



## 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-Avelar<sup>1</sup>, Axel Tiessen-Favier<sup>2†</sup>, Luis Herrera-Estrella<sup>3</sup>, Felipe Ascencio-Valle<sup>1</sup>, César Cardona-Félix<sup>4</sup>, Gracia Alicia Gómez-Anduro<sup>1\*</sup>

<sup>1</sup> 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.

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

<sup>3</sup> 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.

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

\*Corresponding author

E-mail address: [ggomez@cibnor.mx](mailto: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

Gene expression patterns are controlled by *cis* regulatory elements (CRE) present in promoters. To understand how CRE are responsible for transcriptional activity and shape the architecture of a promoter, it is necessary to identify them and associate them with specific responses. Comparative studies of Dof promoters have focused on analyzing these sequences in different species. However, it is unknown how CRE are conserved among different varieties of the same species. In this work, a comparative analysis of the *Dof01* gene promoter sequences of five varieties of maize (*Zea mays*) (B73, Palomero Toluqueño, Teosinte, W22 and Mo17) was performed to understand the architecture of the *Dof01* promoter. Sequence analysis shows that there is a 179 bp indel element that is present in the B73 *Dof01* promoter but absent in the other four varieties. Also, highly conserved CRE can be identified and may be related to the response to low temperatures. Therefore, we analyzed the expression of the *Dof01* gene in B73 and Palomero under conditions of stress from low (4°C) and high (42°C) temperature. The expression data considered that the Palomero *Dof01* gene is over-expressed with respect to the control due to the cold response. This work is the first step for the functional characterization of the *Dof01* gene promoter to identify the regulatory elements associated with the temperature stress response and obtain the minimum promoter necessary for transcriptional regulation.

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

## RESUMEN

Los patrones de expresión de genes son controlados por elementos regulatorios *cis* (CRE) presentes en los promotores. Para entender cómo los CRE son responsables de la actividad transcripcional y forman la arquitectura de un promotor, es necesario identificarlos y asociarlos a respuestas específicas. Los estudios comparativos de promotores Dof se han enfocado en analizar estas secuencias en diferentes especies. Sin embargo, se desconoce cómo los CRE son conservados entre diferentes variedades de una misma especie. En este trabajo se realizó un análisis comparativo de las secuencias del promotor del gen *Dof01* de cinco variedades de maíz (*Zea mays*) (B73, Palomero Toluqueño, Teosinte, W22 y Mo17) para entender la arquitectura del promotor *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 otras cuatro variedades. También, se identificaron CRE altamente conservados y que pueden estar relacionados con la respuesta a bajas temperaturas. Por eso, se analizó la expresión del gen *Dof01* en B73 y Palomero bajo condiciones de estrés por baja (4°C) y alta (42°C) temperatura. Los datos de expresión mostraron que el gen *Dof01* de Palomero se sobre expresa con respecto al control debido a la respuesta

a frío. Este trabajo es el primer paso para la caracterización funcional del promotor del gen *Dof01* para identificar los elementos regulatorios asociados con la respuesta a estrés por temperatura 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 importante 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 la adaptación de las especies (Lemmon *et al.*, 2014). De acuerdo con Creux *et al.* (2008), la identificación de elementos regulatorios conservados entre secuencias de promotores homólogos en especies evolutivamente distantes, puede proveer evidencia sobre redes regulatorias ancestrales (Creux *et al.*, 2008). Sin embargo, también es importante conocer 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 (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).

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 de maíz, con implicaciones significativas para el entendimiento de su organización y evolución (Prasanna, 2012). Las razas ancestrales de este cultivo, como Palomero Toluqueño, 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). Por ejemplo, Palomero Toluqueño está adaptado a climas fríos con temperaturas medias que oscilan los 12.5°C y que frecuentemente descienden de los 6°C, pero que además se considera tolerante a las heladas pero con una pobre adaptación a las altas temperaturas (Eagles & Lothrop, 1994). Aguilar-Rangel *et al.* (2017) realizaron un análisis del transcriptoma de B73 y Palomero Toluqueño bajo condiciones de estrés abiótico, entre ellos, baja y alta temperatura en plántulas. Los autores encontraron identificaron 277 genes que mostraron expresión diferencial entre Palomero Toluqueño y B73, incluyendo genes relacionados con la expresión de factores de transcripción.

Estudiar las regiones regulatorias localizadas en el promotor del gen *Dof01* 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 tal como la temperatura. 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 (Koch *et al.*, 2001; Lockton & Gaut, 2005; Creux *et al.*, 2008; Reineke *et al.*, 2011; Wittkopp & Kalay, 2012; Creux *et al.*, 2013; Das *et al.*, 2019; Powell *et al.*, 2019). 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 CRE presentes, pero no realizan una comparación de la distribución de los CRE entre pares de promotores de genes homólogos.

Por lo que el objetivo del presente trabajo fue realizar un análisis comparativo de las secuencias del promotor *Dof01* en cinco variedades de maíz (se incluyó al ancestro progenitor del maíz actual, el Teosinte (TEO) y una de las razas de maíz ancestrales del centro de México, Palomero Toluqueño (PAL), 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 mediante análisis de expresión.

## **2. MATERIALES Y MÉTODOS**

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

Se accedió a la base de datos de Phytozome V12 (Goodstein *et al.*, 2012) para identificar la secuencia del promotor del gen *Dof01* de maíz B73. Se seleccionaron 1794 pb disponibles, río arriba del codón de inicio para posteriores análisis. Esta secuencia fue empleada para identificar las regiones homólogas a *Dof01* 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 la plataforma PlantPAN (Chow *et al.*, 2016). Además, se identificaron los sitios de inicio de la transcripción (TSS, del inglés Transcription Start Site) y las 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, se utilizaron estos datos para representar de manera gráfica los TFBS, los TSSs y las 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, se realizó un alineamiento múltiple con la herramienta Clustal W. En el alineamiento se identificó la posición de los sitios de unión y se compararon las posiciones de los factores de transcripción que regulan la expresión del gen *Dof01*.

Además, para entender cómo se distribuyen los elementos regulatorios con relación a la posición del TSS, se realizó 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 del gen *Dof01* en B73, se tomaron los datos del atlas de expresión de Hoopes *et al.*, (2018) de 79 tejidos distintos y se extrajeron los valores de expresión con base al

Log2 de *Dof01* y los factores de transcripción que se unen a la secuencia de dicho promotor. Se realizó un test de correlación de Pearson para encontrar los factores de transcripción que más influyen en la expresión de *Dof01* y se seleccionaron 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/oscuridad y  $27 \pm 1^\circ\text{C}$ ). Las plántulas de 9 días de edad se sometieron a los tratamientos de estrés por temperatura baja ( $4^\circ\text{C}$ ) y alta ( $42^\circ\text{C}$ ) 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  $-80^\circ\text{C}$  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 Palomero de 10 días después de la germinación mediante el método de Trizol (ThermoFisher) utilizando las indicaciones del fabricante. Se emplearon tres réplicas biológicas de cada tratamiento. Se sintetizó el cDNA con la enzima transcriptasa reversa ImProm-II (Promega) utilizando 1  $\mu\text{g}$  de RNA total después del tratamiento con DNasa I. El análisis de la expresión de los genes se realizó por PCR tiempo real con un termociclador Corbett RotorGene 6000 (ThermoFisher) utilizando la mezcla de reacción SsoFast Eva Green (BioRad). Para la selección del gen endógeno, se probaron tres candidatos (*Factor de Elongación 1 $\alpha$*  (*EF1  $\alpha$* ),  *$\beta$ -Tubulina* y *Ubiquitina-9*) para determinar su estabilidad (M). El nivel de expresión relativa se calculó por el método de  $2^{-\Delta\Delta\text{ct}}$  (Livak & Schmittgen, 2001) con respecto al gen *EF1  $\alpha$* . 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 *et al.*, (2018), el cual contiene datos de expresión de genes de maíz bajo tratamientos de estrés, entre ellos temperatura baja ( $4^\circ\text{C}$ ) y alta ( $42^\circ\text{C}$ ) en plántulas de 10 días después de la germinación.

