



**UNIVERSIDAD NACIONAL
DE MÉXICO**

INSTITUTO DE BIOTECNOLOGÍA

Estudio de la función de la familia 4 de proteínas
LEA de *Arabidopsis thaliana* en la respuesta a sequía

Tesis que para obtener el grado de
Doctora en Ciencias

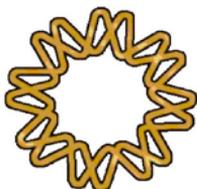
presenta

M. en C. Yadira Olvera Carrillo

Directora de tesis:

Dra. Alejandra Alicia Covarrubias Robles

Diciembre 2010





Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

A los miembros del Comité Tutorial, por sus críticas y comentarios durante el desarrollo experimental de este trabajo de tesis:

Dra. Martha Patricia Coello Cutiño
Dr. Mario Rocha Sosa
Dra. Alejandra A. Covarrubias Robles

A los invitados en el Comité Tutorial Ampliado:

Dra. Martha Vázquez Laslop
Dra. Tzvetanka Dimitrova Dinkova
Dr. Jorge Nieto Sotelo
Dr. Felipe Cruz García

A los miembros del jurado del examen para la obtención del grado, por sus valiosos comentarios sobre el trabajo experimental y las correcciones en el manuscrito de la tesis,

PRESIDENTE	Dra. Gladys Iliana Cassab López
SECRETARIO	Dra. Martha Vázquez Laslop
VOCAL	Dra. Tzvetanka Dimitrova Dinkova
SUPLENTE	Dr. Felipe Cruz García
SUPLENTE	Dra. Alejandra Alicia Covarrubias Robles

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por otorgarme la Beca de Doctorado (No. de Registro 169924), a la DGPA por otorgarme la beca para escribir la tesis y concluir el trabajo experimental (Clave IN222309) y al Sistema Nacional de Investigadores (SNI) por otorgarme el apoyo de Ayudante de Investigador .

Al Dr. Jian-Kan Zhu por aceptarme en su laboratorio de la Universidad de California en Riverside, durante la estancia de investigación que realicé como parte del programa de Becas Mixtas otorgada por CONACYT, durante mi último año de estudios del Doctorado.

Al Dr. Felipe Cruz por su asesoría en el análisis estadístico de los resultados obtenidos en los experimentos fenotípicos con plantas transgénicas, que me ayudó mucho para la publicación del artículo.

A mis compañeros del laboratorio de todos estos años, los primeros que ya se fueron del lab: Sonia, Marina, Caty, Erickita, Erika, Fernando LL. A los nuevos: Cecy, Lucero, Miguel R, Miguel P, Rocío y los genomikids Carlos, Fernando R, Daniel, Mine, Arturo (que le sacaron las primeras canas a Pepe!). A mis alumnos: César, Gloria Amparo y Coral, por continuar con mis proyectos y seguir mis consejos. A los que siempre estuvieron en el lab: Rosy, Pepe (bueno mientras no estaba de post doc!), Pancho, Adrianita y Doña Chuy. A todos ellos por los momentos que hemos pasado juntos, así como por su disposición a ayudarme y por hacer que el laboratorio funcione en un ambiente agradable.

A mi familia, mis amigos Daniela, Ángeles, Sergio, Paul, y todos mis demás compañeros del Instituto de Biotecnología y de la Universidad de California por darme ánimos para continuar en el arduo camino del Doctorado haciendo mi vida más feliz y relajada, por compartir conmigo momentos muy amenos y apoyarme siempre en momentos tristes y difíciles.

A la UNAM por haber sido mi casa durante todos estos años: por haberme formado desde los estudios de Bachillerato en la Escuela Nacional Preparatoria 8 "Miguel E. Schulz" donde me enamoré de la Biología y continué en la carrera de Biología en la Facultad de Ciencias de Ciudad Universitaria, el campus más increíble de la Universidad. Finalmente, en mi formación en el Posgrado, durante los estudios de Maestría y Doctorado en el Instituto de Biotecnología. Gracias por haber hecho de mí lo que soy ahora. Siempre llevaré en mi corazón el recuerdo de mi Universidad!

Esta tesis de Doctorado se llevó a cabo dentro del Programa de Doctorado en Ciencias Bioquímicas en el Departamento de Biología Molecular de Plantas del Instituto de Biotecnología de la UNAM, bajo la asesoría de la Dra. Alejandra A. Covarrubias Robles, a quien quiero dar un agradecimiento muy especial, reconozco y admiro su apoyo, sus enseñanzas y su comprensión.

