf, B.; Rubel, F. World Map of the Köppen-Geiger Climate Classification Updated. *Meteorol. Z.* 2006, 15, 259–263, doi:10.1127/0941-2948/2006/0130.
61. Thornthwaite, C.W. An Approach toward a Rational Classification of Climate. *Geogr. Rev.* 1948, 38, 55, doi:10.2307/210739.
62. Beguería, S.; Vicente-Serrano, S.M. *SPEI: Calculation of the Standardised Precipitation-Evapotranspiration Index*; R Core Team: Indianapolis, IN, USA, 2017.
63. Alboukadel, K.; Mundt, F. *Factoextra: Extract and Visualize the Results of Multivariate Data Analyses*; R Package Version: Vienna, Austria, 2020.
64. Frichot, E.; François, O. LEA: An R Package for Landscape and Ecological Association Studies. *Methods Ecol. Evol.* 2015, 6, 925–929, doi:10.1111/2041-210X.12382.
65. Wang, J.; Zhang, Z. GAPIT Version 3: Boosting Power and Accuracy for Genomic Association and Prediction. *Genomics Proteomics Bioinformatics.* 2021, 19, 629–640, doi:10.1016/j.gpb.2021.08.005.
66. Liu, X.; Huang, M.; Fan, B.; Buckler, E.S.; Zhang, Z. Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLoS Genet.* 2016, 12, e1005767, doi:10.1371/journal.pgen.1005767.
67. LiLin-Yin *CMplot: Circle Manhattan Plot*; R Package Version: Vienna, Austria, 2022.
68. IWGSC The International Wheat Genome Sequencing Consortium. *Science.* 2018, 361, eaar7191, doi:10.1126/science.aar7191.
69. Ouellette, L.A.; Reid, R.W.; Blanchard, S.G.; Brouwer, C.R. LinkageMapView - Rendering High-Resolution Linkage and QTL Maps. *BOINFP.* 2018, 34, 306–307, doi:10.1093/bioinformatics/btx576

## Supplementary Information



**Figure S1.** Contributions to the total variation of each component for the evaluated variables in the PCA. a) Bioclimatic variables related to aridity; b) Temperature related variables, calculated indices, AP, and ELEV.



**Figure S2.** Contributions to component principals 1 and 2 for the evaluated variables in the PCA. a) Bioclimatic variables related to aridity; b) Temperature related variables, calculated indices, AP, and ELEV.



**Table S1.** Detection of associated SNPs for seven climatic variables.

Var	Chr	No.	SNP	Physical position (bp)	<i>P-value</i>	SNPs	Genes
AMT	1A	1	108825112 F 0-19:T>A-19:T>A	277825486	6.60E-09	2	1
		2	108891114 F 0-33:A>G-33:A>G	388440316	5.89E-09		1
	1D	3	108739422 F 0-31:T>A-31:T>A	2113561	1.49E-06	1	1
	2D	4	108968402 F 0-23:C>T-23:C>T	79989524	9.63E-07	1	2
	3D	5	107698139 F 0-48:C>T-48:C>T	97153088	4.42E-07	2	1
		6	108308207 F 0-18:G>A-18:G>A	546672913	2.49E-07		2
	4D	7	109431634 F 0-54:G>A-54:G>A	370691683	5.90E-10	1	3
	5D	8	108305241 F 0-68:A>G-68:A>G	379130055	1.29E-06	1	1
						Total	12
MaxT	1A	1	108825112 F 0-19:T>A-19:T>A	277825486	8.02E-08	2	1
		2	108891114 F 0-33:A>G-33:A>G	388440316	1.32E-16		1
		3	106510612 F 0-30:C>T-30:C>T	502971897	2.05E-07		2
	3A	4	108476623 F 0-32:T>A-32:T>A	507114154	4.55E-07	3	2
		5	108028914 F 0-20:A>G-20:A>G	574516154	1.86E-07		2
	4D	6	109431634 F 0-54:G>A-54:G>A	370691683	1.09E-16	1	3
	5D	7	108652995 F 0-58:G>C-58:G>C	528660566	1.50E-08	1	1
	6A	8	109126792 F 0-49:C>T-49:C>T	5854616	1.23E-06	1	2
	6D	9	108582786 F 0-49:T>G-49:T>G	143959936	1.95E-12	1	2
	7A	10	108830300 F 0-9:T>G-9:T>G	64789408	3.05E-10	2	2
		11	109126469 F 0-35:G>A-35:G>A	662017143	8.41E-07		2
	7B	12	107878167 F 0-16:G>A-16:G>A	107522176	1.30E-06	1	2
						Total	22
AP	1A	1	108256081 F 0-24:C>G-24:C>G	588647447	1.25E-07	1	1
	2A	2	109058266 F 0-7:C>T-7:C>T	640221452	1.32E-07	1	2
	2B	3	107797386 F 0-11:T>G-11:T>G	91596926	1.14E-09	2	1
		4	107593284 F 0-50:T>A-50:T>A	768567114	3.32E-07		1
	4A	5	108145069 F 0-52:C>A-52:C>A	714179675	1.02E-06	1	2
	5B	6	109305833 F 0-34:C>G-34:C>G	526249013	1.56E-06	1	1
	6B	7	109177937 F 0-11:A>G-11:A>G	223441723	1.81E-07	2	1
		8	109523315 F 0-24:C>T-24:C>T	269818400	1.26E-06		0
						Total	9
PS	1D	1	107874524 F 0-37:G>C-37:G>C	485539732	1.89E-07	1	3
	2A	2	108020469 F 0-43:T>C-43:T>C	711971015	1.78E-08	2	2
		3	108514047 F 0-10:A>C-10:A>C	729345577	3.13E-08		2
	2B	4	107794074 F 0-23:T>C-23:T>C	76022529	9.09E-08	2	2
		5	109021888 F 0-34:T>C-34:T>C	795754781	1.02E-07		1
		6	108953425 F 0-21:A>G-21:A>G	50526285	1.78E-06		1
	3B	7	107601308 F 0-23:G>A-23:G>A	741467423	1.38E-09	3	0
		8	108516380 F 0-24:T>C-24:T>C	758205945	5.53E-10		1

Table S1. Continued.

