



CENTRO DE INVESTIGACIONES BIOLÓGICAS
DEL NOROESTE, S.C.

Programa de Estudios de Posgrado

ANÁLISIS COMPARATIVO DE LAS SECUENCIAS DE
PROMOTORES DE LA FAMILIA DE FACTORES DE
TRANSCRIPCIÓN DOF EN MAÍZ

T E S I S

Que para obtener el grado de

Doctor en Ciencias

Uso, Manejo y Preservación de los Recursos Naturales
(Orientación en Biotecnología)

P r e s e n t a

Héctor Daniel Garza Avelar

La Paz, Baja California Sur, septiembre de 2020

ACTA DE LIBERACIÓN DE TESIS

En la Ciudad de La Paz, B. C. S., siendo las 9 horas del día 7 del Mes de septiembre del 2020, se procedió por los abajo firmantes, miembros de la Comisión Revisora de Tesis avalada por la Dirección de Estudios de Posgrado y Formación de Recursos Humanos del Centro de Investigaciones Biológicas del Noroeste, S. C., a liberar la Tesis de Grado titulada:

“ANÁLISIS COMPARATIVO DE LAS SECUENCIAS DE PROMOTORES DE LA FAMILIA DE FACTORES DE TRANSCRIPCIÓN DOF EN MAÍZ”

Presentada por el alumno:

Héctor Daniel Garza Avelar

Aspirante al Grado de DOCTOR EN CIENCIAS EN EL USO, MANEJO Y PRESERVACIÓN DE LOS RECURSOS NATURALES CON ORIENTACIÓN EN Biotecnología

Después de intercambiar opiniones los miembros de la Comisión manifestaron su **APROBACIÓN DE LA TESIS**, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISIÓN REVISORA



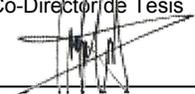
Dra. Gracia Alicia Gómez Anduro
Co-Directora de Tesis



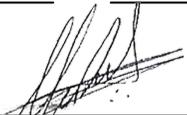
Dr. Luis Herrera Estrella
Co-Director de Tesis



Dr. Felipe de Jesús Ascencio Valle
Co-Tutor



Dr. César Salvador Cardona Félix
Co-Tutor



Dr. Mario Rojas Arzaluz
Co-Tutor



Dra. Gracia Alicia Gómez Anduro,
Directora de Estudios de Posgrado y
Formación de Recursos Humanos

Conformación de comités

Comité Tutorial

Dra. Gracia Alicia Gómez Anduro
Centro de Investigaciones Biológicas del Noroeste, S.C.
Co-Directora de Tesis

Dr. Luis Herrera Estrella
Laboratorio Nacional de Genómica para la Biodiversidad
Co-Director de Tesis

Dr. Felipe de Jesús Ascencio Valle
Centro de Investigaciones Biológicas del Noroeste, S.C.
Co-Tutor

Dr. César Cardona Félix
Centro Interdisciplinario de Ciencias Marinas - IPN
Co-Tutor

Dr. Mario Rojas Arzaluz
BioEra
Co-Tutor

Comité Revisor de Tesis

Dra. Gracia Alicia Gómez Anduro
Dr. Luis Herrera Estrella
Dr. Felipe de Jesús Ascencio Valle
Dr. César Cardona Félix
Dr. Mario Rojas Arzaluz

Jurado de Examen

Dra. Gracia Alicia Gómez Anduro
Dr. Luis Herrera Estrella
Dr. Felipe de Jesús Ascencio Valle
Dr. César Cardona Félix
Dr. Mario Rojas Arzaluz

Suplentes

Dra. Elizabeth Monreal Escalante
Dra. Liliana Carolina Rojo Arreola

Resumen

Uno de los principales mecanismos por los cuales un organismo sufre cambios fenotípicos es la modificación en las secuencias de ADN encargadas de regular la expresión de genes. Durante el proceso de domesticación del maíz, la mano del hombre influyó en la selección de genes y secuencias de promotores que han sido conservados y que pueden estar relacionados a la adaptabilidad de las plantas a diferentes condiciones. Por eso, es necesario identificar qué elementos regulatorios han sido conservados entre variedades silvestres y actuales. En el presente trabajo realizamos un análisis comparativo de las secuencias de promotores que regulan la expresión de la familia de factores de transcripción Dof en Maíz. Para eso, seleccionamos al ancestro silvestre del maíz actual, el Teosinte, a la raza ancestral de Palomero Toluqueño y a un maíz actual, B73. Se determinó la arquitectura de los promotores Dof y se identificaron los elementos regulatorios presentes a lo largo de las secuencias. También, se realizó un análisis de expresión del gen Dof01 de maíz Palomero bajo condiciones de estrés por temperatura baja y alta (4°C y 42°C, respectivamente), comparando los resultados contra datos de expresión de maíz B73 dispuestos en bases de datos. En ellas identificamos los elementos regulatorios *cis* responsables de la arquitectura del promotor, a qué familias de factores de transcripción pertenecen y los perfiles de expresión espacio-temporal de los mismos. Realizamos un análisis de correlación para identificar los factores de transcripción que regulan la expresión del gen Dof01. El análisis de secuencias mostró que hay un elemento de inserción/delección de 179 pb que está presente en el promotor Dof01 de B73, pero ausente en las secuencias de las otras dos variedades. Además, identificamos los elementos regulatorios *cis* que pueden estar relacionados a la respuesta a bajas temperaturas. El análisis comparativo de las secuencias del promotor Dof01 mostró que los elementos regulatorios están altamente conservados. Sin embargo, algunos sitios de unión a factores de transcripción presentes en el promotor de Teosinte no se encuentran en B73 y viceversa. Por otra parte, los datos de expresión del gen Dof01 de PAL muestran una sobreexpresión con respecto al control debido a la respuesta a frío, mientras que los datos de expresión del gen Dof01 de B73 muestran que este gen se reprime con respecto al control. Este trabajo es el primer paso para la caracterización funcional del promotor Dof01 de maíz Palomero y B73, para identificar los elementos regulatorios asociados a la respuesta a estrés por frío y la obtención del promotor mínimo necesario para regular la transcripción.

Palabras clave: Factor de transcripción, Promotor, Maíz, Domesticación



Vo.Bo. Dra. Gracia Gómez Anduro

Summary

One of the main mechanisms by which an organism undergoes phenotypic changes is the modification in the DNA sequences responsible for regulating the expression of genes. During the domestication process of maize, the hand of man influenced the selection of genes and promoter sequences that have been conserved and that may be related to the adaptability of plants to different conditions. For this reason, it is necessary to identify which regulatory elements have been conserved between wild and current varieties. In the present work we carry out a comparative analysis of the promoter sequences that regulate the expression of the Dof family of transcription factors in Corn. For that, we selected the wild ancestor of current maize, Teosinte, the ancestral race of Palomero Toluqueño and a current maize, B73. The architecture of the Dof promoters was determined and the regulatory elements present throughout the sequences were identified. Also, an expression analysis of the Dof01 gene of Palomero maize was performed under conditions of stress due to low and high temperature (4°C and 42°C, respectively), comparing the results against the expression data of B73 maize arranged in databases. In them we identify the *cis*-regulatory elements responsible for the promoter architecture, which families of transcription factors they belong to, and their spatio-temporal expression profiles. We performed a correlation analysis to identify the transcription factors that regulate the expression of the Dof01 gene. Also, we compared the expression data of the Dof01 gene taken from the B73 maize expression atlas under the same temperature stress conditions. Sequence analysis showed that there is a 179 bp insertion / deletion element that is present in the Dof01 promoter of B73, but absent in the sequences of the other two varieties. In addition, we identify the *cis* regulatory elements that may be related to the response to low temperatures. The comparative analysis of the sequences of the Dof01 promoter showed that the regulatory elements are highly conserved. However, some transcription factor binding sites present in the Teosinte promoter are not found in B73 and vice versa. On the other hand, the expression data of the PAL Dof01 gene show overexpression with respect to the control due to the response to cold, while the expression data of the B73 Dof01 gene show that this gene is repressed with respect to the control. This work is the first step for the functional characterization of the Dof01 promoter from Palomero and B73, to identify the regulatory elements associated with the response to cold stress and to obtain the minimum promoter necessary to regulate transcription.

KeyWords: Transcription Factor, Promoter, Maize, Domestication



Vo.Bo. Dra. Gracia Gómez Anduro

Dedicatoria

*A mis padres, por siempre trabajar tan duro para sacarnos adelante
y por formar a la persona en la que me he convertido.*

A Daniela, mi motivación, por estar en los momentos más difíciles.

A David, mi amigo, por tus palabras de apoyo y por siempre levantar mi ánimo.

A Mario, fuiste mi mentor y amigo, siempre confiaste en mí, D.E.P.

Agradecimientos

Al Consejo Nacional de Ciencia y Tecnología por otorgar la beca No. 428977 Para la realización de mis estudios de doctorado.

Al CIBNOR, por ser mí casa durante los últimos diez años

Al personal de Posgrado por siempre estar atentos a las necesidades que como estudiantes tenemos. En especial a Tania Núñez, Horacio Sandoval, Adriana Franzoni, Lic. Claudia Olachea, Lic. Leticia González y Lic. Osvelia Ibarra.

A mi comité tutorial por sus aportaciones en la realización de este trabajo.

Al Dr. Axel Tiessen, por las ideas aportadas a este trabajo, D.E.P.

Al Laboratorio de Biología Molecular de Plantas, en especial al cDr. Julio Hernández González, por su valiosa ayuda y amistad.

Al Laboratorio de Biotecnología Vegetal, en especial al M.C. Mario Arce Montoya, por su valiosa ayuda durante los ensayos experimentales y mentoría durante toda la tesis.

Al Laboratorio de Patogénesis Microbiana y su personal técnico

Al Laboratorio de Genética Molecular, en especial a la M.C. Delia Rojas por su apoyo en la realización de algunos experimentos.

A la Dra. Carolina Garciglia, amiga y compañera de laboratorio de quien siempre recibí apoyo.

A los biorockstars, Paty, Caro, Ivonne, Camilo, Marco y Daniel, por las sesiones de zoom para echarme porras y en quienes he encontrado una muy buena amistad.

Contenido	
Resumen	i
Summary	ii
Dedicatoria	iii
Agradecimientos	iv
Contenido	v
Lista de figuras	vii
Lista de tablas	viii
Abreviaturas	ix
1. INTRODUCCIÓN	1
2. ANTECEDENTES	4
3. JUSTIFICACIÓN	8
4. HIPÓTESIS	9
5. OBJETIVOS	10
5.1. Objetivo general	10
5.2. Objetivos particulares	10
6. MATERIAL Y MÉTODOS	11
6.1. Identificación de secuencias de promotores Dof de maíz	11
6.2. Análisis filogenético de promotores Dof	11
6.3. Identificación de sitios de unión a factores de transcripción	11
6.4. Densidad de la distribución de elementos regulatorios	12
6.5. Relación entre secuencias de promotores y familias de factores de transcripción	12
6.6. Identificación de regiones conservadas entre promotores de genes parálogos	13
6.7. Análisis de correlación	13
6.8. Material vegetal y tratamientos de estrés por temperatura en plántulas de maíz palomero	13
6.9. Aislamiento de RNA y análisis de expresión	14
7. RESULTADOS	15
7.1. Variación en la identidad de secuencias homólogas de promotores Dof .	15
7.2. Los promotores de genes parálogos divergen en funcionalidad	17
7.3. La caja TATA influye en la distribución de elementos regulatorios	19
7.4. Los elementos regulatorios <i>cis</i> no se distribuyen aleatoriamente	21

7.5. Tres familias de factores de transcripción están sobre representadas	22
7.6. Los promotores presentan regiones de elementos <i>cis</i> conservados entre variedades	24
7.7. El promotor del gen Dof01 de B73 presenta un elemento de inserción/delección con respecto a PAL y TEO	27
7.8. Los TFBS presentes en el promotor Dof01 de B73 están relacionados con la respuesta a estrés por temperatura	28
7.9. La respuesta de Dof01 a estrés por temperatura difiere entre B73 y PAL	32
8. DISCUSIÓN	34
9. CONCLUSIONES	44
10. LITERATURA CITADA	45
11. ANEXOS	56
Anexo A	56
Anexo B	58
Anexo C	62
Anexo D	66
Anexo E	99

Lista de figuras

Figura 1. Análisis filogenético de las secuencias de promotores Dof de maíz B73, palomero (PAL) y Teocinte (TEO).....	19
Figura 2. Distribución de densidad de elementos <i>cis</i> en promotores Dof de maíz..	20
Figura 3. Distribución de densidad de elementos regulatorios <i>cis</i> en promotores Dof clasificados por familia según la cadena a la que se unen.....	22
Figura 4. Ocurrencias de TFBS en promotores Dof por familia de factores de transcripción.....	23
Figura 5. Arquitectura de promotores Dof de tres variedades de maíz..	27
Figura 6. Análisis de correlación de Pearson entre la expresión de factores de transcripción y la expresión de Dof01 en B73.....	29
Figura 7. Datos de expresión de Dof01 en relación a otros factores de transcripción bajo tratamientos de estrés abiótico.	31
Figura 8. Expresión del gen Dof01 en plántulas de maíz Palomero bajo tratamientos de estrés por temperatura.	33

Lista de tablas

Tabla I. Alineamiento local entre las secuencias de promotores Dof de B73, palomero (PAL) y Teocinte (TEO).....	15
Tabla II. Nombre y familia a la que pertenecen los factores de transcripción asociados a la expresión de Dof01.	30
Tabla III. Lista de motivos localizados en secuencias de promotores Dof y simbología empleada para su identificación.	56
Tabla IV. Factores de transcripción que reconocen sitios de unión en el promotor Dof01 de B73, PAL y TEO.....	62

Abreviaturas

ADN Ácido Desoxirribonucleico

CRE *cis*-regulatory element (Elemento Regulatorio *cis*)

CDS CoDing Sequence (Secuencia Codificante)

Dof DNA With One Zinc Finger

PAL Palomero Toluqueño

TEO Teosinte

TFBS Transcription Factor Binding Site (Sitio de Unión a Factor de Transcripción)

TSS Transcription Start Site (Sitio de Inicio de la Transcripción)

UTR Untranslated Region (Región no Traducida)

1. INTRODUCCIÓN

La principal razón por la que las especies se adaptan de manera específica al ambiente donde se desarrollan, se debe a cambios evolutivos en la expresión de genes (Tirosh *et al.*, 2009). Estos cambios se deben principalmente a mutaciones que afectan las regiones encargadas de regular la transcripción, como los elementos regulatorios *cis* a lo largo de la secuencia de un promotor. Los cambios en los elementos regulatorios de genes que codifican proteínas funcionalmente conservadas provocan cambios fenotípicos en las especies afectadas (Lemmon *et al.*, 2014; Wittkopp y Kalay, 2012). Las plantas terrestres evolucionaron hace 450 millones de años de las algas ancestrales, cada evento evolutivo se caracteriza por la expansión de genes que caracterizan nuevas rutas metabólicas o el incremento en la diversidad de algunas familias de factores de transcripción (Bowman *et al.*, 2017).

Una de las familias de genes que divergieron antes de las algas verdes y las plantas terrestres es la familia de factores de transcripción Dof (Del inglés DNA with One Zinc Finger) (Shigyo *et al.*, 2007). Estos factores de transcripción son específicos de plantas y poseen un dominio de unión al ácido desoxirribonucleico (ADN) altamente conservado (Yanagisawa, 2004). Las proteínas Dof funcionan como activadores o represores de la transcripción y están involucrados en la regulación de diversos procesos biológicos (Moreno-Risueno *et al.*, 2007; Yanagisawa, 2004). El primer miembro de esta familia de genes fue identificado en maíz (Yanagisawa, 1995; Yanagisawa y Izui, 1990, 1993) y en 2015 se descubrieron 46 miembros que codifican para esta familia de factores de transcripción en el genoma de maíz B73 (Chen y Cao, 2015).

El maíz (*Zea mays* ssp. *Mays*) es el segundo cultivo de mayor importancia a nivel mundial utilizado como fuente de alimento o para la elaboración de diversos productos industriales, como sorbitol, almidón, glucosa, jarabe de alta fructosa, monohidrato de dextrosa, entre otros. Además, el maíz es un organismo modelo para estudios de genética debido a su gran diversidad (Prasanna, 2012; N. Yang *et al.*, 2017). El maíz fue domesticado de su ancestro silvestre el Teosinte (*Z. mays*

ssp. *Mexicana*), hace aproximadamente 9000 años (J. Doebley y Stec, 1993; Prasanna, 2012). Además del Teosinte, existen en la actualidad razas consideradas ancestrales, como Palomero Toluqueño, una raza de polinización abierta que crece bajo condiciones de temperatura por debajo de los 15°C a 2000 msnm (Eagles y Lothrop, 1994). El genoma de maíz ha experimentado dos eventos de duplicación de su genoma, hace aproximadamente 20.5 y 11.4 millones de años (Gaut y Doebley, 1997). De acuerdo con Chen y Cao (2015), el genoma de maíz contiene 12 pares de genes parálogos tipo Dof que surgieron por eventos de duplicación.

Los modelos de evolución de genes predicen que existe una divergencia funcional entre genes duplicados a causa de la pérdida de elementos *cis* de sus promotores (Lockton y Gaut, 2005). Los estudios encaminados en determinar la divergencia funcional en familias de genes, están enfocados en analizar la divergencia en la expresión de genes o determinar cuán conservados están los elementos regulatorios en las regiones promotoras de diferentes especies de plantas (Nicole Marie Creux *et al.*, 2008; Freeling y Subramaniam, 2009; Lockton y Gaut, 2005; Powell *et al.*, 2019; Y. Wang *et al.*, 2012). Sin embargo, no hay estudios enfocados en determinar cómo los elementos regulatorios, presentes en las secuencias de los promotores de una familia de factores de transcripción, son conservados entre diferentes variedades de una misma especie y hasta qué nivel influyó el proceso de domesticación en el grado de conservación o divergencia de elementos regulatorios. Por eso, en el presente trabajo, se realizó un análisis comparativo entre las secuencias de promotores Dof de maíz B73, Teosinte y Palomero Toluqueño. Se determinó la arquitectura de los promotores Dof y se identificaron los elementos regulatorios presentes a lo largo de las secuencias. También, se realizó un análisis de expresión del gen Dof01 de maíz Palomero bajo condiciones de estrés por temperatura baja y alta (4°C y 42°C, respectivamente), comparando los resultados contra datos de expresión de maíz B73 dispuestos en bases de datos. Este estudio tiene como propósito revelar cómo el proceso de

domesticación de maíz influyó en la conservación de elementos regulatorios presentes en las secuencias de los promotores Dof en esta especie.

2. ANTECEDENTES

La domesticación de plantas silvestres ha provisto al hombre de cultivos que sirven principalmente a sus necesidades nutricionales. Se han documentado diversos cambios genéticos asociados a este proceso, los cuales residen principalmente en los elementos regulatorios *cis* que controlan la expresión (Swinnen *et al.*, 2016). Algunos estudios han determinado cómo el proceso de domesticación de cultivos de interés agrícola como el maíz, presentan cambios significativos en la expresión de sus genes con respecto a sus ancestros silvestres (Hufford *et al.*, 2012). De acuerdo a Swinnen y colaboradores (2016), existe una presión de selección de regiones regulatorias de elementos *cis*, ocasionada por el hombre durante el proceso de domesticación. Esto ha ocasionado que exista una menor variación en la expresión, entre variedades silvestres y actuales, de genes que se consideran potencialmente seleccionados durante dicho proceso de domesticación (Swinnen *et al.*, 2016). En 2013 Meyer y Purugganan, determinaron que la mayoría de los cambios en elementos regulatorios *cis* derivados del proceso de domesticación han ocurrido en promotores que dirigen la expresión de factores de transcripción (J. F. Doebley *et al.*, 2006; Meyer y Purugganan, 2013). El proceso de domesticación promueve una evolución rápida del fenotipo a través de una selección artificial (Wright *et al.*, 2005). Las mutaciones o cambios evolutivos que afectan los elementos regulatorios contribuyen a la diversidad fenotípica entre especies y entre variedades de una misma especie (Wittkopp y Kalay, 2012).

La domesticación del maíz a partir de su progenitor silvestre, es considerado un sistema modelo para la investigación del efecto de la domesticación sobre la evolución del genoma y la respuesta a la selección (Swanson-Wagner *et al.*, 2012). En 2012, Swanson-Wagner y colaboradores analizaron la expresión de más de 18,000 genes de más de 50 genotipos de maíz y Teosinte. En este estudio encontraron evidencia de más de 600 genes con niveles de expresión significativamente distintos entre maíz y Teosinte. Por otra parte, encontraron que más de 1,000 genes mostraron evidencia de perfiles de coexpresión alteradas,

atribuyendo dichas diferencias al proceso de domesticación. Esto concuerda con la hipótesis sobre cuáles son los principales cambios en los genomas que son responsables de la evolución de las modificaciones morfológicas en un organismo, atribuyendo dichos cambios a las alteraciones en la expresión de genes que codifican proteínas funcionales conservadas y que tales cambios ocurren en gran medida como alteraciones en las secuencias de elementos regulatorios *cis* (Carroll, 2008). También, se ha encontrado evidencia de una fuerte selección durante el proceso de domesticación del maíz que llevó a su mejoramiento. Además, de encontrar pérdida de elementos regulatorios *cis* como efecto del proceso de domesticación (Hufford *et al.*, 2012).

A pesar de que el tiempo de divergencia es relativamente corto, el maíz exhibe diferencias fenotípicas sustanciales en la arquitectura de la planta, la morfología y el tamaño de la mazorca y la adaptabilidad a diferentes climas con respecto al Teosinte, lo que refleja cambios evolutivos rápidos y pronunciados derivados de las modificaciones en elementos regulatorios que tienen como efecto cambios en la expresión de genes (Huang *et al.*, 2016).

La expresión de un gen es un proceso complejo que es regulado a diferentes niveles, uno de ellos es la transcripción (Petrillo *et al.*, 2014). A nivel transcripcional, los promotores juegan un papel importante. Estos contienen elementos regulatorios *cis* que son conocidos como sitios de unión a factores de transcripción (TFBS, del inglés Transcription Factor Binding Sites). Los TFBS son secuencias de ADN cortas distribuidas a lo largo de la secuencia de un promotor y son reconocidas por factores de transcripción. En los organismos eucariotas superiores, la arquitectura de los genes, así como las propiedades estructurales de los promotores, juegan un papel importante en la determinación de la variación del número de transcritos en un organismo (nivel de expresión) y la expresión de un gen en un tejido específico (amplitud de expresión), lo que finalmente determina el fenotipo (Das y Bansal, 2019). Entender la arquitectura de los promotores de plantas es importante para el entendimiento de la evolución de los promotores (Yamamoto, Ichida, Matsui, *et al.*, 2007). Los elementos regulatorios

cis se encuentran usualmente agrupados en regiones de los promotores donde los factores de transcripción pueden unirse como homo o heterodímeros para modular la transcripción de un gen (Nicky M. Creux *et al.*, 2013). Además, son usualmente compartidos por los promotores de genes coexpresados debido a una trans-regulación en común (Vandepoele *et al.*, 2009). Otra característica de las secuencias de elementos *cis* es su conservación en promotores ortólogos de diferentes especies y géneros, en comparación con las secuencias no codificantes que los flanquean (Freeling y Subramaniam, 2009). La presencia de elementos regulatorios *cis* conservados en secuencias de promotores en especies evolutivamente distantes puede brindar información sobre redes regulatorias ancestrales (Nicole Marie Creux *et al.*, 2008).

Las proteínas que tienen el dominio de unión al ADN Dof, son factores de transcripción específicos de plantas (Yanagisawa, 2002). Las proteínas Dof actúan como activadores o represores transcripcionales con múltiples roles en una variedad de procesos biológicos asociados con la asimilación del carbono, la maduración y germinación de semillas, la respuesta a hormonas y otros procesos relacionados a estrés biótico y abiótico (Cai *et al.*, 2013; Le Hir y Bellini, 2013; Ma *et al.*, 2015; H. Wang *et al.*, 2017). La primer proteína con dominio Dof fue caracterizada en maíz (Yanagisawa, 1995; Yanagisawa y Izui, 1990, 1993), desde entonces, han sido identificadas desde microalgas como *Chlamydomonas reinhardtii* hasta el musgo *Physcomitrella patens* (Moreno-Risueno *et al.*, 2007; X. Yang *et al.*, 2006). De acuerdo a Shigyo y colaboradores (2007), los factores de transcripción Dof surgieron previos a la divergencia de las algas verdes y los ancestros de las plantas terrestres. Estos genes han evolucionado a través de duplicaciones intensivas de un gen ancestral después de la divergencia del musgo y los linajes de angiospermas (Shigyo *et al.*, 2007)

Estudiar las regiones regulatorias localizadas en los promotores de los genes Dof de diferentes razas y/o variedades de maíz, puede proveer información sobre la evolución de los elementos regulatorios y la respuesta a condiciones de estrés. Actualmente existen estudios en los que se compara el nivel de conservación de

elementos regulatorios en genes duplicados de una misma especie o en familias de genes de diferentes especies de plantas (Nicky M. Creux *et al.*, 2013; Nicole Marie Creux *et al.*, 2008; Das y Bansal, 2019; Koch *et al.*, 2001; Lockton y Gaut, 2005; Powell *et al.*, 2019; Reineke *et al.*, 2011; Wittkopp y Kalay, 2012). Sin embargo, no se han realizado estudios en los cuales se analice cómo los elementos regulatorios son conservados entre distintas variedades de una misma especie. Por otro lado, Chen & Cao (2015) identificaron 46 genes Dof en el genoma de maíz B73 y analizaron sus perfiles de expresión espacio-temporal, identificaron cuáles son los genes parálogos y su ubicación en los cromosomas. También, analizaron 1000 pb de las secuencias de los promotores e identificaron los elementos regulatorios *cis* presentes, pero no realizan una comparación de la distribución de estos entre pares de promotores de genes homólogos.

3. JUSTIFICACIÓN

Se ha demostrado que los factores de transcripción Dof en maíz intervienen en diversos procesos relacionados con estrés biótico y abiótico en las plantas. Para ampliar el conocimiento sobre cómo se regula la expresión de estos genes, es necesario entender cómo el proceso de domesticación influyó en la conservación de elementos regulatorios en las secuencias de los promotores responsables de la actividad transcripcional de los mismos. De esta manera, se podrán identificar los elementos que actúan como respuesta a los estímulos medioambientales y que puedan ser considerados como interruptores moleculares en el desarrollo de arreglos de ADN para dirigir la expresión de genes que codifiquen proteínas con potencial aplicación biotecnológica.

4. HIPÓTESIS

Si los elementos regulatorios, que se localizan en secuencias de promotores de genes homólogos, son altamente conservados, entonces al comparar las secuencias de promotores Dof entre variedades silvestres, ancestrales y actuales de maíz, será posible identificar cómo el proceso de domesticación de este cultivo afectó a los elementos regulatorios presentes en las secuencias de promotores Dof en diferentes variedades.

5. OBJETIVOS

5.1. Objetivo general

Identificar cambios en los elementos regulatorios de las secuencias de los promotores Dof de tres variedades de maíz como efecto del proceso de domesticación de este cultivo.

5.2. Objetivos particulares

5.2.1. Identificar secuencias de los promotores Dof en los genomas de maíz B73, Palomero Toluqueño y Teosinte.

5.2.2. Determinar las relaciones filogenéticas entre las secuencias de los promotores Dof de maíz de las tres variedades a nivel de secuencia y conservación de elementos regulatorios.

5.2.3. Caracterizar la arquitectura de los promotores Dof para identificar los elementos regulatorios conservados durante el proceso de domesticación.

5.2.4. Identificar los factores de transcripción que potencialmente regulan la expresión del gen Dof01 en maíz B73 bajo condiciones normales de cultivo y por estrés abiótico.

5.2.5. Comparar la expresión del gen Dof01 bajo condiciones de estrés por temperatura entre B73 y Palomero Toluqueño.

6. MATERIAL Y MÉTODOS

6.1. Identificación de secuencias de promotores Dof de maíz

Se accedió a la base de datos Phytozome V12 (Goodstein *et al.*, 2012) para obtener las secuencias de los 46 promotores Dof reportados en el genoma de maíz B73. Se seleccionaron hasta 2000 pb río arriba del codón de inicio de la traducción (CDS). Se realizó un alineamiento local (BLAST) para la identificación de los promotores Dof en las bases de datos de genomas de maíz Palomero Toluqueño (PAL) y Teosinte (TEO).

6.2. Análisis filogenético de promotores Dof

Se tomaron las secuencias de los promotores Dof de B73, así como las secuencias de PAL y TEO que tuvieran al menos un porcentaje de cobertura, con respecto a su secuencia homóloga en B73, de al menos 25%. Se realizó un análisis filogenético entre las secuencias de los promotores Dof de las tres variedades de maíz. Se utilizó el software MEGA 7.0 (Tamura *et al.*, 2013) para construir un árbol filogenético con el método de Neighbor-Joining. El número de réplicas de bootstrap se ajustó a 1000 con un modelo gamma distribuido. Se estableció el método de eliminación pareada para todas las posiciones de nucleótidos ausentes (gaps) o datos incompletos. El árbol consenso fue exportado a formato Newick el cual se trabajó sobre la herramienta en línea iTOL (Letunic y Bork, 2016).