RESÚMEN

Dentro de los diferentes tipos de estrés que las plantas enfrentan durante su ciclo de vida, el déficit hídrico causa grandes pérdidas en los cultivos de importancia económica. Las proteínas LEA (por Late Embryogenesis Abundant) son proteínas altamente hidrofílicas que se expresan en los tejidos vegetativos y reproductivos de plantas sujetas a déficit hídrico, así como en la etapa de maduración y secado de las semillas. Se ha propuesto que actúan como protectores de la maquinaria celular y están involucradas en la tolerancia de las plantas ante estrés por déficit hídrico. En esta tesis se caracterizaron los patrones de expresión de transcritos y proteínas AtLEA4 en plantas de diferentes etapas del desarrollo bajo condiciones control y déficit hídrico. Se llevó a cabo un análisis del grupo LEA4 en la respuesta a sequía, utilizando un enfoque por pérdida de función. Se analizaron fenotípicamente plantas transgénicas homocigotas con disminución en los niveles de proteína de los miembros que conforman el grupo LEA4 en plantas de *Arabidopsis thaliana*, encontrando fenotipos de sensibilidad durante la germinación en condiciones control y en estrés iónico/osmótico; en plantas adultas bajo tratamientos de sequía, mostraron una menor recuperación de biomasa total y un menor número de botones florales. Aún en condiciones de irrigación óptima se encontró una menor producción total de semillas por planta en las mutantes, por lo que se propone que las proteínas LEA4 pueden tener también un papel en el desarrollo del fruto o la semilla. Se llevó a cabo un análisis filogenético de proteínas LEA4 de diferentes géneros vegetales y se subclasificaron en dos subgrupos, que se originaron a partir de una duplicación génica y hubo una posterior especialización funcional. Se concluye que las proteínas LEA4 están involucradas en el proceso de adaptación de las plantas superiores ante déficit hídrico y que la función de estas proteínas no es redundante con otros grupos de proteínas LEA.

Palabras clave: proteínas LEA, adaptación al estrés abiótico, sequía.

INDICE

CAPITULO 1

INTRODUCCIÓN

I.	El estrés por déficit hídrico y su efecto en las plantas.	7
II.	El agua y su importancia en las relaciones hídricas de la célula vegetal.	12
III.	Evolución de la tolerancia a la desecación.	14
IV.	Regulación de la expresión genética en respuesta a déficit hídrico.	17
	IV.1 El ácido abscísico (ABA) y su importancia como señalizador en plantas.	
	IV.2 Activación transcripcional de genes en respuesta a déficit hídrico.	
	IV.3 Regulación a nivel post-transcripcional en la respuesta a déficit hídrico.	
V.	Proteínas protectoras de la maquinaria celular.	26
	V.1 Proteínas de choque térmico (HSP).	
	V.2 Proteínas de embriogénesis tardía (LEA).	
	V.3 Clasificación de las proteínas LEA.	
	V.4 Sobre-expresión de genes que codifican para proteínas tipo LEA.	
VI.	La genética reversa: una herramienta en el análisis funcional de genes de respuesta a déficit hídrico.	33

ANTECEDENTES	35
---------------------	-----------

DESCRIPCIÓN DE LA ESPECIE EN ESTUDIO	39
---	-----------

MATERIALES Y MÉTODOS	40
-----------------------------	-----------

I.	Análisis bioinformático de la familia de proteínas LEA4: clasificación, motivos conservados y predicción de estructura secundaria.
II.	Material vegetal y condiciones de crecimiento.
III.	Tratamientos de estrés abiótico durante la germinación, en plántulas de dos semanas, y en plantas adultas de 6 a 7 semanas de edad.
IV.	Construcciones, vectores, cepas bacterianas y transformación de plantas.
V.	Extracción de DNA genómico y experimentos tipo Southern.
VI.	Extracción de RNA y experimentos tipo northern y RT-PCR.
VII.	Extracción de proteína y experimentos tipo western.
VIII.	Obtención de anticuerpos dirigidos contra las proteínas AtLEA4-1 y AtLEA4-2.

HIPÓTESIS	57
------------------	-----------

OBJETIVOS	57
------------------	-----------

RESULTADOS	58
-------------------	-----------

I.	Análisis estructural y filogenético del grupo LEA4: estructura secundaria, motivos conservados y sub-clasificación.
----	---

- II. Análisis de los patrones de acumulación de transcritos y proteínas de la familia AtLEA4 durante la embriogénesis y el establecimiento de la plántula, así como en tejidos vegetativos y reproductivos sujetos a déficit hídrico.
- III. Sensibilidad de plantas mutantes en los genes *AtLEA4* en la respuesta a estrés iónico/osmótico durante la germinación y en tratamientos de sequía en plantas adultas.

DISCUSION	79
CONCLUSIONES Y PERSPECTIVAS	80

CAPITULO 2

INTRODUCCIÓN

- I. Análisis estructura-función en las proteínas LEA. **88**
- II. Las interacciones proteína-proteína y el estudio de las proteínas desordenadas: el caso de las proteínas LEA.