Var	Chr	No.	SNP	Physical position (bp)	<i>P</i> -value	SNPs	Genes
PS	4B	9	106772473 F 0-11:T>G-11:T>G	39759168	4.41E-07	1	1
	5A	10	107001591 F 0-37:T>C-37:T>C	619468262	3.81E-07	1	2
	5B	11	108348543 F 0-58:A>G-58:A>G	548120559	2.48E-08	2	1
		12	109119258 F 0-68:T>C-68:T>C	577227418	4.69E-09		2
	6B	13	109354013 F 0-58:A>G-58:A>G	485290761	2.50E-08	1	2
	7A	14	108619258 F 0-17:C>T-17:C>T	552606647	1.35E-08	1	1
	7B	15	108981313 F 0-8:C>T-8:C>T	650581291	5.17E-08	1	1
						Total	22
HITmead	1A	1	108891114 F 0-33:A>G-33:A>G	388440316	9.11E-18	1	1
	1B	2	108853213 F 0-26:C>T-26:C>T	201700273	4.52E-07	1	1
	1D	3	107880401 F 0-66:C>T-66:C>T	413939387	1.53E-07	1	2
	2B	4	107488994 F 0-17:T>C-17:T>C	29129965	1.44E-08	1	3
	3B	5	108146856 F 0-65:T>A-65:T>A	764282419	2.77E-07	1	1
	3D	6	107698139 F 0-48:C>T-48:C>T	97153088	1.25E-08	1	1
	4A	7	109242168 F 0-13:A>G-13:A>G	598521397	4.39E-08	1	2
	4B	8	108145974 F 0-24:T>C-24:T>C	56276689	3.44E-07	1	1
		9	109363420 F 0-34:G>A-34:G>A	62303691	1.34E-08		2
	4D	10	108773963 F 0-37:C>T-37:C>T	97959582	1.59E-06	3	2
		11	109431634 F 0-54:G>A-54:G>A	370691683	1.50E-08		3
	5B	12	109240982 F 0-38:G>C-38:G>C	68925700	2.02E-12	2	1
		13	108773279 F 0-51:C>T-51:C>T	530916378	8.04E-07		2
	7A	14	107178221 F 0-24:T>G-24:T>G	628906923	1.54E-06	1	2
						Total	24
HITmax	1A	1	108825112 F 0-19:T>A-19:T>A	277825486	8.48E-09	2	1
		2	108891114 F 0-33:A>G-33:A>G	388440316	1.94E-11		1
	2B	3	108024073 F 0-47:C>T-47:C>T	38321752	3.81E-09	1	2
		4	108476623 F 0-32:T>A-32:T>A	507114154	1.68E-08		2
	3A	5	108028914 F 0-20:A>G-20:A>G	574516154	7.31E-11	1	2
		6	109122135 F 0-40:A>G-40:A>G	535641207	9.91E-07		1
	3D	7	108308207 F 0-18:G>A-18:G>A	546672913	4.37E-07	2	2
		8	109610030 F 0-20:C>T-20:C>T	575639014	6.69E-08		2
	4A	9	109242168 F 0-13:A>G-13:A>G	598521397	1.94E-09	1	2
	4D	10	109431634 F 0-54:G>A-54:G>A	370691683	4.78E-09	1	3
	6D	11	108582786 F 0-49:T>G-49:T>G	143959936	1.42E-09	1	2
						Total	20
DI	2B	1	108980638 F 0-33:T>C-33:T>C	584134131	2.76E-08	1	1
	2D	2	107489027 F 0-34:C>T-34:C>T	302776397	1.41E-07	1	1
	5B	3	109305833 F 0-34:C>G-34:C>G	526249013	4.31E-07	1	1
	7A	4	108206445 F 0-49:C>A-49:C>A	67617920	5.04E-08	2	1
		5	107952026 F 0-63:G>C-63:G>C	498861613	3.70E-07		1
	7D	6	109035950 F 0-29:C>T-29:C>T	88318125	1.89E-07	1	1
						Total	6

**Table S2.** Other genes and proteins identified by GEA with seven climatic variables.

Chr	SNP	Gene	Variable	Protein	Function
<b>Morphological responses</b>					
2B	107488994 F 0-17:T>C-17:T>C	TraesCS2B02G060700	HITmead	Fbox ( <i>FBX</i> ).	Regulation of multiple developmental processes.
2B	108980638 F 0-33:T>C-33:T>C	TraesCS2B02G410500	DI	Ovate Family Proteins ( <i>OFP</i> ).	Morphological regulation in response to stress signals.
4B	106772473 F 0-11:T>G-11:T>G	TraesCS4B02G051500	PS	Transcription factor SCREAM2 ( <i>SCRM2</i> ).	Constitutive stomatal differentiation in the epidermis.
4B	108145974 F 0-24:T>C-24:T>C	TraesCS4B02G064000	HITmead	Growth regulating factor 1 ( <i>GIF1</i> ).	Regulation of cell expansion in foliar tissues and cotyledons.
4D	108773963 F 0-37:C>T-37:C>T	TraesCS4D02G117100	HITmead	Fbox ( <i>FBX</i> ).	Regulation of multiple developmental processes.
6B	109177937 F 0-11:A>G-11:A>G	TraesCS6B02G190600	AP	ADP-ribosylation factor GTPase ( <i>AGD12</i> ).	Calcium-dependent proteins that affect gravitropism in <i>AT</i> .
6B	109354013 F 0-58:A>G-58:A>G	TraesCS6B02G269700	PS	Similar to OCL1 homeobox ( <i>ROC5</i> ).	Modulates the curling of the leaves.
<b>Responses to biotic stress</b>					
2A	108514047 F 0-10:A>C-10:A>C	TraesCS2A02G499400	PS	Pelota homolog ( <i>PELO</i> ).	Defense responses to bacteria.
3A	108028914 F 0-20:A>G-20:A>G	TraesCS3A02G329100	MaxT and HITmax	rRNA N-glycosidase ( <i>RIP</i> ).	Antibacterial and antifungal activity.
3B	108146856 F 0-65:T>A-65:T>A	TraesCS3B02G523300	HITmead	Disease resistance ( <i>NBS-LRR</i> ).	Disease resistance genes in plants.
3D	109610030 F 0-20:C>T-20:C>T	TraesCS3D02G475300	HITmax	Protein LIFEGUARD 4 ( <i>LFG4</i> ).	Defense response to pathogens.
4A	108145069 F 0-52:C>A-52:C>A	TraesCS4A02G447200	AP	Disease resistance ( <i>NBS-LRR</i> ).	Disease resistance genes in plants.
5B	108773279 F 0-51:C>T-51:C>T	TraesCS5B02G349800	HITmead	Flavin-containing monooxygenase ( <i>FMO</i> ).	Defense response to pathogens.
5B	109119258 F 0-68:T>C-68:T>C	TraesCS5B02G400100	PS	Cyclic nucleotide-gated ion channel 2 ( <i>CNGC2</i> ).	Defense response to pathogens.

Table S2. Continued.