### 3. RESULTADOS

#### 3.1. Segmento de inserción/delección en el promotor *Dof01*

Se realizó un alineamiento local (BLAST) utilizando la secuencia del promotor *Dof01* de B73 para buscar e identificar las respectivas secuencias homólogas en cuatro genomas de diferentes variedades de maíz (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) a lo largo de toda la secuencia del promotor. La secuencia de Mo17 (1584 pb) tiene una identidad del 99.08%, en ella se encuentran 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 río arriba y 17 río abajo 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*.

#### 3.2. Arquitectura del promotor *Dof01*

En la Fig. 1 se representó de manera gráfica cada una de las cinco secuencias del promotor *Dof01* para comparar la ubicación de los elementos regulatorios. 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. Se aprecia 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 es de 238 pb, 233 pb y 206 pb, respectivamente. También, se representó 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, se aprecia que está presente una caja TATA, la cual no está en la secuencia de PAL pero sí en las otras cuatro variedades. Además, esta región posee cinco TFBS que no se localizan en el resto de las secuencias. La caja TATA

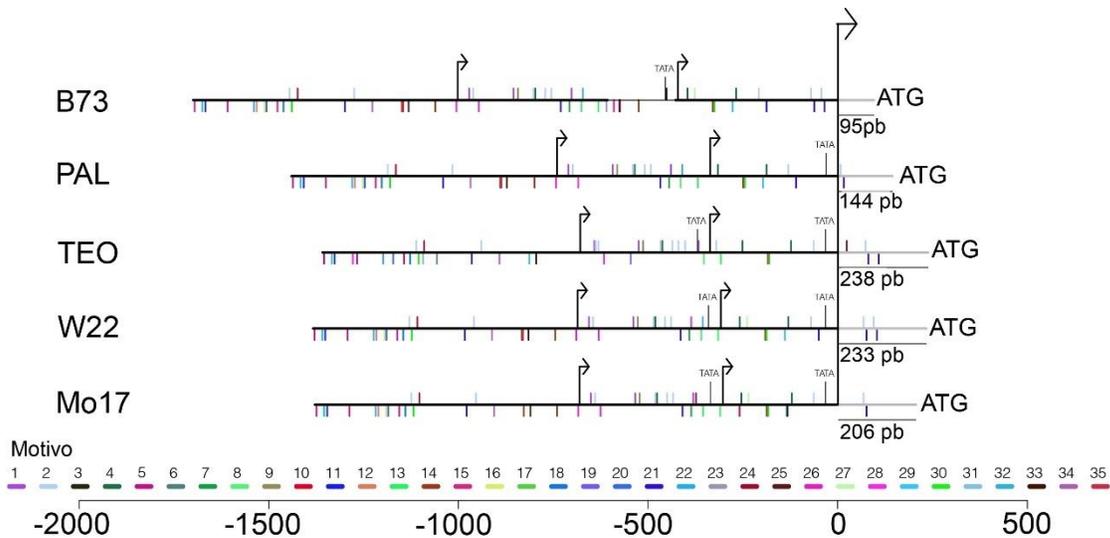
que de la posición -35 pb en las secuencias de PAL, TEO, W22 y Mo17 no está presente en la secuencia de B73, esto puede ser motivo de la diferencia en la respuesta a estrés por frío/calor observada entre B73 y PAL provocada por un cambio significativo en la actividad del promotor. En total se identificaron 35 motivos distintos en las cinco secuencias.

**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.

**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.

Query	Sujeto	% Identidad	Longitud de alineamiento (pb)	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

. Las secuencias consenso de cada motivo se muestran en la Fig. S2.



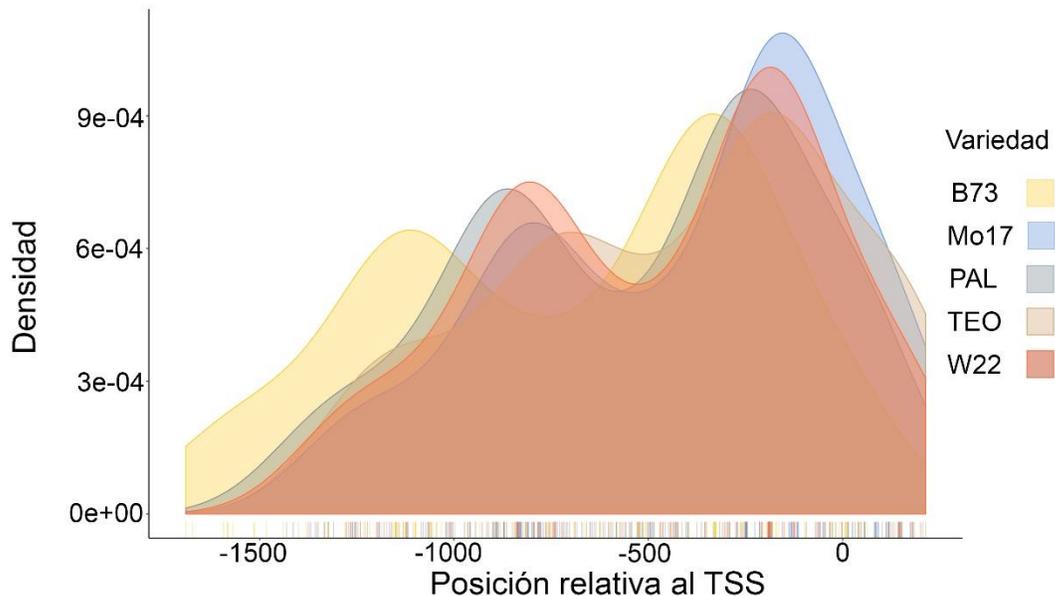
**Figure 1.** Representation of the architecture of the *Dof01* promoter of five maize varieties. The horizontal black lines represent the size of the sequence of each promoter located in their respective genomes. The horizontal gray line represents the 5'-UTR. The larger vertical arrow represents the 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 start codon ATG are also indicated. The lower scale represents the size of the sequence where zero (0) is the position of the main TSS and in negative the region upstream of this, while in positive the regions downstream.

**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 indican también las posiciones de las cajas TATA y el codón de inicio ATG. La escala inferior representa el tamaño de la secuencia donde cero (0) es la posición del TSS principal y en negativo la región río arriba de este, mientras que en positivo las regiones río abajo.