6.3. Identificación de sitios de unión a factores de transcripción

Las regiones promotoras de los genes Dof se analizaron con la herramienta en línea *multiple promoter tool* de PlantPAN 3.0 (Chow *et al.*, 2016). Se identificaron los probables sitios de unión de factores de transcripción distribuidos a lo largo de las secuencias de los promotores. Además, con la herramienta *TSS Plant* de Softberry (Shahmuradov *et al.*, 2017) se identificaron elementos regulatorios como múltiples sitios de inicio de la transcripción (TSS) y cajas TATA. Con la finalidad de entender las relaciones que existen entre los promotores de ambas variedades de maíz, según el número y tipo de factores de transcripción que reconocen sus

secuencias, se construyó un mapa de calor con ayuda del software R y el paquete *ggpubr* (Kassambara, 2020).

6.4. Densidad de la distribución de elementos regulatorios

La posición de los elementos regulatorios localizados en las secuencias se ajustó de manera relativa a la posición del TSS más cercano al CDS. Se graficó la distribución de densidad de elementos regulatorios con ayuda del software R (R Foundation for Statistical Computing, 2008) y el paquete *ggpubr* para cada promotor analizado. Se realizó una clasificación de los elementos regulatorios distribuidos según la hebra en la que se encuentran, la variedad de maíz en la que se unen, así como la clasificación del promotor según la presencia o ausencia de caja TATA. También, se graficó la distribución de densidad de las diferentes familias de factores de transcripción que reconocen sitios de unión en los promotores analizados y estos se clasificaron según la preferencia de unión de acuerdo con la hebra de ADN.

6.5. Relación entre secuencias de promotores y familias de factores de transcripción

Se construyó un mapa de calor para entender cómo se relacionan entre sí las secuencias de los promotores analizados según el número y tipo de familias de factores de transcripción que identifican sitios de unión en sus secuencias. Se utilizaron los datos obtenidos de la herramienta *multiple promoter tool* de PlantPAN 3.0 para identificar tanto la posición de unión en el promotor como la familia a la que pertenecen múltiples factores de transcripción identificados que reconocen un mismo sitio. El mapa de calor se construyó con el Software R y el paquete *ComplexHeatmap* (Gu *et al.*, 2016) utilizando los datos normalizados del número de ocurrencias de cada factor de transcripción en cada promotor. Se utilizó el método de “k-means” para el agrupamiento de los datos de promotores y factores de transcripción y la construcción de los dendrogramas correspondientes.

6.6. Identificación de regiones conservadas entre promotores de genes parálogos

Se realizó un análisis comparativo entre las secuencias de los promotores que logramos identificar en las tres variedades. Se representaron de manera gráfica las arquitecturas de los promotores Dof de PAL y TEO con respecto a su similitud a B73. Se utilizó la herramienta en línea *cross species* de PlantPAN 3.0 (Chang *et al.*, 2008) para identificar regiones conservadas entre las secuencias. En el gráfico se incluyeron los diferentes motivos relacionados a los TFBS, los múltiples TSSs y cajas TATA localizadas y las regiones que corresponden a las regiones no traducidas (UTR, del inglés Untranslated Region) 5' de cada promotor.

6.7. Análisis de correlación

Para comprender cómo los factores de transcripción afectan la expresión de Dof01 en B73, tomamos un conjunto de datos de expresión del atlas de Hoopes (Hoopes *et al.*, 2018). Extrajimos los valores de expresión transformados (\log_2) de Dof01 y los factores de transcripción relacionados con TFBS en la secuencia del promotor Dof01, en 79 tejidos diferentes. Hicimos una prueba de correlación de Pearson para encontrar los factores de transcripción que más afectan la expresión de Dof01 en B73 y seleccionamos aquellos con un valor absoluto superior a 0.65.

6.8. Material vegetal y tratamientos de estrés por temperatura en plántulas de maíz palomero

Se obtuvieron semillas de maíz Palomero Toluqueño, genotipo EDMX-2233 proporcionadas por el Centro Internacional de Mejoramiento del Maíz y el Trigo (CIMMYT). Las semillas se esterilizaron superficialmente sumergiéndolas en una solución de etanol al 75% durante 10 minutos, seguidas de hipoclorito de sodio al 10% (v/v) y se lavaron tres veces con agua destilada estéril. Posteriormente se germinaron en papel filtro estéril dentro de cajas de policarbonato en condiciones de cultivo estándar (16/8 h luz/oscuridad y $27 \pm 1^\circ\text{C}$). Se tomaron plántulas de nueve días y se sometieron a tratamientos de estrés por temperaturas baja (4°C) y alta (42°C) en cámaras bioclimáticas durante 12 h. Las plántulas utilizadas como control se mantuvieron en condiciones de cultivo estándar.

6.9. Aislamiento de RNA y análisis de expresión

Se extrajo ARN total de plántulas de 10 días de edad usando el reactivo de Trizol (ThermoFisher) siguiendo las instrucciones del fabricante. Se utilizaron tres réplicas biológicas con tres plántulas cada una para cada tratamiento. El ADNc se sintetizó con el sistema de transcriptasa inversa ImProm-II (Promega). El análisis de la expresión se realizó mediante PCR en tiempo real en un termociclador Corbett RotorGene 6000 (ThermoFisher) utilizando SsoFast Eva Green Supermix (BioRad). El gen *EF1α* se usó como control endógeno y el nivel de expresión relativa se calculó con el método $2^{-\Delta\Delta ct}$ (Livak y Schmittgen, 2001). Se realizaron pruebas de Tukey y análisis de varianza para determinar las diferencias significativas entre las medias de los tratamientos con respecto al control. Los resultados del análisis de expresión se representaron con el uso del software R V3.5 utilizando el paquete ggpubr. Los resultados se compararon con los datos reportados en el atlas de expresión de genes de maíz bajo tratamientos de estrés (Hoopes *et al.*, 2019) para la expresión del gen Dof01 de B73, donde las plantas de maíz B73 se sometieron a estrés por frío (4°C) y calor (42°C).

7. RESULTADOS

7.1. Variación en la identidad de secuencias homólogas de promotores Dof

Se obtuvieron un total de 134 secuencias de promotores Dof entre las tres variedades analizadas (46 en B73, 43 en PAL y 45 en TEO). Se realizó un alineamiento local BLAST para comparar los porcentajes de cobertura e identidad de las secuencias de PAL y TEO con respecto a la de B73 (Tabla I). Se obtuvo un porcentaje de cobertura de las secuencias de maíz Palomero con respecto a las de B73 que va desde el 2% en Dof45 hasta un 96% en Dof12. Los porcentajes de identidad tuvieron una variación que va desde el 86.43% en Dof28 hasta un 99.57% en Dof04. No se identificaron secuencias correspondientes a los promotores Dof03, Dof34 y Dof36. Por otra parte, entre las secuencias de Teosinte identificadas, el porcentaje mínimo de cobertura está dado por la secuencia del promotor Dof41 con un 2%, mientras que Dof03, 04, 08, 09, 11, 18, 37, 39 y 43 tuvieron un 100% de cobertura con respecto a sus secuencias homólogas en B73. La identidad entre las secuencias de B73 y TEO varió entre 79.26% en Dof12 y un 99.8% en Dof24 y Dof39.

Tabla I. Alineamiento local entre las secuencias de promotores Dof de B73, palomero (PAL) y Teocinte (TEO). Se identificaron a los promotores de las variedades de PAL y TEO con respecto a su secuencia homóloga en B73. Se nombró a los promotores de PAL y TEO de acuerdo con su secuencia homóloga en B73. Se indican el porcentaje de identidad y el porcentaje de cobertura (Cob) de cada secuencia con respecto a B73. En color rojo se marcan las secuencias descartadas para posteriores análisis. Las secuencias sombreadas en color gris son aquellas que están representadas por las tres variedades, nombre de identificación del gen Dof para el parental B73 (ID).

Nombre B73	ID	Nombre PAL	Identidad	Cob	Nombre TEO	Identidad	Cob
ZmB73_Dof01	AC155434.2_FGT006	ZmPAL_Dof01	97.86	40	ZmTEO_Dof01	92.89	90
ZmB73_Dof02	GRMZM2G009406	ZmPAL_Dof02	93.36	24	ZmTEO_Dof02	93.7	53
ZmB73_Dof03	GRMZM2G146283	ZmPAL_Dof03			ZmTEO_Dof03	98.26	100
ZmB73_Dof04	GRMZM2G327189	ZmPAL_Dof04	99.57	26	ZmTEO_Dof04	92.49	100
ZmB73_Dof05	GRMZM5G880268	ZmPAL_Dof05	95.26	10	ZmTEO_Dof05	90.98	84
ZmB73_Dof06	GRMZM2G089949	ZmPAL_Dof06	96.27	24	ZmTEO_Dof06	98.65	93
ZmB73_Dof07	GRMZM2G378490	ZmPAL_Dof07	96.03	52	ZmTEO_Dof07		

ZmB73_Dof08	GRMZM2G463525	ZmPAL_Dof08	96.72	52	ZmTEO_Dof08	97.28	100
ZmB73_Dof09	GRMZM2G449950	ZmPAL_Dof09	95.37	46	ZmTEO_Dof09	98.36	100
ZmB73_Dof10	GRMZM2G435475	ZmPAL_Dof10	90.65	54	ZmTEO_Dof10	89.23	99
ZmB73_Dof11	GRMZM2G123900	ZmPAL_Dof11	98.77	48	ZmTEO_Dof11	98.01	100
ZmB73_Dof12	GRMZM2G178767	ZmPAL_Dof12	87.28	96	ZmTEO_Dof12	79.26	15
ZmB73_Dof13	GRMZM2G456452	ZmPAL_Dof13	88.81	21	ZmTEO_Dof13	86.54	7
ZmB73_Dof14	GRMZM2G134545	ZmPAL_Dof14	94.48	15	ZmTEO_Dof14	92.02	89
ZmB73_Dof15	GRMZM2G451771	ZmPAL_Dof15	91.53	85	ZmTEO_Dof15	92.29	53
ZmB73_Dof16	GRMZM2G089850	ZmPAL_Dof16	93.99	58	ZmTEO_Dof16	96.65	6
ZmB73_Dof17	GRMZM2G114998	ZmPAL_Dof17	97.71	47	ZmTEO_Dof17	88.73	99
ZmB73_Dof18	GRMZM2G131897	ZmPAL_Dof18	91.29	36	ZmTEO_Dof18	96.41	100
ZmB73_Dof19	GRMZM2G135703	ZmPAL_Dof19	91.98	86	ZmTEO_Dof19	93.41	97
ZmB73_Dof20	GRMZM2G371058	ZmPAL_Dof20	90.18	16	ZmTEO_Dof20	88.4	59
ZmB73_Dof21	GRMZM2G162749	ZmPAL_Dof21	94.9	66	ZmTEO_Dof21	93.74	85
ZmB73_Dof22	GRMZM2G171852	ZmPAL_Dof22	93.87	85	ZmTEO_Dof22	95.35	99
ZmB73_Dof23	GRMZM2G010290	ZmPAL_Dof23	98.78	44	ZmTEO_Dof23	92.25	40
ZmB73_Dof24	GRMZM2G084130	ZmPAL_Dof24	96.16	42	ZmTEO_Dof24	99.81	26
ZmB73_Dof25	GRMZM2G108865	ZmPAL_Dof25	94.79	95	ZmTEO_Dof25	95.69	97
ZmB73_Dof26	GRMZM2G061292	ZmPAL_Dof26	96.9	30	ZmTEO_Dof26	96.89	55
ZmB73_Dof27	GRMZM2G017470	ZmPAL_Dof27	92.74	47	ZmTEO_Dof27	94.03	91
ZmB73_Dof28	GRMZM2G093725	ZmPAL_Dof28	86.43	40	ZmTEO_Dof28	94.58	52
ZmB73_Dof29	GRMZM2G140694	ZmPAL_Dof29	95.94	50	ZmTEO_Dof29	91.28	97
ZmB73_Dof30	AC233935.1_FGT005	ZmPAL_Dof30	99.25	52	ZmTEO_Dof30	96.45	89
ZmB73_Dof31	GRMZM2G144188	ZmPAL_Dof31	93.05	64	ZmTEO_Dof31	95.42	29
ZmB73_Dof32	GRMZM2G011832	ZmPAL_Dof32	94.69	49	ZmTEO_Dof32	94.62	96
ZmB73_Dof33	GRMZM2G144172	ZmPAL_Dof33	94.73	44	ZmTEO_Dof33	91.87	75
ZmB73_Dof34	GRMZM2G064655	ZmPAL_Dof34			ZmTEO_Dof34	99.62	98
ZmB73_Dof35	GRMZM2G179069	ZmPAL_Dof35	97.42	32	ZmTEO_Dof35	88.79	83
ZmB73_Dof36	GRMZM2G137502	ZmPAL_Dof36			ZmTEO_Dof36	95.76	96
ZmB73_Dof37	GRMZM2G394973	ZmPAL_Dof37	97.32	31	ZmTEO_Dof37	93.32	100
ZmB73_Dof38	GRMZM2G042218	ZmPAL_Dof38	91.66	89	ZmTEO_Dof38	94.51	92
ZmB73_Dof39	GRMZM2G045678	ZmPAL_Dof39	97.36	42	ZmTEO_Dof39	99.8	100
ZmB73_Dof40	GRMZM2G138455	ZmPAL_Dof40	90.94	29	ZmTEO_Dof40	90.97	82
ZmB73_Dof41	GRMZM2G142718	ZmPAL_Dof41	100	49	ZmTEO_Dof41	91.84	2
ZmB73_Dof42	GRMZM2G394941	ZmPAL_Dof42	95.51	4	ZmTEO_Dof42	97.59	54
ZmB73_Dof43	GRMZM2G589696	ZmPAL_Dof43	89.59	39	ZmTEO_Dof43	94.11	100
ZmB73_Dof44	AC209819.3_FGT009	ZmPAL_Dof44	99.4	31	ZmTEO_Dof44	96.07	88
ZmB73_Dof45	GRMZM2G082490	ZmPAL_Dof45	97.62	2	ZmTEO_Dof45	91.22	84
ZmB73_Dof46	GRMZM2G176063	ZmPAL_Dof46	99.89	44	ZmTEO_Dof46	94.87	93

7.2. Los promotores de genes parálogos divergen en funcionalidad

Se obtuvieron un total de 93 secuencias de promotores Dof de las cuales 46 fueron de B73, 16 de PAL y 29 de TEO (secuencias sombreadas en color gris en Tabla I). Con las secuencias seleccionadas se construyó el árbol filogenético (

Figura 1). En el árbol se agrupan las secuencias conforme a su similitud, por ejemplo, las secuencias homólogas entre variedades son parte del mismo clado. Sin embargo, es preciso mencionar que algunas de las secuencias de promotores Dof de B73 que dirigen la expresión de genes parálogos, pueden no encontrarse en el mismo grupo de miembros. Por ejemplo, el promotor Dof05 de Teosinte está agrupado junto a las secuencias de Dof23 de B73 y PAL en lugar de agruparse con Dof05 de B73. Por su parte ZmB73_Dof05 está agrupado con ZmB73_Dof42, mientras que la secuencia que dirige la expresión del gen parálogo en B73, Dof35, se encuentra agrupada en un clado separado.

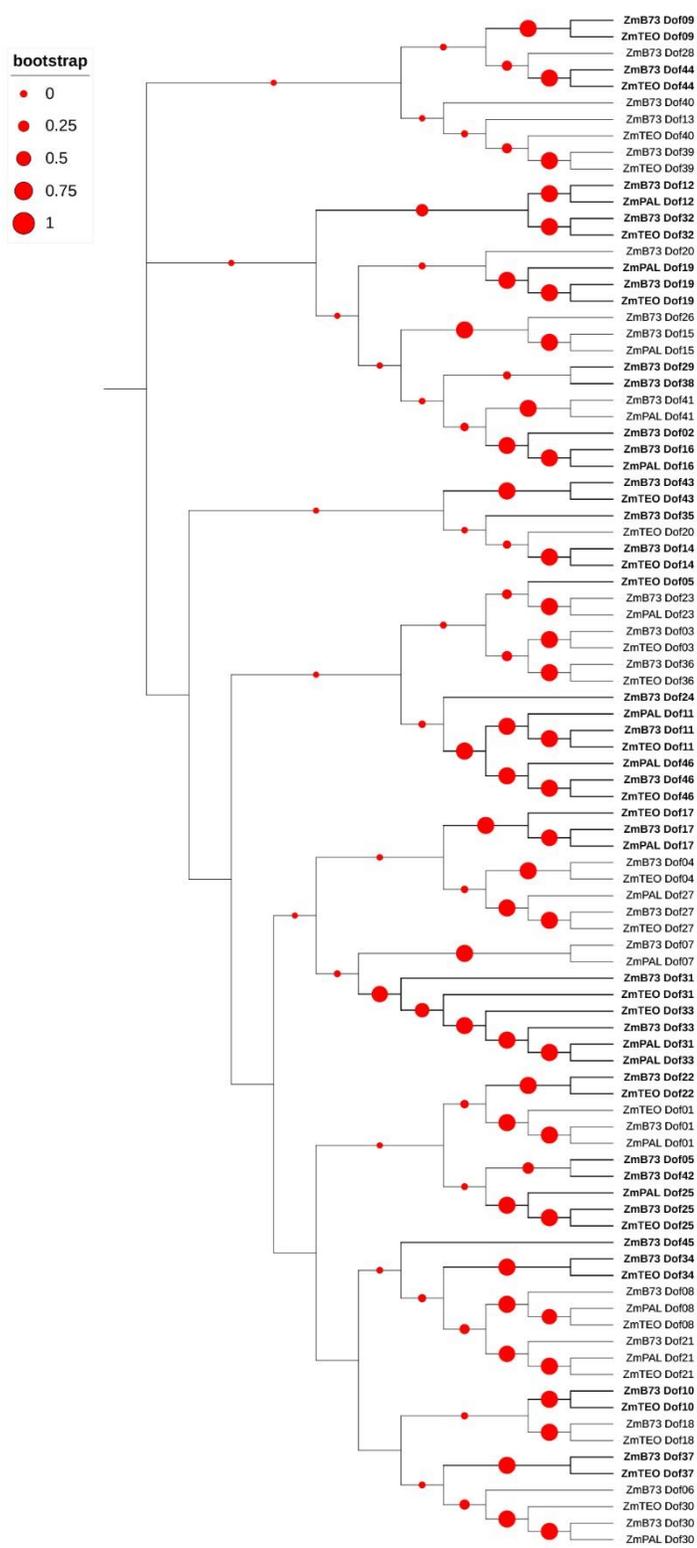


Figura 1. Análisis filogenético de las secuencias de promotores Dof de maíz B73, palomero (PAL) y Teocinte (TEO). Los promotores de genes parálogos están indicados con letra negrita y ramas más gruesas. El tamaño del círculo rojo en cada nodo representa el soporte de réplicas de bootstrap del análisis, el cual está indicado en la leyenda en la parte superior izquierda.

7.3. La caja TATA influye en la distribución de elementos regulatorios

Se realizó un gráfico de densidad de distribución de elementos regulatorios presentes en los promotores Dof (Figura 2). En él, podemos determinar la probabilidad de encontrar un elemento *cis* a lo largo de la secuencia de los promotores. Los gráficos fueron clasificados según la variedad de maíz (B73, PAL o TEO), el tipo de promotor (TATA+ o TATA-) y la hebra en donde se localiza el elemento regulatorio. La posición de cada uno de los elementos regulatorios se representa con barras de color en la parte inferior de cada gráfico. La escala inferior representa la distancia relativa al TSS más cercano al ATG. La escala lateral izquierda, representa la probabilidad o función de densidad.

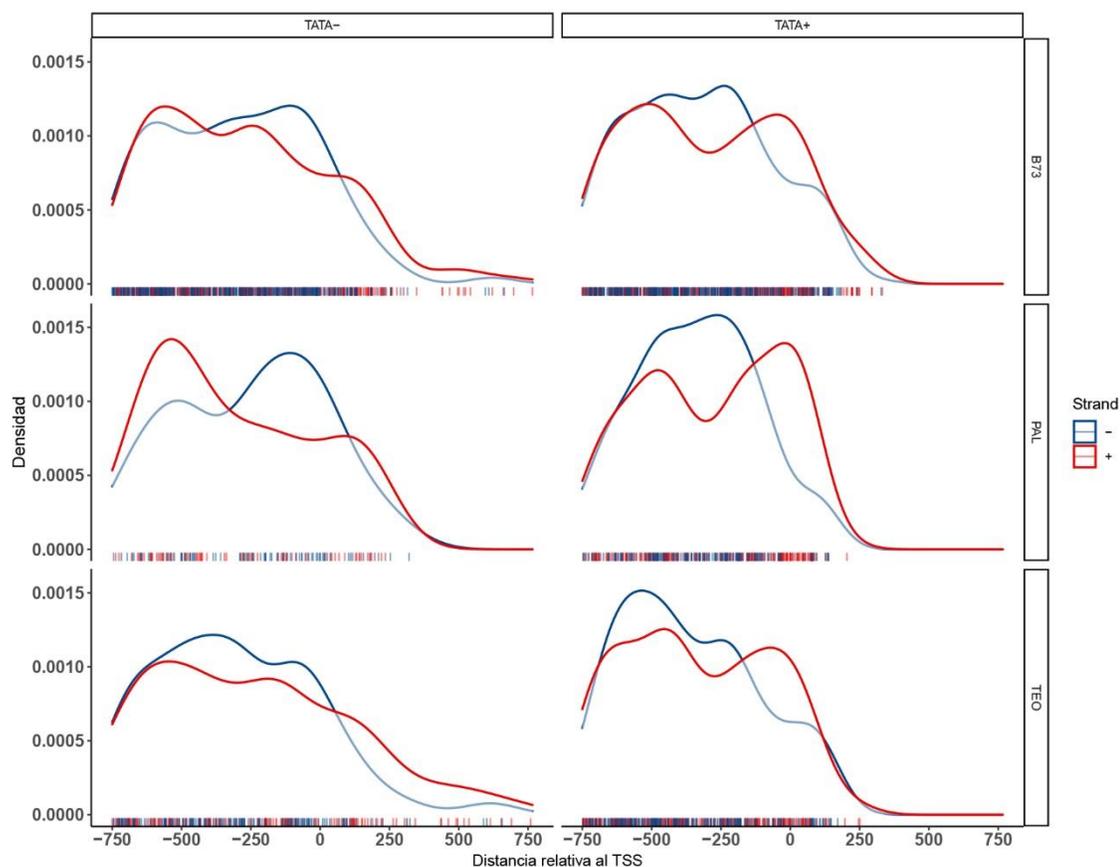


Figura 2. Distribución de densidad de elementos *cis* en promotores Dof de maíz. Los gráficos representan la probabilidad de localizar un elemento *cis* en una posición determinada. La escala inferior indica la posición relativa en la secuencia de los promotores con respecto a la posición del TSS (0) más cercano al ATG.

Las secuencias de los promotores clasificados como TATA- presentan una distribución más amplia de elementos regulatorios a lo largo de toda la secuencia, extendiéndose incluso hasta posiciones que superan los 500 pb río arriba del TSS en B73 y TEO. En el caso de las secuencias de maíz Palomero, la distribución de elementos regulatorios alcanza picos de máxima densidad, contrario a lo observado en B73 y TEO. Sin embargo, la densidad de elementos *cis* en palomero es claramente menor a lo localizado en las otras dos variedades. Por otra parte, los elementos regulatorios presentes en promotores tipo TATA+, se distribuyen en áreas más estrechas formando picos de mayor densidad. En las tres variedades se observan estos picos de máxima densidad en la hebra positiva en torno al TSS y disminuyen después de los 250 pb río arriba. Sin embargo, en la hebra anti

sentido, se puede observar que el mayor pico de densidad se alcanza después de los 500 pb río arriba del TSS. Además, en el caso de B73 y PAL, estos picos tienen un área de distribución mayor con respecto a sus contrapartes en las hebras positivas del DNA.

7.4. Los elementos regulatorios *cis* no se distribuyen aleatoriamente

Se realizó un gráfico de distribución de densidad de los sitios de unión que son reconocidos por factores de transcripción según la familia a la que pertenecen (**Figura 3**). Algunas de las familias de factores de transcripción tienen preferencia de unión a ciertas posiciones relativas a la posición del TSS, mostrando picos de máxima densidad. También, el comportamiento en la distribución de estos elementos regulatorios responde a la hebra de unión al DNA. Las familias AT-Hook, B3, bHLH y GATA, se distribuyen de forma similar en ambas cadenas, indicando que estos elementos *cis* no tienen una preferencia de unión que difiera entre ambos sentidos. Sin embargo, el resto de las familias de factores de transcripción mantienen un comportamiento distinto en la distribución según la hebra a la que se unen.

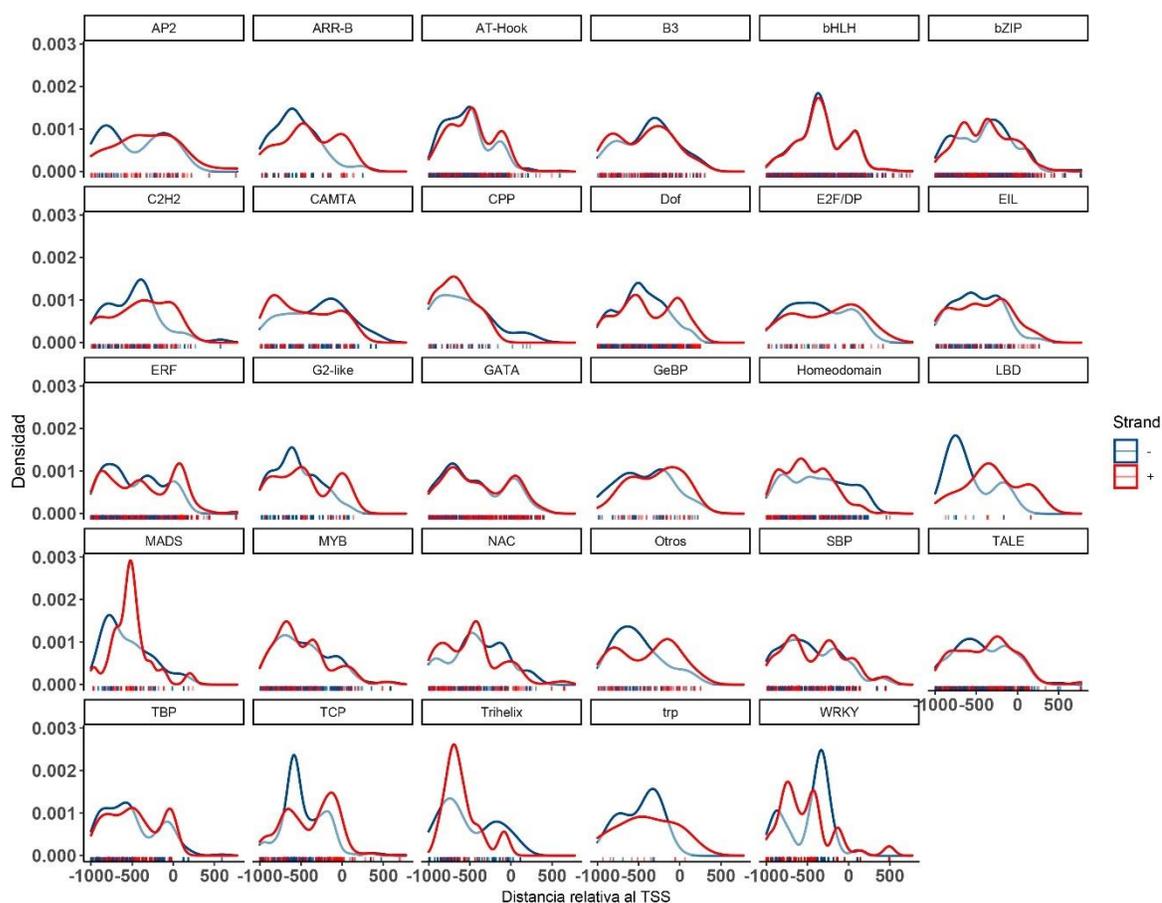


Figura 3. Distribución de densidad de elementos regulatorios *cis* en promotores Dof clasificados por familia según la cadena a la que se unen. En azul se muestra la hebra antisentido y en rojo la hebra sentido.

7.5. Tres familias de factores de transcripción están sobre representadas

Se identificaron un total de 29 familias de factores de transcripción distintos, las cuales reconocen 31,540 sitios de unión en los 93 promotores analizados. En la **Figura 4** se muestra la relación entre promotores de acuerdo al número y tipo de factores de transcripción que se unen a ellos. Los factores de transcripción están clasificados según la familia a la que pertenecen y están agrupados en el dendrograma superior en tres principales grupos (A-C) según su similitud con respecto a la media de los datos de la columna correspondiente. Además, en la parte superior de cada columna del heatmap se presenta, como gráfico de barras, la suma de ocurrencias que tiene cada familia de factores de transcripción en los promotores. Los factores de transcripción con mayor presencia en los promotores

Dof pertenecen a las familias bHLH, ERF y bZIP, con 4140, 3714 y 3258 ocurrencias, respectivamente, todas pertenecientes al grupo A. Por otra parte, los factores de transcripción menos representados pertenecen a las familias LBD y trp con 27 y 31 ocurrencias entre los 93 promotores, del grupo B y C, respectivamente.