Chr	SNP	Gene	Variable	Protein	Function
<b>Photosynthesis</b>					
2A	108020469 F 0-43:T>C-43:T>C	TraesCS2A02G467700	PS	(S)-2-hydroxy-acid oxidase ( <i>GLO1</i> ).	Photorespiration and regulation of photosynthesis.
2D	108968402 F 0-23:C>T-23:C>T	TraesCS2D02G136200	AMT	Magnesium-chelatase subunit, chloroplastic ( <i>CHLH</i> ).	Chlorophyll biosynthesis.
4D	109363420 F 0-34:G>A-34:G>A	TraesCS4D02G086900	HITmead	Rieske (2Fe-2S) domain ( <i>PSB33</i> ).	Assembly and stabilization of the PSII.
5D	107001591 F 0-37:T>C-37:T>C	TraesCS5A02G437100	PS	Photosystem II reaction center L ( <i>psbL</i> ).	The PSII core proteins.
7A	108830300 F 0-9:T>G-9:T>G	TraesCS7A02G107200	MaxT	Gamma carbonic anhydrase 1 ( $\gamma$ CA1).	Photorespiration.
7D	109035950 F 0-29:C>T-29:C>T	TraesCS7D02G138100	DI	Cytochrome b6-f complex subunit 5 ( <i>PETG</i> ).	Electron transfer between PSII and PSI.
<b>Flowering</b>					
1A	108891114 F 0-33:A>G-33:A>G	TraesCS1A02G220300	AMT, MaxT, HITmead and HITmax.	CONSTANS-like ( <i>COL</i> ).	Regulation of floral development.
2A	109058266 F 0-7:C>T-7:C>T	TraesCS2A02G390500	AP	Casein kinase II subunit beta ( <i>CK2</i> ).	Circadian Regulation.
2B	107488994 F 0-17:T>C-17:T>C	TraesCS2B02G059500	HITmead	Flowering promoting factor 1 ( <i>FPF1</i> ).	Control of flowering time.
3B	108953425 F 0-21:A>G-21:A>G	TraesCS3B02G081800	PS	Adenylyl-sulfate kinase ( <i>APK</i> ).	Circadian Regulation.
<b>Nutrient assimilation</b>					
2A	109058266 F 0-7:C>T-7:C>T	TraesCS2A02G390800	AP	Metal-nicotianamine transporter ( <i>YSL1</i> ).	Iron uptake and transport.
3A	108028914 F 0-20:A>G-20:A>G	TraesCS3A02G329700	MaxT and HITmax	Purple acid phosphatase ( <i>PAP</i> ).	Nutritional deficiencies.
7B	107878167 F 0-16:G>A-16:G>A	TraesCS7B02G093800	MaxT	Starch synthase, chloroplastic/amyloplastic ( <i>wSs2a-3</i> ).	Starch biosynthesis.

## CONCLUSIONES GENERALES

El cambio climático afecta la producción de cultivos de cereales en muchas partes del mundo debido a los aumentos en la frecuencia, intensidad e impactos de los fenómenos meteorológicos extremos, por lo que las tasas de adaptación de las especies deben ser aceleradas para asegurar su supervivencia.

En este sentido, la caracterización de una gran colección de variedades locales adaptadas a amplios regímenes climáticos ha permitido la identificación de regiones genómicas que revelan un amplio marco genético cuantitativo para la adaptabilidad del trigo local.

Nuestros resultados muestran el efecto de la selección natural sobre la estructura de las regiones genómicas en las poblaciones locales de trigo a lo largo del tiempo y la posibilidad de captar la variabilidad genética debida al ambiente a través de estudios de asociación (GWAS) con variables fenotípicas, como el hábito de crecimiento y con variables climáticas.

Observamos que las adaptaciones locales están presentes a lo largo de los 21 cromosomas del trigo en múltiples regiones genómicas. Estas regiones proporcionan información valiosa para comprender las funciones potenciales de nuevos genes en características de importancia agronómica, como la expresión del hábito de crecimiento y la tolerancia al estrés provocados por la sequía y las altas temperaturas.

En cuanto al hábito de crecimiento identificamos 20 regiones genómicas en 11 cromosomas con genes candidatos, cuyo estudio y caracterización puede ser base en la predicción genómica de dicha característica.

También encontramos 89 genes vitales que incluyen varios puntos críticos sobre los mecanismos de respuesta al estrés abiótico, destacando un número considerable de genes de señalización mediados por hormonas vegetales, procesos reguladores de la pared celular, cambios morfo-fisiológicos, fotosíntesis, floración y algunos mecanismos de respuesta al estrés abiótico.

Los hallazgos reportados constituyen un nuevo recurso para los bancos de germoplasma en la clasificación de accesiones en cuanto a su hábito de crecimiento y en la selección de colectas que portan alelos vinculados a respuestas climáticas específicas.

Los genes identificados podrían ser útiles en estudios genéticos funcionales enfocados en desarticular los mecanismos genéticos subyacentes a la tolerancia al estrés ambiental, como sequías prolongadas y altas temperaturas. Además, los loci significativos identificados requieren una validación adicional. Tanto los genes identificados y los SNP asociados después de la validación podrían utilizarse en la selección asistida por marcadores (SAM), para acelerar las ganancias genéticas en la adaptación del trigo harinero a los cambios climáticos.

## REFERENCIAS

- Acevedo, E., Silva, P., & Silva, H. (2002). Wheat growth and physiology. En *Bread wheat: Improvement and Production* (B.C. Curtis; S. Rajaram and H. Gómez Macpherson). F.A.O. Corporate Repository. <https://www.fao.org/3/y4011e/y4011e06.htm>
- Akter, N., & Islam, M. R. (2017). Heat stress effects and management in wheat. A review. *Agronomy for Sustainable Development*, 37(5), 37. <https://doi.org/10.1007/s13593-017-0443-9>
- Alqudah, A. M., Sallam, A., Stephen Baenziger, P., & Börner, A. (2020). GWAS: Fast-forwarding gene identification and characterization in temperate Cereals: lessons from Barley – A review. *Journal of Advanced Research*, 22, 119-135. <https://doi.org/10.1016/j.jare.2019.10.013>
- Anderson, C. A., Pettersson, F. H., Clarke, G. M., Cardon, L. R., Morris, A. P., & Zondervan, K. T. (2010). Data quality control in genetic case-control association studies. *Nature Protocols*, 5(9), 1564-1573. <https://doi.org/10.1038/nprot.2010.116>
- Barrett, B., Bayram, M., & Kidwell, K. (2002). Identifying AFLP and microsatellite markers for vernalization response gene *Vrn-B1* in hexaploid wheat using reciprocal mapping populations. *Plant Breeding*, 121(5), 400-406. <https://doi.org/10.1046/j.1439-0523.2002.732319.x>
- Beales, J., Turner, A., Griffiths, S., Snape, J. W., & Laurie, D. A. (2007). A *Pseudo-Response Regulator* is misexpressed in the photoperiod insensitive *Ppd-D1a* mutant of wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, 115(5), 721-733. <https://doi.org/10.1007/s00122-007-0603-4>
- Bendix, C., Marshall, C. M., & Harmon, F. G. (2015). Circadian clock genes universally control key agricultural traits. *Molecular Plant*, 8(8), 1135-1152. <https://doi.org/10.1016/j.molp.2015.03.003>
- Bush, W. S., & Moore, J. H. (2012). Chapter 11: Genome-Wide Association Studies. *PLoS Computational Biology*, 8(12), e1002822. <https://doi.org/10.1371/journal.pcbi.1002822>
- Challa, S., & Neelapu, N. R. R. (2018). Genome-Wide Association Studies (GWAS) for abiotic stress tolerance in plants. En *Biochemical, Physiological and Molecular Avenues for Combating Abiotic Stress Tolerance in Plants* (pp. 135-150). Elsevier. <https://doi.org/10.1016/B978-0-12-813066-7.00009-7>
- Charmet, G. (2011). Wheat domestication: Lessons for the future. *Comptes Rendus Biologies*, 334(3), 212-220. <https://doi.org/10.1016/j.crv.2010.12.013>