### 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). Se identificó 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 disminuye hacia los 500 pb y vuelve a incrementar cerca de los -850 pb. Se encontraron dos

TSS alternativos cuya función pudiera estar relacionada a un transcrito alternativo 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.

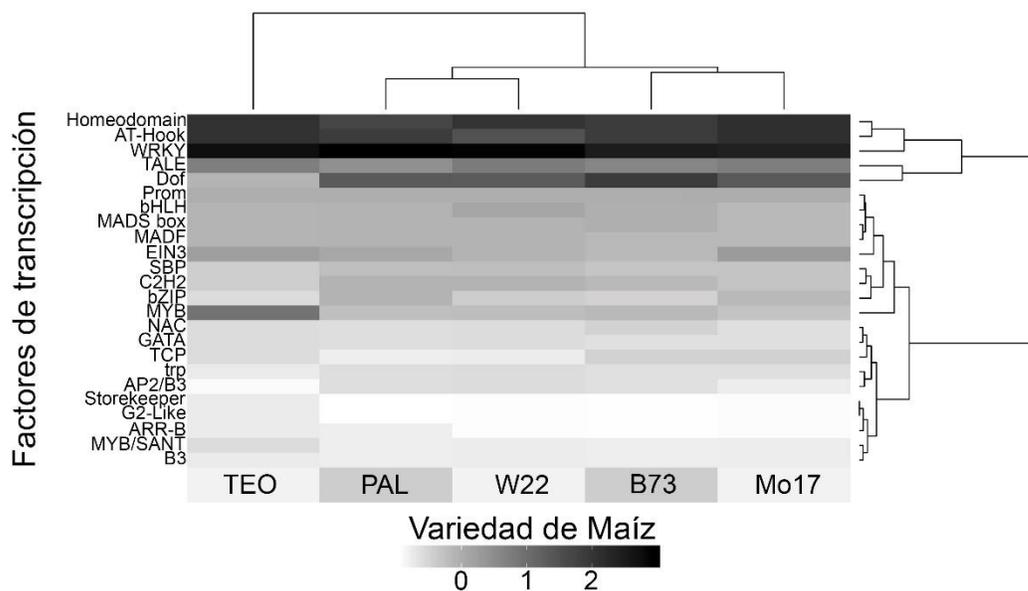


**Figure 2.** Density of distribution of *cis* elements in the *Dof01* promoter of the five maize varieties. Two peaks of maximum density are located near the main TSS and an alternate TSS. The higher the peak, the greater the likelihood that a transcription factor will join that region. The lower scale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

**Figura 2.** Densidad de la distribución de los 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 de densidad representa la probabilidad de unión de un factor de transcripción a una posición específica.

Adicionalmente, se encontraron 318 factores de transcripción distintos que pertenecen a 23 familias que reconocen 655 potenciales sitios de unión, cuya actividad podría ser 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, se construyó un mapa de calor (Fig. 3) con ayuda del software R. En la Fig. 3 se observa 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. Se observa 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 reconocidos por tres factores de transcripción distintos, los cuales son *Dof02* (GRMZM2G009406), *Dof03* (GRMZM2G146283) y *Dof35* (GRMZM2G179069). Dieciocho de estos están en B73, 15 en Mo17, 13 en PAL y W22 y sólo cinco en TEO.



**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.

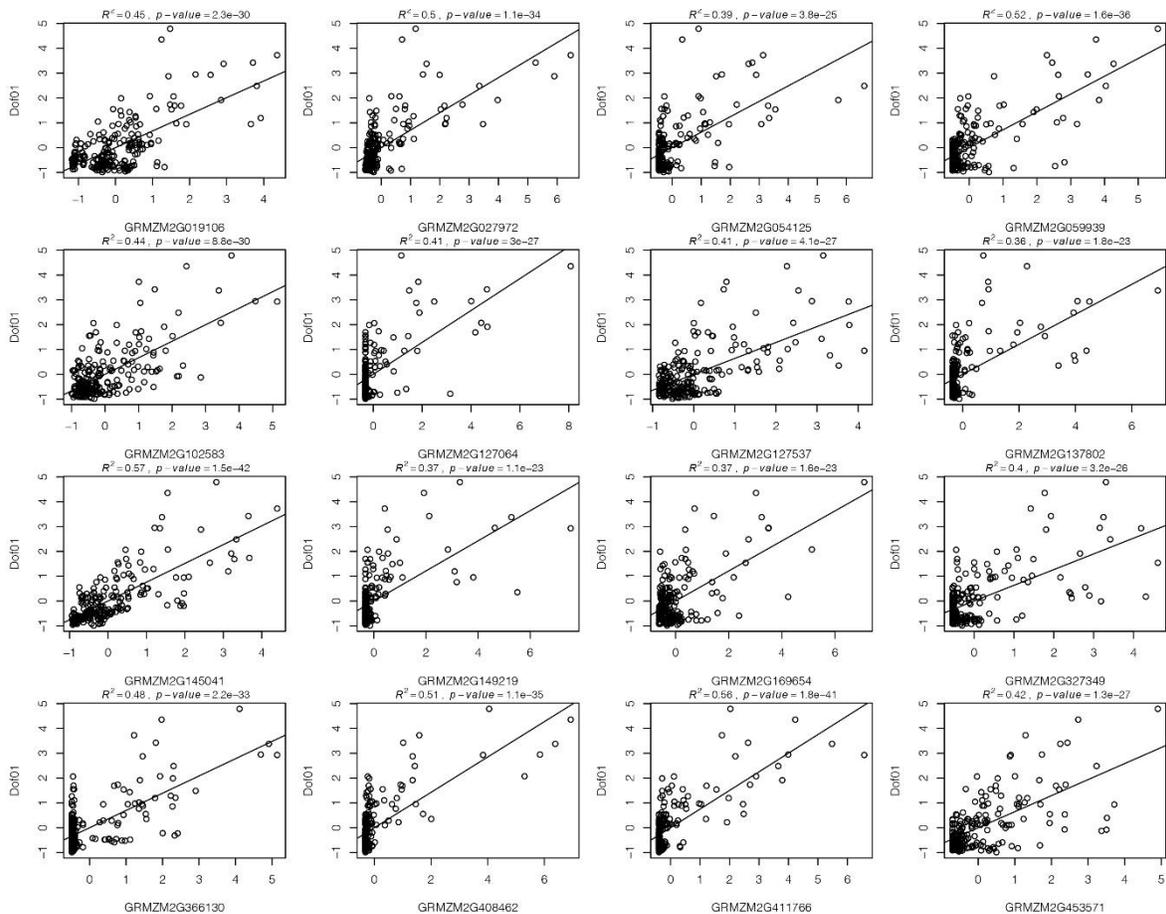
**Figura 3.** Relaciones entre el tipo y número de factores de transcripción, clasificados por familia, que se unen a la secuencia del promotor *Dof01* de las cinco variedades de maíz. La escala inferior representa el número de desviaciones

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.