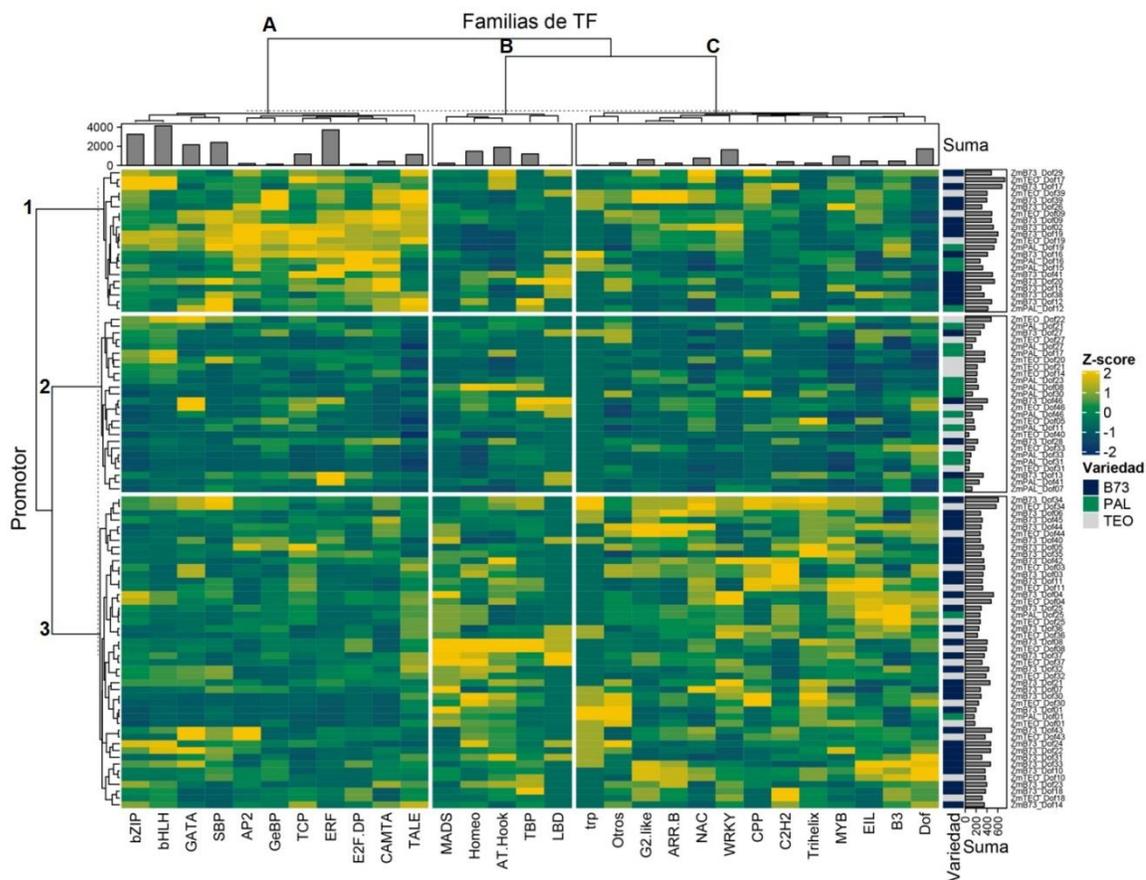


Figura 4. Ocurrencias de TFBS en promotores Dof por familia de factores de transcripción.

Por otra parte, los promotores se dividen en tres principales grupos (1-3) según su relación en cuanto al tipo y número de factores de transcripción que reconocen sitios de unión en sus secuencias. El grupo 3 es el más abundante y contiene 46 secuencias de promotores Dof, de las cuales 29 pertenecen a B73. Sólo 15 secuencias de TEO y dos de PAL forman parte de este grupo. El segundo grupo más abundante (2) está conformado por 12 secuencias de PAL, 10 de TEO y sólo

cuatro de B73, mientras que el grupo con menos miembros (1) se compone de 13 secuencias de B73, cuatro de PAL y cuatro de TEO. El promotor Dof17 de TEO (grupo 1) es el que mayor número de ocurrencias de factores de transcripción posee con un total de 728, mientras que el promotor con el menor número de ocurrencias corresponde a Dof40 de TEO con un total de 67 (grupo 2).

Debido a la formación de grupos de promotores y de factores de transcripción, se pueden identificar nueve clusters. El cluster 1A se caracteriza por aquellos promotores del grupo 1 que contienen el mayor número de ocurrencias por arriba de la media de cada factor de transcripción de las familias bZIP, bHLH, GATA, SBP, AP2, GeBP, TCP, ERF, E2F DP, CAMTA y TALE. En cambio, el grupo 3 de promotores tiene un mayor número de ocurrencias de sitios de unión para los factores de transcripción del grupo B y C. El grupo 2 de promotores se caracteriza por tener ocurrencias de factores de transcripción que están en torno a la media de ocurrencias entre todos los promotores.

7.6. Los promotores presentan regiones de elementos *cis* conservados entre variedades

Se realizó la comparación entre las secuencias de promotores Dof de las tres variedades de maíz y se representó de manera gráfica su arquitectura (

Figura 5). Las comparaciones se realizaron entre cada promotor de Palomero y Teosinte con respecto a su secuencia homóloga en B73. De manera gráfica se representó la posición y el tipo de motivo localizado, así como los diferentes TSSs y cajas TATA distribuidas a lo largo de las secuencias. También, se señala la longitud del 5' UTR de cada promotor y las regiones que se encuentran conservadas entre pares de secuencias homólogas. En las diferentes arquitecturas, se puede apreciar que existen regiones de elementos regulatorios que presentan el mismo orden entre las diferentes variedades. La principal distinción, es la ubicación relativa de los elementos *cis* con respecto al TSS principal, debido a la presencia de regiones de inserción/delección (indels). Por ejemplo, en la

Figura 5A se presenta una región de 179 pb en la secuencia del promotor Dof01 de B73 que está ausente en las secuencias de Palomero y Teosinte. Otra

diferencia que podemos localizar en las secuencias es el número de TSSs distribuidos, así como las cajas TATA asociados a estos. Por ejemplo, en la región antes mencionada, existe una caja TATA en B73 que no está presente en la secuencia de PAL pero que si se localiza en TEO. Sin embargo, tanto PAL como TEO, contienen una caja TATA asociada al TSS principal que no está presente en B73.

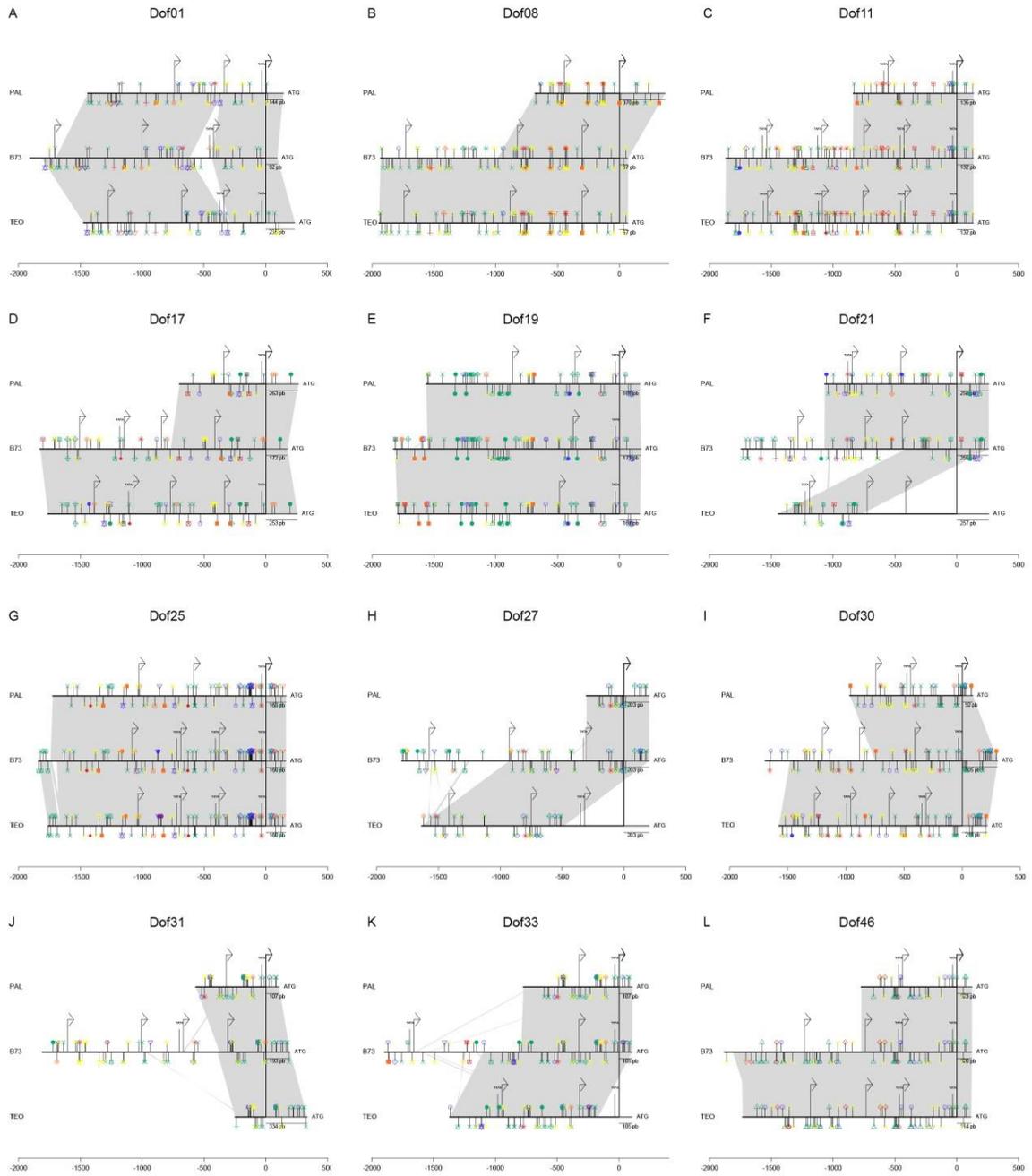


Figura 5. Arquitectura de promotores Dof de tres variedades de maíz. Únicamente se representaron las secuencias que se obtuvieron de las tres variedades (B73, PAL y TEO). La flecha de mayor tamaño representa al TSS más cercano al ATG. El resto de flechas representan los TSSs alternativos localizados en cada secuencia. Las líneas verticales con figuras de colores en su extremo superior, representan los diferentes motivos o TFBS distribuidos en los promotores. Las regiones sombreadas de color gris, representan las regiones que están conservadas entre pares de secuencias. A) Dof01; B) Dof08; C) Dof11; D) Dof17; E) Dof19; F) Dof21; G) Dof25; H) Dof27; I) Dof30; J) Dof31; K) Dof33; L) Dof46.

De la misma manera, es posible apreciar que las regiones conservadas entre los pares de secuencias analizados se encuentran desplazados de manera significativa. Por ejemplo, la región que comprende desde los 203 pb hasta los -1000 pb en B73 tiene un alto porcentaje de similitud con la región que va desde los -500 hasta -1600 pb en TEO (

Figura 5H). En esta región, se aprecian el mismo número, tipo y posición de elementos *cis* compartidos entre ambas variedades, sin embargo, las posiciones de los TSSs son diferentes.

En el **Anexo A** se encuentra la Tabla IV donde se relacionan los diferentes motivos localizados en los promotores analizados, así como la simbología implementada en los gráficos de arquitectura de los promotores.

7.7.El promotor del gen Dof01 de B73 presenta un elemento de inserción/delección con respecto a PAL y TEO

Se realizó un análisis detallado de la arquitectura del promotor Dof01 a través de un alineamiento múltiple de las secuencias homólogas de Dof01 en B73, PAL y TEO (**Anexo B**). En el alineamiento se identifican todas las regiones conservadas entre las tres secuencias de promotores analizadas, así como las familias de factores de transcripción que reconocen sitios de unión en ellas. En el alineamiento se tomó en cuenta únicamente la hebra sentido. En total se identificaron 55 TFBS, de los cuales 22 son compartidos entre las tres secuencias, 10 se encuentran presentes entre B73 y PAL, tres entre B73 y TEO y siete de estos se comparten entre PAL y TEO. En el caso de B73, este presenta tres TFBS

que están presentes en la región indel de 179 pb, mientras que en PAL existen 2 TFBS que sólo están presentes en su secuencia. Por otra parte, Teosinte contiene ocho TFBS que no comparte con las otras dos secuencias homólogas. Los diferentes factores de transcripción que pertenecen a las familias identificadas y que reconocen sitios de unión en la secuencia del promotor Dof01 en las tres variedades, se enlistan en la Tabla IV contenida en el **Anexo C**. Se identificaron un total de 250 factores de transcripción distintos que reconocen sitios de unión tan sólo en la hebra sentido en las tres secuencias del promotor Dof01 analizadas.

7.8. Los TFBS presentes en el promotor Dof01 de B73 están relacionados con la respuesta a estrés por temperatura

Para entender cómo los factores de transcripción que se unen al promotor Dof01 de B73 influyen en la expresión de este factor de transcripción, se realizó un análisis de correlación de Pearson (

Figura 6). El análisis de correlación muestra cómo está correlacionada la expresión de los 9 factores de transcripción que potencialmente influyen más en la expresión de Dof01 en B73.

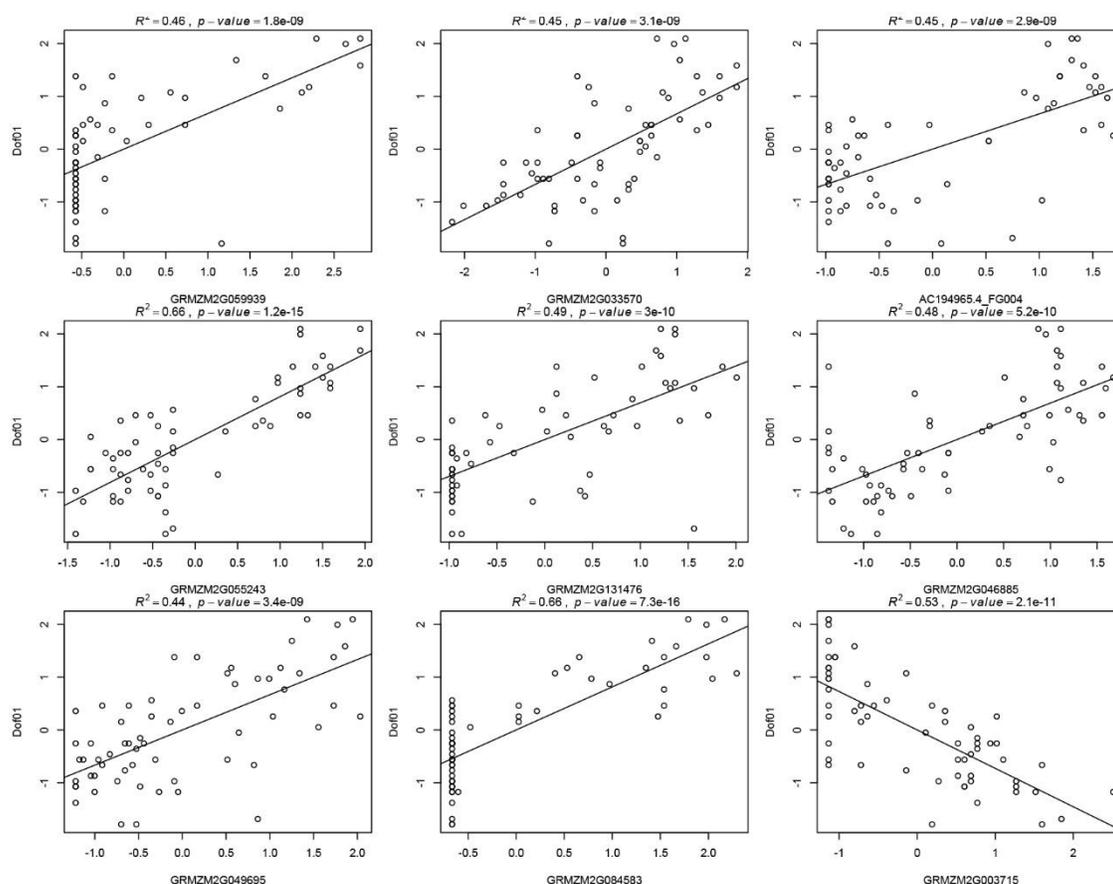


Figura 6. Análisis de correlación de Pearson entre la expresión de factores de transcripción y la expresión de Dof01 en B73.

En los diferentes gráficos se aprecia que la correlación entre la expresión de los diferentes factores de transcripción con respecto a la expresión de Dof01 es positiva y significativa, excepto con el gen GRMZM2G003715, el cual presenta una correlación negativa con la expresión de Dof01. Este gen pertenece a la familia de factores de transcripción NAC y se ha identificado como el gen ZmNAC61, de acuerdo a la base de datos de grassius.org. En la Tabla II se especifican los nombres de los diferentes factores de transcripción que se identificaron mediante el análisis de correlación. También se muestra su identificador y la familia a la que pertenecen.

Tabla II. Nombre y familia a la que pertenecen los factores de transcripción asociados a la expresión de Dof01.

TF	Familia	Nombre
GRMZM2G059939	AP2/B3	ZmEREB162
GRMZM2G033570	EIN3	ZmEIL1
AC194965.4_FG004	GATA	ZmGATA12
GRMZM2G055243	Homeodomain	ZmHB72
GRMZM2G131476	Homeodomain	ZmHB121
GRMZM2G046885	MADS box	ZmMADS73
GRMZM2G049695	Myb/SANT	ZmMYBR24
GRMZM2G084583	Myb/SANT	ZmMYB38
GRMZM2G003715	NAC/NAM	ZmNAC61

También se comparó la expresión de Dof01 en respuesta a estrés abiótico, con relación a la expresión de cada uno de los factores de transcripción mencionados en la Tabla II. En la **Figura 7** se muestra el nivel de expresión del gen Dof01 de B73 en respuesta a diferentes factores de estrés abiótico como sequía, salinidad y temperatura. En la **Figura 7** se representan diferentes partes de plantas de maíz que fueron analizadas bajo los tratamientos antes mencionados. El color está asociado al nivel de expresión de Dof01 con respecto a la expresión de los otros genes analizados, donde el color rojo indica una sobre expresión, el color azul indica una sub regulación y el color amarillo representa que no hay cambios en la expresión del gen Dof01.

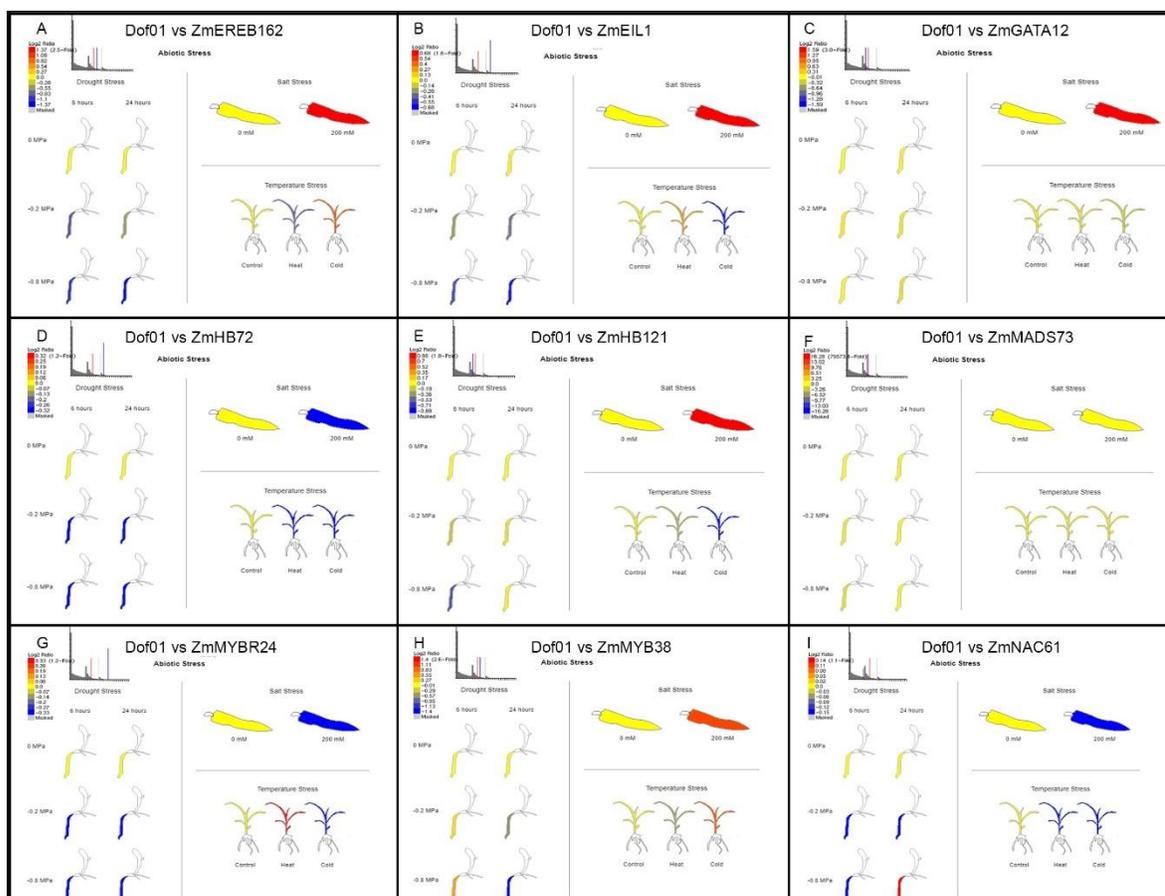


Figura 7. Datos de expresión de Dof01 en relación a otros factores de transcripción bajo tratamientos de estrés abiótico.

En la Figura 7A, se muestra la expresión de Dof01 en relación al gen ZmERE162. La expresión de Dof01 se ve reducida hasta 2.5 veces en el tejido de raíz bajo los tratamientos de sequía. Así mismo, el tratamiento de estrés por temperatura alta muestra una disminución de la expresión de Dof01. Por el contrario, la expresión de este gen se ve aumentada hasta 2.5 veces debido al tratamiento de salinidad y frío.

En comparación con la expresión del gen ZmEIL1, Dof01 muestra una disminución en su expresión debido a los tratamientos de sequía y frío, pero aumenta debido a los tratamientos de calor y salinidad Figura 7B.

Con respecto al gen ZmGATA12, la expresión de Dof01 no se ve afectada más que por el tratamiento de salinidad, donde su expresión aumenta hasta 3 veces (Figura 7C).

Los niveles de expresión de Dof01 se ven reducidos hasta 1.2 veces con respecto al gen ZmHB72 en todos los tratamientos (Figura 7D). Por otra parte, el gen Dof01 es sobre expresado debido a salinidad con respecto al gen ZmHB121, pero su expresión disminuye en los tratamientos de estrés por sequía y frío (Figura 7E). En el caso de la respuesta de Dof01 con respecto al gen ZmMADS73 (Figura 7F), no parece presentar cambios bajo ninguno de los tratamientos.

Al comparar la expresión del gen Dof01 con respecto al gen ZmMYBR24, se aprecia que el primero disminuye hasta 1.2 veces debido a los tratamientos de salinidad, sequía y temperatura baja, pero aumenta debido a temperatura alta (Figura 7G). En cuanto a la comparación con el gen ZmMYB38, las principales respuestas se ven debido a los tratamientos de salinidad y frío, donde los niveles de expresión de Dof01 aumentan, mientras que, debido al tratamiento de sequía, disminuyen. Por último, la expresión de Dof01 se ve disminuida hasta 1.1 en todos los tratamientos, excepto a las 24 horas de tratamiento de sequía a -0.8 MPa, donde el gen se sobre expresa.

7.9. La respuesta de Dof01 a estrés por temperatura difiere entre B73 y PAL

Para comprender cómo responden los genes Dof al estrés por temperatura, sometimos las plantas de PAL a los tratamientos de frío (4°C) y calor (42°C). Analizamos la expresión del gen Dof01 en los tejidos de hoja, tallo y raíz de plántulas de 10 días de edad y comparamos estos resultados con los datos de expresión de estrés por frío y calor de B73 del atlas de Hoopes. En la **Figura 8** se muestra la expresión del gen Dof01 en los tejidos de hoja, tallo y raíz en condiciones de estrés por temperatura baja y alta comparados contra las condiciones normales de cultivo (control). En el tejido de hoja, Dof01 se sobre expresa significativamente con respecto al control debido al tratamiento con frío, mientras que se reprime significativamente debido al efecto del estrés por calor. En el caso del tallo, no se encontraron diferencias estadísticamente significativas

con respecto al control. Por otro lado, en el tejido de raíz, el gen Dof01 se sobre expresa significativamente más de diez veces con respecto al control, así como por el efecto del tratamiento de estrés por calor, cuya expresión está ligeramente reprimida con respecto al tratamiento control, pero sin ser significativa estadísticamente.

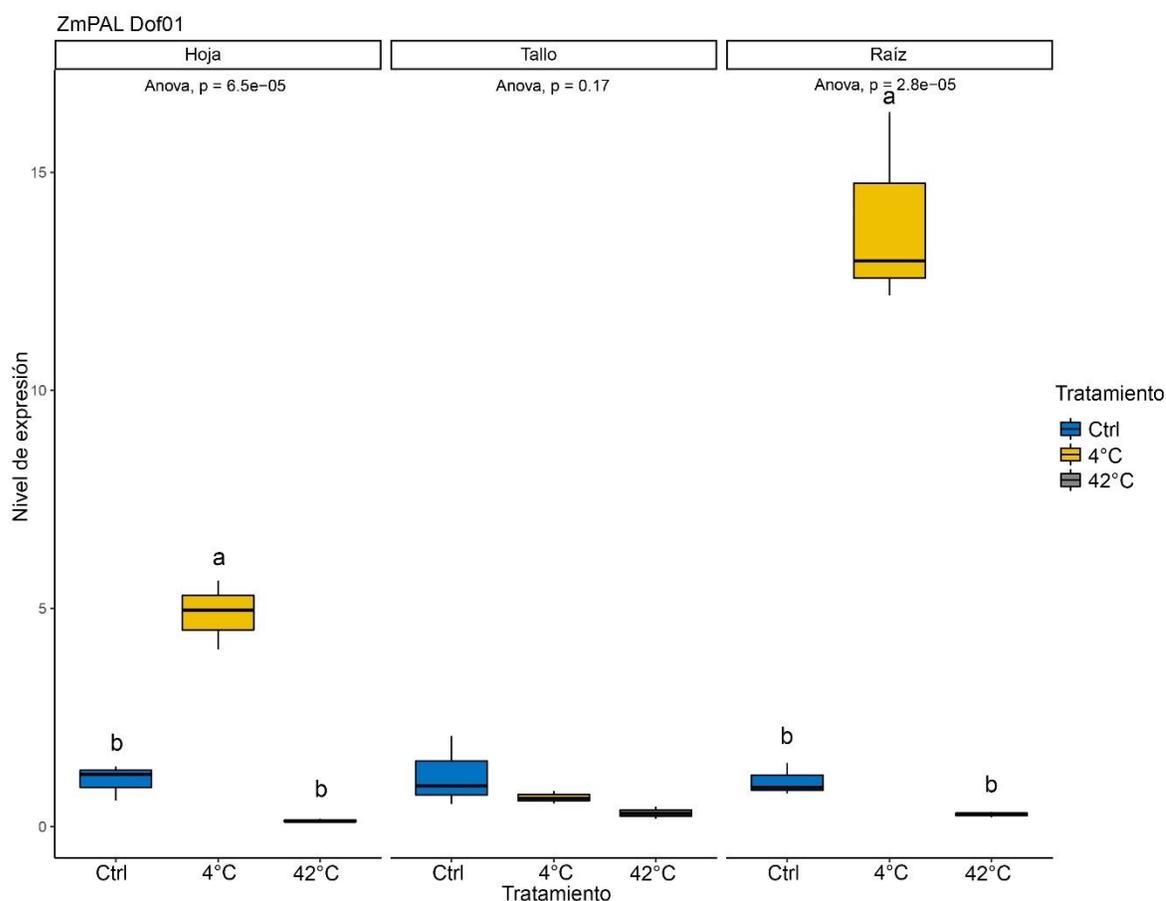


Figura 8. Expresión del gen Dof01 en plántulas de maíz Palomero bajo tratamientos de estrés por temperatura.

Si comparamos el comportamiento de la expresión de Dof01 reportado en B73 bajo las mismas condiciones de cultivo, tal como se muestra en la Figura 7, se puede apreciar que el gen Dof01 se comporta de manera contraria a lo que muestra el mismo gen en las plántulas de maíz Palomero. El gen Dof01 en B73 se reprime en condiciones de crecimiento a 4°C mientras que se sobre expresa debido a temperatura de 42°C.

8. DISCUSIÓN

Determinar en qué tejido y cómo un gen es transcrito son dos de las incógnitas que surgen cuando se habla de regulación de la expresión. Una pequeña aproximación, es entender cuáles son las relaciones que existen entre promotores, factores de transcripción y la distribución de elementos regulatorios a lo largo de las secuencias no codificantes que regulan la expresión de un gen (Yu *et al.*, 2016). La regulación de la expresión está sujeta a las interacciones entre los elementos regulatorios *cis* y *trans* con la maquinaria basal de la transcripción, liderada principalmente por la RNA pol II (Butler y Kadonaga, 2002; Smale y Kadonaga, 2003). Las diferentes combinaciones de elementos regulatorios dan lugar a diferentes niveles de respuesta en la expresión de un gen, ya sea para transcribirlo o reprimirlo (Gonzalez, 2016). Estos tipos de respuesta variarán según el órgano, tejido, etapa de desarrollo y al ambiente en el que habita un organismo. Estas diferencias son las que dan lugar a los fenotipos y los cambios evidentes que vemos entre variedades de una misma especie a través del curso de la evolución y como efecto de la mano del hombre (Tirosh *et al.*, 2009).