- Chen, J., Zavala, C., Ortega, N., Petroli, C., Franco, J., Burgueño, J., Costich, D. E., & Hearne, S. J. (2016). The development of quality control genotyping approaches: A case study using elite maize lines. *PLOS ONE*, *11*(6), e0157236. <https://doi.org/10.1371/journal.pone.0157236>
- Chen, L., Du, Y., Lu, Q., Chen, H., Meng, R., Cui, C., Lu, S., Yang, Y., Chai, Y., Li, J., Liu, L., Qi, X., Li, H., Mishina, K., Yu, F., & Hu, Y.-G. (2018). The photoperiod-insensitive allele *Ppd-D1a* promotes earlier flowering in *Rht12* dwarf plants of bread wheat. *Frontiers in Plant Science*, *9*, 1312. <https://doi.org/10.3389/fpls.2018.01312>
- Collevatti, R. G., Novaes, E., Silva-Junior, O. B., Vieira, L. D., Lima-Ribeiro, M. S., & Grattapaglia, D. (2019). A genome-wide scan shows evidence for local adaptation in a widespread keystone Neotropical forest tree. *Heredity*, *123*(2), 117-137. <https://doi.org/10.1038/s41437-019-0188-0>
- Contreras-Moreira, B., Serrano-Notivoli, R., Mohammed, N. E., Cantalapiedra, C. P., Beguería, S., Casas, A. M., & Igartua, E. (2019). Genetic association with high-resolution climate data reveals selection footprints in the genomes of barley landraces across the Iberian Peninsula. *Molecular Ecology*, *28*(8), 1994-2012. <https://doi.org/10.1111/mec.15009>
- Corrado, G., & Rao, R. (2017). Towards the genomic basis of local adaptation in landraces. *Diversity*, *9*(4), 51. <https://doi.org/10.3390/d9040051>
- Crossa, J., Jarquín, D., Franco, J., Pérez-Rodríguez, P., Burgueño, J., Saint-Pierre, C., Vikram, P., Sansaloni, C., Petroli, C., Akdemir, D., Sneller, C., Reynolds, M., Tattaris, M., Payne, T., Guzman, C., Peña, R. J., Wenzl, P., & Singh, S. (2016). Genomic prediction of gene bank wheat landraces. *G3 Genes/Genomes/Genetics*, *6*(7), 1819-1834. <https://doi.org/10.1534/g3.116.029637>
- Daloz, A. S., Rydsaa, J. H., Hodnebrog, Ø., Sillmann, J., van Oort, B., Mohr, C. W., Agrawal, M., Emberson, L., Stordal, F., & Zhang, T. (2021). Direct and indirect impacts of climate change on wheat yield in the Indo-Gangetic plain in India. *Journal of Agriculture and Food Research*, *4*, 100132. <https://doi.org/10.1016/j.jafr.2021.100132>
- Danquah, A., de Zelicourt, A., Colcombet, J., & Hirt, H. (2014). The role of ABA and MAPK signaling pathways in plant abiotic stress responses. *Biotechnology Advances*, *32*(1), 40-52. <https://doi.org/10.1016/j.biotechadv.2013.09.006>
- Danyluk, J., Kane, N. A., Breton, G., Limin, A. E., Fowler, D. B., & Sarhan, F. (2003). *TaVRT-1*, a putative transcription factor associated with vegetative to reproductive transition in cereals. *Plant Physiology*, *132*(4), 1849-1860. <https://doi.org/10.1104/pp.103.023523>
- De La Torre, A. R., Wilhite, B., & Neale, D. B. (2019). Environmental Genome-Wide Association reveals climate adaptation is shaped by subtle to moderate allele frequency shifts in loblolly pine. *Genome Biology and Evolution*, *11*(10), 2976-2989. <https://doi.org/10.1093/gbe/evz220>
- Dreisigacker, S., Burgueño, J., Pacheco, A., Molero, G., Sukumaran, S., Rivera-Amado, C., Reynolds, M., & Griffiths, S. (2021). Effect of flowering time-related genes on biomass, harvest index, and grain yield in CIMMYT elite spring bread wheat. *Biology*, *10*(9), 855. <https://doi.org/10.3390/biology10090855>
- Dubcovsky, J., Loukoianov, A., Fu, D., Valarik, M., Sanchez, A., & Yan, L. (2006). Effect of photoperiod on the regulation of wheat vernalization genes *VRN1* and *VRN2*.

- Plant Molecular Biology*, 60(4), 469-480. <https://doi.org/10.1007/s11103-005-4814-2>
- EL Sabagh, A., Islam, M. S., Hossain, A., Iqbal, M. A., Mubeen, M., Waleed, M., Reginato, M., Battaglia, M., Ahmed, S., Rehman, A., Arif, M., Athar, H.-U.-R., Ratnasekera, D., Danish, S., Raza, M. A., Rajendran, K., Mushtaq, M., Skalicky, M., Brestic, M., ... Abdelhamid, M. T. (2022). Phytohormones as growth regulators during abiotic stress tolerance in plants. *Frontiers in Agronomy*, 4, 765068. <https://doi.org/10.3389/fagro.2022.765068>
- Elias, J. C. F., Gonçalves-Vidigal, M. C., Ariani, A., Valentini, G., Martiniano-Souza, M. da C., Vaz Bisneta, M., & Gepts, P. (2021). Genome-Environment Association analysis for bio-climatic variables in common bean (*Phaseolus vulgaris* L.) from Brazil. *Plants*, 10(8), 1572. <https://doi.org/10.3390/plants10081572>
- Feldman, M., & Levy, A. A. (2012). Genome evolution due to allopolyploidization in wheat. *Genetics*, 192(3), 763-774. <https://doi.org/10.1534/genetics.112.146316>
- Gates, D. J., Runcie, D., Janzen, G. M., Navarro, A. R., Willcox, M., Sonder, K., Snodgrass, S. J., Rodríguez-Zapata, F., Sawers, R. J. H., Rellán-Álvarez, R., Buckler, E. S., Hearne, S., Hufford, M. B., & Ross-Ibarra, J. (2019). Single-gene resolution of locally adaptive genetic variation in Mexican maize [Preprint]. *Evolutionary Biology*. <https://doi.org/10.1101/706739>
- Giraldo, P., Benavente, E., Manzano-Agugliaro, F., & Gimenez, E. (2019). Worldwide research trends on wheat and barley: A bibliometric comparative analysis. *Agronomy*, 9(7), 352. <https://doi.org/10.3390/agronomy9070352>
- Gull, A., Ahmad Lone, A., & Ul Islam Wani, N. (2019). Biotic and abiotic stresses in plants. En A. Bosco de Oliveira (Ed.), *Abiotic and Biotic Stress in Plants*. IntechOpen. <https://doi.org/10.5772/intechopen.85832>
- Guo, X., & Elston, Robert C. (1999). Linkage information content of polymorphic genetic markers. *Human Heredity*, 49(2), 112-118. <https://doi.org/10.1159/000022855>
- Gupta, P., Kabbaj, H., El Hassouni, K., Maccaferri, M., Sanchez-Garcia, M., Tuberosa, R., & Bassi, F. M. (2020). Genomic regions associated with the control of flowering time in durum wheat. *Plants*, 9(12), 1628. <https://doi.org/10.3390/plants9121628>
- Gurgul, A., Semik, E., Pawlina, K., Szmatoła, T., Jasielczuk, I., & Bugno-Poniewierska, M. (2014). The application of genome-wide SNP genotyping methods in studies on livestock genomes. *Journal of Applied Genetics*, 55(2), 197-208. <https://doi.org/10.1007/s13353-014-0202-4>
- Guzmán, C., Ammar, K., Govindan, V., & Singh, R. (2019). Genetic improvement of wheat grain quality at CIMMYT. *Frontiers of Agricultural Science and Engineering*, 6(3), 265. <https://doi.org/10.15302/J-FASE-2019260>
- Haider, N. (2013). The origin of the B-genome of bread wheat (*Triticum aestivum* L.). *Russian Journal of Genetics*, 49(3), 263-274. <https://doi.org/10.1134/S1022795413030071>
- Hellwege, J. N., Keaton, J. M., Giri, A., Gao, X., Velez Edwards, D. R., & Edwards, T. L. (2017). Population stratification in Genetic Association Studies. *Current Protocols in Human Genetics*, 95(1). <https://doi.org/10.1002/cphg.48>
- Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., Li, C., Zhu, C., Lu, T., Zhang, Z., Li, M., Fan, D., Guo, Y., Wang, A., Wang, L., Deng, L., Li, W., Lu, Y., Weng, Q., ... Han, B. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature Genetics*, 42(11), 961-967. <https://doi.org/10.1038/ng.695>