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

Se realizó un análisis de correlación de Pearson entre los valores de expresión de los genes que codifican para los factores de transcripción que potencialmente se unen al promotor *Dof01* y los valores de expresión de este gen en B73. La Fig. 4 muestra un gráfico de dispersión de la correlación de la expresión de *Dof01* con la expresión de 16 factores de transcripción. Se encontró una mayor correlación con el gen que codifica para la proteína relacionada a reveille 1 (GRMZM2G145041) con un coeficiente de determinación ( $R^2$ ) de 0.57, la cual posee un dominio de unión al DNA tipo MYB. Diez de los 16 genes correlacionados pertenecen a la familia de factores de transcripción WRKY (GRMZM2G411766; GRMZM2G408462; GRMZM2G027972; GRMZM2G102583; GRMZM2G453571; GRMZM2G127064; GRMZM2G327349; GRMZM2G054125; GRMZM2G149219; GRMZM2G137802) con  $R^2$  de 0.56, 0.51, 0.5, 0.44, 0.42, 0.41, 0.4, 0.39, 0.37 y 0.36, respectivamente. Dos de los 16 factores de transcripción señalados contienen el dominio HD-ZIP (GRMZM2G366130 y GRMZM2G127537) con un  $R^2$  de 0.48 y 0.41, respectivamente. El factor de transcripción con el dominio de unión al DNA tipo AP2/B3 (GRMZM2G059939) tuvo un coeficiente  $R^2$  de 0.52. El gen que codifica para el factor de transcripción con dominio RAV (GRMZM2G169654) tuvo un  $R^2$  de 0.37. Mientras que el FT tipo bZIP (GRMZM2G019106) tuvo un coeficiente de determinación de 0.45.

La Fig. S3 muestra la expresión de los genes *Dof01*, bZIP y MYB a través de 79 tejidos de maíz B73. Estos genes se expresan principalmente en hoja, mientras que aquellos en los que la expresión disminuye, están relacionados al desarrollo de la semilla.



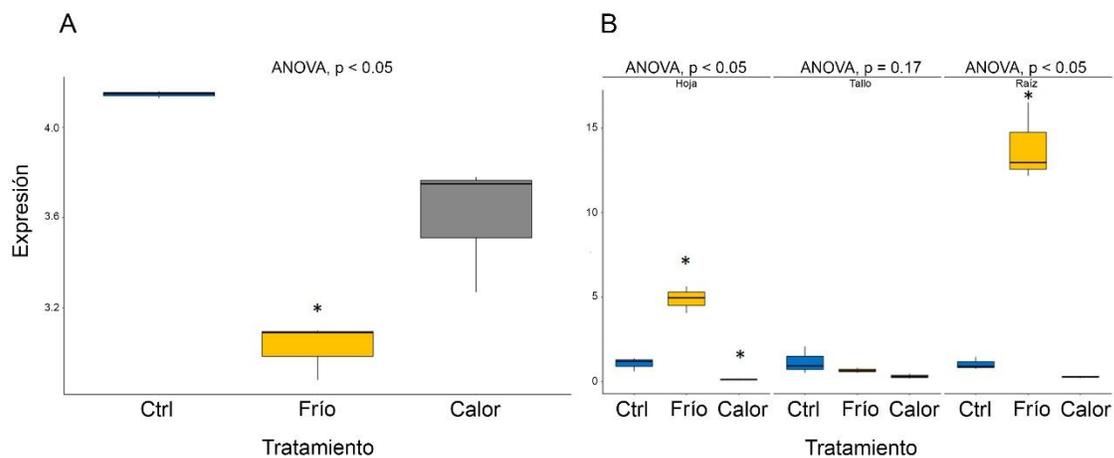
**Figure 4.** Pearson's correlation analysis between the expression of the *Dof01* gene of B73 and the transcription factors involved in the regulation of its expression. The coefficient of determination of the linear regression as well as the p-value are shown at the top of each scatter plot. The identification number of each gene analyzed is shown at the bottom of each graph. The left and bottom axis shows the value of the absolute expression of the transformed data in 79 different tissues.

**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. El eje inferior y lateral izquierdo, muestran el valor de la expresión absoluta de los datos transformados de los factores de transcripción en 79 tejidos distintos.

### 3.5. El gen *Dof01* se expresa diferencialmente en maíz palomero y B73 bajo estrés por frío

En la Fig. 5 se muestran los datos de expresión relativa del gen *Dof01* de plántulas de maíz B73 (Fig.5A) y PAL (Fig. 5B) de nueve días de edad bajo tratamientos de estrés por temperatura baja y alta. De acuerdo a los datos de expresión relativa de

maíz B73 obtenidos del atlas de Hoopes, se aprecia que el gen *Dof01* se encuentra significativamente reprimido por efecto de ambos tratamientos de estrés con respecto al tratamiento control. En el caso de maíz palomero (Fig. 5B), la expresión del gen *Dof01* se induce 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 induce significativamente más de diez veces a consecuencia del estrés por frío con respecto al control. Mientras que, debido al estrés por calor, no se encontraron diferencias significativas.



**Figure 5.** Relative 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 whole plant and B) PAL in leaf, stem and root tissues is shown.

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

#### 4. DISCUSIÓN

En este trabajo se realizó una caracterización general de la arquitectura del promotor *Dof01* en cinco variedades de maíz (B73, Palomero Toluqueño, Teosinte, W22 y Mo17). Se identificaron elementos regulatorios que están conservados entre los cinco promotores que pueden ser 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 sobre el 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 común de todas las razas de maíz, el cual fue domesticado hace alrededor de 9,000 a 10,000 años. Lemmon *et al.* (2014), analizaron elementos regulatorios en *cis* y *trans* entre Teosinte y B73. 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). En este trabajo se observó que la mayor divergencia en elementos regulatorios se presenta entre el promotor *Dof01* de TEO y B73.

La variedad W22 ha servido como planta modelo 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 genéticos y moleculares. 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 el caso del promotor *Dof01*, se encontró que 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 al ambiente y aquellos elementos *cis* 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). Se observó que el ancestro del maíz, Teosinte, difiere en número y tipo de elementos regulatorios con respecto a B73 en la secuencia del promotor *Dof01*. La pérdida o ganancia de estos elementos regulatorios *cis* es una de las principales causas de la diversidad genética en plantas, 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, se debe estudiar en conjunto estos tres elementos de una manera en la que se provea nueva información sobre el funcionamiento de la regulación de la expresión de genes.

De acuerdo con Lockton y Gaut (2005), los promotores homólogos comparten elementos regulatorios *cis* incluso a través de distancias evolutivamente considerables. En la Fig. 1 se observa 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 con 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 los 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 expresión del gen *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).

Las diferencias en los patrones de expresión del gen *Dof01* en ambas variedades de maíz puede deberse a los diferentes elementos *cis* localizados en los promotores de los dos 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 (Fig. S3) y se encuentra en la misma posición en ambos promotores. Por otra parte, el otro elemento *cis* localizado en el promotor y relacionado a la respuesta a frío es AP2/B3, el cual cuenta con un sitio de unión en B73 y dos sitios en PAL. Este factor de transcripción se expresa principalmente en hoja, por lo que puede ser el principal elemento regulador de la respuesta a frío en el promotor *Dof01* y el cual podría servir para contrarrestar el efecto de regulador negativo del factor de transcripción MYB. Se requieren estudios de caracterización funcional de los promotores e identificación de promotor mínimo para determinar qué elementos regulatorios están jugando un papel más importante en la regulación de la expresión de *Dof01*, así como determinar sus posibles aplicaciones en el campo de la biotecnología para la generación de plantas que regulen la expresión de genes en respuesta al frío.