En este trabajo, se realizó un análisis comparativo entre las secuencias de los promotores que regulan la expresión de la familia de factores de transcripción Dof. Si bien, se cree que esta familia surgió durante las primeras formas de vida vegetales como las microalgas (Bowman *et al.*, 2017; Moreno-Risueno *et al.*, 2007), a lo largo de la historia de la evolución se han multiplicado los miembros pertenecientes a ella y cada paso de duplicación conlleva a la adquisición, pérdida, mantenimiento o complementariedad de una función (Shigyo *et al.*, 2007). Para entender cómo el proceso de domesticación del maíz influyó en la conservación y funcionalidad de las secuencias no codificantes que regulan la expresión de los genes Dof, se tomaron tres variedades de maíz que han tenido destinos evolutivos diferentes: Teosinte, Palomero Toluqueño y B73. Se cree que el maíz, tal como lo conocemos, se originó a través de los años por la domesticación del Teosinte, considerado como el ancestro silvestre.

Una de las principales limitantes de este trabajo fue la disponibilidad de información adecuada para realizar análisis más detallados sobre los promotores de las tres variedades para los 46 miembros de la familia de factores de transcripción Dof. A pesar de que los genomas de maíz B73 y Palomero Toluqueño se secuenciaron por primera vez en 2009 (Schnable *et al.*, 2009; Vielle-Calzada *et al.*, 2009), el interés comercial que presenta la variedad híbrida ha llevado a estudiar más a fondo su genoma y a establecerlo como un punto de referencia en los estudios de genómica. A la fecha, la versión más reciente del genoma de maíz B73 (V4) difiere por mucho, en calidad y amplitud a la única versión de maíz Palomero Toluqueño. Las secuencias de los promotores Dof en la raza ancestral de maíz, disponibles en el genoma secuenciado, son de 43 de las 46 secuencias de maíz B73 utilizadas como referencia. Sin embargo, la información útil disponible de estas se mantuvo con coberturas menores al 50% del total de la longitud de cada secuencia comparada (2000 nt), encontrando únicamente fragmentos de secuencias que no son aptos para los análisis realizados. Sin embargo, es interesante observar que el porcentaje de identidad entre las secuencias comparadas se mantiene por encima del 90%. Este parámetro es un indicador de la divergencia que existe entre cada par de secuencias analizadas y que aquellas secuencias que son homólogas difieren entre sí en un porcentaje muy bajo. Una de las principales cuestiones que surgen al analizar este tipo de información es si ¿este tipo de diferencias afectan a regiones potencialmente conocidas como reguladoras de la expresión? O ¿afectan a regiones intergénicas que no están asociadas a este proceso?

Con respecto a Teosinte, su genoma fue secuenciado en el año 2017 (N. Yang *et al.*, 2017) y desde entonces también cuenta con una única versión. Sin embargo, la tecnología implementada para su secuenciación es más avanzada que la utilizada con Palomero. Este hecho permitió contar con un genoma más completo y de mejor calidad, lo que a su vez hizo posible la identificación de 45 de las 46 secuencias de promotores Dof que se buscaron utilizando a B73 como referencia. En este aspecto, los porcentajes de cobertura superan en su mayoría

al 70%, permitiendo encontrar fragmentos de secuencias de promotores Dof más largas. Incluso, son nueve los casos en los que se contó con un 100% de cobertura de las secuencias con respecto a su par homólogo en B73. En cuanto a su porcentaje de identidad, también fue posible determinar que este es mayor al 90%, lo que conlleva al mismo planteamiento sobre la conservación en la funcionalidad de sus secuencias regulatorias.

Al realizar los análisis filogenéticos es posible apreciar que prácticamente todas las secuencias homólogas se agrupan entre sí. Sin embargo, un hecho interesante es la división que existe entre, lo que se considera como secuencias parálogas en B73. Las secuencias parálogas fueron descritas en 2015 por Chen y Cao, quienes determinaron que en el genoma de maíz B73 existen 46 genes Dof y entre ellos se encuentran 12 pares de secuencias parálogas que surgieron por eventos de duplicación del genoma. Este tipo de eventos evolutivos ocurrieron millones de años atrás (Chen y Cao, 2015; Gaut y Doebley, 1997), posterior a los eventos de especiación del maíz, por lo que las tres variedades analizadas deben contar con el mismo grupo de genes parálogos. Sin embargo, es posible que sus secuencias regulatorias o promotores hayan tenido una mayor divergencia durante estos eventos de duplicación, dando lugar posteriormente a las diferentes variedades y fenotipos que conocemos actualmente del maíz, cada uno adaptado a condiciones de crecimiento distintas.

Dos de los elementos más importantes en el proceso de transcripción son el sitio de inicio de la transcripción y la caja TATA. Algunos estudios indican que los genes pueden ser regulados por múltiples TSS (Mejía-Guerra *et al.*, 2015; Schor *et al.*, 2017; Yamamoto *et al.*, 2011), cada uno dando lugar a transcritos alternos que no son precisamente funcionales y por lo tanto no llegan a ser traducidos. En 2015, Mejía-Guerra y colaboradores realizaron un análisis completo del genoma de maíz B73 y Mo17 para identificar los TSS predominantes, aquellos que se encargan de la transcripción primaria del gen que regulan, y los TSS alternos (Mejía-Guerra *et al.*, 2015). En este estudio encontraron que la mayoría de los elementos regulatorios se encuentran distribuidos en torno al TSS y que esta característica

está asociada a la presencia de cajas TATA. Además, asocian la relación que existe entre el número de elementos *cis*, su densidad en torno al TSS y el tipo de expresión de un gen (constitutiva o inducible). De acuerdo a los resultados obtenidos en este trabajo, el comportamiento en la distribución de elementos regulatorios a lo largo de los promotores Dof, está asociado al tipo de promotor (TATA+/TATA-). Esta característica puede estar directamente relacionada con el tipo y nivel de respuesta ante estímulos medio ambientales en los promotores analizados. Sin embargo, no se dispone de datos de expresión de las tres variedades que pueda servir para un análisis comparativo de la respuesta de los genes bajo circunstancias similares. Otros reportes, como el de Yamamoto y colaboradores (2011), mencionan que los promotores tienen ciertas características asociadas a la presencia de la caja TATA, como por ejemplo, la distancia desde el TSS al CDS, la densidad de elementos regulatorios, el nivel de expresión y el tipo de expresión (constitutiva o inducible) (Yamamoto *et al.*, 2011).

La comparación entre secuencias de promotores Dof de las tres variedades analizadas demostró que el tipo de promotor afecta principalmente a la densidad de distribución de elementos regulatorios a lo largo de su secuencia y que estos pudieran predominar de manera distinta según la hebra que se analice. También, se encontró una asociación entre la posición en la que se une un factor de transcripción y la cadena de DNA. De la misma manera que reportan Yamamoto y colaboradores (2007), algunos elementos regulatorios tienen preferencia según el sentido de la secuencias del motivo que reconocen y la distancia al TSS (Yamamoto, Ichida, Matsui, *et al.*, 2007; Yamamoto *et al.*, 2011). También, fue posible identificar que en maíz Palomero y en Teosinte, la densidad de elementos regulatorios a lo largo del fragmento de secuencia analizado fue mayor en promotores tipo TATA+ que promotores TATA-, en comparación con las secuencias de maíz B73 donde no se observa una clara diferencia en la densidad de elementos regulatorios según el tipo de promotor. Estos resultados corroboran lo encontrado en otras especies como *Arabidopsis* y arroz, donde se afirma que ciertos motivos tienen preferencias por posiciones específicas relativas al TSS

(Srivastava *et al.*, 2014; Yamamoto, Ichida, Abe, *et al.*, 2007; Yamamoto, Ichida, Matsui, *et al.*, 2007). Esto se debe principalmente a que, dependiendo el tipo de promotor, el tipo, número y distribución de elementos regulatorios va a ser distinto debido a los diferentes mecanismos de activación de la transcripción, según varíe la expresión de manera espacio-temporal (Nicole Marie Creux *et al.*, 2008; Yamamoto *et al.*, 2011).

De acuerdo a la base de datos de PlantTFDB, el número de familias de factores de transcripción encontrados en plantas es de 58. En los análisis realizados a las secuencias de promotores Dof de maíz se identificaron un total de 29 familias de factores de transcripción cuyos miembros reconocen sitios de unión en las secuencias de los promotores y que potencialmente son los responsables de la regulación de la expresión de los genes Dof. Las familias que reconocen un mayor número de motivos o sitios de unión en las secuencias de los promotores Dof son bHLH, bZIP y ERF, quienes poseen el mayor número factores de transcripción que pertenecen a estas familias, 308, 216 y 204 proteínas, respectivamente, identificados en maíz.

La familia de factores de transcripción bHLH es una familia que ha sido bien caracterizada en organismos eucariotas, principalmente en mamíferos y que normalmente se une a los promotores en forma de homodímeros o heterodímeros (Toledo-Ortiz *et al.*, 2003). De acuerdo a Zhang y colaboradores (2018), los miembros de la súper familia de factores de transcripción bHLH caracterizados en maíz, están implicados en múltiples procesos fisiológicos relacionados al crecimiento y desarrollo de las plantas (Zhang *et al.*, 2018). Sin embargo, en estudios realizados en *Arabidopsis* y arroz, también se les conoce por su participación en procesos como la germinación, regulación de la floración y respuestas a estrés (Zhang *et al.*, 2018). Este tipo de respuestas también están asociadas a la actividad de los factores de transcripción Dof, por lo que seguramente, los miembros de la familia bHLH juegan un papel determinante durante la activación de genes Dof para la regulación de procesos asociados a los antes mencionados.

Los factores de transcripción de la familia bZIP han sido ampliamente caracterizados principalmente en *Arabidopsis*, donde se ha encontrado su relación con procesos como la defensa de la planta ante patógenos, señalización mediada por luz y condiciones de estrés, maduración de la semilla y desarrollo de la flor (Jakoby *et al.*, 2002).

Por su parte, los miembros de la familia de factores de transcripción ERF han sido relacionados con la regulación de genes asociados a procesos como crecimiento y desarrollo de la planta, así como a la respuesta a diferentes estímulos medioambientales, transducción de señales mediadas por hormonas, respuesta a estrés biótico y abiótico y regulación del metabolismo en diferentes especies de plantas como *Arabidopsis*, tabaco y arroz (Nakano *et al.*, 2006).

La relación que guardan los promotores Dof analizados en las tres variedades, con respecto al contenido en el número y tipo de factores de transcripción que se unen a sus secuencias se ve reflejada en el dendrograma representado en la **Figura 4**. En este, es posible apreciar que los pares de secuencias homólogas entre variedades se agrupan entre sí, salvo algunos casos en PAL o TEO que pueden verse influenciados por el tamaño de secuencia empleado durante el análisis en comparación con las de B73. Sin embargo, esta aproximación indica que, a pesar de las diferencias entre el porcentaje de similitud que existen entre las secuencias, el número y tipo de elementos regulatorios, en este caso, familias de factores de transcripción que reconocen sitios de unión en los promotores Dof se mantiene en gran medida entre las tres variedades.

Para lograr entender más a detalle las diferencias y similitudes entre secuencias de promotores homólogas, se representó la arquitectura de 12 tripletes de promotores Dof homólogos. Aquí fue posible identificar el grado de conservación de elementos *cis* que existen, así como las diferencias entre las secuencias analizadas. La principal característica que muestran estas secuencias es la distribución de elementos regulatorios en posiciones relativas al TSS distintas entre sí, pero manteniendo grupos de elementos regulatorios en el mismo orden. Esto ha sido documentado por Creux y colaboradores (2013), quienes identificaron

módulos de elementos regulatorios *cis* conservados en secuencias homólogas de genes que codifican el gen de la celulosa sintasa (*CesA*) de diferentes especies de eucalipto (Nicky M. Creux *et al.*, 2013). Otros trabajos donde se caracteriza la arquitectura de promotores de genes homólogos, reportan la presencia de módulos de elementos regulatorios en un orden y posición similar entre secuencias, resaltando la importancia del mantenimiento de dichas estructuras que son parte de la expresión basal (Nicole Marie Creux *et al.*, 2008; Freeling y Subramaniam, 2009; Powell *et al.*, 2019; Ravel *et al.*, 2014). Las distribuciones de elementos regulatorios en los promotores Dof, así como la separación entre módulos de elementos regulatorios, son un indicador de la importancia de dichos elementos *cis* para las respuestas basales en la expresión de estos genes y que han sido conservados durante el proceso de domesticación.

Una de las secuencias de promotores que presenta un alto porcentaje de similitud entre las tres variedades pero que cuenta con una región indel es Dof01. El primer miembro de esta familia de factores de transcripción identificados en maíz presenta una distribución de elementos *cis* similar entre las tres variedades, con diferencias puntuales como la presencia/ausencia de cajas TATA asociadas a determinados TSS. La región de inserción que está presente en B73 pero ausente en PAL y TEO puede deberse a un elemento transponible. Los elementos transponibles (TEs) tienen una influencia significativa en la regulación de genes ya que pueden proveer de sitios de unión a factores de transcripción en las regiones promotoras donde se inserten (Bennetzen, 2005; Feschotte, 2008; Hirsch y Springer, 2017; Mejía-Guerra *et al.*, 2015). Otra de las consecuencias que llegan a generar los TEs son la incorporación de sitios de metilación del DNA o la modificación de la estructura de la cromatina, lo que genera alteraciones en el proceso de transcripción al limitar el acceso de factores de transcripción o de la RNA polimerasa (Hirsch y Springer, 2017). Los elementos transponibles fueron descubiertos en maíz y se considera que los genes son “pequeñas islas que flotan en un mar de transposones” (Mita y Boeke, 2016). En maíz, se han realizado trabajos que atribuyen un efecto en la regulación de genes como respuesta a

condiciones de estrés por temperaturas altas (50°C), bajas (5°C), salinidad y radiación UV debido a la presencia de TEs en regiones regulatorias (Makarevitch *et al.*, 2015). Esto puede explicar las principales diferencias en la respuesta a estrés por temperatura entre B73 y PAL debido a la región indel que modifica la arquitectura de estos promotores.

El análisis de las secuencias del promotor Dof01 de las tres variedades donde se identifican los factores de transcripción que reconocen sitios de unión (**Anexo B**), mostró marcadas diferencias entre las secuencias de Dof01 de B73 y TEO en cuanto al número, tipo y posición de factores de transcripción. Sin embargo, la mayoría de los factores de transcripción que potencialmente se unen a estas secuencias se encuentran conservados, por lo que puede inferirse que potencialmente estos factores de transcripción mantienen procesos de regulación de la transcripción de Dof01 básicos como respuesta a estímulos similares y aquellos factores de transcripción en los que difieren pudieran estar relacionados a condiciones asociadas al entorno de desarrollo de cada organismo (Powell *et al.*, 2019).

La localización de factores de transcripción que se unen a la secuencia del promotor Dof01 se complementó con un análisis de correlación entre la expresión del gen Dof01 en B73 y la expresión de estos factores de transcripción en la misma variedad. Entre los elementos con mayor correlación se encontraron los genes ZmMYB38 y ZmHB72 con la mayor correlación positiva y el gen ZmNAC61 con una correlación negativa. En el caso del factor de transcripción ZmMYB38 ha sido asociado con la represión de la expresión de genes estructurales requeridos para la biosíntesis de antocianinas (Ambawat *et al.*, 2013) y de la biosíntesis de lignina (Tak *et al.*, 2017). Por su parte, el factor de transcripción ZmHB72, también conocido como KNOX2 juega un papel importante en la represión del desarrollo de los gametofitos (Furumizu *et al.*, 2015). Por último, ha sido reportado que el gen ZmMYB38 desempeña un rol como regulador negativo en la expresión de genes asociados a la vía fenilpropanoide para la síntesis de flavonoides, así como de diversos genes estructurales en plantas de maíz (Agarwal *et al.*, 2016). La

correlación descrita entre la expresión de factores de transcripción y su posible papel en la regulación de la expresión del gen Dof01 no queda completamente entendida, ya que dicha correlación no implica causalidad. Es necesario realizar la construcción de redes regulatorias como respuesta a condiciones específicas para entender las distintas interacciones que existen entre factores de transcripción. Sin embargo, podemos realizar aproximaciones al entendimiento entre las relaciones antes mencionadas. La información pública disponible en bases de datos permite la elaboración de suposiciones en torno al comportamiento en la expresión de factores de transcripción y la respuesta a condiciones de estrés. En este caso, se analizó la expresión del gen Dof01 relativa a los factores de transcripción anteriormente descritos, tomando en cuenta condiciones de estrés abiótico como salinidad, sequía y temperaturas alta y baja. El principal factor que mayormente afectó la expresión de Dof01 fue la temperatura.

Para entender cómo afectan las diferencias estructurales entre promotores a la expresión de un gen en diferentes variedades, se realizó el análisis de expresión del gen Dof01 en maíz Palomero sometido a condiciones de estrés por baja y alta temperatura. Los resultados mostraron que el comportamiento en la expresión de Dof01 difiere en ambas variedades, donde en PAL, Dof01 se sobre expresa, pero en B73 se reprime significativamente principalmente por baja temperatura. Estos resultados indican que las diferencias entre las secuencias de ambos promotores en cuanto al tipo, número y posición de los factores de transcripción, son los posibles factores que influyen en las diferencias en la actividad del promotor como respuesta a un mismo estímulo.

Es importante estudiar más a fondo la relación que existe entre las diferentes combinaciones de elementos que son responsables de regular la transcripción. Normalmente se visualiza la región promotora como una estructura lineal donde los factores de transcripción acceden fácilmente a los sitios de unión. Sin embargo, estas son meras representaciones gráficas para facilitar la comprensión de la arquitectura de regiones no codificantes regulatorias. La estructura de la cromatina juega un papel importante en la compactación del DNA que le permite a

los factores de transcripción que se unen en regiones distales (miles de bases río arriba del TSS), acceder a los sitios de unión e interactuar a su vez con la maquinaria basal de la transcripción y con la región proximal del promotor (Hernandez-Garcia y Finer, 2014). Otro punto a considerar, son las implicaciones epigenéticas asociadas a la metilación de bases en regiones conocidas como islas CpG, las cuales también juegan un papel importante en la modificación de la actividad de un promotor y por lo tanto en los niveles de expresión sin modificar las secuencias de DNA (Haberle y Lenhard, 1996).

Este trabajo sienta las bases para la identificación de promotores como candidatos para su caracterización funcional. Primeramente, es necesario tener claras las necesidades para el uso de promotores en una industria biotecnológica o para futuras investigaciones, con la finalidad de dirigir la búsqueda de promotores que se ajusten a dichas necesidades. Por ejemplo, identificar aquellos promotores que regulen la expresión de genes de manera tejido específica, bajo ciertas condiciones de estrés o durante cierta etapa de desarrollo de un organismo. Posteriormente, aislar y caracterizar promotores de manera funcional que cumplan con las características deseadas y determinar qué elementos regulatorios son los responsables de dichas características. Esto puede llevar a la creación de promotores sintéticos que muestren patrones de expresión de genes que normalmente no se ven en la naturaleza al utilizar combinaciones de elementos regulatorios *cis* que le den a un promotor las características deseadas.

9. CONCLUSIONES

El proceso de domesticación del maíz influyó en la diferenciación de la arquitectura de promotores Dof. Los elementos regulatorios identificados en estos promotores son altamente conservados entre las tres variedades de maíz analizadas. Sin embargo, aquellos elementos regulatorios *cis* en los que difieren las secuencias de promotores homólogos están asociados a respuestas específicas en cada variedad. Este trabajo proporciona una visión general de la funcionalidad de los promotores y sirve como trabajo previo para la identificación de secuencias de promotores candidatas para la caracterización funcional con potenciales aplicaciones biotecnológicas en la regulación de la expresión de proteínas heterólogas en células vegetales. En ese sentido, se propone al promotor Dof01 de B73 y PAL, como candidatos para la caracterización e identificación del promotor mínimo mediante ensayos de funcionalidad bajo tratamientos de estrés por temperatura.

10. LITERATURA CITADA

- Agarwal, T., Grotewold, E., Doseff, A. I., y Gray, J. (2016). MYB31/MYB42 Syntelogs Exhibit Divergent Regulation of Phenylpropanoid Genes in Maize, Sorghum and Rice. *Scientific Reports*, 6(1), 1-17. <https://doi.org/10.1038/srep28502>
- Ambawat, S., Sharma, P., Yadav, N. R., y Yadav, R. C. (2013, julio 1). MYB transcription factor genes as regulators for plant responses: An overview. *Physiology and Molecular Biology of Plants*, Vol. 19, pp. 307-321. <https://doi.org/10.1007/s12298-013-0179-1>
- Bennetzen, J. L. (2005). Transposable elements, gene creation and genome rearrangement in flowering plants. *Current Opinion in Genetics & Development*, 15(6), 621-627. <https://doi.org/10.1016/J.GDE.2005.09.010>
- Bowman, J. L., Kohchi, T., Yamato, K. T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., Adam, C., Aki, S. S., Althoff, F., Araki, T., Arteaga-Vazquez, M. A., Balasubramanian, S., Barry, K., Bauer, D., Boehm, C. R., Briginshaw, L., Caballero-Perez, J., Catarino, B., Chen, F., Chiyoda, S., Chovatia, M., Davies, K. M., Delmans, M., Demura, T., Dierschke, T., Dolan, L., Dorantes-Acosta, A. E., Eklund, D. M., Florent, S. N., Flores-Sandoval, E., Fujiyama, A., Fukuzawa, H., Galik, B., Grimanelli, D., Grimwood, J., Grossniklaus, U., Hamada, T., Haseloff, J., Hetherington, A. J., Higo, A., Hidakawa, Y., Hundley, H. N., Ikeda, Y., Inoue, K., Inoue, S.-I., Ishida, S., Jia, Q., Kakita, M., Kanazawa, T., Kawai, Y., Kawashima, T., Kennedy, M., Kinoshita, K., Kinoshita, T., Kohara, Y., Koide, E., Komatsu, K., Kopischke, S., Kubo, M., Kyoizuka, J., Lagercrantz, U., Lin, S.-S., Lindquist, E., Lipzen, A. M., Lu, C.-W., De Luna, E., Martienssen, R. A., Minamino, N., Mizutani, M., Mizutani, M., Mochizuki, N., Monte, I., Mosher, R., Nagasaki, H., Nakagami, H., Naramoto, S., Nishitani, K., Ohtani, M., Okamoto, T., Okumura, M., Phillips, J., Pollak, B., Reinders, A., Rövekamp, M., Sano, R., Sawa, S., Schmid, M. W., Shirakawa, M., Solano, R., Spunde, A., Suetsugu, N., Sugano, S., Sugiyama, A., Sun, R., Suzuki, Y., Takenaka, M., Takezawa, D., Tomogane, H., Tsuzuki, M.,

- Ueda, T., Umeda, M., Ward, J. M., Watanabe, Y., Yazaki, K., Yokoyama, R., Yoshitake, Y., Yotsui, I., Zachgo, S., y Schmutz, J. (2017). Insights into Land Plant Evolution Garnered from the *Marchantia polymorpha* Genome. *Cell*, 171(2), 287-304.e15. <https://doi.org/10.1016/j.cell.2017.09.030>
- Butler, J. E., y Kadonaga, J. T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev*, 16. <https://doi.org/10.1101/gad.1026202>
- Cai, X., Zhang, Y., Zhang, C., Zhang, T., Hu, T., Ye, J., Zhang, J., Wang, T., Li, H., y Ye, Z. (2013). Genome-wide Analysis of Plant-specific Dof Transcription Factor Family in Tomato. *Journal of Integrative Plant Biology*, 55(6), 552-566. <https://doi.org/10.1111/jipb.12043>
- Carroll, S. B. (2008, julio 11). Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution. *Cell*, Vol. 134, pp. 25-36. <https://doi.org/10.1016/j.cell.2008.06.030>
- Chang, W.-C., Lee, T.-Y., Huang, H.-D., Huang, H.-Y., y Pan, R.-L. (2008). PlantPAN: Plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. *BMC Genomics*, 9(1), 561. <https://doi.org/10.1186/1471-2164-9-561>
- Chen, Y., y Cao, J. (2015). Comparative Analysis of Dof Transcription Factor Family in Maize. *Plant Molecular Biology Reporter*, 33(5), 1245-1258. <https://doi.org/10.1007/s11105-014-0835-9>
- Chow, C.-N., Zheng, H.-Q., Wu, N.-Y., Chien, C.-H., Huang, H.-D., Lee, T.-Y., Chiang-Hsieh, Y.-F., Hou, P.-F., Yang, T.-Y., y Chang, W.-C. (2016). PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. *Nucleic Acids Research*, 44(D1), D1154-D1160. <https://doi.org/10.1093/nar/gkv1035>
- Creux, Nicky M., De Castro, M. H., Ranik, M., Maleka, M. F., y Myburg, A. A. (2013). Diversity and cis-element architecture of the promoter regions of cellulose synthase genes in *Eucalyptus*. *Tree Genetics & Genomes*, 9(4), 989-1004. <https://doi.org/10.1007/s11295-013-0611-2>

- Creux, Nicole Marie, Ranik, M., Berger, D. K., y Myburg, A. A. (2008). Comparative analysis of orthologous cellulose synthase promoters from *Arabidopsis*, *Populus* and *Eucalyptus*: evidence of conserved regulatory elements in angiosperms. *New Phytologist*, 179(3), 722-737. <https://doi.org/10.1111/j.1469-8137.2008.02517.x>
- Das, S., y Bansal, M. (2019). Variation of gene expression in plants is influenced by gene architecture and structural properties of promoters. *PLOS ONE*, 14(3), e0212678. <https://doi.org/10.1371/journal.pone.0212678>
- Doebley, J. F., Gaut, B. S., y Smith, B. D. (2006, diciembre 29). The Molecular Genetics of Crop Domestication. *Cell*, Vol. 127, pp. 1309-1321. <https://doi.org/10.1016/j.cell.2006.12.006>
- Doebley, J., y Stec, A. (1993). Inheritance of the morphological differences between maize and teosinte: Comparison of results for two F2 populations. *Genetics*. Recuperado de <http://www.mendeley.com/research/inheritance-morphological-differences-between-maize-teosinte-comparison-results-two-f2-populations>
- Eagles, H. A., y Lothrop, J. E. (1994). Highland Maize from Central Mexico—Its Origin, Characteristics, and Use in Breeding Programs. *Crop Science*, 34(1), 11. <https://doi.org/10.2135/cropsci1994.0011183X003400010002x>
- Feschotte, C. (2008, mayo). Transposable elements and the evolution of regulatory networks. *Nature Reviews Genetics*, Vol. 9, pp. 397-405. <https://doi.org/10.1038/nrg2337>
- Freeling, M., y Subramaniam, S. (2009). Conserved noncoding sequences (CNSs) in higher plants. *Current Opinion in Plant Biology*, 12(2), 126-132. <https://doi.org/10.1016/J.PBI.2009.01.005>
- Furumizu, C., Alvarez, J. P., Sakakibara, K., y Bowman, J. L. (2015). Antagonistic Roles for KNOX1 and KNOX2 Genes in Patterning the Land Plant Body Plan Following an Ancient Gene Duplication. *PLoS Genetics*, 11(2), 1-24. <https://doi.org/10.1371/journal.pgen.1004980>
- Gaut, B. S., y Doebley, J. F. (1997). DNA sequence evidence for the segmental

- allotetraploid origin of maize. *Proceedings of the National Academy of Sciences*, 94(13), 6809-6814. <https://doi.org/10.1073/pnas.94.13.6809>
- Gonzalez, D. H. (2016). Introduction to Transcription Factor Structure and Function. En *Plant Transcription Factors* (pp. 3-11). <https://doi.org/10.1016/B978-0-12-800854-6.00001-4>
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., y Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, 40(D1), D1178-D1186. <https://doi.org/10.1093/nar/gkr944>
- Gu, Z., Eils, R., y Schlesner, M. (2016). Genome analysis Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, 32(18), 2847-2849. <https://doi.org/10.1093/bioinformatics/btw313>
- Haberle, V., y Lenhard, B. (1996). Seminars in cell & developmental biology. En 23. Recuperado de <https://spiral.imperial.ac.uk:8443/handle/10044/1/29890>
- Hernandez-Garcia, C. M., y Finer, J. J. (2014). Identification and validation of promoters and cis-acting regulatory elements. *Plant Science*, 217-218, 109-119. <https://doi.org/10.1016/J.PLANTSCI.2013.12.007>
- Hirsch, C. D., y Springer, N. M. (2017). Transposable element influences on gene expression in plants. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1860(1). <https://doi.org/10.1016/j.bbagr.2016.05.010>
- Hoopes, G. M., Hamilton, J. P., Wood, J. C., Esteban, E., Pasha, A., Vaillancourt, B., Provar, N. J., y Buell, C. R. (2018). An Updated Gene Atlas for Maize Reveals Organ-Specific and Stress-Induced Genes. *The Plant Journal*, tpj.14184. <https://doi.org/10.1111/tpj.14184>
- Huang, J., Gao, Y., Jia, H., y Zhang, Z. (2016). Characterization of the teosinte transcriptome reveals adaptive sequence divergence during maize domestication. *Molecular Ecology Resources*, 16(6), 1465-1477. <https://doi.org/10.1111/1755-0998.12526>
- Hufford, M. B., Xu, X., Van Heerwaarden, J., Pyhäjärvi, T., Chia, J. M., Cartwright, R. A., Elshire, R. J., Glaubitz, J. C., Guill, K. E., Kaepler, S. M., Lai, J., Morrell,