- Hyles, J., Bloomfield, M. T., Hunt, J. R., Trethowan, R. M., & Trevaskis, B. (2020). Phenology and related traits for wheat adaptation. *Heredity*, *125*(6), 417-430. <https://doi.org/10.1038/s41437-020-0320-1>
- Ikegawa, S. (2012). A short history of the Genome-Wide Association Study: Where we were and where we are going. *Genomics & Informatics*, *10*(4), 220. <https://doi.org/10.5808/GI.2012.10.4.220>
- Jalmi, S. K., & Sinha, A. K. (2015). ROS mediated MAPK signaling in abiotic and biotic stress-striking similarities and differences. *Frontiers in Plant Science*, *6*. <https://doi.org/10.3389/fpls.2015.00769>
- Jaradat, A. (2013). Wheat Landraces: A mini review. *Emirates Journal of Food and Agriculture*, *25*(1), 20. <https://doi.org/10.9755/ejfa.v25i1.15376>
- Kamran, A., Iqbal, M., & Spaner, D. (2014). Flowering time in wheat (*Triticum aestivum* L.): A key factor for global adaptability. *Euphytica*, *197*(1), 1-26. <https://doi.org/10.1007/s10681-014-1075-7>
- Khumalo, T. P., Barnard, A., Maphosa, L., & Tsilo, T. J. (2017). Impact of growth habit and architecture genes on adaptation and performance of bread wheat. In R. Wanyera & J. Owuoch (Eds.), *Wheat Improvement, Management and Utilization*. InTech. <https://doi.org/10.5772/67116>
- Kippes, N., Debernardi, J. M., Vasquez-Gross, H. A., Akpınar, B. A., Budak, H., Kato, K., Chao, S., Akhunov, E., & Dubcovsky, J. (2015). Identification of the *VERNALIZATION 4* gene reveals the origin of spring growth habit in ancient wheats from South Asia. *Proceedings of the National Academy of Sciences*, *112*(39). <https://doi.org/10.1073/pnas.1514883112>
- Kulkarni, M., Soolanayakanahally, R., Ogawa, S., Uga, Y., Selvaraj, M. G., & Kagale, S. (2017). Drought response in wheat: Key genes and regulatory mechanisms controlling root system architecture and transpiration efficiency. *Frontiers in Chemistry*, *5*, 106. <https://doi.org/10.3389/fchem.2017.00106>
- Kumar, A., Sharma, S., Chunduri, V., Kaur, A., Kaur, S., Malhotra, N., Kumar, A., Kapoor, P., Kumari, A., Kaur, J., Sonah, H., & Garg, M. (2020). Genome-wide identification and characterization of heat shock protein family reveals role in development and stress conditions in *Triticum aestivum* L. *Scientific Reports*, *10*(1), 7858. <https://doi.org/10.1038/s41598-020-64746-2>
- Kumar, S., Sharma, V., Chaudhary, S., Tyagi, A., Mishra, P., Priyadarshini, A., & Singh, A. (2012). Genetics of flowering time in bread wheat *Triticum aestivum*: Complementary interaction between vernalization-insensitive and photoperiod-insensitive mutations imparts very early flowering habit to spring wheat. *Journal of Genetics*, *91*(1), 33-47. <https://doi.org/10.1007/s12041-012-0149-3>
- Lamaoui, M., Jemo, M., Datla, R., & Bekkaoui, F. (2018). Heat and drought stresses in crops and approaches for their mitigation. *Frontiers in Chemistry*, *6*, 26. <https://doi.org/10.3389/fchem.2018.00026>
- Laurie, C. C., Doheny, K. F., Mirel, D. B., Pugh, E. W., Bierut, L. J., Bhangale, T., Boehm, F., Caporaso, N. E., Cornelis, M. C., Edenberg, H. J., Gabriel, S. B., Harris, E. L., Hu, F. B., Jacobs, K. B., Kraft, P., Landi, M. T., Lumley, T., Manolio, T. A., McHugh, C., ... for the GENEVA Investigators. (2010). Quality control and quality assurance in genotypic data for genome-wide association studies. *Genetic Epidemiology*, *34*(6), 591-602. <https://doi.org/10.1002/gepi.20516>