## **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., Chávez Montes R. A., González-Segovia E., Ross-Ibarra J., Simpson J. K. & Sawers R. 2017. Allele specific expression analysis identifies regulatory variation associated with stress-related genes in the Mexican highland maize landrace Palomero Toluqueño. *PeerJ*. (5): e3737.

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): plx025.

Chen Y. & Cao J. 2015. Comparative analysis of dof transcription factor family in maize. *Plant Molecular Biology Reporter*. 33(5):1245–1258.

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): 8122–8147.

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): D1154–D1160.

Creux N.M., De Castro M.H., Ranik M., Maleka M. F. & Myburg A. A. 2013. Diversity and *cis*-element architecture of the promoter regions of cellulose synthase genes in *Eucalyptus*. *Tree Genetics & Genomes*. (9): 989–1004.

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): 722-737.

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): e0212678.

Eagles H.A. & Lothrop J.E. 1994. Highland maize from central Mexico—its origin, characteristics, and use in breeding programs. *Crop Science*. (34): 11-19.

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(Database issue): D1178–D1186.

Hoopes G. M., Hamilton J. P., Wood J. C., Esteban E., Pasha A., Vaillancourt B., Provart N. J. & Buell C. R. 2019. An updated gene atlas for maize reveals organ-specific and stress-induced genes. *The Plant journal: for cell and molecular biology*. 97(6): 1154–1167.

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): D1040–D1045.

Koch M. A., Weisshaar B., Kroymann J., Haubold B. & 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.

Lemmon Z. H., Bukowski R., Sun Q. & Doebley, J. F. 2014. The role of *cis* regulatory evolution in maize domestication. *PLoS genetics*. 10(11): e1004745.

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): 402-408.

Lockton S. & Gaut B. S. 2005. Plant conserved non-coding sequences and paralogue evolution. *Trends in Genetics*. 21(1): 60-65.

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): 33.

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): 3309–3320.

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): 379–390.

Powell R. V., Willett C. R., Goertzen L. R. & Rashotte A. M. 2019. Lineage specific conservation of *cis*-regulatory elements in Cytokinin Response Factors. *Scientific Reports*. 9(1): 13387.

Prasanna B. M. 2012. Diversity in global maize germplasm: characterization and utilization. *Journal of Biosciences*. 37(5): 843 – 855.

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. R., 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. & Wilson R. K. 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science*. 326(5956): 1112-1115.

Reineke A. R., Bornberg-Bauer E. & Gu J. 2011. Evolutionary divergence and limits of conserved non-coding sequence detection in plant genomes. *Nucleic acids research*. 39(14): 6029–6043.

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): e65.

Springer N. M., Anderson S. N., Andorf C. M., Ahern K. R., Bai F., Barad O., Barbazuk W. B., Bass H. 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): 1282-1288.

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): 2199–2210.

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): 299.

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*. 326(5956): 1078.

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): e0170210.

- 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): 23072.
- Wittkopp P. J. & Kalay G. 2012. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics*. 13(1): 59-69.
- 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): 574.
- 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): 333–342.
- 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.
- 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): 16028-16036.
- Yu C.-P., Lin J.-J. & Li W.-H. 2016. Positional distribution of transcription factor binding sites in *Arabidopsis thaliana*. *Scientific Reports*. 6(1): 25164.
- 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): 3921-3928.