- P. L., Shannon, L. M., Song, C., Springer, N. M., Swanson-Wagner, R. A., Tiffin, P., Wang, J., Zhang, G., Doebley, J., McMullen, M. D., Ware, D., Buckler, E. S., Yang, S., y Ross-Ibarra, J. (2012). Comparative population genomics of maize domestication and improvement. *Nature Genetics*, *44*(7), 808-811. <https://doi.org/10.1038/ng.2309>
- Jakoby, M., Weisshaar, B., Dröge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T., y Parcy, F. (2002, marzo 1). bZIP transcription factors in Arabidopsis. *Trends in Plant Science*, Vol. 7, pp. 106-111. [https://doi.org/10.1016/S1360-1385\(01\)02223-3](https://doi.org/10.1016/S1360-1385(01)02223-3)
- Kassambara, A. (2020). ggpubr package | R Documentation. *R package version 0.3.0*. Recuperado de <https://www.rdocumentation.org/packages/ggpubr/versions/0.1.1>
- Koch, M. A., Weisshaar, B., Kroymann, J., Haubold, B., y Mitchell-Olds, T. (2001). Comparative Genomics and Regulatory Evolution: Conservation and Function of the Chs and Apetala3 Promoters. *Molecular Biology and Evolution*, *18*(10), 1882-1891. <https://doi.org/10.1093/oxfordjournals.molbev.a003729>
- Le Hir, R., y Bellini, C. (2013). The Plant-Specific Dof Transcription Factors Family: New Players Involved in Vascular System Development and Functioning in Arabidopsis. *Frontiers in Plant Science*, *4*(May), 164. <https://doi.org/10.3389/fpls.2013.00164>
- Lemmon, Z. H., Bukowski, R., Sun, Q., y Doebley, J. F. (2014). The Role of *cis* Regulatory Evolution in Maize Domestication. *PLoS Genetics*, *10*(11), e1004745. <https://doi.org/10.1371/journal.pgen.1004745>
- Letunic, I., y Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, *44*(W1), W242-W245. <https://doi.org/10.1093/nar/gkw290>
- Livak, K. J., y Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2- $\Delta\Delta$ CT Method. *Methods*, *25*(4), 402-408. <https://doi.org/10.1006/meth.2001.1262>
- Lockton, S., y Gaut, B. S. (2005). Plant conserved non-coding sequences and

- paralogue evolution. *Trends in Genetics*, 21(1), 60-65. <https://doi.org/10.1016/j.tig.2004.11.013>
- Ma, J., Li, M.-Y., Wang, F., Tang, J., y Xiong, A.-S. (2015). Genome-wide analysis of Dof family transcription factors and their responses to abiotic stresses in Chinese cabbage. *BMC Genomics*, 16(1), 33. <https://doi.org/10.1186/s12864-015-1242-9>
- Makarevitch, I., Waters, A. J., West, P. T., Stitzer, M., Hirsch, C. N., Ross-Ibarra, J., y Springer, N. M. (2015). Transposable Elements Contribute to Activation of Maize Genes in Response to Abiotic Stress. *PLoS Genetics*, 11(1), e1004915. <https://doi.org/10.1371/journal.pgen.1004915>
- Mejía-Guerra, M. K., Li, W., Galeano, N. F., Vidal, M., Gray, J., Doseff, A. I., y Grotewold, E. (2015). Core Promoter Plasticity Between Maize Tissues and Genotypes Contrasts with Predominance of Sharp Transcription Initiation Sites. *The Plant Cell*, 27(12), 3309-3320. <https://doi.org/10.1105/tpc.15.00630>
- Meyer, R. S., y Purugganan, M. D. (2013, diciembre 18). Evolution of crop species: Genetics of domestication and diversification. *Nature Reviews Genetics*, Vol. 14, pp. 840-852. <https://doi.org/10.1038/nrg3605>
- Mita, P., y Boeke, J. D. (2016, abril 1). How retrotransposons shape genome regulation. *Current Opinion in Genetics and Development*, Vol. 37, pp. 90-100. <https://doi.org/10.1016/j.gde.2016.01.001>
- Moreno-Risueno, M. Á., Martínez, M., Vicente-Carbajosa, J., y Carbonero, P. (2007). The family of DOF transcription factors: from green unicellular algae to vascular plants. *Molecular Genetics and Genomics*, 277(4), 379-390. <https://doi.org/10.1007/s00438-006-0186-9>
- Nakano, T., Suzuki, K., Fujimura, T., y Shinshi, H. (2006). Genome-wide analysis of the ERF gene family in arabidopsis and rice. *Plant Physiology*, 140(2), 411-432. <https://doi.org/10.1104/pp.105.073783>
- Petrillo, E., Herz, M. A. G., Barta, A., Kalyna, M., y Kornblihtt, A. R. (2014, octubre 1). Let there be light: Regulation of gene expression in plants. *RNA Biology*, Vol. 11, pp. 1215-1220. <https://doi.org/10.4161/15476286.2014.972852>

- Powell, R. V., Willett, C. R., Goertzen, L. R., y Rashotte, A. M. (2019). Lineage specific conservation of *cis*-regulatory elements in Cytokinin Response Factors. *Scientific Reports*, 9(1), 13387. <https://doi.org/10.1038/s41598-019-49741-6>
- Prasanna, B. M. (2012). Diversity in global maize germplasm: Characterization and utilization. *Journal of Biosciences*, 37(5), 843-855. <https://doi.org/10.1007/s12038-012-9227-1>
- R Foundation for Statistical Computing. (2008). R: A language and environment for statistical computing. En *Vienna, Austria* (Vol. 2).
- Ravel, C., Fiquet, S., Boudet, J., Dardevet, M., Vincent, J., Merlino, M., Michard, R., y Martre, P. (2014). Conserved *cis*-regulatory modules in promoters of genes encoding wheat high-molecular-weight glutenin subunits. *Frontiers in Plant Science*, 5, 621. <https://doi.org/10.3389/fpls.2014.00621>
- Reineke, A. R., Bornberg-Bauer, E., y Gu, J. (2011). Evolutionary divergence and limits of conserved non-coding sequence detection in plant genomes. *Nucleic acids research*, 39(14), 6029-6043. <https://doi.org/10.1093/nar/gkr179>
- Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T. A., Minx, P., Reily, A. D., Courtney, L., Kruchowski, S. S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., Rock, S. M., Belter, E., Du, F., Kim, K., Abbott, R. M., Cotton, M., Levy, A., Marchetto, P., Ochoa, K., Jackson, S. M., Gillam, B., Chen, W., Yan, L., Higginbotham, J., Cardenas, M., Waligorski, J., Applebaum, E., Phelps, L., Falcone, J., Kanchi, K., Thane, T., Scimone, A., Thane, N., Henke, J., Wang, T., Ruppert, J., Shah, N., Rotter, K., Hodges, J., Ingenthron, E., Cordes, M., Kohlberg, S., Sgro, J., Delgado, B., Mead, K., Chinwalla, A., Leonard, S., Crouse, K., Collura, K., Kudrna, D., Currie, J., He, R., Angelova, A., Rajasekar, S., Mueller, T., Lomeli, R., Scara, G., Ko, A., Delaney, K., Wissotski, M., Lopez, G., Campos, D., Braidotti, M., Ashley, E., Golser, W., Kim, H., Lee, S., Lin, J., Dujmic, Z., Kim, W., Talag, J., Zuccolo, A., Fan, C., Sebastian, A., Kramer, M., Spiegel, L., Nascimento, L., Zutavern, T., Miller, B., Ambroise, C., Muller, S., Spooner, W., Narechania, A., Ren, L., Wei, S., Kumari, S., Faga, B., Levy, M. J.,

- McMahan, L., Van Buren, P., Vaughn, M. W., Ying, K., Yeh, C.-T., Emrich, S. J., Jia, Y., Kalyanaraman, A., Hsia, A.-P., Barbazuk, W. B., Baucom, R. S., Brutnell, T. P., Carpita, N. C., Chaparro, C., Chia, J.-M., Deragon, J.-M., Estill, J. C., Fu, Y., Jeddelloh, J. A., Han, Y., Lee, H., Li, P., Lisch, D. R., Liu, S., Liu, Z., Nagel, D. H., McCann, M. C., SanMiguel, P., Myers, A. M., Nettleton, D., Nguyen, J., Penning, B. W., Ponnala, L., Schneider, K. L., Schwartz, D. C., Sharma, A., Soderlund, C., Springer, N. M., Sun, Q., Wang, H., Waterman, M., Westerman, R., Wolfgruber, T. K., Yang, L., Yu, Y., Zhang, L., Zhou, S., Zhu, Q., Bennetzen, J. L., Dawe, R. K., Jiang, J., Jiang, N., Presting, G. G., Wessler, S. R., Aluru, S., Martienssen, R. A., Clifton, S. W., McCombie, W. R., Wing, R. A., y Wilson, R. K. (2009). The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science*, 326(5956), 1112 LP - 1115. <https://doi.org/10.1126/science.1178534>
- Schor, I. E., Degner, J. F., Harnett, D., Cannavò, E., Casale, F. P., Shim, H., Garfield, D. A., Birney, E., Stephens, M., Stegle, O., y Furlong, E. E. M. (2017). Promoter shape varies across populations and affects promoter evolution and expression noise. *Nature Genetics*, 49(4), 550-558. <https://doi.org/10.1038/ng.3791>
- Shahmuradov, I. A., Umarov, R. K., y Solovyev, V. V. (2017). TSSPlant: a new tool for prediction of plant Pol II promoters. *Nucleic Acids Research*, 45(8), gkw1353. <https://doi.org/10.1093/nar/gkw1353>
- Shigyo, M., Tabei, N., Yoneyama, T., y Yanagisawa, S. (2007). Evolutionary processes during the formation of the plant-specific dof transcription factor family. *Plant and Cell Physiology*, 48(1), 179-185. <https://doi.org/10.1093/pcp/pcl044>
- Smale, S. T., y Kadonaga, J. T. (2003). The RNA Polymerase II Core Promoter. *Annual Review of Biochemistry*, 72(1), 449-479. <https://doi.org/10.1146/annurev.biochem.72.121801.161520>
- Srivastava, R., Rai, K. M., Srivastava, M., Kumar, V., Pandey, B., Singh, S. P., Bag, S. K., Singh, B. D., Tuli, R., y Sawant, S. V. (2014). Distinct Role of Core Promoter Architecture in Regulation of Light-Mediated Responses in Plant

- Genes. *Molecular Plant*, 7(4), 626-641. <https://doi.org/10.1093/MP/SST146>
- Swanson-Wagner, R., Briskine, R., Schaefer, R., Hufford, M. B., Ross-Ibarra, J., Myers, C. L., Tiffin, P., y Springer, N. M. (2012). Reshaping of the maize transcriptome by domestication. *Proceedings of the National Academy of Sciences of the United States of America*, 109(29), 11878-11883. <https://doi.org/10.1073/pnas.1201961109>
- Swinnen, G., Goossens, A., y Pauwels, L. (2016). Lessons from Domestication: Targeting *Cis*-Regulatory Elements for Crop Improvement. *Trends in Plant Science*, Vol. 21. <https://doi.org/10.1016/j.tplants.2016.01.014>
- Tak, H., Negi, S., y Ganapathi, T. R. (2017). Overexpression of MusaMYB31, a R2R3 type MYB transcription factor gene indicate its role as a negative regulator of lignin biosynthesis in banana. *PLOS ONE*, 12(2), e0172695. <https://doi.org/10.1371/journal.pone.0172695>
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., y Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, 30(12), 2725-2729. <https://doi.org/10.1093/molbev/mst197>
- Tirosh, I., Barkai, N., y Verstrepen, K. J. (2009). Promoter architecture and the evolvability of gene expression. *Journal of Biology*, 8(11), 95. <https://doi.org/10.1186/jbiol204>
- Toledo-Ortiz, G., Huq, E., y Quail, P. H. (2003). The Arabidopsis basic/helix-loop-helix transcription factor family. *Plant Cell*, 15(8), 1749-1770. <https://doi.org/10.1105/tpc.013839>
- Vandepoele, K., Quimbaya, M., Casneuf, T., De Veylder, L., y Van de Peer, Y. (2009). Unraveling transcriptional control in Arabidopsis using *cis*-regulatory elements and coexpression networks. *Plant physiology*, 150(2), 535-546. <https://doi.org/10.1104/pp.109.136028>
- Vielle-Calzada, J.-P., Martínez de la Vega, O., Hernández-Guzmán, G., Ibarra-Laclette, E., Alvarez-Mejía, C., Vega-Arreguín, J. C., Jiménez-Moraila, B., Fernández-Cortés, A., Corona-Armenta, G., Herrera-Estrella, L., y Herrera-Estrella, A. (2009). The Palomero genome suggests metal effects on

- domestication. *Science (New York, N.Y.)*, 326(5956), 1078. <https://doi.org/10.1126/science.1178437>
- Wang, H., Zhao, S., Gao, Y., y Yang, J. (2017). Characterization of dof transcription factors and their responses to osmotic stress in poplar (*populus trichocarpa*). *PLoS ONE*, 12(1), e0170210. <https://doi.org/10.1371/journal.pone.0170210>
- Wang, Y., Wang, X., y Paterson, A. H. (2012). Genome and gene duplications and gene expression divergence: a view from plants. *Annals of the New York Academy of Sciences*, 1256(1), 1-14. <https://doi.org/10.1111/j.1749-6632.2011.06384.x>
- Wittkopp, P. J., y Kalay, G. (2012). Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics*, 13(1), 59-69. <https://doi.org/10.1038/nrg3095>
- Wright, S. I., Bi, I. V., Schroeder, S. C., Yamasaki, M., Doebley, J. F., McMullen, M. D., y Gaut, B. S. (2005). The effects of artificial selection on the maize genome. *Science*, 308(5726), 1310-1314. <https://doi.org/10.1126/science.1107891>
- Yamamoto, Y. Y., Ichida, H., Abe, T., Suzuki, Y., Sugano, S., y Obokata, J. (2007). Differentiation of core promoter architecture between plants and mammals revealed by LDSS analysis. *Nucleic Acids Research*, 35(18), 6219-6226. <https://doi.org/10.1093/nar/gkm685>
- Yamamoto, Y. Y., Ichida, H., Matsui, M., Obokata, J., Sakurai, T., Satou, M., Seki, M., Shinozaki, K., y Abe, T. (2007). Identification of plant promoter constituents by analysis of local distribution of short sequences. *BMC Genomics*, 8(1), 1-23. <https://doi.org/10.1186/1471-2164-8-67>
- Yamamoto, Y. Y., Yoshioka, Y., Hyakumachi, M., y Obokata, J. (2011). Characteristics of core promoter types with respect to gene structure and expression in *Arabidopsis thaliana*. *DNA research: an international journal for rapid publication of reports on genes and genomes*, 18(5), 333-342. <https://doi.org/10.1093/dnares/dsr020>
- Yanagisawa, S. (1995). A novel DNA-binding domain that may form a single zinc finger motif. *Nucleic Acids Research*, 23(17), 3403-3410.

- <https://doi.org/10.1093/nar/23.17.3403>
- Yanagisawa, S. (2002). The Dof family of plant transcription factors. *Trends in Plant Science*, 7(12), 555-560. [https://doi.org/http://dx.doi.org/10.1016/S1360-1385\(02\)02362-2](https://doi.org/http://dx.doi.org/10.1016/S1360-1385(02)02362-2)
- Yanagisawa, S. (2004). Dof Domain Proteins: Plant-Specific Transcription Factors Associated with Diverse Phenomena Unique to Plants. *Plant and Cell Physiology*, 45(4), 386-391. <https://doi.org/10.1093/pcp/pch055>
- Yanagisawa, S., y Izui, K. (1990). Multiple interactions between tissue-specific nuclear proteins and the promoter of the phosphoenolpyruvate carboxylase gene for C4 photosynthesis in *Zea mays*. *MGG Molecular & General Genetics*, 224(3), 325-332. <https://doi.org/10.1007/BF00262425>
- Yanagisawa, S., y Izui, K. (1993). Molecular cloning of two DNA-binding proteins of maize that are structurally different but interact with the same sequence motif. *Journal of Biological Chemistry*, 268(21), 16028-16036. Recuperado de <http://www.jbc.org/content/268/21/16028.abstract>
- Yang, N., Xu, X.-W., Wang, R.-R., Peng, W.-L., Cai, L., Song, J.-M., Li, W., Luo, X., Niu, L., Wang, Y., Jin, M., Chen, L., Luo, J., Deng, M., Wang, L., Pan, Q., Liu, F., Jackson, D., Yang, X., Chen, L.-L., y Yan, J. (2017). Contributions of *Zea mays* subspecies mexicana haplotypes to modern maize. *Nature communications*, 8(1), 1874. <https://doi.org/10.1038/s41467-017-02063-5>
- Yang, X., Tuskan, G. A., y Cheng, M. Z.-M. (2006). Divergence of the Dof gene families in poplar, *Arabidopsis*, and rice suggests multiple modes of gene evolution after duplication. *Plant physiology*, 142(3), 820-830. <https://doi.org/10.1104/pp.106.083642>
- Yu, C.-P., Lin, J.-J., y Li, W.-H. (2016). Positional distribution of transcription factor binding sites in *Arabidopsis thaliana*. *Scientific Reports*, 6(1), 25164. <https://doi.org/10.1038/srep25164>
- Zhang, T., Lv, W., Zhang, H., Ma, L., Li, P., Ge, L., y Li, G. (2018). Genome-wide analysis of the basic Helix-Loop-Helix (bHLH) transcription factor family in maize. *BMC Plant Biology*, 18, 235. <https://doi.org/10.1186/s12870-018-1441-z>

11. ANEXOS

Anexo A

Tabla III. Lista de motivos localizados en secuencias de promotores Dof y simbología empleada para su identificación.

Motivo	Nombre	Símbolo	Motivo	Nombre	Símbolo
1	ANAERO1CONSENSUS		21	ARELIKEGHPGDFR2	
2	ANAERO3CONSENSUS		22	ARECOREZMGAPC4	
3	ANAERO4CONSENSUS		23	GCBP2ZMGAPC4	
4	DOFCOREZM		24	QELEMENTZMZM13	
5	MYBPLANT		25	TATA-300MOTIFZMZEIN	
6	MYBPZM		26	INTRONLOWER	
7	POLASIG3		27	TATA-314MOTIFZMSBE1	
8	TATAPVTRNALEU		28	OPAQUE2ZMB32	
9	ABREAZMRAB28		29	CGTGTSPHZMC1	
10	ABREBZMRAB28		30	INTRONUPPER	
11	ANAERO2CONSENSUS		31	OCCENHANMOTIFAT	
12	DRE1COREZMRAB17		32	SPHZMC1	
13	DRE2COREZMRAB17		33	C1MOTIFZMBZ2	
14	DRECRTCOREAT		34	CONSERVED11NTZMATP1	
15	GCAACREPEATZMZEIN		35	ABRECE3ZMRAB28	
16	GRAZMRAB17		36	ANAEROBICC/SZMGAPC4	

17	IDRSZMFER1		37	RYREPEAT4	
18	MNF1ZMPPC1		38	TATA-284MOTIFZMSBE1	
19	OCTAMOTIF2		39	PSREGIONZMZM13	
20	23BPZM27KDAZEIN		40	SPHCOREZMC1	

Anexo B

ZmTEO_Dof01	-----CCCCCCCCA CCCCAAAAA AAAAAACCGGTGCTGCTGTCTTCAGTCCCTCCCATC	54
ZmB73_Dof01	CACCCCCCCTCCC CCCCAAAAA AAAAAACCGGTGCTGCTGTCTTCAGTCCCTCCCATC	60
ZmPAL_Dof01	-----	0
	AT-Hook	
ZmTEO_Dof01	AGTTGTTCTTCATCATGTCCTTTATGTTCCCTTT TGACT ATTTGATTC GAGAAAAA AAG	114
ZmB73_Dof01	AGTTGTTCTTCATCATGTCCTTTATGTTCCCTTT TGACT ATTTGATTC GAGAAA -AAAAG	119
ZmPAL_Dof01	-----	0
	TALE AT-Hook	
ZmTEO_Dof01	TTGCGACTTTCATGATCCA AAGCATATCTTTTCCCTTTTCCAATACTT GCAACTGCTCTG	174
ZmB73_Dof01	TTGCGACTTTCATGATCCA AAGCATATCTTTTCCCTTTTCCAATACTT GCAACTGCTCTG	179
ZmPAL_Dof01	--GCGACTTTCATGATCCA AAGCATATCTTTTCCCTTTTCCAATACTT GCAACTGCTCTG	58
	***** GATA	
ZmTEO_Dof01	GCTGGTGGCTGCCGCCCTGTTCCTCGTTAGAAACTTGCAGATAGTC TTAGTAGTACTGC	234
ZmB73_Dof01	GCTGGTGGCTGCCGCCCTGTTCCTCGTTAGAAACTTGCAGATAGTC TTAGTAGTACTGC	239
ZmPAL_Dof01	GCTGGTGGCTGCCGCCCTGTTCCTCGTTAGAAACTTGCAGATAGTC TTAGTAGTACTGC	118
	***** SBP/bHLH	
ZmTEO_Dof01	AAGTGC AAGATATGGTCCGATTGTCCCTCTAGCTGCGGATTGTGTTTCTTAATCTA---	291
ZmB73_Dof01	AAGTGC AAGATATGGTCCGATTGTCCCTCTAGCTGCGGATTGTGTTTCTTAATCTATGG	299
ZmPAL_Dof01	AAGTGC AAGATATGGTCCGATTGTCCCTCTAGCTGCGGATTGTGTTTCTTAATCTATGG	178

ZmTEO_Dof01	----TGGAGACTAGGTTAGGTTAGCTACACATCATGGCAACCTACTTTTCCGTTTAGCTG	347
ZmB73_Dof01	ACTAT GGAGGGTAGGTTA GGTTAGCTACACATCATGGCAACCTACTTTTCCGTTAAGCTG	359
ZmPAL_Dof01	ACTAT GGAGGGTAGGTTA GGTTAGCTACACATCATGGCAACCTACTTTTCCGTTAAGCTG	238
	***** MYB	
ZmTEO_Dof01	GTTGG TGCAATTGCA AAAGGCTGACCTTGGGGCTGTGTTTGGCAACAACCTAGTAGCCGTA	407
ZmB73_Dof01	GTTGGTCAATTGCAAAAGGCTGACCTTGGGGCTGTGTTTGGCAACAACCTAGTAGCCGTA	419
ZmPAL_Dof01	GTTGGTCAATTGCAAAAGGCTGACCTTGGGGCTGTGTTTGGCAACAACCTAGTAGCCGTA	298
	***** AT-Hook TALE	
ZmTEO_Dof01	TGACTTACTAGGACTCAGTT TGTCTCT GAAACAACCGATGCTTCGGTTTTTGTG AGAATAT	467
ZmB73_Dof01	TGACTTACGAGGGTTCAATT TGTCTCT GAAACAACCGATGCTTCGGTTTTTGTGAGGATAT	479
ZmPAL_Dof01	TGACTTACGAGGGTTCAATT TGTCTCT GAAACAACCGATGCTTCGGTTTTTGTGAGGATAT	358
	***** TALE ARF G2-like	
ZmTEO_Dof01	GCAGAGAGAAGATCAGCT ACTCGTAGTGATAGGCGAATCCCTTTCCCTCGCTTGAGTA	527
ZmB73_Dof01	GCAGAGAGAAGATCAGCT ACACGTAGTGATAGGCGAATCCCTTTC-CCCGCTTGAGTA	538
ZmPAL_Dof01	GCAGAGAGAAGATCAGCT ACACGTAGTGATAGGCGAATCCCTTTC-CCCGCTTGAGTA	417
	***** GATA/bZIP	
ZmTEO_Dof01	CTCATTAAAGCCA TTTCTTGATGTGGCT ACCCGTCGGTCTGTGGAAGTAGTGTCTTA	587
ZmB73_Dof01	CTCATTAAAGCCA TTTCTTGATGTGGCT ACCCGTCGGTCTGTGGAAGTAGTGTCTTA	598
ZmPAL_Dof01	CTCATTAAAGCCA TTTCTTGATGTGGCT ACCCGTCGGTCTGTGGAAGTAGTGTCTTA	477
	***** Dof	

ZmTEO_Dof01	GTTATTTGCCAGTAAATGTCATAGACATATTTTACAGCATCTTACAATCATATTTTCGTTC	1641
ZmB73_Dof01	GTTATTTGCCAGTAAATGTAATAGACATATTTTACATCATCTTACAATCATATTTTCCTTC	1844
ZmPAL_Dof01	-----TGTCAGACATATTTTACAGCATCTTACAATCACATTTTCCTTC	1513
	* * * * *	
	ARR-B	
ZmTEO_Dof01	TATCTCAATACATGTATTTACTTAGTCTTCGTGTATTTGTCAGGATGTATGTGAAGATA	1701
ZmB73_Dof01	TATCTCAATACATCTATTTACTTAGTCTTCGTGTATTTGTCAGGATTTATGTGAAGATA	1904
ZmPAL_Dof01	TATCTCAATACATGTATTTACTTAGTCTTCGTGTATTTGTCAGGATGTATGTGAAGATA	1573

	bZIP/EIL/NAC	EIL
ZmTEO_Dof01	GAGGGGACATG	1712
ZmB73_Dof01	GAGGGGACATG	1915
ZmPAL_Dof01	GAGGGGACATG	1584

Anexo C

Tabla IV. Factores de transcripción que reconocen sitios de unión en el promotor Dof01 de B73, PAL y TEO.

TF ID	TF family	TF ID	TF family	TF ID	TF family
GRMZM2G059939	AP2; B3	GRMZM2G008898	bHLH	GRMZM2G336408	bHLH
GRMZM2G159592	AP2; B3	GRMZM2G024530	bHLH	GRMZM2G350165	bHLH
GRMZM2G169654	AP2; B3	GRMZM2G027563	bHLH	GRMZM2G378653	bHLH
GRMZM2G004878	AT-Hook	GRMZM2G030762	bHLH	GRMZM2G385543	bHLH
GRMZM2G008456	AT-Hook	GRMZM2G046493	bHLH	GRMZM2G407119	bHLH
GRMZM2G089562	AT-Hook	GRMZM2G067654	bHLH	AC155610.2_FG005	bZIP
GRMZM2G106133	AT-Hook	GRMZM2G074438	bHLH	GRMZM2G002075	bZIP
GRMZM2G123887	AT-Hook	GRMZM2G080168	bHLH	GRMZM2G011932	bZIP
GRMZM2G127034	AT-Hook	GRMZM2G083504	bHLH	GRMZM2G019106	bZIP
GRMZM2G150950	AT-Hook	GRMZM2G092091	bHLH	GRMZM2G074373	bZIP
GRMZM2G159926	AT-Hook	GRMZM2G101350	bHLH	GRMZM2G117851	bZIP
GRMZM2G164489	AT-Hook	GRMZM2G116785	bHLH	GRMZM2G129247	bZIP
GRMZM2G028980	B3	GRMZM2G137358	bHLH	GRMZM2G132868	bZIP
GRMZM2G035405	B3	GRMZM2G137541	bHLH	GRMZM2G153144	bZIP
GRMZM2G073750	B3	GRMZM2G144275	bHLH	GRMZM2G157722	bZIP
GRMZM2G086949	B3	GRMZM2G145579	bHLH	GRMZM2G161009	bZIP
GRMZM2G160005	B3	GRMZM2G171818	bHLH	GRMZM2G164622	bZIP
GRMZM2G378580	B3	GRMZM2G176289	bHLH	GRMZM2G168079	bZIP
AC197246.3_FG002	bHLH	GRMZM2G178182	bHLH	GRMZM2G171370	bZIP
AC212449.4_FG005	bHLH	GRMZM2G180406	bHLH	GRMZM2G478417	bZIP
AC233869.1_FG003	bHLH	GRMZM2G180452	bHLH	AC187157.4_FG005	bZIP

Continuación...

TF ID	TF family	TF ID	TF family	TF ID	TF family
GRMZM2G003509	bZIP	GRMZM2G101058	GATA	GRMZM2G034113	Homeodomain
GRMZM2G109987	bZIP	GRMZM2G110295	GATA	GRMZM2G038252	Homeodomain
GRMZM2G178102	bZIP	GRMZM2G113098	GATA	GRMZM2G041462	Homeodomain
GRMZM2G469551	bZIP	GRMZM2G114775	GATA	GRMZM2G044752	Homeodomain
AC185108.3_FG011	C2H2	GRMZM2G135381	GATA	GRMZM2G047715	Homeodomain
GRMZM2G009406	Dof	GRMZM2G138967	GATA	GRMZM2G055243	Homeodomain
GRMZM2G146283	Dof	GRMZM2G140669	GATA	GRMZM2G060507	Homeodomain
GRMZM2G179069	Dof	GRMZM2G163200	GATA	GRMZM2G061101	Homeodomain
MNB1A_MAIZE	Dof	GRMZM2G325850	GATA	GRMZM2G068672	Homeodomain
AC234203.1_FG011	EIN3	GRMZM2G379005	GATA	GRMZM2G069274	Homeodomain
GRMZM2G003558	EIN3	GRMZM2G397616	GATA	GRMZM2G069365	Homeodomain
GRMZM2G006752	EIN3	GRMZM2G404973	GATA	GRMZM2G094241	Homeodomain
GRMZM2G033570	EIN3	AC218972.3_FG005	Homeodomain	GRMZM2G105834	Homeodomain
GRMZM2G151811	EIN3	AC235534.1_FG007	Homeodomain	GRMZM2G106276	Homeodomain
GRMZM2G317584	EIN3	GRMZM2G001289	Homeodomain	GRMZM2G116658	Homeodomain
AC194965.4_FG004	GATA	GRMZM2G002225	Homeodomain	GRMZM2G117164	Homeodomain
AC202864.3_FG002	GATA	GRMZM2G004957	Homeodomain	GRMZM2G118063	Homeodomain
GRMZM2G031983	GATA	GRMZM2G017087	Homeodomain	GRMZM2G122897	Homeodomain
GRMZM2G039586	GATA	GRMZM2G026643	Homeodomain	GRMZM2G123140	Homeodomain
GRMZM2G044576	GATA	GRMZM2G028041	Homeodomain	GRMZM2G126239	Homeodomain
GRMZM2G054615	GATA	GRMZM2G032332	Homeodomain	GRMZM2G126646	Homeodomain

Continuación...