- Ledesma-Ramírez, L., Solís-Moya, E., Iturriaga, G., Sehgal, D., Reyes-Valdes, M. H., Montero-Tavera, V., Sansaloni, C. P., Burgueño, J., Ortiz, C., Aguirre-Mancilla, C. L., Ramírez-Pimentel, J. G., Vikram, P., & Singh, S. (2019). GWAS to identify genetic loci for resistance to yellow rust in wheat pre-breeding lines derived from diverse exotic crosses. *Frontiers in Plant Science*, *10*, 1390. <https://doi.org/10.3389/fpls.2019.01390>
- Li, H., Peng, Z., Yang, X., Wang, W., Fu, J., Wang, J., Han, Y., Chai, Y., Guo, T., Yang, N., Liu, J., Warburton, M. L., Cheng, Y., Hao, X., Zhang, P., Zhao, J., Liu, Y., Wang, G., Li, J., & Yan, J. (2013). Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nature Genetics*, *45*(1), 43-50. <https://doi.org/10.1038/ng.2484>
- Li, H., Vikram, P., Singh, R. P., Kilian, A., Carling, J., Song, J., Burgueno-Ferreira, J. A., Bhavani, S., Huerta-Espino, J., Payne, T., Sehgal, D., Wenzl, P., & Singh, S. (2015). A high density GBS map of bread wheat and its application for dissecting complex disease resistance traits. *BMC Genomics*, *16*(1), 216. <https://doi.org/10.1186/s12864-015-1424-5>
- Li, L., Liu, B., Olsen, K. M., & Wendel, J. F. (2015). A re-evaluation of the homoploid hybrid origin of *Aegilops tauschii*, the donor of the wheat D-subgenome. *New Phytologist*, *208*(1), 4-8. <https://doi.org/10.1111/nph.13294>
- Li, Y., Xiong, H., Guo, H., Zhou, C., Xie, Y., Zhao, L., Gu, J., Zhao, S., Ding, Y., & Liu, L. (2020). Identification of the vernalization gene *VRN-B1* responsible for heading date variation by QTL mapping using a RIL population in wheat. *BMC Plant Biology*, *20*(1), 331. <https://doi.org/10.1186/s12870-020-02539-5>
- Lodhi, S. S., Maryam, S., Rafique, K., Shafique, A., Yousaf, Z. A., Talha, A. M., Gul, A., & Amir, R. (2020). Overview of the prospective strategies for conservation of genomic diversity in wheat landraces. En *Climate Change and Food Security with Emphasis on Wheat* (pp. 293-309). Elsevier. <https://doi.org/10.1016/B978-0-12-819527-7.00021-2>
- Lopes, M. S., El-Basyoni, I., Baenziger, P. S., Singh, S., Royo, C., Ozbek, K., Aktas, H., Ozer, E., Ozdemir, F., Manickavelu, A., Ban, T., & Vikram, P. (2015). Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. *Journal of Experimental Botany*, *66*(12), 3477-3486. <https://doi.org/10.1093/jxb/erv122>
- López-Hernández, F., & Cortés, A. J. (2019). Last-generation Genome–Environment Associations reveal the genetic basis of heat tolerance in common bean (*Phaseolus vulgaris* L.). *Frontiers in Genetics*, *10*, 954. <https://doi.org/10.3389/fgene.2019.00954>
- Maccaferri, M., Sanguineti, M. C., Corneti, S., Ortega, J. L. A., Salem, M. B., Bort, J., DeAmbrogio, E., del Moral, L. F. G., Demontis, A., El-Ahmed, A., Maalouf, F., Machlab, H., Martos, V., Moragues, M., Motawaj, J., Nachit, M., Nserallah, N., Ouabbou, H., Royo, C., ... Tuberosa, R. (2008). Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. *Genetics*, *178*(1), 489-511. <https://doi.org/10.1534/genetics.107.077297>
- Nakano, Y., & Kobayashi, Y. (2020). Genome-wide Association Studies of agronomic traits consisting of field- and molecular-based phenotypes. *Reviews in Agricultural Science*, *8*(0), 28-45. [https://doi.org/10.7831/ras.8.0\\_28](https://doi.org/10.7831/ras.8.0_28)

- Nazim Ud Dowla, M. A. N., Edwards, I., O'Hara, G., Islam, S., & Ma, W. (2018). Developing wheat for improved yield and adaptation under a changing climate: Optimization of a few key genes. *Engineering*, 4(4), 514-522. <https://doi.org/10.1016/j.eng.2018.06.005>
- Nezhadahmadi, A., Prodhani, Z. H., & Faruq, G. (2013). Drought tolerance in wheat. *The Scientific World Journal*, 2013, 1-12. <https://doi.org/10.1155/2013/610721>
- Ni, Z., Li, H., Zhao, Y., Peng, H., Hu, Z., Xin, M., & Sun, Q. (2018). Genetic improvement of heat tolerance in wheat: Recent progress in understanding the underlying molecular mechanisms. *The Crop Journal*, 6(1), 32-41. <https://doi.org/10.1016/j.cj.2017.09.005>
- Ochagavía, H., Prieto, P., Zikhali, M., Griffiths, S., & Slafer, G. A. (2019). Earliness *Per Se* by temperature interaction on wheat development. *Scientific Reports*, 9(1), 2584. <https://doi.org/10.1038/s41598-019-39201-6>
- Pasam, R. K., Sharma, R., Malosetti, M., van Eeuwijk, F. A., Haseneyer, G., Kilian, B., & Graner, A. (2012). Genome-wide association studies for agronomical traits in a world wide spring barley collection. *BMC Plant Biology*, 12(1), 16. <https://doi.org/10.1186/1471-2229-12-16>
- Pavlidis, P., Jensen, J. D., & Stephan, W. (2010). Searching for footprints of positive selection in Whole-Genome SNP data from nonequilibrium populations. *Genetics*, 185(3), 907-922. <https://doi.org/10.1534/genetics.110.116459>
- Peleg, Z., Fahima, T., Korol, A. B., Abbo, S., & Saranga, Y. (2011). Genetic analysis of wheat domestication and evolution under domestication. *Journal of Experimental Botany*, 62(14), 5051-5061. <https://doi.org/10.1093/jxb/err206>
- Peng, J., Ronin, Y., Fahima, T., Röder, M. S., Li, Y., Nevo, E., & Korol, A. (2003). Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat. *Proceedings of the National Academy of Sciences*, 100(5), 2489-2494. <https://doi.org/10.1073/pnas.252763199>
- Pixley, K. V., Salinas-Garcia, G. E., Hall, A., Kropff, M., Ortiz, C., Bouvet, L. C., Suhalia, A., Vikram, P., & Singh, S. (2018). CIMMYT's Seeds of Discovery Initiative: Harnessing biodiversity for food security and sustainable development. *Indian Journal of Plant Genetic Resources*, 31(1), 1. <https://doi.org/10.5958/0976-1926.2018.00001.3>
- Pluess, A. R., Frank, A., Heiri, C., Lalagüe, H., Vendramin, G. G., & Oddou-Muratorio, S. (2016). Genome–environment association study suggests local adaptation to climate at the regional scale in *Fagus sylvatica*. *New Phytologist*, 210(2), 589-601. <https://doi.org/10.1111/nph.13809>
- Qaseem, M. F., Qureshi, R., Muqaddasi, Q. H., Shaheen, H., Kousar, R., & Röder, M. S. (2018). Genome-wide association mapping in bread wheat subjected to independent and combined high temperature and drought stress. *PLOS ONE*, 13(6), e0199121. <https://doi.org/10.1371/journal.pone.0199121>
- Rahman, S., Islam, S., Yu, Z., She, M., Nevo, E., & Ma, W. (2020). Current progress in understanding and recovering the wheat genes lost in evolution and domestication. *International Journal of Molecular Sciences*, 21(16), 5836. <https://doi.org/10.3390/ijms21165836>
- Rasheed, A., & Xia, X. (2019). From markers to genome-based breeding in wheat. *Theoretical and Applied Genetics*, 132(3), 767-784. <https://doi.org/10.1007/s00122-019-03286-4>