# SUPPLEMENTARY MATERIAL

	<b>P1 F      GATA</b>	
ZmTEO_Dof01	<b>GCGACTTTCATGATCCAAGC</b> ATATCTTTTCCCCTTTTCCAATACTTGCAACTGCTCTGGC	60
ZmPAL_Dof01	<b>GCGACTTTCATGATCCAAGC</b> ATATCTTTTCCCCTTTTCCAATACTTGCAACTGCTCTGGC	60
ZmMo17_Dof01	<b>GCGACTTTCATGATCCAAGC</b> ATATCTTTTCCCCTTTTCCAATACTTGCAACTGCTCTGGC	60
ZmB73_Dof01	<b>GCGACTTTCATGATCCAAGC</b> ATATCTTTTCCCCTTTTCCAATACTTGCAACTGCTCTGGC	60
ZmW22_Dof01	<b>GCGACTTTCATGATCCAAGC</b> ATATCTTTTCCCCTTTTCCAATACTTGCAACTGCTCTGGC	60
	*****	
	<b>SBP      bHLH</b>	
ZmTEO_Dof01	TGGTGGCTGCCGCCCTGTCTCGTTAGAAACTTGCAGATAGTCTTAGTAGTACTGCAA	120
ZmPAL_Dof01	TGGTGGCTGCCGCCCTGTCTCGTTAGAAACTTGCAGATAGTCTTAGTAGTACTGCAA	120
ZmMo17_Dof01	TGGTGGCTGCCGCCCTGTCTCGTTAGAAACTTGCAGATAGTCTTAGTAGTACTGCAA	120
ZmB73_Dof01	TGGTGGCTGCCGCCCTGTCTCGTTAGAAACTTGCAGATAGTCTTAGTAGTACTGCAA	120
ZmW22_Dof01	TGGTGGCTGCCGCCCTGTCTCGTTAGAAACTTGCAGATAGTCTTAGTAGTACTGCAA	120
	*****	
	<b>ZF-HD</b>	
ZmTEO_Dof01	<b>GTGCAAGATATGGTCCGATTGTC</b> CCCTAGCTGCGGATTGTTGTTT <b>CTTAATCTA</b> -----	175
ZmPAL_Dof01	<b>GTGCAAGATATGGTCCGATTGTC</b> CCCTAGCTGCGGATTGTTGTTT <b>CTTAATCTA</b> -----	180
ZmMo17_Dof01	<b>GTGCAAGATATGGTCCGATTGTC</b> CCCTAGCTGCGGATTGTTGTTT <b>CTTAATCTA</b> -----	180
ZmB73_Dof01	<b>GTGCAAGATATGGTCCGATTGTC</b> CCCTAGCTGCGGATTGTTGTTT <b>CTTAATCTA</b> -----	180
ZmW22_Dof01	<b>GTGCAAGATATGGTCCGATTGTC</b> CCCTAGCTGCGGATTGTTGTTT <b>CTTAATCTA</b> -----	180
	*****	
	<b>MYB      MYB-related      P2 F</b>	
ZmTEO_Dof01	-- <b>GGAGACTAGGTTAGGTTAGCT</b> TACACATCATGGCAACCTACTTTTCCGTTTAGCTGGT	233
ZmPAL_Dof01	<b>TATGGAGGGTAGGTTAGGTTAGCT</b> TACACATCATGGCAA <b>CTACTTTTCCGTTAAGCTGGT</b>	240
ZmMo17_Dof01	<b>TATGGAGGGTAGGTTAGGTTAGCT</b> TACACATCATGGCAA <b>CTACTTTTCCGTTAAGCTGGT</b>	240
ZmB73_Dof01	<b>TATGGAGGGTAGGTTAGGTTAGCT</b> TACACATCATGGCAA <b>CTACTTTTCCGTTAAGCTGGT</b>	240
ZmW22_Dof01	<b>TATGGAGGGTAGGTTAGGTTAGCT</b> TACACATCATGGCAA <b>CTACTTTTCCGTTAAGCTGGT</b>	240
	*****	
	<b>AT-HOOK      TALE</b>	
ZmTEO_Dof01	TGG <b>TGCAATTGCA</b> AAAGGCT <b>AGACCT</b> TGGGGCTGTGTTT <b>TGCAACA</b> ACTAGTAGCCGTATG	293
ZmPAL_Dof01	TGGTTCAATTGCAAAAGGCT <b>AGACCT</b> TGGGGCTGCGTTT <b>TGCAACA</b> ACTAGTAGCCGTATG	300
ZmMo17_Dof01	TGGTTCAATTGCAAAAGGCT <b>AGACCT</b> TGGGGCTGCGTTT <b>TGCAACA</b> ACTAGTAGCCGTATG	300
ZmB73_Dof01	TGGTTCAATTGCAAAAGGCT <b>AGACCT</b> TGGGGCTGCGTTT <b>TGCAACA</b> ACTAGTAGCCGTATG	300
ZmW22_Dof01	TGGTTCAATTGCAAAAGGCT <b>AGACCT</b> TGGGGCTGCGTTT <b>TGCAACA</b> ACTAGTAGCCGTATG	300
	*****	
	<b>TALE      ARF/B3      G2-like</b>	
ZmTEO_Dof01	<b>ACTTACGAGGTTCAATT</b> TGCTCT <b>TGAACA</b> ACCGATGCTTCGGTTTT <b>TGTGAGA</b> ATATGC	353
ZmPAL_Dof01	<b>ACTTACGAGGTTCAATT</b> TGCTCT <b>TGAACA</b> ACCGATGCTTCGGTTTT <b>TGTGAGA</b> ATATGC	360
ZmMo17_Dof01	<b>ACTTACGAGGTTCAATT</b> TGCTCT <b>TGAACA</b> ACCGATGCTTCGGTTTT <b>TGTGAGA</b> ATATGC	360
ZmB73_Dof01	<b>ACTTACGAGGTTCAATT</b> TGCTCT <b>TGAACA</b> ACCGATGCTTCGGTTTT <b>TGTGAGA</b> ATATGC	360
ZmW22_Dof01	<b>ACTTACGAGGTTCAATT</b> TGCTCT <b>TGAACA</b> ACCGATGCTTCGGTTTT <b>TGTGAGA</b> ATATGC	360
	*****	
	<b>GATA      bZIP</b>	
ZmTEO_Dof01	AGAGAGAAGATCAGCTACTCGTAGTGGATAGGCGAATCCCTTCCCTCGCTT <b>GAGTACT</b>	413
ZmPAL_Dof01	AGAGAGAAGATCAGCTACTCGTAGTGGATAGGCGAATCCCTTCCCTCGCTT <b>GAGTACT</b>	419
ZmMo17_Dof01	AGAGAGAAGATCAGCTACTCGTAGTGGATAGGCGAATCCCTTCCCTCGCTT <b>GAGTACT</b>	419
ZmB73_Dof01	AGAGAGAAGATCAGCTACTCGTAGTGGATAGGCGAATCCCTTCCCTCGCTT <b>GAGTACT</b>	419
ZmW22_Dof01	AGAGAGAAGATCAGCTACTCGTAGTGGATAGGCGAATCCCTTCCCTCGCTT <b>GAGTACT</b>	419
	*****	
	<b>Dof</b>	
ZmTEO_Dof01	<b>CATTAAAGCCA</b> TTTCTTGATGTGGCTACCCGTCGGTTC <b>TGTGGA</b> AGTAGTGT <b>TTCTAAA</b>	473
ZmPAL_Dof01	<b>CATTAAAGCCA</b> TTTCTTGATGTGGCTACCCGTCGGTTC <b>TGTGGA</b> AGTAGTGT <b>TTCTAAA</b>	479
ZmMo17_Dof01	<b>CATTAAAGCCA</b> TTTCTTGATGTGGCTACCCGTCGGTTC <b>TGTGGA</b> AGTAGTGT <b>TTCTAAA</b>	479
ZmB73_Dof01	<b>CATTAAAGCCA</b> TTTCTTGATGTGGCTACCCGTCGGTTC <b>TGTGGA</b> AGTAGTGT <b>TTCTAAA</b>	479
ZmW22_Dof01	<b>CATTAAAGCCA</b> TTTCTTGATGTGGCTACCCGTCGGTTC <b>TGTGGA</b> AGTAGTGT <b>TTCTAAA</b>	479
	*****	
	<b>SBP</b>	
ZmTEO_Dof01	CAGCTTGGATTAGGTAGAGCTCTATTCTATCCCTTGTC <b>CAAGTTAGCCTTGTGATGCAGA</b>	533
ZmPAL_Dof01	CAGCTTGGG <b>TACGTAGGGCTCT</b> ATTCTATCCCTTGTC <b>CAAGTTAGCCTTGTGATGCAGA</b>	538
ZmMo17_Dof01	CAGCTTGGG <b>TACGTAGGGCTCT</b> ATTCTATCCCTTGTC <b>CAAGTTAGCCTTGTGATGCAGA</b>	538
ZmB73_Dof01	CAGCTTGGG <b>TACGTAGGGCTCT</b> ATTCTATCCCTTGTC <b>CAAGTTAGCCTTGTGATGCAGA</b>	538
ZmW22_Dof01	CAGCTTGGG <b>TACGTAGGGCTCT</b> ATTCTATCCCTTGTC <b>CAAGTTAGCCTTGTGATGCAGA</b>	538
	*****	