TF ID	TF family	TF ID	TF family	TF ID	TF family
GRMZM2G126808	Homeodomain	GRMZM2G097059	MADS box	GRMZM2G049695	Myb/SANT
GRMZM2G127537	Homeodomain	GRMZM2G099522	MADS box	GRMZM2G050305	Myb/SANT
GRMZM2G130442	Homeodomain	GRMZM2G110153	MADS box	GRMZM2G056400	Myb/SANT
GRMZM2G131476	Homeodomain	GRMZM2G129034	MADS box	GRMZM2G056407	Myb/SANT
GRMZM2G134260	Homeodomain	GRMZM2G147716	MADS box	GRMZM2G057408	Myb/SANT
GRMZM2G135447	Homeodomain	GRMZM2G148693	MADS box	GRMZM2G060834	Myb/SANT
GRMZM2G142962	Homeodomain	GRMZM2G152862	MADS box	GRMZM2G070865	Myb/SANT
GRMZM2G145690	Homeodomain	GRMZM2G159397	MADS box	GRMZM2G077789	Myb/SANT
GRMZM2G148074	Homeodomain	GRMZM2G160565	MADS box	GRMZM2G078636	Myb/SANT
GRMZM2G159431	Homeodomain	GRMZM2G160687	MADS box	GRMZM2G082264	Myb/SANT
GRMZM2G351330	Homeodomain	GRMZM2G171365	MADS box	GRMZM2G084583	Myb/SANT
GRMZM2G366130	Homeodomain	GRMZM2G359952	MADS box	GRMZM2G089244	Myb/SANT
GRMZM2G370332	Homeodomain	GRMZM2G475390	MADS box	GRMZM2G099797	Myb/SANT
GRMZM2G433591	Homeodomain	AC197146.3 FG002	Myb/SANT	GRMZM2G104789	Myb/SANT
GRMZM2G452178	Homeodomain	AC209374.4 FG003	Myb/SANT	GRMZM2G108959	Myb/SANT
GRMZM2G003514	MADS box	GRMZM2G000818	Myb/SANT	GRMZM2G123308	Myb/SANT
GRMZM2G018589	MADS box	GRMZM2G015021	Myb/SANT	GRMZM2G124715	Myb/SANT
GRMZM2G046885	MADS box	GRMZM2G034110	Myb/SANT	GRMZM2G126834	Myb/SANT
GRMZM2G052890	MADS box	GRMZM2G034563	Myb/SANT	GRMZM2G135052	Myb/SANT
GRMZM2G071620	MADS box	GRMZM2G041415	Myb/SANT	GRMZM2G139284	Myb/SANT
GRMZM2G072582	MADS box	GRMZM2G045748	Myb/SANT	GRMZM2G160838	Myb/SANT

Continuación...

TF ID	TF family	TF ID	TF family	TF ID	TF family
GRMZM2G162434	Myb/SANT	GRMZM2G058588	SBP	GRMZM2G148467	SBP
GRMZM2G170148	Myb/SANT	GRMZM2G061734	SBP	GRMZM2G156621	SBP
GRMZM2G175827	Myb/SANT	GRMZM2G065451	SBP	GRMZM2G156756	SBP
GRMZM2G315506	Myb/SANT	GRMZM2G067624	SBP	GRMZM2G163813	SBP
GRMZM2G370425	Myb/SANT	GRMZM2G081127	SBP	GRMZM2G168229	SBP
GRMZM2G400489	Myb/SANT	GRMZM2G097275	SBP	GRMZM2G169270	SBP
GRMZM2G405094	Myb/SANT	GRMZM2G098557	SBP	GRMZM2G307588	SBP
GRMZM2G419239	Myb/SANT	GRMZM2G101499	SBP	GRMZM2G390470	SBP
GRMZM2G431156	Myb/SANT	GRMZM2G101511	SBP	GRMZM2G414805	SBP
GRMZM2G003715	NAC; NAM	GRMZM2G126018	SBP	GRMZM2G460544	SBP
GRMZM2G163914	NAC; NAM	GRMZM2G126827	SBP	GRMZM2G149238	TBP
GRMZM2G036297	SBP	GRMZM2G133646	SBP	GRMZM2G161418	TBP
GRMZM2G040785	SBP	GRMZM2G138421	SBP		

Anexo D



Mexican Journal of Biotechnology 2020, 5(2):135-161

ORIGINAL RESEARCH



Conservation of *cis*-regulatory elements in maize *Dof01* promoter: a key regulator of the cold stress response

Elementos regulatorios *cis* conservados en el promotor *Dof01* de maíz: regulador clave de la respuesta a estrés por frío

Héctor Daniel Garza-Avelar¹, Axel Tiessen-Favier^{2†}, Luis Herrera-Estrella³, Felipe Ascencio-Valle¹, César Cardona-Félix⁴, Gracia Alicia Gómez-Anduro^{1*}

¹ Centro de Investigaciones Biológicas del Noroeste (CIBNOR) Av. Instituto Politécnico Nacional 195, Col. Playa Palo de Santa Rita, La Paz, B.C.S. 23096, México.

² Departamento de Ingeniería Genética, CINVESTAV, Unidad Irapuato, Km. 9.6 Libramiento Norte, 36821 Irapuato, México.

³ Laboratorio Nacional de Genómica para la Biodiversidad, Unidad de Genómica Avanzada del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Guanajuato, México.

⁴ CONACyT–Instituto Politécnico Nacional, Centro Interdisciplinario de Ciencias Marinas, La Paz, México.

*Corresponding author

E-mail address: ggomez@cibnor.mx (G. Gómez-Anduro)

Article history:

Received: 26 February 2020 / Received in revised form: 23 April 2020 / Accepted: 24 April 2020 / Published online: 4 May 2020.

<https://doi.org/10.29267/mxjb.2020.5.2.135>

Supplementary material:156-161.

†In Memory of Dr. Axel Tiessen Favier

ABSTRACT

Transcription regulation is a key process in adapting organisms to the environment. It is necessary to identify the regulatory elements responsible for the transcriptional activity of a promoter, which shape its architecture, to associate them with a specific type of response. For that, it is essential to analyze sequences of promoters of homologous genes of different landraces. In the present work, we perform a comparative analysis of the sequences of the Dof01 promoter of five maize (*Zea mays*) landraces (B73, Palomero Toluqueño [PAL], Teosinte [TEO], W22 and Mo17). We identify *cis*-regulatory elements (CREs) responsible for promoter architecture and to which families of transcription factors they belong and their spatio-temporal expression profiles. In addition, we subject seedlings of PAL to conditions of stress due to low (4°C) and high (42°C) temperatures. Sequence analysis showed that there is an insertion/deletion element of 179 bp that is present in the Dof01 promoter of B73, but absent in the sequences of the other four varieties. Also, we identify CREs that may be related to the response at low temperatures. The analysis of the five sequences of the Dof01 promoter showed that the regulatory elements are highly conserved. On the other hand, the expression data of the PAL Dof01 gene show an overexpression with respect to control due to cold response. This work is the first step for the functional characterization of the Dof01 promoter, to identify the regulatory elements associated with the cold stress response and obtain the minimum promoter necessary to regulate transcription.

Keywords: *cis*-regulatory elements, Dof, promoter, transcription factors, transcription factor binding sites, transcription start site

RESÚMEN

La regulación de la transcripción es un proceso clave en el cual es necesario conocer la arquitectura de los promotores y los elementos regulatorios responsables de su actividad para asociarlos a respuestas específicas. En el presente trabajo, realizamos un análisis comparativo de las secuencias del promotor Dof01 de cinco variedades de maíz (*Zea mays*) (B73, Palomero Toluqueño [PAL], Teosinte [TEO], W22 y Mo17). Identificamos elementos regulatorios *cis* (CREs) responsables de la arquitectura del promotor y las familias de factores de transcripción a los que pertenecen, así como sus perfiles de expresión espacio-temporal. Sometimos plántulas de PAL a condiciones de estrés por temperatura baja (4°C) y alta (42°C). El análisis de secuencias mostró que hay un elemento de inserción/delección de 179 pb que está presente en el promotor Dof01 de B73 pero ausente en las otras cuatro variedades. También, identificamos CREs que pueden estar relacionados con la respuesta a bajas temperaturas. El análisis de secuencias mostró que los CREs están altamente conservados. Los datos de expresión muestran que el gen Dof01 de PAL se sobreexpresa con respecto al control debido a la respuesta a frío. Este trabajo es el primer paso para la caracterización funcional del promotor Dof01, para identificar los elementos regulatorios asociados con la respuesta a estrés por frío y obtener el promotor mínimo necesario para regular la transcripción.

Palabras clave: Dof, elementos regulatorios *cis*, factores de transcripción, promotor, sitio de inicio de la transcripción, sitio de unión a factores de transcripción

1. INTRODUCCIÓN

La unión de factores de transcripción (FTs) a sus sitios blanco (TFBS, del inglés Transcription Factor Binding Sites) es el primer paso en el proceso de expresión de genes. Por eso es interesante saber dónde y cómo los TFBS están distribuidos a lo largo de la secuencia de un promotor (Yu *et al.*, 2016). Las diferencias en la expresión de genes entre linajes divergentes son generadas por las modificaciones en los elementos regulatorios *cis* y se cree que son una de las principales causas de la evolución y adaptación de las especies (Lemmon *et al.*, 2014). De acuerdo a Creux y colaboradores (2008), identificar elementos regulatorios conservados entre secuencias de promotores homólogos en especies evolutivamente distantes, puede proveernos evidencia sobre redes regulatorias ancestrales (Creux *et al.*, 2008). Sin embargo, también es importante saber cómo los elementos regulatorios están conservados entre diferentes variedades de la misma especie, lo que ocasiona niveles de adaptación específica al ambiente en el cual viven. A pesar de la adición de mecanismos de respuesta a estrés durante el curso de la domesticación del maíz, aún existen razas susceptibles a condiciones ambientales extremas (Ahanger *et al.*, 2017). Las alteraciones en la expresión de genes y la subsecuente producción de proteínas específicas durante la tolerancia a frío juegan un papel importante en la distribución y supervivencia de las plantas, así como en su rendimiento.

Las investigaciones sobre la regulación transcripcional pueden llevarnos a entender las relaciones entre elementos regulatorios *cis*, la arquitectura de los promotores, los factores de transcripción y la respuesta a estrés en plantas. Una categoría funcional de genes de interés para estudiar la respuesta a estrés son los factores de transcripción. Esto debido a su capacidad de modular la expresión de un gen bajo una amplia variedad de condiciones ambientales de manera espacio-temporal. Las proteínas con dominio Dof (del inglés DNA with One zinc Finger) son factores de transcripción específicos de plantas con un dominio de unión al DNA altamente conservado que actúan como activadores o represores transcripcionales involucrados en diversos procesos biológicos específicos de las

plantas (Yanagisawa, 2004). La familia de genes Dof ha sido estudiada desde las algas verdes unicelulares como *Chlamydomonas reinhardtii* hasta el musgo *Physcomitrella patens* (Moreno-Risueno *et al.*, 2007), pero su región promotora no ha sido estudiada a detalle.

El maíz (*Zea mays*) es el segundo cultivo agronómico más importante como fuente de alimento y además es considerado un organismo modelo con una amplia diversidad genética (Prasanna, 2012). La secuenciación del genoma de este cultivo (Schnable *et al.*, 2009; Vielle-Calzada *et al.*, 2009) es un acontecimiento importante en la investigación del genoma de maíz, con implicaciones significativas para el entendimiento de su organización y evolución (Prasanna, 2012). Las razas de este cultivo representan una fuente genética invaluable para estudiar la regulación de genes, porque están adaptadas a condiciones ambientales que son consideradas estresantes en la agricultura convencional (Aguilar-Rangel *et al.*, 2017). Estudiar las regiones regulatorias del promotor Dof01 en diferentes razas y/o variedades de maíz, nos puede dar información sobre la evolución de los elementos regulatorios y la respuesta a condiciones de estrés tal como la temperatura. Por eso, en el presente trabajo realizamos un análisis comparativo de las secuencias del promotor Dof01 en cinco variedades de maíz, donde incluimos al ancestro progenitor del maíz actual, el Teosinte y una de las razas de maíz ancestrales del centro de México, Palomero Toluqueño, así como tres cultivos híbridos empleados en estudios de genética comparativa como B73, W22 y Mo17, para identificar los cambios en los elementos regulatorios, así como sus implicaciones en la respuesta de los cultivos a condiciones de estrés por temperatura.

2. MATERIALES Y MÉTODOS

2.1. Identificación de la secuencia del promotor Dof01 en cinco genomas de maíz

Accedimos a la base de datos de Phytozome V12 (Goodstein *et al.*, 2012) para identificar la secuencia del promotor Dof01 de maíz B73. Seleccionamos 1794 pb disponibles, río arriba del codón de inicio, como secuencia del promotor. Esta

secuencia fue empleada para identificar las secuencias del promotor Dof01 homólogas en los genomas de cuatro variedades de maíz (TEO, PAL, W22 y Mo17) mediante el algoritmo BLAST.

2.2. Análisis de los elementos regulatorios *cis* en las secuencias del promotor Dof01

Las secuencias del promotor Dof01 de las cinco variedades de maíz se emplearon para la búsqueda de TFBS utilizando la herramienta multiple promoter analysis tool de PlantPAN (Chow *et al.*, 2016). Además, identificamos los sitios de inicio de transcripción (TSS, del inglés Transcription Start Site) y cajas TATA con ayuda de la herramienta en línea plant promoter tool de Softberry (Shahmuradov *et al.*, 2017). Las posiciones de los elementos regulatorios encontrados se ajustaron de manera relativa a la posición del TSS (posición 0).

Para entender cómo están conservados los elementos regulatorios, utilizamos estos datos para representar de manera gráfica los TFBS, TSSs y cajas TATA presentes a lo largo de las cinco secuencias del promotor Dof01 analizadas.

Para identificar qué factores de transcripción reconocen los sitios de unión localizados en las secuencias, realizamos un alineamiento múltiple con la herramienta ClustalW. En el alineamiento identificamos la posición de los sitios de unión y comparamos las posiciones de los factores de transcripción que regulan la expresión de Dof01.

Además, para entender cómo se distribuyen los elementos regulatorios con relación a la posición del TSS, realizamos un gráfico de distribución de densidad con la ayuda del paquete ggpubr del software R.

2.3. Análisis de correlación entre la expresión del gen Dof01 y otros factores de transcripción

Para entender cómo los factores de transcripción afectan la expresión de Dof01 en B73, tomamos los datos del atlas de expresión de Hoopes (Hoopes *et al.*, 2018) de 79 tejidos distintos y extrajimos los valores de expresión Log2 de Dof01 y los factores de transcripción que se unen a la secuencia del promotor Dof01. Realizamos un test de correlación de Pearson para encontrar los factores de

transcripción que más afectan la expresión de Dof01 y seleccionamos aquellos con valor absoluto mayor a 0.6 para estudiarlos más a detalle.

2.4. Material vegetal y tratamientos de estrés

Se emplearon semillas de maíz Palomero Toluqueño (genotipo EDMX-2233), proporcionadas por el CIMMYT (Centro Internacional para el Mejoramiento del Maíz y Trigo), las cuales se esterilizaron superficialmente por inmersión en una solución de etanol al 75% por 10 minutos, seguido de hipoclorito de sodio al 10% por 5 minutos y lavadas con abundante agua destilada estéril en tres ocasiones. Las semillas se germinaron sobre papel filtro dentro de cajas magenta en una cámara climática bajo condiciones de cultivo controladas (16/8 h luz/obscuridad y $27 \pm 1^\circ\text{C}$). Las plántulas de 9 días de edad se sometieron a los tratamientos de estrés por temperatura baja (4°C) y alta (42°C) en una cámara bioclimática durante 12 h. Las plántulas utilizadas como control se mantuvieron en las condiciones estándar. Se tomaron muestras de los tejidos de hoja, tallo y raíz, se congelaron inmediatamente con nitrógeno líquido y se almacenaron a -80°C hasta su procesamiento

2.5. Aislamiento de RNA y análisis de expresión del gen Dof01 en maíz Palomero Toluqueño

Se realizó la extracción de RNA total de los tejidos de plántulas de maíz de 10 días de edad mediante el método de Trizol (ThermoFisher) utilizando las indicaciones del fabricante. Se emplearon tres réplicas biológicas de cada tratamiento. Se sintetizó cDNA con la enzima transcriptasa reversa ImProm-II (Promega) utilizando 1 μg de RNA después del tratamiento con DNAsa I. El análisis de la expresión de genes se realizó por PCR tiempo real con un termociclador Corbett RotorGene 6000 (ThermoFisher) utilizando la mezcla de reacción SsoFast Eva Green (BioRad). Como gen endógeno se utilizó el Factor de Elongación 1-alfa (EF1 α) y el nivel de expresión relativa se calculó por el método de $2^{-\Delta\Delta\text{ct}}$ (Livak & Schmittgen, 2001). Se realizaron pruebas de ANOVA y Tukey post-hoc para determinar las diferencias significativas entre las medias de los tratamientos con respecto al control. Los resultados del análisis de expresión se graficaron con ayuda del

paquete ggpubr del software R V3.5 (R Foundation for Statistical Computing, 2008).

Por otra parte, el análisis de expresión del gen Dof01 de maíz B73 se tomó del atlas de expresión de Hoopes, el cual contiene datos de expresión de genes de maíz bajo tratamientos de estrés, entre ellos temperatura baja (4°C) y alta (42°C) en plántulas de la misma edad.

3. RESULTADOS

3.1. Segmento de inserción/delección en el promotor Dof01

Realizamos un alineamiento local (BLAST) utilizando la secuencia del promotor Dof01 de B73 para buscar e identificar las respectivas secuencias homólogas en cuatro genomas de maíz distintos (Tabla 1). El alineamiento mostró que las secuencias identificadas tienen un porcentaje de cobertura del 90%. Esto se debe principalmente al fragmento de 179 pb presente en la secuencia del promotor Dof01 de B73 el cual está ausente en las otras cuatro secuencias analizadas. El promotor Dof01 de W22 (1616 pb) tiene el mayor porcentaje de identidad con respecto a B73 con 99.72% y sólo cuatro nucleótidos desiguales. La secuencia del promotor de PAL (1584 pb) mostró una identidad del 99.26%. La región río arriba del segmento de inserción respecto a B73 contiene cinco nucleótidos desiguales, mientras que la región localizada río abajo contiene 22 y un total de 14 segmentos no cubiertos (gaps) en a lo largo de toda la secuencia del promotor. La secuencia de Mo17 (1584 pb) tiene una identidad del 99.08%, en ella podemos encontrar un total de 22 nucleótidos desiguales y 9 segmentos de gaps presentes en la región proximal del promotor. La secuencia del promotor Dof01 de TEO (1596 pb) contiene un total de 64 nucleótidos desiguales (47 en la región río arriba y 17 en la región río debajo de la región de inserción con respecto a B73) y 21 segmentos cortos de inserción/delección entre ambas secuencias analizadas. La Fig. S1 muestra el alineamiento de las cinco secuencias del promotor Dof01, en las cuales se especifica la localización de los sitios donde se unen los factores de transcripción que regulan la expresión del gen Dof01.

Tabla 1. Alineamiento local BLAST de la secuencia del promotor Dof01 de las cinco variedades de maíz. Se muestran dos segmentos alineados de cada variedad con respecto a B73. Se indica el porcentaje de identidad, así como la longitud total del alineamiento, el número de nucleótidos desiguales, los gaps y el valor de error.

Table 1. BLAST local alignment of the sequences of the Dof01 promoters of five varieties of maize. Two aligned segments of each of the varieties are shown with respect to B73. The percentage of identity of each segment is indicated, as well as the total length of the alignment, the number of mismatches, the gaps and the e-value.

Query	Sujeto	% Identidad	Longitud de alineamiento	Desigualdades	Gap	Error
B73	W22	99.723	1083	3	0	0
B73	W22	99.812	533	1	0	0
B73	PAL	99.263	1086	5	1	0
B73	PAL	90.056	533	22	14	0
B73	Mo17	99.077	1083	10	0	0
B73	Mo17	89.963	538	12	9	0
B73	TEO	92.576	1091	47	15	0
B73	TEO	94.61	538	17	6	0

3.2. Arquitectura del promotor Dof01

Representamos de manera gráfica las cinco secuencias del promotor Dof01 estudiadas para comparar las posiciones de los elementos regulatorios. La Fig. 1 muestra las secuencias de los promotores y los motivos de los TFBS que se localizaron. Las secuencias están alineadas con respecto a la posición del TSS más cercano al codón de inicio de cada gen (flecha grande). También se identifica la posición de TSS alternativos (flechas cortas) y las respectivas cajas TATA asociadas a cada TSS. Podemos apreciar que la gran mayoría de los TFBS están

conservados entre las distintas variedades. La secuencia del promotor Dof01 de B73 tiene una región UTR 5' más corta (95 pb), seguida de la secuencia de PAL (144 pb), mientras que esta característica en las secuencias del promotor Dof01 de TEO, W22 y Mo17 son de 238 pb, 233 pb y 206 pb, respectivamente. Representamos la región de inserción/delección presente en la secuencia del promotor de B73 la cual se localiza en las posiciones -616 a -437 pb río arriba del TSS (línea negra delgada). En esta región, podemos ver que está presente una caja TATA, la cual no está presente en la secuencia de PAL pero sí está en las otras cuatro variedades. Además, esta región posee cinco TFBS que no están presentes en las otras cuatro variedades. La caja TATA que se localiza en la posición -35 pb en las secuencias de PAL, TEO, W22 y Mo17 no está presente en la secuencia de B73, esto puede provocar un cambio significativo en la actividad del promotor. En total se localizaron 35 motivos distintos en las cinco secuencias. Las secuencias consenso de cada motivo se muestran en la Fig. S2.

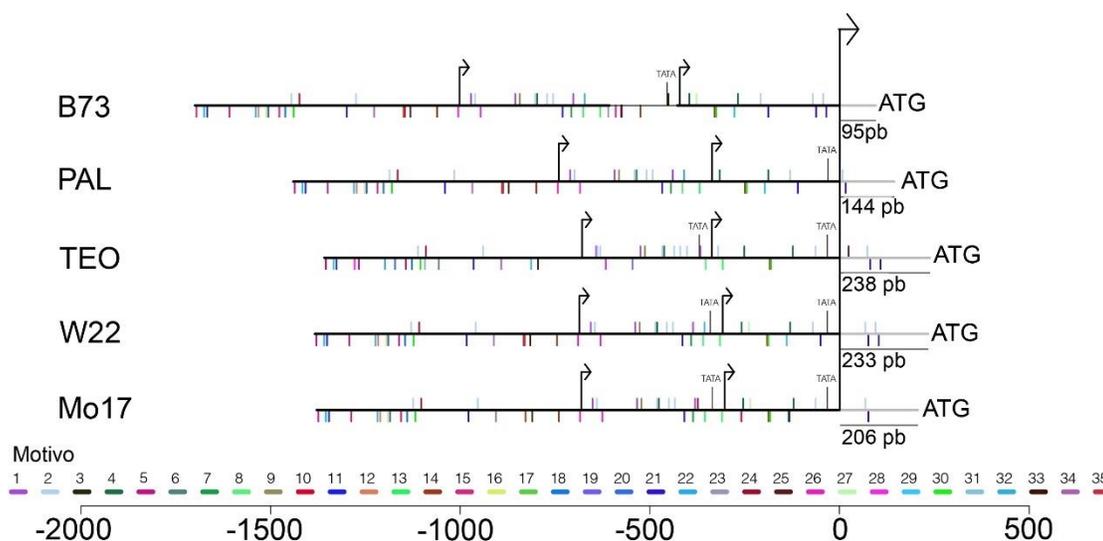


Figura 1. Representación de la arquitectura del promotor Dof01 de cinco variedades de maíz. Las líneas negras horizontales representan el tamaño de la secuencia de cada promotor localizados en sus respectivos genomas. La línea gris horizontal representa la región UTR 5'. La flecha vertical mayor representa el TSS principal mientras que las flechas verticales de menor tamaño representan TSS

alternos. La posición de los elementos *cis* (motivos) se representa con código de color. Se indica también las posiciones de las cajas TATA y el ATG. La escala inferior representa el tamaño de la secuencia donde cero (0) indica la posición del TSS principal y en negativo la región río arriba de este, mientras que en positivo las regiones río abajo.

Figure 1. Representation of the architecture of the Dof01 promoter of five maize varieties. The horizontal black lines represent the size of the sequence of each promoter located in their respective genomes. The horizontal gray line represents the 5'-UTR. The larger vertical arrow represents the main TSS while the smaller vertical arrows represent alternate TSS. The position of *cis* elements (motifs) is represented by color code. The positions of the TATA boxes and the ATG are also indicated. The lower scale represents the size of the sequence where zero (0) indicates the position of the main TSS and in negative the region upstream of this, while in positive the regions downstream.

3.3. Los factores de transcripción se unen a una posición relativa con respecto al TSS

Se graficó la densidad de elementos regulatorios y su distribución a lo largo de las secuencias de los cinco promotores (Fig. 2). Encontramos que hay un pico de máxima densidad entre los 50 y 200 pb río arriba del TSS, cuya distribución muestra una forma de campana. La densidad de elementos regulatorios cae hacia los 500 pb y vuelve a incrementar cerca de los -850 pb. Encontramos dos TSS alternativos cuya función pudiera estar relacionada a un transcrito alterno no funcional. El segundo pico de mayor densidad de elementos regulatorios se localiza alrededor de 200 pb río arriba del TSS alternativo en la posición -700 pb aproximadamente.

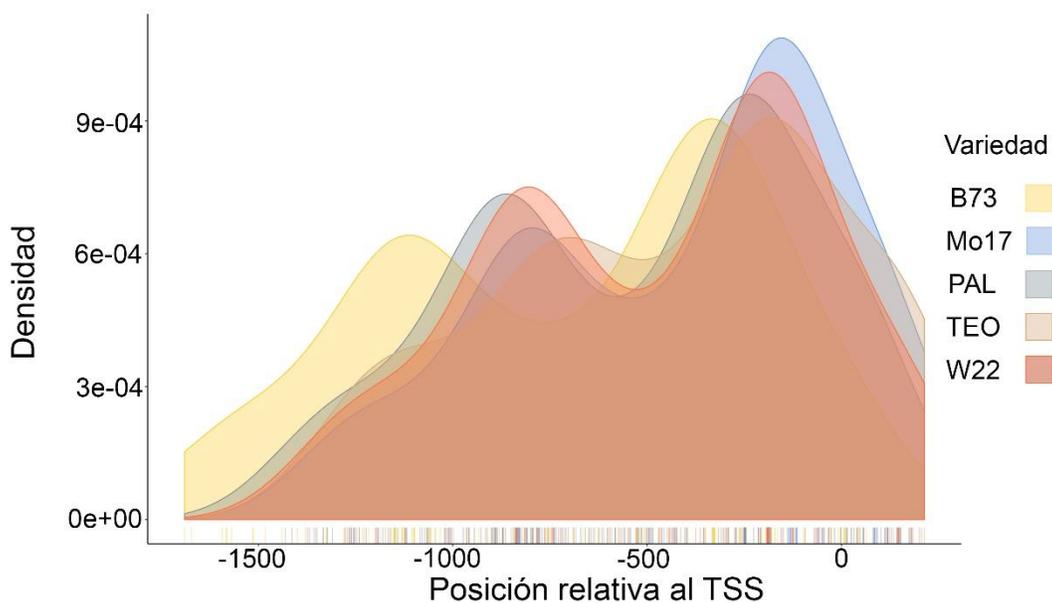


Figura 2. Densidad de distribución de elementos *cis* en el promotor Dof01 de las cinco variedades de maíz. Se localizan dos picos de máxima densidad cerca del TSS principal y un TSS alternativo. Mientras mayor es la altura del pico, mayor es la probabilidad de que un factor de transcripción se una a esa región. La escala inferior representa la posición relativa con respecto al TSS principal. La escala lateral representa la probabilidad.

Figure 2. Density of distribution of *cis* elements in the Dof01 promoter of the five maize varieties. Two peaks of maximum density are located near the main TSS and an alternate TSS. The higher the peak, the greater the likelihood that a transcription factor will join that region. The lower scale represents the relative position with respect to the main TSS. The lateral scale represents the probability.

Adicionalmente, encontramos 318 factores de transcripción distintos que pertenecen a 23 familias que reconocen 655 sitios de unión, cuya actividad es responsable de regular la expresión de Dof01 en las cinco variedades de maíz. Para visualizar el grado de similitud entre los promotores y los factores de transcripción que regulan su actividad, construimos un mapa de calor (Fig. 3) con ayuda del software R. En la Fig. 3 podemos observar una clara separación de los promotores que pertenecen a razas ancestrales como Teosinte, el cual se encuentra totalmente separado del resto, seguido de Palomero, cuya similitud en

términos de número y tipo de elementos regulatorios es más cercana a la de W22. Podemos observar que el TF más abundante pertenece a la familia WRKY con 110 sitios de unión. Teosinte contiene dos sitios de unión únicos que no se encontraron en las otras cuatro secuencias, los cuales pertenecen a las familias G2-Like y Storekeeper, con nueve y cuatro factores de transcripción que reconocen el mismo sitio, respectivamente. Una de las familias menos abundantes es AP2/B3, la cual no está presente en TEO y cuyo sitio de unión es blanco de tres factores de transcripción en Dof01. Uno de los factores de transcripción con el mayor número de sitios de unión pertenece a la familia Dof. Se localizaron un total de 64 sitios de unión en los cinco promotores, con 18 incidencias en B73, 15 en Mo17, 13 en PAL y W22 y sólo cinco en TEO. Sólo tres factores de transcripción que pertenecen a la familia Dof fueron localizados en el promotor Dof01, los cuales son Dof02 (GRMZM2G009406), Dof03 (GRMZM2G146283) y Dof35 (GRMZM2G179069).