- Reed, E., Nunez, S., Kulp, D., Qian, J., Reilly, M. P., & Foulkes, A. S. (2015). A guide to genome-wide association analysis and post-analytic interrogation. *Statistics in Medicine*, *34*(28), 3769-3792. <https://doi.org/10.1002/sim.6605>
- Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*(17), 4348-4370. <https://doi.org/10.1111/mec.13322>
- Robbana, C., Kehel, Z., Ben Naceur, M., Sansaloni, C., Bassi, F., & Amri, A. (2019). Genome-Wide genetic diversity and population structure of Tunisian durum wheat landraces based on DArTseq technology. *International Journal of Molecular Sciences*, *20*(6), 1352. <https://doi.org/10.3390/ijms20061352>
- Rodriguez, M., Scintu, A., Posadinu, C. M., Xu, Y., Nguyen, C. V., Sun, H., Bitocchi, E., Bellucci, E., Papa, R., Fei, Z., Giovannoni, J. J., Rau, D., & Attene, G. (2020). GWAS Based on RNA-Seq SNPs and High-Throughput Phenotyping Combined with Climatic Data Highlights the Reservoir of Valuable Genetic Diversity in Regional Tomato Landraces. *Genes*, *11*(11), 1387. <https://doi.org/10.3390/genes11111387>
- Romero-Navarro, J. A., Willcox, M., Burgueño, J., Romay, C., Swarts, K., Trachsel, S., Preciado, E., Terron, A., Delgado, H. V., Vidal, V., Ortega, A., Banda, A. E., Montiel, N. O. G., Ortiz-Monasterio, I., Vicente, F. S., Espinoza, A. G., Atlin, G., Wenzl, P., Hearne, S., & Buckler, E. S. (2017). A study of allelic diversity underlying flowering-time adaptation in maize landraces. *Nature Genetics*, *49*(3), 476-480. <https://doi.org/10.1038/ng.3784>
- Royo, C., Dreisigacker, S., Ammar, K., & Villegas, D. (2020). Agronomic performance of durum wheat landraces and modern cultivars and its association with genotypic variation in vernalization response (*Vrn-1*) and photoperiod sensitivity (*Ppd-1*) genes. *European Journal of Agronomy*, *120*, 126129. <https://doi.org/10.1016/j.eja.2020.126129>
- Saint Pierre, C., Burgueño, J., Crossa, J., Fuentes Dávila, G., Figueroa López, P., Solís Moya, E., Ireta Moreno, J., Hernández Muela, V. M., Zamora Villa, V. M., Vikram, P., Mathews, K., Sansaloni, C., Sehgal, D., Jarquin, D., Wenzl, P., & Singh, S. (2016). Genomic prediction models for grain yield of spring bread wheat in diverse agro-ecological zones. *Scientific Reports*, *6*(1), 27312. <https://doi.org/10.1038/srep27312>
- Sakuma, S., Salomon, B., & Komatsuda, T. (2011). The domestication syndrome genes responsible for the major changes in plant form in the *Triticeae* crops. *Plant and Cell Physiology*, *52*(5), 738-749. <https://doi.org/10.1093/pcp/pcr025>
- Sallam, A., Alqudah, A. M., Dawood, M. F. A., Baenziger, P. S., & Börner, A. (2019). Drought stress tolerance in wheat and barley: Advances in physiology, breeding and genetics research. *International Journal of Molecular Sciences*, *20*(13), 3137. <https://doi.org/10.3390/ijms20133137>
- Sansaloni, C., Franco, J., Santos, B., Percival-Alwyn, L., Singh, S., Petroli, C., Campos, J., Dreher, K., Payne, T., Marshall, D., Kilian, B., Milne, I., Raubach, S., Shaw, P., Stephen, G., Carling, J., Pierre, C. S., Burgueño, J., Crosa, J., ... Pixley, K. (2020). Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints. *Nature Communications*, *11*(1), 4572. <https://doi.org/10.1038/s41467-020-18404-w>

- Scherer, A., & Christensen, G. B. (2016). Concepts and relevance of Genome-Wide Association Studies. *Science Progress*, 99(1), 59-67. <https://doi.org/10.3184/003685016X14558068452913>
- Sehgal, D., Vikram, P., Sansaloni, C. P., Ortiz, C., Pierre, C. S., Payne, T., Ellis, M., Amri, A., Petroli, C. D., Wenzl, P., & Singh, S. (2015). Exploring and mobilizing the gene bank biodiversity for wheat improvement. *PLOS ONE*, 10(7), e0132112. <https://doi.org/10.1371/journal.pone.0132112>
- Shaw, P. D., Raubach, S., Hearne, S. J., Dreher, K., Bryan, G., McKenzie, G., Milne, I., Stephen, G., & Marshall, D. F. (2017). Germinate 3: Development of a common platform to support the distribution of experimental data on crop wild relatives. *Crop Science*, 57(3), 1259-1273. <https://doi.org/10.2135/cropsci2016.09.0814>
- Shenoda, J. E., Sanad, M. N. M. E., Rizkalla, A. A., El-Assal, S., Ali, R. T., & Hussein, M. H. (2021). Effect of long-term heat stress on grain yield, pollen grain viability and germinability in bread wheat (*Triticum aestivum* L.) under field conditions. *Heliyon*, 7(6), e07096. <https://doi.org/10.1016/j.heliyon.2021.e07096>
- Singh, S., Vikram, P., Sehgal, D., Burgueño, J., Sharma, A., Singh, S. K., Sansaloni, C. P., Joynson, R., Brabbs, T., Ortiz, C., Solis-Moya, E., Govindan, V., Gupta, N., Sidhu, H. S., Basandrai, A. K., Basandrai, D., Ledesma-Ramires, L., Suaste-Franco, M. P., Fuentes-Dávila, G., ... Pixley, K. V. (2018). Harnessing genetic potential of wheat germplasm banks through impact-oriented-prebreeding for future food and nutritional security. *Scientific Reports*, 8(1), 12527. <https://doi.org/10.1038/s41598-018-30667-4>
- Spor, A., Roucou, A., Mounier, A., Bru, D., Breuil, M.-C., Fort, F., Vile, D., Roumet, P., Philippot, L., & Violle, C. (2020). Domestication-driven changes in plant traits associated with changes in the assembly of the rhizosphere microbiota in tetraploid wheat. *Scientific Reports*, 10(1), 12234. <https://doi.org/10.1038/s41598-020-69175-9>
- 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>
- Tang, Y., Liu, X., Wang, J., Li, M., Wang, Q., Tian, F., Su, Z., Pan, Y., Liu, D., Lipka, A. E., Buckler, E. S., & Zhang, Z. (2016). GAPIT Version 2: An Enhanced Integrated Tool for Genomic Association and Prediction. *The Plant Genome*, 9(2). <https://doi.org/10.3835/plantgenome2015.11.0120>
- Tehseen, M. M., Tonk, F. A., Tosun, M., Amri, A., Sansaloni, C. P., Kurtulus, E., Yazbek, M., Al-Sham'aa, K., Ozseven, I., Safdar, L. B., Shehadeh, A., & Nazari, K. (2021). Genome-wide association study of resistance to PstS2 and Warrior races of *Puccinia striiformis* f. *Sp. Tritici* (stripe rust) in bread wheat landraces. *The Plant Genome*, 14(1). <https://doi.org/10.1002/tpg2.20066>
- Trethowan, R. M., Morgunov, A., He, Z., De Pauw, R., Crossa, J., Warburton, M., Baytasov, A., Zhang, C., Mergoum, M., & Alvarado, G. (2006). The global adaptation of bread wheat at high latitudes. *Euphytica*, 152(3), 303-316. <https://doi.org/10.1007/s10681-006-9217-1>
- Trevaskis, B. (2010). The central role of the *VERNALIZATION1* gene in the vernalization response of cereals. *Functional Plant Biology*, 37(6), 479. <https://doi.org/10.1071/FP10056>