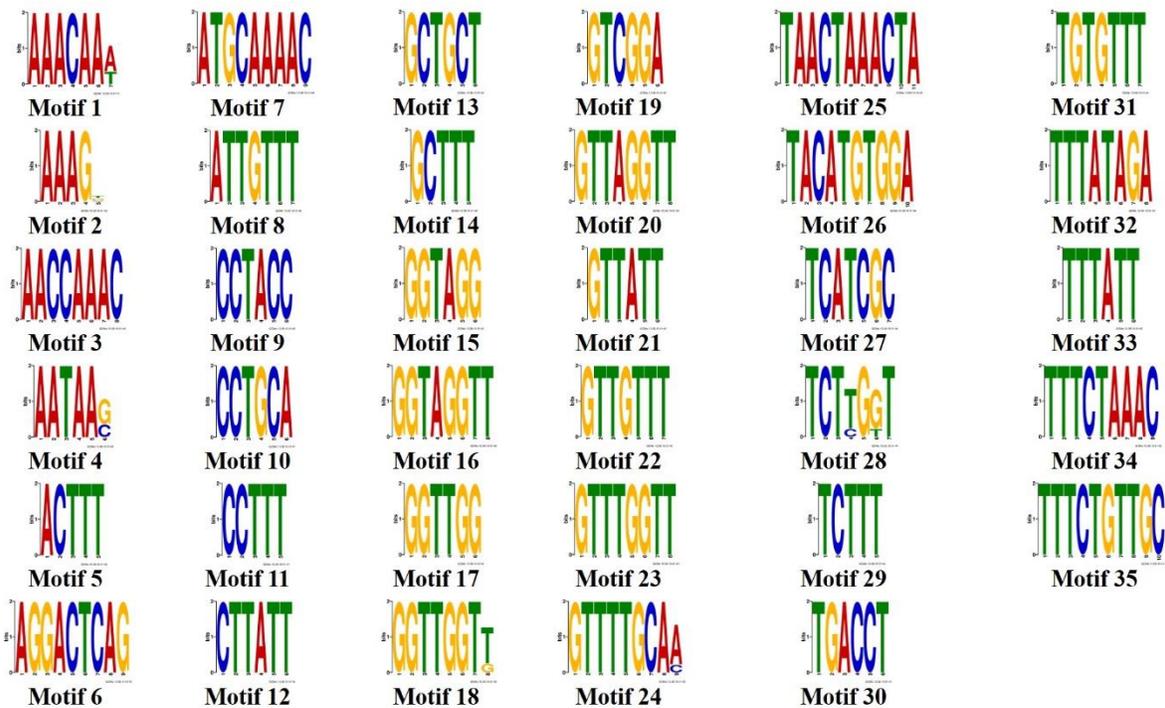
	<b>AT-HOOK</b>	<b>ERF/RAV</b>				
ZmTEO_Dof01	ATTGCCAAATTTCTTTCTGCTGCAAGTTTTTATTTTGTAGTTTCTTGATTG---GGTG		590			
ZmPAL_Dof01	ATTGCAAAATGCTTTTCTGTTGCAA GTTTTTATTTTGTAGTTTCTTGATTGGTGGGTG		598			
ZmMo17_Dof01	ATTGCAAAATGCTTTTACTGTTGCAAGTTTTTATTTTGTAGTTTCTTGATTG---GGTG		595			
ZmB73_Dof01	ATTGCAAAATGCTTTTCTGTTGCAA GTTTTTATTTTGTAGTTTCTTGATTG---GGTG		595			
ZmW22_Dof01	ATTGCAAAATGCTTTTCTGTTGCAA GTTTTTATTTTGTAGTTTCTTGATTG---GGTG		595			
	***** **					
	<b>MYB/AT-HOOK</b>					
ZmTEO_Dof01	TTGGTAGCCAGTTTGGATAAAATGATGGGGATCCTCTATAAGTTTTTAACCTTGATGC		650			
ZmPAL_Dof01	TTGGTAGCCAGTTTGGATAGGTTGATGGGGATCCTCTATAAGTTTTTAACCTTGATGC		658			
ZmMo17_Dof01	TTGGTAGCCAGTTTGGATAGGTTGATGGGGATCCTCTATAAGTTTTTAACCTTGATGC		655			
ZmB73_Dof01	TTGGTAGCCAGTTTGGATAGGTTGATGGGGATCCTCTATAAGTTTTTAACCTTGATGC		655			
ZmW22_Dof01	TTGGTAGCCAGTTTGGATAGGTTGATGGGGATCCTCTATAAGTTTTTAACCTTGATGC		655			
	***** **					
	<b>HD-ZIP</b>	<b>AT-HOOK</b>	<b>bZIP</b>	<b>AT-HOOK</b>		
ZmTEO_Dof01	TATGATTTCGCGGAGCCAGAAAAATTTGAGACCATA-----CATTTGGATTGGGTTGAATA		704			
ZmPAL_Dof01	TATGATTTCGCGGAGCCAAAAATTTGAGACCACCTCAGTACATGTGGATTGGGGTGAATA		718			
ZmMo17_Dof01	TATGATTTCGCGGAGCCAAAAATTTGAGACCACCTCAGTACATGTGGATTGGGGTGAATA		715			
ZmB73_Dof01	TATGATTTCGCGGAGCCAAAAATTTGAGACCACCTCAGTACATGTGGATTGGGGTGAATA		715			
ZmW22_Dof01	TATGATTTCGCGGAGCCAAAAATTTGAGACCACCTCAGTACATGTGGATTGGGGTGAATA		715			
	***** **					
	<b>Dof</b>		<b>TALE</b>			
ZmTEO_Dof01	TTAGAGTACCCAAACAAAGCCTAAAGGTATTTGTGACATCTTGGTTTGGATCCACATGCAG		764			
ZmPAL_Dof01	CTAGAGTGCCCAACAATAATCCTAAAGTACTATGAAATCTTGGTTTGGATCCACATGCAG		778			
ZmMo17_Dof01	CTAGAGTGCCCAACAATAATCCTAAAGTACTATGAAATCTTGGTTTGGATCCACATGCAG		775			
ZmB73_Dof01	CTAGAGTGCCCAACAATAATCCTAAAGTACTATGAAATCTTGGTTTGGATCCACATGCAG		775			
ZmW22_Dof01	CTAGAGTGCCCAACAATAATCCTAAAGTACTATGAAATCTTGGTTTGGATCCACATGCAG		775			
	***** **					
	<b>P3 F</b>		<b>HD-ZIP HD/WOX</b>			
ZmTEO_Dof01	TTATGGCACCTTGGATACAGTTGTTCTGAGCATAGATTTGATTGAGGTCGGATTTCCCA		824			
ZmPAL_Dof01	TTATGGTACCTTGGATACAGTTGTTCTGAGCATGGATTTGATTGAGGTCGAAATTTCCCA		838			
ZmMo17_Dof01	TTATGGTACCTTGGATACAGTTGTTCTGAGCATGGATTTGATTGAGGTCGAAATTTCCCA		835			
ZmB73_Dof01	TTATGGTACCTTGGATACAGTTGTTCTGAGCATGGATTTGATTGAGGTCGAAATTTCCCA		835			
ZmW22_Dof01	TTATGGTACCTTGGATACAGTTGTTCTGAGCATGGATTTGATTGAGGTCGAAATTTCCCA		835			
	***** **					
	<b>AT-HOOK</b>	<b>trp</b>	<b>HD-ZIP</b>			
ZmTEO_Dof01	GAATACTGAAACAAAACGGGCCATACCACATGCGGGCTACTAGAGTTAATGATCCTATGC		884			
ZmPAL_Dof01	AAATACTGAAACAAAACGGGCCATACCACATGCGGGCTACTAGAGTTAATGATCCTAT--		896			
ZmMo17_Dof01	AAATACTGAAACAAAACGGGCCATACCACATGCGGGCTACTAGAGTTAATGATCCTAT--		893			
ZmB73_Dof01	AAATACTGAAACAAAACGGGCCATACCACATGCGGGCTACTAGAGTTAATGATCCTAT--		893			
ZmW22_Dof01	AAATACTGAAACAAAACGGGCCATACCACATGCGGGCTACTAGAGTTAATGATCCTAT--		893			
	***** **					
	<b>TALE</b>		<b>C2H2</b>		<b>AT-HOOK</b>	
ZmTEO_Dof01	TATGACAAGGAATAACAACCATAAGATGTAACACTCAAAGTACTAAACGTAAAAAGTAC		944			
ZmPAL_Dof01	---GACAAAGGAATAACAACCATAAGATGTAACACTCAAAGTACTAAACGTAAAAAGTAC		953			
ZmMo17_Dof01	---GACAAAGGAATAACAACCATAAGATGTAACACTCAAAGTACTAAATGTA AAAAGTAC		950			
ZmB73_Dof01	---GACAAAGGAATAACAACCATAAGATGTAACACTCAAAGTACTAAATGTA AAAAGTAC		950			
ZmW22_Dof01	---GACAAAGGAATAACAACCATAAGATGTAACACTCAAAGTACTAAACGTAAAAAGTAC		950			
	***** **					
	<b>SBP</b>	<b>MYB</b>	<b>EIN3</b>			
ZmTEO_Dof01	AGTGTAAACAAAAAGAAGGATATTACATCGTGCATATGGCTATACAAAACAATGTTCTGT		1004			
ZmPAL_Dof01	AGTGTAAACAAAGCAGAAGTTATTACATCGTGCATATGGCTATGCAAAAACAATGTTTAC		1013			
ZmMo17_Dof01	AGTGTAAATCAAGCAGAAGTTATTACATCATGCATATGGCTATGCAAAAACAATGTTTGC		1010			
ZmB73_Dof01	AGTGTAAACAAAGTAGAAGTTATTACATCGTGCATATGGCTATGCAAAAACAATGTTTAC		1010			
ZmW22_Dof01	AGTGTAAACAAAGTAGAAGTTATTACATCGTGCATATGGCTATGCAAAAACAATGTTTAC		1010			
	***** **					
	<b>C2H2</b>					
ZmTEO_Dof01	ATTGTTTAC-----AGAGTGGAGTCTAAAATGAGATACAAAGATGAGTAAGCT		1052			
ZmPAL_Dof01	ACTATAGACTACATGTTTATAGAGTGAAGTTTAAAATGAGATATGAGGATGAGTAAGCT		1073			
ZmMo17_Dof01	ACTATAGACTACATGTTTATACATAGTGGAGTTTAAAATGAGATACGAGGATGAGTAAGCT		1070			
ZmB73_Dof01	ACTATAGACTACATGTTTATAGAGTGAAGTTTAAAATGAGATATGAGGACGAGTAAGCT		1070			
ZmW22_Dof01	ACTATAGACTACATGTTTATAGAGTGAAGTTTAAAATGAGATATGAGGACGAGTAAGCT		1070			
	* * * * *					