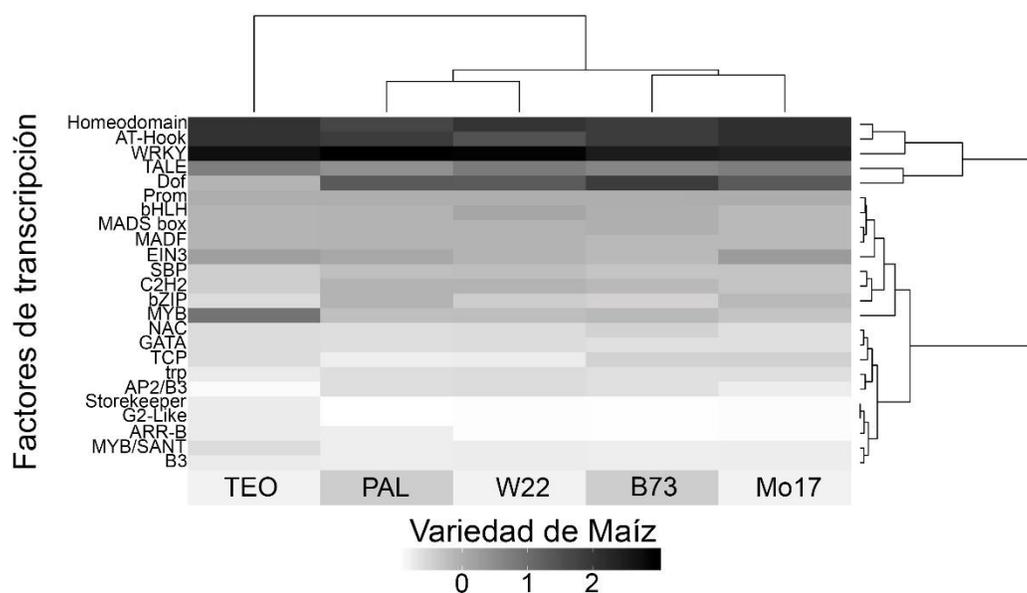


Figura 3. Relaciones entre el tipo y número de factores de transcripción, clasificados por familia, que se unen a la secuencia del promotor Dof01 de las cinco variedades de maíz. La escala inferior representa el número de desviaciones estándar de los datos escalados, a mayor intensidad de color, mayor es el número de sitios de unión del factor de transcripción. A la izquierda se muestra el nombre

de la familia de factores de transcripción. El dendrograma de la derecha muestra la relación entre el número y tipo de factores de transcripción localizados. El dendrograma superior muestra la relación entre los promotores de acuerdo a la incidencia de factores de transcripción que se unen a sus secuencias.

Figure 3. Relationships between the type and number of transcription factors, classified by family, that bind to the Dof01 promoter sequence of the five maize varieties. The lower scale represents the number of standard deviations of the scaled data, the higher the color intensity, the greater the number of binding sites of the transcription factor. On the left is the name of the family of transcription factors. The dendrogram on the right shows the relationship between the number and type of localized transcription factors. The upper dendrogram shows the relationship between promoters according to the incidence of transcription factors that bind their sequences.

3.4. Factores de transcripción relacionados a la respuesta a estrés por frío se unen a la secuencia del promotor Dof01

Para comprender cómo los factores de transcripción que se unen a Dof01 en maíz B73 afectan su expresión, realizamos un análisis de correlación de Pearson. La Fig. 4 muestra un gráfico de dispersión de la correlación de la expresión de Dof01 con la expresión de los 16 factores de transcripción mayormente correlacionados. La mayor correlación la encontramos con el gen que codifica para la proteína relacionada a reveille 1, la cual posee un dominio de unión al DNA tipo MYB (GRMZM2G145041). Este gen está relacionado a la regulación de la expresión de genes asociados al reloj circadiano y a hojas libres de auxinas en *Arabidopsis* (Rawat et al., 2009) y como regulador negativo de la tolerancia a congelamiento y la aclimatación a frío (Meissner et al., 2013). Cinco de los nueve genes correlacionados pertenecen a la familia de factores de transcripción WRKY (GRMZM2G411766; GRMZM2G127064; GRMZM2G027972; GRMZM2G408462; GRMZM2G102583). Los TFs WRKY tienen el mayor número de miembros entre todas las familias de factores de transcripción. Otra familia correlacionada a la expresión de Dof01 en B73 es la que contiene el dominio HD-ZIP (del inglés

Homeobox Domain-Leucine Zipper) (GRMZM2G366130). La familia HD-ZIP está relacionada a procesos como la diferenciación y mantenimiento celular, así como a la regulación y transporte de lípidos en plantas (Chew, Hrmova, & Lopato, 2013). El factor de transcripción que regula la expresión de Dof01 en B73 y que contienen el dominio de unión al DNA tipo AP2/B3 (GRMZM2G059939), está asociado con varias funciones involucradas en procesos fisiológicos y de desarrollo en las plantas. El factor de transcripción tipo RAV que tiene este dominio ha sido estudiado en organismos como *Brassica napus* (canola) y es inducido por factores como el frío, NaCl y PEG (Polietilen glicol) (Zhuang, Sun, Zhou, Xiong, & Zhang, 2011). Finalmente, el factor de transcripción tipo bZIP (GRMZM2G019106) es un miembro de la familia de reguladores transcripcionales asociados con procesos como la germinación y maduración de semillas, inducción y desarrollo floral, fotomorfogénesis y también, está involucrado en la respuesta a estrés y señalización hormonal. La Fig. S3 muestra la expresión de los genes Dof01, bZIP y MYB a través de 79 tejidos de maíz. Los tejidos en los que estos genes son expresados principalmente son hoja, mientras que aquellos en los que la expresión disminuye, están relacionados al desarrollo de la semilla.

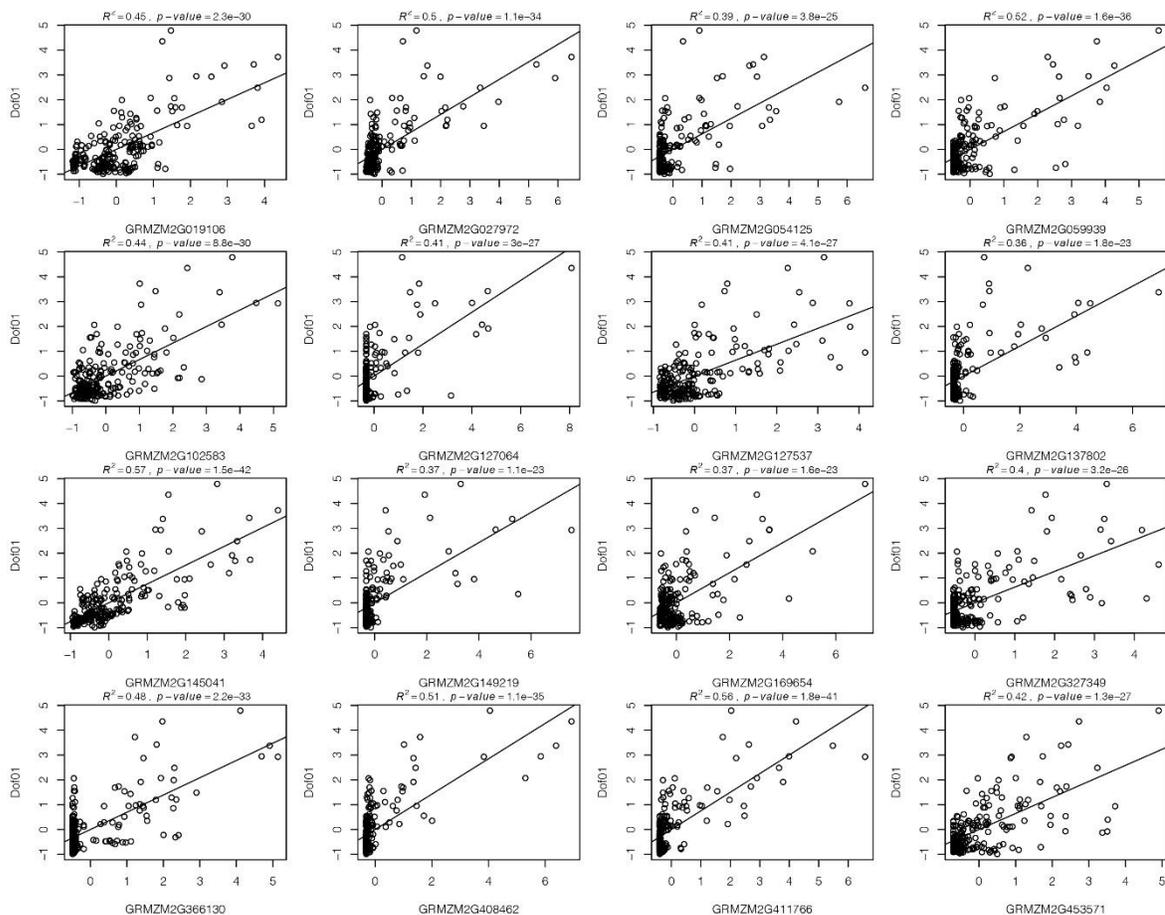


Figura 4. Análisis de correlación de Pearson entre la expresión del gen Dof01 de B73 y los factores de transcripción involucrados en la regulación de su expresión. En la parte superior de cada gráfico de dispersión se muestra el coeficiente de determinación de la regresión lineal, así como el valor p. En la parte inferior de cada gráfico se muestra el número identificador de cada gen analizado. La escala izquierda muestra el valor de la expresión absoluta de los datos transformados de Dof01 en 79 tejidos distintos, mientras que la escala inferior muestra los datos de expresión del factor de transcripción en cuestión en los mismos tejidos.

Figure 4. Pearson's correlation analysis between the expression of the Dof01 gene of B73 and the transcription factors involved in the regulation of its expression. The coefficient of determination of the linear regression as well as the p-value are shown at the top of each scatter plot. The identification number of each gene analyzed is shown at the bottom of each graph. The left scale shows the value of

the absolute expression of the transformed Dof01 data in 79 different tissues, while the lower scale shows the expression data of the transcription factor in question in the same tissues.

3.5. El gen Dof01 de maíz palomero se expresa de manera contraria al de B73 bajo estrés por frío

Para entender cómo el gen Dof01 responde al estrés por temperatura, sometimos plántulas de PAL de 9 días de edad a tratamientos de frío (4°C) y calor (42°C). Analizamos la expresión del gen Dof01 en hoja, tallo y raíz en plántulas de 10 días de edad y comparamos los resultados con los descritos en el atlas de expresión de B73 elaborado por Hoopes y colaboradores (Hoopes *et al.*, 2018). La Fig. 5A muestra los valores de expresión del gen Dof01 analizados en B73. Podemos apreciar que el gen Dof01 se encuentra significativamente reprimido por efecto del frío y el calor con respecto al control. En el caso de PAL (Fig. 5B), el gen Dof01 se sobre expresa significativamente con respecto al control debido al tratamiento por frío, mientras que es reprimido significativamente debido al efecto del estrés por calor en hoja. En el caso del tallo, no se encontraron diferencias significativas con respecto al control bajo ninguno de los tratamientos. Por otra parte, en el tejido de raíz, el gen Dof01 se sobre expresa significativamente más de diez veces a consecuencia del estrés por frío con respecto al control. Mientras que, debido al estrés por calor, su expresión es ligeramente menor con respecto al control, sin mostrar evidencias significativas.

Las diferencias en los patrones de expresión del gen Dof01 en ambas razas de maíz puede deberse a los diferentes elementos *cis* localizados en los promotores de ambos genes. En ambos casos, el gen *reveille 1* (MYB) se encuentra presente, el cual es un regulador negativo de la expresión por frío y que es expresado principalmente en los tejidos de hoja, anteras y estilos y se encuentra en la misma posición en ambos promotores. Por otra parte, el otro elemento *cis* relacionado a la respuesta a frío localizado en el promotor es AP2/B3, el cual fue localizado en un sitio de unión en B73 y en dos sitios en PAL. Este factor de transcripción se expresa principalmente en hoja, por lo que podemos deducir que puede ser el

principal elemento regulador de la respuesta a frío en el promotor Dof01 y el cual sirve para contrarrestar el efecto de regulador negativo del factor de transcripción MYB. Sin embargo, se requieren de estudios de caracterización funcional de los promotores para determinar qué elementos regulatorios están jugando un papel más importante en la regulación de la expresión de Dof01.

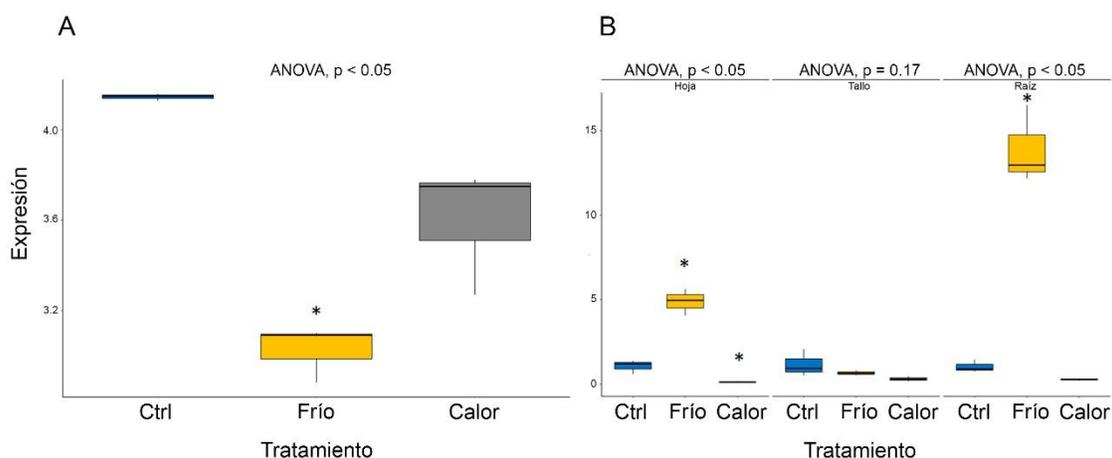


Figura 5. Datos de expresión del gen Dof01 en maíz B73 y Palomero. Se muestra el nivel de expresión relativa con respecto al control de los tratamientos de frío y calor en A) B73 y B) PAL en los tejidos de hoja, tallo y raíz.

Figure 5. Expression data of the Dof01 gene in B73 and Palomero. The level of relative expression with respect to the control of cold and heat treatments in A) B73 and B) PAL in leaf, stem and root tissues is shown.

4. DISCUSIÓN

En el presente trabajo realizamos una caracterización general de la arquitectura del promotor Dof01 en cinco variedades de maíz (B73, Palomero Toluqueño, Teosinte, W22 y Mo17). Identificamos todos los elementos regulatorios que están conservados entre los cinco promotores y que son responsables de la actividad de regulación de la transcripción. El entendimiento de las principales similitudes y diferencias que existen entre las cinco secuencias, principalmente B73 y PAL, que regulan la respuesta en condiciones de estrés por temperatura como el frío y calor, son claves para entender más acerca del proceso de regulación de la expresión.

Palomero Toluqueño es considerada una raza ancestral (Vega-Arreguín *et al.*, 2009) el cual está adaptado a climas fríos y se ha reportado que su genoma es hasta 28% menor comparado con B73 (Vielle-Calzada *et al.*, 2009) y contiene grandes regiones en los cromosomas con baja variabilidad nucleotídica enriquecida en genes involucrados en la respuesta de las plantas al ambiente.

Por otra parte, Teosinte (*Zea mays* ssp. *Parviglumis*) es considerado el ancestro en común de todas las razas de maíz, el cual fue domesticado hace alrededor de 9,000 a 10,000 años. En 2014, Lemmon y colaboradores, analizaron elementos regulatorios en *cis* y *trans* entre Teosinte y B73. Ellos encontraron que aproximadamente el 70% de alrededor de 70,000 genes presentaron evidencia de divergencia regulatoria, sugiriendo una selección de elementos regulatorios *cis* durante el proceso de domesticación, un proceso que favorece la sobre regulación de la expresión de genes (Lemmon *et al.*, 2014). Observamos que la mayor divergencia en elementos regulatorios se presenta entre el promotor Dof01 de TEO y B73.

La variedad de maíz W22 ha servido como plataforma base para estudios de genética reversa en maíz y su genoma es utilizado como referencia para estudios de genómica funcional (Springer *et al.*, 2018). Mientras que Mo17 es una de las variedades híbridas más conocidas y modernas de maíz. El híbrido generado por la cruce entre B73 y Mo17 es la línea más comúnmente usada para estudios de genética y moleculares (Stupar & Springer, 2006). Stupar y Springer (2006)

reportaron que la mayor variación en la expresión de genes entre B73 y Mo17 se debe a la variación en los elementos regulatorios *cis*, sugiriendo que existe un mecanismo para los patrones de expresión en híbridos de estas dos variedades y que los cambios potenciales en elementos *cis* puede deberse a la variación en las secuencias (posiblemente en regiones promotoras), a los estados alterados de la cromatina o a diferencias en la estabilidad del RNA.

En nuestro trabajo encontramos, que para el caso del promotor Dof01, la mayoría de los elementos *cis* presentes en las cinco variedades están conservados. Dichos elementos regulatorios pueden estar involucrados en la respuesta basal de las plantas con el ambiente y aquellos que están presentes en unas variedades, pero no en otras, pueden estar asociados con respuestas específicas intrínsecas de cada variedad. Los modelos de evolución de genes predicen que la divergencia funcional entre genes duplicados (sub-funcionalización) es originada por la pérdida de elementos regulatorios (Lockton & Gaut, 2005). Los elementos regulatorios *cis* analizados en el promotor Dof01 de las cinco variedades de maíz, nos llevó a observar que el ancestro del maíz, Teosinte, difiere en número y tipo de elementos regulatorios con respecto a B73. La pérdida o ganancia de estos elementos regulatorios *cis* es una de las principales causas de la diversidad genética en plantas, por lo tanto, ocasionando modificaciones en su fenotipo. Además, el fragmento de 179 pb presente en la secuencia del promotor Dof01 de B73, puede deberse a la inserción de un elemento transponible, lo que da origen a una neo/sub-funcionalización del gen. Lo anterior puede explicar las diferencias entre la expresión del gen Dof01 en B73 y PAL debido a los tratamientos por frío y calor. Los promotores son reguladores clave y juegan un rol muy importante en la tolerancia a estrés abiótico. Su arquitectura es el componente principal que determina el número de transcritos (nivel de expresión) y la expresión de un gen en un tejido específico (amplitud de la expresión), lo cual, eventualmente forma el fenotipo (Das & Bansal, 2019). Por eso, es necesario entender la asociación entre la secuencia del promotor, los elementos regulatorios *cis* y los factores de transcripción (Creux *et al.*, 2008). Por lo tanto, debemos estudiar en conjunto estos

tres elementos de una manera en la que podamos proveer de nueva información sobre el funcionamiento de la regulación de la expresión de genes.

De acuerdo a Lockton y Gaut (2005), los promotores homólogos comparten elementos regulatorios *cis* incluso a través de distancias evolutivamente considerables. En la Fig. 1 podemos ver cómo, a pesar del hecho de que los elementos regulatorios están conservados, también lo está su ubicación con respecto al TSS. Además, es posible identificar qué familias de factores de transcripción están asociadas a la presencia de otras (Fig. 3).

En la mayoría de los organismos eucariotas, los elementos regulatorios *cis* están significativamente enriquecidos alrededor del sitio de inicio de la transcripción (Mejia-Guerra, *et al.*, 2018). El análisis de la distribución de elementos *cis* en los promotores Dof mostró que existen picos de máxima densidad de elementos regulatorios en torno a los TSS localizados en los promotores (Fig. 2) entre los -200 pb y los -800 pb. Sin embargo, es necesario determinar si la densidad de elementos regulatorios está asociada a la presencia de cajas TATA en los promotores Dof de maíz, ya que hay reportes de que el número de TFBS en promotores con caja TATA es hasta dos veces menor que en promotores sin caja TATA, asegurando que los promotores tipo TATA están asociados a las respuestas a estrés, mientras que los promotores sin caja TATA están asociados a respuestas constitutivas (Yamamoto *et al.*, 2011).

El análisis de la expresión de Dof01 en B73 y PAL mostró que la respuesta ante las condiciones de estrés por frío y calor es distinta entre ambas variedades, principalmente debido al frío. Diversos autores han señalado que los factores de transcripción Dof están asociados con respuestas a estrés por frío, calor, salinidad, sequía, estrés osmótico, ácido abscísico, al virus del mosaico de la sandía y mildiu (Chen & Cao, 2015; Ma *et al.*, 2015; Wang *et al.*, 2017; Wen *et al.*, 2016; Wu *et al.*, 2016).

Con este trabajo, concluimos que el promotor Dof01 es un candidato ideal para estudios de caracterización funcional y posibles aplicaciones en el campo de la biotecnología para la generación de plantas que expresen genes en respuesta a

frío. Se requieren más estudios que profundicen en determinar y caracterizar funcionalmente la región promotora mínima de este promotor y así determinar el papel que cumplen algunos elementos regulatorios presentes.

AGRADECIMIENTOS

Agradecemos al Consejo Nacional de Ciencia y Tecnología (CONACYT) – beca 483627 y a las unidades del LANGEBIO y CINVESTAV-IRAPUATO, así como al CIBNOR y el personal técnico que apoyó esta investigación Araceli Oropeza, Julio Hernández González y Mario Arce Montoya.

CONFLICTO DE INTERESES

Los autores declaran que no existe conflicto de intereses

REFERENCIAS

- Aguilar-Rangel, M. R., Montes, R. A. C., González-Segovia, E., Ross-Ibarra, J., Simpson, J., & Sawers, R. J. H. (2017). Allele specific expression analysis identifies regulatory variation associated with stress-related genes in the Mexican highland maize landrace Palomero Toluqueño. *BioRxiv*, 152397.
- Ahanger, M. A., Akram, N. A., Ashraf, M., Alyemeni, M. N., Wijaya, L., & Ahmad, P. (2017). Plant responses to environmental stresses-from gene to biotechnology. *AoB PLANTS*, 9(4).
- Chen, Y., & Cao, J. (2015). Comparative analysis of dof transcription factor family in maize. *Plant Molecular Biology Reporter*, 33(5).
- Chew, W., Hrmova, M., & Lopato, S. (2013). Role of Homeodomain leucine zipper (HD-Zip) IV transcription factors in plant development and plant protection from deleterious environmental factors. *International Journal of Molecular Sciences*, 14(4).
- Chow, C.-N., Zheng, H.-Q., Wu, N.-Y., Chien, C.-H., Huang, H.-D., Lee, T.-Y., Chiang-Hsieh, Y. -F., Hou, P. -F., Yang, T.-Y., & Chang, W.-C. (2016). PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. *Nucleic Acids Research*, 44(D1).
- Creux, N. M., Ranik, M., Berger, D. K., & Myburg, A. A. (2008). Comparative analysis of orthologous cellulose synthase promoters from *Arabidopsis*, *Populus* and *Eucalyptus*: evidence of conserved regulatory elements in angiosperms. *New Phytologist*, 179(3).
- Das, S., & Bansal, M. (2019). Variation of gene expression in plants is influenced by gene architecture and structural properties of promoters. *PloS One*, 14(3).
- Eagles, H. A., & Lothrop, J. E. (1994). Highland maize from central Mexico—its origin, characteristics, and use in breeding programs. *Crop Science*, 34(1).
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, 40(D1).

- Hoopes, G. M., Hamilton, J. P., Wood, J. C., Esteban, E., Pasha, A., Vaillancourt, B., Provart, N. J., & Buell, C. R. (2018). An updated gene atlas for maize reveals organ-specific and stress-induced genes. *The Plant Journal*, *tpj.14184*.
- Jin, J., Tian, F., Yang, D.-C., Meng, Y.-Q., Kong, L., Luo, J., & Gao, G. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research*, *45(D1)*.
- Lemmon, Z. H., Bukowski, R., Sun, Q., & Doebley, J. F. (2014). The Role of *cis* regulatory evolution in maize domestication. *PLoS Genetics*, *10(11)*.
- Livak, K. J., & Schmittgen, T. D. (2001). analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods*, *25(4)*.
- Lockton, S., & Gaut, B. S. (2005). Plant conserved non-coding sequences and paralogue evolution. *Trends in Genetics*, *21(1)*.
- Ma, J., Li, M.-Y., Wang, F., Tang, J., & Xiong, A.-S. (2015). Genome-wide analysis of Dof family transcription factors and their responses to abiotic stresses in Chinese cabbage. *BMC Genomics*, *16(1)*.
- Mejía-Guerra, M. K., Li, W., Galeano, N. F., Vidal, M., Gray, J., Doseff, A. I., & Grotewold, E. (2015). Core promoter plasticity between maize tissues and genotypes contrasts with predominance of sharp transcription initiation sites. *The Plant Cell*, *27(12)*.
- Moreno-Risueno, M. Á., Martínez, M., Vicente-Carbajosa, J., & Carbonero, P. (2007). The family of DOF transcription factors: from green unicellular algae to vascular plants. *Molecular Genetics and Genomics*, *277(4)*.
- Prasanna, B. M. (2012). Diversity in global maize germplasm: Characterization and utilization. *Journal of Biosciences*, *37(5)*.
- Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., & Wilson, R. K. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science*, *326(5956)*.
- Shahmuradov, I. A., Umarov, R. K., & Solovyev, V. V. (2017). TSSPlant: a new tool for prediction of plant Pol II promoters. *Nucleic Acids Research*, *45(8)*.

Springer, N. M., Anderson, S. N., Andorf, C. M., Ahern, K. R., Bai, F., Barad, O., Barbazuk, W. B., Bass, Hank W., Baruch, K., Ben-Zvi, G., Buckler, E. S., Bukowski, R., Campbell, M. S., Cannon, E. K. S., Chomet, P., Dawe, R. K., Davenport, R., Dooner, H. K., Du, L. H., Du, C., Easterling, K. A., Gault, C., Guan, J. C., Hunter, C. T., Jander, G., Jiao, Y., Koch, K. E., Kol, G., Köllner, T. G., Kudo, T., Li, Q., Lu, F., Mayfield-Jones, D., Mei, W., McCarty, D. R., Noshay, J. M., Portwood, J. L., Ronen, G., Settles, A. M., Shem-Tov, D., Shi, J., Soifer, I., Stein, J. C., Stitzer, M. C., Suzuki, M., Vera, D. L., Vollbrecht, E., Vrebalov, J. T., Ware, D., Wei, S., Wimalanathan, K., Woodhouse, M. R., Xiong, W., & Brutnell, T. P. (2018). The maize W22 genome provides a foundation for functional genomics and transposon biology. *Nature Genetics*, 50(9).

Stupar, R. M., & Springer, N. M. (2006). *Cis*-transcriptional variation in maize inbred lines b73 and mo17 leads to additive expression patterns in the f1 hybrid. *Genetics*, 173(4).

Vega-Arreguín, J. C., Ibarra-Laclette, E., Jiménez-Moraila, B., Martínez, O., Vielle-Calzada, J. P., Herrera-Estrella, L., & Herrera-Estrella, A. (2009). Deep sampling of the Palomero maize transcriptome by a high throughput strategy of pyrosequencing. *BMC Genomics*, 10.

Vielle-Calzada, J.-P., Martínez de la Vega, O., Hernández-Guzmán, G., Ibarra-Laclette, E., Alvarez-Mejía, C., Vega-Arreguín, J. C., Jiménez-Moraila, B., Fernández-Cortés, A., Corona-Armenta, G., Herrera-Estrella, L., & Herrera-Estrella, A. (2009). The Palomero genome suggests metal effects on domestication. *Science (New York, N.Y.)*, 326(5956).

Wang, H., Zhao, S., Gao, Y., & Yang, J. (2017). Characterization of dof transcription factors and their responses to osmotic stress in poplar (*populus trichocarpa*). *PLoS ONE*, 12(1).

Wen, C., Cheng, Q., Zhao, L., Mao, A., Yang, J., Yu, S., Weng, Y., & Xu, Y. (2016). Identification and characterization of Dof transcription factors in the cucumber genome. *Scientific Reports*, 6(1).

Wu, Z., Cheng, J., Cui, J., Xu, X., Liang, G., Luo, X., Chen, X., Tang, X., Hu, K. & Qin, C. (2016). Genome-wide identification and expression profile of dof transcription factor gene family in pepper (*Capsicum annuum* L.). *Frontiers in Plant Science*, 7.