- Trevaskis, B. (2015). Wheat gene for all seasons. *Proceedings of the National Academy of Sciences*, *112*(39), 11991-11992. <https://doi.org/10.1073/pnas.1516398112>
- Trevaskis, B. (2018). Developmental pathways are blueprints for designing successful crops. *Frontiers in Plant Science*, *9*, 745. <https://doi.org/10.3389/fpls.2018.00745>
- Trevaskis, B., Hemming, M. N., Dennis, E. S., & Peacock, W. J. (2007). The molecular basis of vernalization-induced flowering in cereals. *Trends in Plant Science*, *12*(8), 352-357. <https://doi.org/10.1016/j.tplants.2007.06.010>
- Turkmen, A. S., Yuan, Y., & Billor, N. (2019). Evaluation of methods for adjusting population stratification in genome-wide association studies: Standard versus categorical principal component analysis. *Annals of Human Genetics*, *83*(6), 454-464. <https://doi.org/10.1111/ahg.12339>
- Turuspekov, Y., Baibulatova, A., Yermekbayev, K., Tokhetova, L., Chudinov, V., Sereda, G., Ganal, M., Griffiths, S., & Abugalieva, S. (2017). GWAS for plant growth stages and yield components in spring wheat (*Triticum aestivum* L.) harvested in three regions of Kazakhstan. *BMC Plant Biology*, *17*(S1), 190. <https://doi.org/10.1186/s12870-017-1131-2>
- United Nations, Department of Economic and Social Affairs, & Population Division. (2019). World population prospects 2019: Highlights. (ST/ESA/SER.A/423).
- Valizadeh, J., Ziaei, S. M., & Mazlounzadeh, S. M. (2014). Assessing climate change impacts on wheat production (a case study). *Journal of the Saudi Society of Agricultural Sciences*, *13*(2), 107-115. <https://doi.org/10.1016/j.jssas.2013.02.002>
- Verma, V., Ravindran, P., & Kumar, P. P. (2016). Plant hormone-mediated regulation of stress responses. *BMC Plant Biology*, *16*(1), 86. <https://doi.org/10.1186/s12870-016-0771-y>
- Vikram, P., Franco, J., Burgueño-Ferreira, J., Li, H., Sehgal, D., Saint Pierre, C., Ortiz, C., Sneller, C., Tattaris, M., Guzman, C., Sansaloni, C. P., Ellis, M., Fuentes-Davila, G., Reynolds, M., Sonder, K., Singh, P., Payne, T., Wenzl, P., Sharma, A., ... Singh, S. (2016). Unlocking the genetic diversity of creole wheats. *Scientific Reports*, *6*(1), 23092. <https://doi.org/10.1038/srep23092>
- Weale, M. E. (2010). Quality control for Genome-Wide Association Studies. En M. R. Barnes & G. Breen (Eds.), *Genetic Variation* (Vol. 628, pp. 341-372). Humana Press. [https://doi.org/10.1007/978-1-60327-367-1\\_19](https://doi.org/10.1007/978-1-60327-367-1_19)
- Whittal, A., Kaviani, M., Graf, R., Humphreys, G., & Navabi, A. (2018). Allelic variation of vernalization and photoperiod response genes in a diverse set of North American high latitude winter wheat genotypes. *PLOS ONE*, *13*(8), e0203068. <https://doi.org/10.1371/journal.pone.0203068>
- Xiao, D., Bai, H., & Liu, D. (2018). Impact of future climate change on wheat production: A simulated case for China's wheat system. *Sustainability*, *10*(4), 1277. <https://doi.org/10.3390/su10041277>
- Xu, Y., Zhan, C., & Huang, B. (2011). Heat shock proteins in association with heat tolerance in grasses. *International Journal of Proteomics*, *2011*, 1-11. <https://doi.org/10.1155/2011/529648>
- Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S., & Dubcovsky, J. (2006). The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proceedings of the National Academy of Sciences*, *103*(51), 19581-19586. <https://doi.org/10.1073/pnas.0607142103>

- Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., Bennetzen, J. L., Echenique, V., & Dubcovsky, J. (2004). The Wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science*, *303*(5664), 1640-1644. <https://doi.org/10.1126/science.1094305>
- Yang, F. P., Zhang, X. K., Xia, X. C., Laurie, D. A., Yang, W. X., & He, Z. H. (2009). Distribution of the photoperiod insensitive *Ppd-D1a* allele in Chinese wheat cultivars. *Euphytica*, *165*(3), 445-452. <https://doi.org/10.1007/s10681-008-9745-y>
- Yoshida, T., Nishida, H., Zhu, J., Nitcher, R., Distelfeld, A., Akashi, Y., Kato, K., & Dubcovsky, J. (2010). *Vrn-D4* is a vernalization gene located on the centromeric region of chromosome 5D in hexaploid wheat. *Theoretical and Applied Genetics*, *120*(3), 543-552. <https://doi.org/10.1007/s00122-009-1174-3>
- Zhang, X. K., Xiao, Y. G., Zhang, Y., Xia, X. C., Dubcovsky, J., & He, Z. H. (2008). Allelic variation at the vernalization genes *Vrn-A1*, *Vrn-B1*, *Vrn-D1*, and *Vrn-B3* in Chinese wheat cultivars and their association with growth habit. *Crop Science*, *48*(2), 458-470. <https://doi.org/10.2135/cropsci2007.06.0355>
- Zhao, J., Boerwinkle, E., & Xiong, M. (2005). An entropy-based statistic for Genomewide Association Studies. *The American Journal of Human Genetics*, *77*(1), 27-40. <https://doi.org/10.1086/431243>
- Zikhali, M., & Griffiths, S. (2015). The effect of earliness per se (*Eps*) genes on flowering time in bread wheat. In Y. Ogihara, S. Takumi, & H. Handa (Eds.), *Advances in Wheat Genetics: From Genome to Field* (pp. 339-345). Springer Japan. [https://doi.org/10.1007/978-4-431-55675-6\\_39](https://doi.org/10.1007/978-4-431-55675-6_39)
- Zikhali, M., Wingen, L. U., & Griffiths, S. (2016). Delimitation of the *Earliness per se D1* (*Eps-D1*) flowering gene to a subtelomeric chromosomal deletion in bread wheat (*Triticum aestivum*). *Journal of Experimental Botany*, *67*(1), 287-299. <https://doi.org/10.1093/jxb/erv458>