	EIN3	
ZmTEO_Dof01	GAGTACTTGCTTGGCTCCTTGGATTTCTAAAGTTCTGTTATTTGCCAGGATTTCTGAAGT	1462
ZmPAL_Dof01	GAGTACTTGCTTGGCTCCCAGGATTTCTAAAGTTCTGTTATTTGCCAGTA-----	1469
ZmMo17_Dof01	GAGTACTTGCTTGGCTCCTTGGATTTCTAAAGT-----	1450
ZmB73_Dof01	<u>ATGTACTTGCTTGGCTCCTTGGATTTCTAAAGTTCTGTTATTTGCTAGGATTTCTAAAGT</u>	1660
ZmW22_Dof01	<u>ATGTACTTGCTTGGCTCGTTGGATTTCTAAAGTTCTGTTATTTGCTAGGATTTCTAAAGT</u>	1482
	*****	
	MYB/SANT	
ZmTEO_Dof01	TCTGTTATTTGCCAGTAAATGTCATAGACATATTTTACAGCATCTTACAATCATATTTCC	1522
ZmPAL_Dof01	-----AA--TGTGACACATATTTTACAGCATCTTACAATCATATTTCC	1510
ZmMo17_Dof01	TATGTTATTTGCCAGTAAATGTCATAGACATATTTTACATCATCTTACAATCATATTTCC	1510
ZmB73_Dof01	TCTGTTATTTGCCAGTAAATGTAATAGACATATTTTACATCATCTTACAATCATATTTCC	1720
ZmW22_Dof01	TCTGTTATTTGCCAGTAAATGTAATAGACATATTTTACATCATCTTACAATCATATTTCC	1542
	* * *	
	bZIP/EIN3/NAC	
	EIN3	
ZmTEO_Dof01	TTCTATCTCAATACATGTATTTACTTAGTCTTCGTTATTTGTCAGGATGTATGTGAAG	1582
ZmPAL_Dof01	TTCTATCTCAATACATGTATTTACTTAGTCTTCGTTATTTGTCAGGATGTATGTGAAG	1570
ZmMo17_Dof01	TTCTATCTCAATACATGTATTTACTTAGTCTTCGTTATTTGTCAGGATGTATGTGAAG	1570
ZmB73_Dof01	TTCTATCTCAATACATGTATTTACTTAGTCTTCGTTATTTGTCAGGATGTATGTGAAG	1780
ZmW22_Dof01	TTCTATCTCAATACATGTATTTACTTAGTCTTCGTTATTTGTCAGGATGTATGTGAAG	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
	*****	

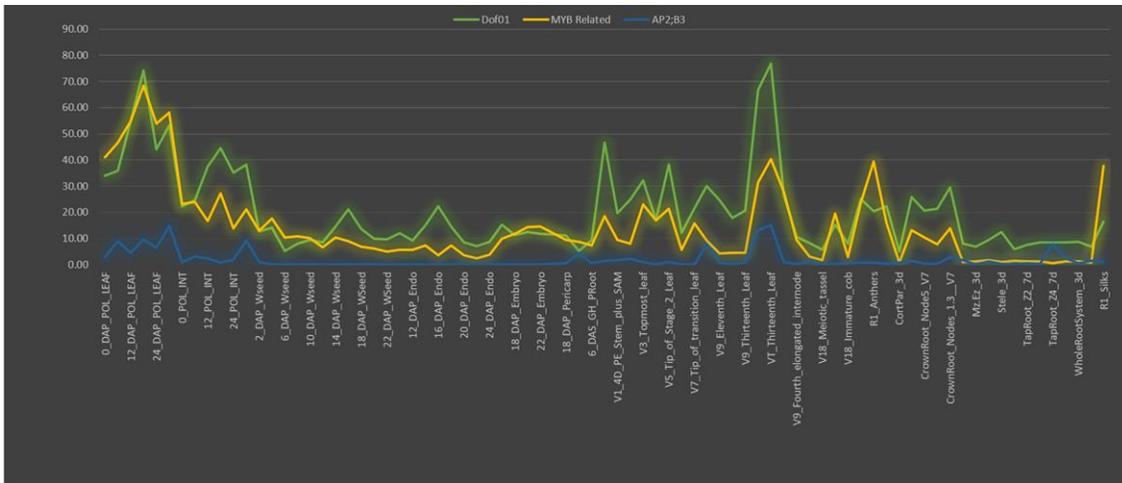
**Figure S1.** Multiple alignment of the *Dof01* promoter sequence of the five maize varieties. Text highlighted in gray indicates the location of the motifs that are recognized by transcription factors. The asterisks below each line indicate that the position of each nucleotide is conserved in all five 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.

**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 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 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 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.

**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.