- Yamamoto, Y. Y., Yoshioka, Y., Hyakumachi, M., & Obokata, J. (2011). Characteristics of core promoter types with respect to gene structure and expression in *Arabidopsis thaliana*. *DNA Research : An International Journal for Rapid Publication of Reports on Genes and Genomes*, 18(5).
- Yanagisawa, S. (2004). Dof domain proteins: plant-specific transcription factors associated with diverse phenomena unique to plants. *Plant and Cell Physiology*, 45(4).
- Yanagisawa, S., & Izui, K. (1993). Molecular cloning of two DNA-binding proteins of maize that are structurally different but interact with the same sequence motif. *Journal of Biological Chemistry*, 268(21).
- Yu, C.-P., Lin, J.-J., & Li, W.-H. (2016). Positional distribution of transcription factor binding sites in *Arabidopsis thaliana*. *Scientific Reports*, 6(1).
- Zhuang, J., Sun, C.-C., Zhou, X.-R., Xiong, A.-S., & Zhang, J. (2011). Isolation and characterization of an AP2/ERF-RAV transcription factor BnaRAV-1-HY15 in *Brassica napus* L. HuYou15. *Molecular Biology Reports*, 38(6).

MATERIAL SUPLEMENTARIO

Figura S1. Alineamiento múltiple de la secuencia del promotor Dof01 de las cinco variedades de maíz. El texto resaltado en color gris indica la ubicación de los motivos que son reconocidos por factores de transcripción. Los asteriscos debajo de cada línea indican que la posición de cada nucleótido está conservada en las cinco secuencias. Los factores de transcripción se nombran justo arriba de su sitio de unión y sólo se muestran los que están en la hebra sentido. Al final de cada línea se muestra la longitud de la secuencia. Los guiones presentes en cada secuencia muestran la ubicación de las inserciones/deleciones.

Figure S1. Multiple alignment of the Dof01 promoter sequence of the five maize varieties. Text highlighted in gray indicates the location of the motifs that are recognized by transcription factors. The asterisks below each line indicate that the position of each nucleotide is conserved in all five sequences. Transcription factors are named just above their binding site and only those in the sense strand are shown. At the end of each line the length of the sequence is shown. The dashed lines present in each sequence show the location of the indels.

	P1 F	GATA	
ZmTEO_Dof01	GCGACTTTTCATGATCCAAGC	ATATCTTTTCCCCTTTTCCAATACTTGCAACTGCTCTGGC	60
ZmPAL_Dof01	GCGACTTTTCATGATCCAAGC	ATATCTTTTCCCCTTTTCCAATACTTGCAACTGCTCTGGC	60
ZmMo17_Dof01	GCGACTTTTCATGATCCAAGC	ATATCTTTTCCCCTTTTCCAATACTTGCAACTGCTCTGGC	60
ZmB73_Dof01	GCGACTTTTCATGATCCAAGC	ATATCTTTTCCCCTTTTCCAATACTTGCAACTGCTCTGGC	60
ZmW22_Dof01	GCGACTTTTCATGATCCAAGC	ATATCTTTTCCCCTTTTCCAATACTTGCAACTGCTCTGGC	60

		SBP bHLH	
ZmTEO_Dof01	TGGTGGCTGCCGCCCTGTCTCGTTAGAAACTTTGCAGATAGTC	TTAGTAGTACTGCAA	120
ZmPAL_Dof01	TGGTGGCTGCCGCCCTGTCTCGTTAGAAACTTTGCAGATAGTC	TTAGTAGTACTGCAA	120
ZmMo17_Dof01	TGGTGGCTGCCGCCCTGTCTCGTTAGAAACTTTGCAGATAGTC	TTAGTAGTACTGCAA	120
ZmB73_Dof01	TGGTGGCTGCCGCCCTGTCTCGTTAGAAACTTTGCAGATAGTC	TTAGTAGTACTGCAA	120
ZmW22_Dof01	TGGTGGCTGCCGCCCTGTCTCGTTAGAAACTTTGCAGATAGTC	TTAGTAGTACTGCAA	120

		ZF-HD	
ZmTEO_Dof01	GTGCAAGATATGGTCCGATTGTCCCTCTAGCTGCGGATTGTTGTTT	CTTAATCTA----	175
ZmPAL_Dof01	GTGCAAGATATGGTCCGATTGTCCCTCTAGCTGCGGATTGTTGTTT	CTTAATCTA----	180
ZmMo17_Dof01	GTGCAAGATATGGTCCGATTGTCCCTCTAGCTGCGGATTGTTGTTT	CTTAATCTA----	180
ZmB73_Dof01	GTGCAAGATATGGTCCGATTGTCCCTCTAGCTGCGGATTGTTGTTT	CTTAATCTA----	180
ZmW22_Dof01	GTGCAAGATATGGTCCGATTGTCCCTCTAGCTGCGGATTGTTGTTT	CTTAATCTA----	180

	MYB MYB-related	P2 F	
ZmTEO_Dof01	--TGGAGACTAGGTTAGGTTAGCTACACATCATGGCAACCTACTTTCCGTTAGCTGGT		233
ZmPAL_Dof01	TATGGAGGGTAGGTTAGGTTAGCTACACATCATGGCAACCTACTTTCCGTTAGCTGGT		240
ZmMo17_Dof01	TATGGAGGGTAGGTTAGGTTAGCTACACATCATGGCAACCTACTTTCCGTTAGCTGGT		240
ZmB73_Dof01	TATGGAGGGTAGGTTAGGTTAGCTACACATCATGGCAACCTACTTTCCGTTAGCTGGT		240
ZmW22_Dof01	TATGGAGGGTAGGTTAGGTTAGCTACACATCATGGCAACCTACTTTCCGTTAGCTGGT		240

	AT-HOOK	TALE	
ZmTEO_Dof01	TGGTGAATTGCAAAAGGCTGACCTGGGGCTGTGTTTTGCAACAACCTAGTAGCCGATG		293
ZmPAL_Dof01	TGGTTCAAATTGCAAAAGGCTGACCTGGGGCTGTGTTTTGCAACAACCTAGTAGCCGATG		300
ZmMo17_Dof01	TGGTTCAAATTGCAAAAGGCTGACCTGGGGCTGTGTTTTGCAACAACCTAGTAGCCGATG		300
ZmB73_Dof01	TGGTTCAAATTGCAAAAGGCTGACCTGGGGCTGTGTTTTGCAACAACCTAGTAGCCGATG		300
ZmW22_Dof01	TGGTTCAAATTGCAAAAGGCTGACCTGGGGCTGTGTTTTGCAACAACCTAGTAGCCGATG		300

	TALE	ARF/B3	G2-like
ZmTEO_Dof01	ACTTACTAGGACTCAGTTTGTCTCTGAACAACCGATGCTTCGGTTTTTGTGAGGATATGC		353
ZmPAL_Dof01	ACTTACGAGGGTTCAAATTTGTCTCTGAACAACCGATGCTTCGGTTTTTGTGAGGATATGC		360
ZmMo17_Dof01	ACTTACGAGGGTTCAAATTTGTCTCTGAACAACCGATGCTTCGGTTTTTGTGAGGATATGC		360
ZmB73_Dof01	ACTTACGAGGGTTCAAATTTGTCTCTGAACAACCGATGCTTCGGTTTTTGTGAGGATATGC		360
ZmW22_Dof01	ACTTACGAGGGTTCAAATTTGTCTCTGAACAACCGATGCTTCGGTTTTTGTGAGGATATGC		360

	GATA	bZIP	
ZmTEO_Dof01	AGAGAGAAGATCAGCTACTCGTAGTGGATAGGCGAATTCCTTTCCCTCGCTTGAGTACT		413
ZmPAL_Dof01	AGAGAGAAGATCAGCTACACGTAGTGGATAGGCGAATTCCTTTTC-CCCGCTTGAGTACT		419
ZmMo17_Dof01	AGAGAGAAGATCAGCTACACGTAGTGGATAGGCGAATTCCTTTTC-CCCGCTTGAGTACT		419
ZmB73_Dof01	AGAGAGAAGATCAGCTACACGTAGTGGATAGGCGAATTCCTTTTC-CCCGCTTGAGTACT		419
ZmW22_Dof01	AGAGAGAAGATCAGCTACACGTAGTGGATAGGCGAATTCCTTTTC-CCCGCTTGAGTACT		419

	Dof		
ZmTEO_Dof01	CATTAAAGCCATTTCTTGATGTGGCTACCCGTCGGTTCTGTGGAAGTAGTGTTCCTAAA		473
ZmPAL_Dof01	CATTAAAGCCATTTCTTGATGTGGCTACCCGTCGGTTCTGTGGAAGTAGTGTTCCTAAA		479
ZmMo17_Dof01	CATTAAAGCCATTTCTTGATGTGGCTACCCGTCGGTTCTGTGGAAGTAGTGTTCCTAAA		479
ZmB73_Dof01	CATTAAAGCCATTTCTTGATGTGGCTACCCGTCGGTTCTGTGGAAGTAGTGTTCCTAAA		479
ZmW22_Dof01	CATTAAAGCCATTTCTTGATGTGGCTACCCGTCGGTTCTGTGGAAGTAGTGTTCCTAAA		479

	SBP		
ZmTEO_Dof01	CAGCTTGGG-TACGTAGGGCTCTATTCTATCCCTTGTCAGTTTAGCCTTGATGTCAGA		533
ZmPAL_Dof01	CAGCTTGGG-TACGTAGGGCTCTATTCTATCCCTTGTCAGTTTAGCCTTGATGTCAGA		538
ZmMo17_Dof01	CAGCTTGGG-TACGTAGGGCTCTATTCTATCCCTTGTCAGTTTAGCCTTGATGTCAGA		538
ZmB73_Dof01	CAGCTTGGG-TACGTAGGGCTCTATTCTATCCCTTGTCAGTTTAGCCTTGATGTCAGA		538
ZmW22_Dof01	CAGCTTGGG-TACGTAGGGCTCTATTCTATCCCTTGTCAGTTTAGCCTTGATGTCAGA		538

	AT-HOOK	ERF/RAV			
ZmTEO_Dof01	ATTGCCAAATTTCTTTTCTGCTGCAAGTTTTTATTTTGTAGTTTCTTGATTG---GGTG		590		
ZmPAL_Dof01	ATTGCCAAATTTCTTTTCTGCTGCAAGTTTTTATTTTGTAGTTTCTTGATTGTTGGTGGGTG		598		
ZmMo17_Dof01	ATTGCCAAATTTCTTTTCTGCTGCAAGTTTTTATTTTGTAGTTTCTTGATTG---GGTG		595		
ZmB73_Dof01	ATTGCCAAATTTCTTTTCTGCTGCAAGTTTTTATTTTGTAGTTTCTTGATTG---GGTG		595		
ZmW22_Dof01	ATTGCCAAATTTCTTTTCTGCTGCAAGTTTTTATTTTGTAGTTTCTTGATTG---GGTG		595		

	MYB/AT-HOOK				
ZmTEO_Dof01	TTGGTAGCCAGTTTGGATAAATTTGATGGGGATCCTCTATAAGGTTTTAACCTTGATGC		650		
ZmPAL_Dof01	TTGGTAGCCAGTTTGGATAGGTTGATGGGGATCCTCTATAAGCTTTTAACCTTGATGC		658		
ZmMo17_Dof01	TTGGTAGCCAGTTTGGATAGGTTGATGGGGATCCTCTATAAGCTTTTAACCTTGATGC		655		
ZmB73_Dof01	TTGGTAGCCAGTTTGGATAGGTTGATGGGGATCCTCTATAAGCTTTTAACCTTGATGC		655		
ZmW22_Dof01	TTGGTAGCCAGTTTGGATAGGTTGATGGGGATCCTCTATAAGCTTTTAACCTTGATGC		655		

	HD-ZIP	AT-HOOK	bZIP	AT-HOOK	
ZmTEO_Dof01	TATGATTGCGGGAGCCAGAAAAATTTGAGACCATA-----CATTGGATTGGGGTGAATA		704		
ZmPAL_Dof01	TATGATTGCGGGAGCCAGAAAAATTTGAGACCACCTCAGTACATGTTGGATTGGGGTGAATA		718		
ZmMo17_Dof01	TATGATTGCGGGAGCCAGAAAAATTTGAGACCACCTCAGTACATGTTGGATTGGGGTGAATA		715		
ZmB73_Dof01	TATGATTGCGGGAGCCAGAAAAATTTGAGACCACCTCAGTACATGTTGGATTGGGGTGAATA		715		
ZmW22_Dof01	TATGATTGCGGGAGCCAGAAAAATTTGAGACCACCTCAGTACATGTTGGATTGGGGTGAATA		715		

	Dof		TALE		
ZmTEO_Dof01	TTAGAGTACCCAAACAAAGCCTAAAGGTATTTGTGACATCTTGGTTGGATCCACATGCAG		764		
ZmPAL_Dof01	CTAGAGTGCCCAAAACAAATCCTAAAGGTACTATGAAATCTTGGTTGGATCCACATGCAG		778		
ZmMo17_Dof01	CTAGAGTGCCCAAAACAAATCCTAAAGGTACTATGAAATCTTGGTTGGATCCACATGCAG		775		
ZmB73_Dof01	CTAGAGTGCCCAAAACAAATCCTAAAGGTACTATGAAATCTTGGTTGGATCCACATGCAG		775		
ZmW22_Dof01	CTAGAGTGCCCAAAACAAATCCTAAAGGTACTATGAAATCTTGGTTGGATCCACATGCAG		775		

	P3 F		HD-ZIP HD/WOX		
ZmTEO_Dof01	TTATGGCACCTTGGATACAGTTTCTGAGCATAGATTGATTGAGGTCGGAATTTCCCA		824		
ZmPAL_Dof01	TTATGGTACCTTGGATAAGTTTCTGAGCATAGATTGATTGAGGTCGGAATTTCCCA		838		
ZmMo17_Dof01	TTATGGTACCTTGGATAAGTTTCTGAGCATAGATTGATTGAGGTCGGAATTTCCCA		835		
ZmB73_Dof01	TTATGGTACCTTGGATAAGTTTCTGAGCATAGATTGATTGAGGTCGGAATTTCCCA		835		
ZmW22_Dof01	TTATGGTACCTTGGATAAGTTTCTGAGCATAGATTGATTGAGGTCGGAATTTCCCA		835		

	AT-HOOK	trp	HD-ZIP		
ZmTEO_Dof01	GAATACTGAAACAAAACGGGCTACCCACTATGCGGCTACTAGAGTTAATGATCCATGC		884		
ZmPAL_Dof01	AAATACTGAAACAAAACGGGCTACCCACTATGCGGCTACTAGAGTTAATGATCCATGC		896		
ZmMo17_Dof01	AAATACTGAAACAAAACGGGCTACCCACTATGCGGCTACTAGAGTTAATGATCCATGC		893		
ZmB73_Dof01	AAATACTGAAACAAAACGGGCTACCCACTATGCGGCTACTAGAGTTAATGATCCATGC		893		
ZmW22_Dof01	AAATACTGAAACAAAACGGGCTACCCACTATGCGGCTACTAGAGTTAATGATCCATGC		893		

	TALE		C2H2		AT-HOOK
ZmTEO_Dof01	TATGACAAGGAATAACAACCATAGATGTAACACTCAAAGTACTAAAACGTAAAAAGTAC		944		
ZmPAL_Dof01	---GACAAGGAATAACAACCATAGATGTAACACTCAAAGTACTAAAACGTAAAAAGTAC		953		
ZmMo17_Dof01	---GACAAGGAATAACAACCATAGATGTAACACTCAAAGTACTAAAACGTAAAAAGTAC		950		
ZmB73_Dof01	---GACAAGGAATAACAACCATAGATGTAACACTCAAAGTACTAAAACGTAAAAAGTAC		950		
ZmW22_Dof01	---GACAAGGAATAACAACCATAGATGTAACACTCAAAGTACTAAAACGTAAAAAGTAC		950		

	SBP		MYB		EIN3
ZmTEO_Dof01	AGTGTAAACAAAAAGGATATTACATCGTGATATGGCTATACAAAAAATGTTCTGT		1004		
ZmPAL_Dof01	AGTGTAAACAAAGCAGAAGTTATTACATCGTGATATGGCTATGCAAAAACATGTTTAC		1013		
ZmMo17_Dof01	AGTGTAAACAAAGCAGAAGTTATTACATCGTGATATGGCTATGCAAAAACATGTTTGC		1010		
ZmB73_Dof01	AGTGTAAACAAAGTAGAAGTTATTACATCGTGATATGGCTATGCAAAAACATGTTTAC		1010		
ZmW22_Dof01	AGTGTAAACAAAGTAGAAGTTATTACATCGTGATATGGCTATGCAAAAACATGTTTAC		1010		

	C2H2				
ZmTEO_Dof01	ATTGTTTAC-----AGAGTGGAGTCTAAAATGAGATACAAAGATGAGTAAGCT		1052		
ZmPAL_Dof01	ACTATAGACTACATTTGTTTATAGAGTGAAGTTTAAAAATGAGATATGAGGATGAGTAAGCT		1073		
ZmMo17_Dof01	ACTATAGACTACATTTGTTTATAGAGTGAAGTTTAAAAATGAGATATGAGGATGAGTAAGCT		1070		
ZmB73_Dof01	ACTATAGACTACATTTGTTTATAGAGTGAAGTTTAAAAATGAGATATGAGGACGAGTAAGCT		1070		
ZmW22_Dof01	ACTATAGACTACATTTGTTTATAGAGTGAAGTTTAAAAATGAGATATGAGGACGAGTAAGCT		1070		
	* * * *		* * * * *		

ZmTEO_Dof01	G-----	1053
ZmPAL_Dof01	GCTAGATATATA-----	1085
ZmMo17_Dof01	GCTAGATATATAGATAGTCTC-----	1091
ZmB73_Dof01	GCTAGATATATAGGGGGTGTGGTTGTAGGGATTAAACTTTAGTCCCTCCACTTTATT	1130
ZmW22_Dof01	GCTAGATATATAG-----	1083
	*	
	MADS box	
ZmTEO_Dof01	-----	1053
ZmPAL_Dof01	-----	1085
ZmMo17_Dof01	-----	1091
ZmB73_Dof01	CCATTTTAGTTCTCAAATTGTTAAATAGGGAACTAAAATAGAGCTTTAACTTCCATATT	1190
ZmW22_Dof01	-----	1083
	MYB-related AT-HOOK	
ZmTEO_Dof01	-----C	1054
ZmPAL_Dof01	-----	1085
ZmMo17_Dof01	-----	1091
ZmB73_Dof01	TAGTAATTTAGGGACTAAAATAGAATGAAATAGAGGGACTAAAAATTAGTCCCTATAAAC	1250
ZmW22_Dof01	-----	1083
	HD-ZIP SBP	
ZmTEO_Dof01	TAGATATATAGATAGTCTCAATGATAACATACCGGAGAGACTAGTTCCTGCAATAAGTA	1114
ZmPAL_Dof01	-----GTCTCAATGATAACATACCGGAGAGATTAGTTCCTGCAATAAGTA	1131
ZmMo17_Dof01	-----AATGATAACATACTGGAGAGA-TTAGTTCCTGCAATAAGTA	1131
ZmB73_Dof01	CAAAACACCCCATAGTCTCAATGATAACATACCGGAGAGATTAGTTCCTGCAATAAGTA	1310
ZmW22_Dof01	-----ATAGTCTCAATGATAACATACCGGAGAGATTAGTTCCTGCAATAAGTA	1132
	***** * * * *	
ZmTEO_Dof01	CAGTAACACCACCATCGCTCATATCTTGATGGAATATGATGTACCACTGAAGTATATCTG	1174
ZmPAL_Dof01	CAGTAACACCACCATCGCTCATATCTTGATGGAATATGATGTACCACTGAAGTATATCTG	1191
ZmMo17_Dof01	CAGTAACACCATCATCGCTCATATCTTGATGGAATATGATGTACCACTGAAGTATATCTG	1191
ZmB73_Dof01	CAGTAACACCATCATCGCTCATATCTTGATGGAATATGATGTACCACTGAAGTATATCTG	1370
ZmW22_Dof01	CAGTAACACCATCATCGCTCATATCTTGATGGAATATGATGTACCACTGAAGTATATCTG	1192
	***** * * * *	
	TALE AT-HOOK P4 F	
ZmTEO_Dof01	CTTTGACCTTGGCAAACACTAGTCTGTGAAAAATGCTGATTAAGGCGTTCCCTCTTTTACA	1234
ZmPAL_Dof01	CTTTGACCTTGGCAAACACTAGTCTGTGAAAAATGCTGATTAAGGCGTTCCCTCTTTTACA	1251
ZmMo17_Dof01	CTTTGACCTTGGCAAACACTAGTCTGTGAAAAATGCTGATTAAGGCGTTCCCTCTTTTATT	1251
ZmB73_Dof01	CTTTGACCTTGGCAAACACTAGTCTGTGAAAAATGCTGATTAAGGCGTTCCCTCTTTTACA	1430
ZmW22_Dof01	CTTTGACCTTGGCAAACACTAGTCTGTGAAAAATGCTGATTAAGGCGTTCCCTCTTTTACA	1252
	***** * * * *	
ZmTEO_Dof01	AATAAGATG-----CCAAGGTTTGCTATAGCTGCCAAGATTTTGTAGCTTGTGTAAGGAGG	1289
ZmPAL_Dof01	AATAAGATGGCA-----AGGTTTGTAGCTGCCAAGATTTTGTAGCTTGTGTAAGGAGG	1304
ZmMo17_Dof01	TTACAATAAGATGTC AAG--TTTGCTAGATGCCAAGATTTTGTAGCTTGTGTAAGGAGG	1309
ZmB73_Dof01	AATAAG-----ATGTCAAGGTTTGCTATAGCTGCCAAGATTTTGTAGCTTGTGTAAGGAGG	1485
ZmW22_Dof01	AATAAG-----ATGTCAAGGTTTGCTATAGCTGCCAAGATTTTGTAGCTTGTGTAAGGAGG	1307
	* * * * *	
	AT-HOOK/HD TALE AT-HOOK	
ZmTEO_Dof01	GAATGAAAGGTAGTCACTGACTGTAG-----TGGAGAGGATATATAAATTATGCGTCTT	1342
ZmPAL_Dof01	GAATGAAAGATAATCACFGACTGTAGTTATTAGTGAAGAGGATATATAAATTATGTCCTT	1364
ZmMo17_Dof01	GAATGAAAGATAATCACFGACTGTA-----GTGAAGAGGATATATAAATTATGTCCTT	1362
ZmB73_Dof01	GAATGAAAGATAAATCACTGACTGTAGTTATTAGTGAAGAGGATATATAAATTATGTCCTT	1545
ZmW22_Dof01	GAATGAAAGATAATCACFGACTGTAGTTATTAGTGAAGAGGATATATAAATTATGTCCTT	1367
	***** * * * *	
ZmTEO_Dof01	CTAAGCAAACACGGAGTAAATGTCTTAGATATAATCATATAACTAACTATATAAGCATC	1402
ZmPAL_Dof01	CCAAGCAAACACGGAGTAAATGTCTTAGATATAATCATATAA-----CTATATAAGCATC	1419
ZmMo17_Dof01	CCAAGCAAACACGGAGTAAATGTCTTAGATATAATCATATAA-----CTATATAAGCATC	1417
ZmB73_Dof01	CTAAGCAAACACGGAGTAAATGTCTTAGATATAAGCATATAA-----CTATATAAGCATC	1600
ZmW22_Dof01	CTAAGCAAACACGGAGTAAATGTCTTAGATATAAGCATATAA-----CTATATAAGCATC	1422
	* * * * *	

EIN3		
ZmTEO_Dof01	GAGTACTTGCTTGGCTCCTTGGATTCTAAAGTTCTGTTATTTGCCAGGATTTCTGAAGT	1462
ZmPAL_Dof01	GAGTACTTGCTTGGCTCCCAGGATTTCTAAAGTTCTGTTATTTGCCAGTA-----	1469
ZmMo17_Dof01	GAGTACTTGCTTGGCTCCTTGGATTCTAAAGT-----	1450
ZmB73_Dof01	<u>ATGTACTT</u> GCTTGGCTCCTTGGATTCTAAAGTTCTGTTATTTGCTAGGATTTCTAAAGT	1660
ZmW22_Dof01	<u>ATGTACTT</u> GCTTGGCTCGTTGGATTCTAAAGTTCTGTTATTTGCTAGGATTTCTAAAGT	1482

MYB/SANT		
ZmTEO_Dof01	TCTGTTATTTGCCAGTAAATGTCATAGACATATTTTACAGCATCTTACAATCATATTTTCG	1522
ZmPAL_Dof01	-----AA--TGTCAGACATATTTTACAGCATCTTACAATCACATTTCC	1510
ZmMo17_Dof01	TATGTTATTTGCCAGTAAATGTCATAGACATATTTTACATCATCTTACAATCATATTTCC	1510
ZmB73_Dof01	TCTGTTATTTGCCAGTAAATGTAATAGACATATTTTACATCATCTTACAATCATATTTCC	1720
ZmW22_Dof01	TCTGTTATTTGCCAGTAAATGTAATAGACATATTTTACATCATCTTACAATCATATTTCC	1542
	* * * * *	
bZIP/EIN3/NAC		EIN3
ZmTEO_Dof01	TTCTATCTCAATACATGTATTTACTTAGTCTTCGTGTATTTGTCAGGATGTATGTGAAG	1582
ZmPAL_Dof01	TTCTATCTCAATACATGTATTTACTTAGTCTTCGTGTATTTGTCAGGATGTATGTGAAG	1570
ZmMo17_Dof01	TTCTATCTCAATACATGTATTTACTTAGTCTTCGTGTATTTGTCAGGATGTATGTGAAG	1570
ZmB73_Dof01	TTCTATCTCAATACATCTATTTACTTAGTCTTCGTGTATTTGTCAGGATGTATGTGAAG	1780
ZmW22_Dof01	TTCTATCTCAATACATCTATTTACTTAGTCTTCGTGTATTTGTCAGGATGTATGTGAAG	1602

P5 R		
ZmTEO_Dof01	<u>ATAGAGGGGACATG</u>	1596
ZmPAL_Dof01	<u>ATAGAGGGGACATG</u>	1584
ZmMo17_Dof01	<u>ATAGAGGGGACATG</u>	1584
ZmB73_Dof01	<u>ATAGAGGGGACATG</u>	1794
ZmW22_Dof01	<u>ATAGAGGGGACATG</u>	1616

Figura S2. Secuencia consenso de los motivos localizados en la secuencia del promotor Dof01 de las cinco variedades de maíz. El tamaño de cada letra representa el grado de conservación del nucleótido en cada posición. Más de un nucleótido puede estar compartiendo una misma posición sin afectar la unión de un factor de transcripción.

Figure S2. Consensus sequence of the motifs located in the sequence of the Dof01 promoter of the five maize varieties. The size of each letter represents the degree of conservation of the nucleotide in each position. More than one nucleotide may be sharing the same position without affecting the binding of a transcription factor.

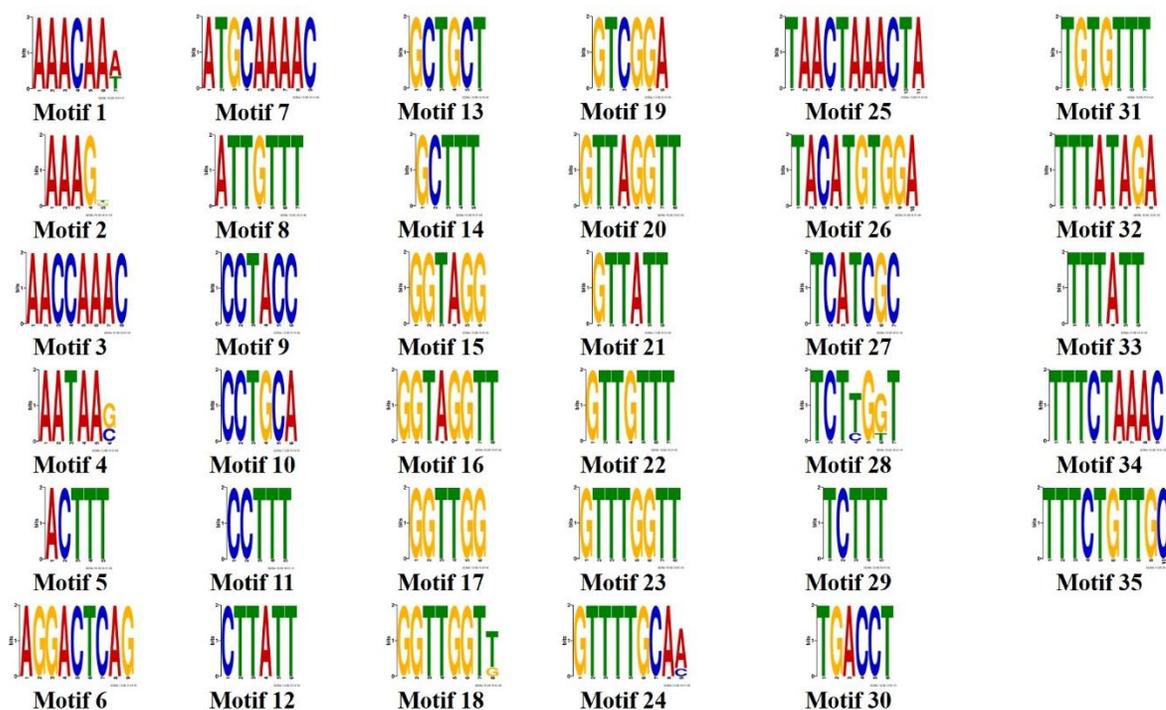


Figura S3. Expresión del gen Dof01 y de los factores de transcripción MYB y AP2/B3 en 79 tejidos de maíz B73. En la parte inferior del gráfico se enumeran los tejidos en los que se midió la expresión. Valores de expresión absoluta obtenidos del atlas de expresión de maíz.

Figure S3. Expression of the Dof01 gene and the MYB and AP2/B3 transcription factors in 79 B73 maize tissues. The tissues in which the expression was measured are listed at the bottom of the graph. Absolute expression values obtained from the maize expression atlas.