MATERIAL Y MÉTODOS **95**

- I. Material vegetal y tratamientos de sequía con plantas de tres semanas de edad.
- II. Obtención de plantas transgénicas que sobre-expresan la proteína de fusión AtLEA4-5-TAPa.
- III. Espectrometría de masas.
- IV. Ensayos de dos híbridos en levadura.

RESULTADOS Y DISCUSIÓN **99**

- I. La proteína AtLEA4-5 tiene interacciones débiles con otras proteínas y los dominios de interacción pueden estar relacionados a la estructura secundaria de las proteínas interaccionantes.
- II. La purificación de complejos proteicos por TAP ("Tandem Affinity Purification") mostró una variedad de proteínas localizadas en diferentes compartimentos celulares que pueden interaccionar con la AtLEA4-5.

REFERENCIAS

116

INTRODUCCIÓN

I. EL ESTRÉS POR DÉFICIT HÍDRICO Y SU EFECTO EN LAS PLANTAS

Se puede considerar al estrés como cualquier factor externo que ejerce una influencia negativa en un organismo vivo. Puede ser biótico, cuando se debe a la influencia de otros organismos, o abiótico cuando surge de un exceso o un déficit en las condiciones físicas o químicas del ambiente. El estrés desencadena una amplia variedad de respuestas en las plantas, tales como cambios en la tasa de crecimiento, alteraciones en la expresión genética y en el metabolismo, entre otros (Bray *et al.*, 2000). El estrés por déficit hídrico en las plantas ocurre cuando el aporte de agua desde el suelo es menor que la tendencia de la planta a perder agua por evaporación-transpiración (Hall, 1993). Éste es uno de los tipos de estrés abiótico más comunes que las plantas enfrentan durante su ciclo de vida y tiene un gran impacto en su desarrollo, crecimiento y productividad. El déficit hídrico se puede producir por la sequía, pero también existen otros tipos de estrés que provocan el mismo efecto de deshidratación en las células, como el congelamiento o el estrés osmótico que podría ser provocado por una alta concentración de sales en el medio (Bray, 1997).

Los efectos del déficit hídrico se pueden medir en relación a la planta completa por la disminución en su tasa de crecimiento (acumulación de biomasa) o en procesos primarios de asimilación, como la fijación de CO₂, entre otros parámetros (Taiz y Zeiger, 1991). Los primeros efectos en algunas plantas son la inhibición temporal del crecimiento de los órganos aéreos (hojas y tallos) y el cierre de estomas para disminuir la pérdida de agua por evaporación-transpiración. Cuando la deshidratación es más severa se empiezan a afectar los componentes celulares: hay daños en las estructuras subcelulares, el volumen de la célula y la forma de la membrana plasmática cambian, ocurren cambios en la concentración de solutos, que a su vez provocan una ruptura de los gradientes de potencial hídrico, lo que causa pérdida de la turgencia y de la integridad de las membranas y la desnaturalización de ciertas proteínas (Ingram y Bartels, 1996; Bray *et al.*, 2000).

El déficit hídrico también es un componente normal de algunos procesos del desarrollo, como lo es en el caso de la mayoría de las plantas con flores, cuyas semillas pasan por un proceso de desecación en la última etapa de su maduración al cual son resistentes (más de un 90% de pérdida de agua) para promover una reproducción exitosa en la siguiente generación (Bray, 1997). El estudio de los mecanismos que se han seleccionado para sobrevivir a la desecación en esta etapa de desarrollo de las plantas es de fundamental importancia, ya que la semilla es la estructura protectora, portadora y dispersora del material genético de la mayoría de las plantas cultivadas, y ocupa una posición crítica dentro del ciclo de vida de las plantas superiores, ya que sus características fisiológicas y bioquímicas determinan en gran parte el éxito en el establecimiento de un individuo de la siguiente generación (Bewley y Black, 1994).

Las respuestas de las plantas a los diferentes tipos de estrés ambiental son diversas y complejas, aunque en general los efectos son deletéreos y resultan en una disminución en el número o la calidad de las semillas. Algunas veces pueden ocurrir diferentes tipos de estrés de manera simultánea (por ejemplo, sequía y temperaturas elevadas), y el efecto de éstos en el desarrollo de la semilla puede ser mayor para la suma que para cada estrés por separado. La duración y severidad del estrés y la etapa de desarrollo de la planta en la cual ocurre, son parámetros que deben considerarse al analizar los efectos del déficit hídrico en la respuesta de la planta (Bewley y Black, 1994).

Las plantas se pueden clasificar, de acuerdo a la capacidad para mantener su desarrollo en relación a la disponibilidad de agua en el ambiente, en **mesófitas**, que son las que requieren un aporte continuo y constante de agua para su desarrollo óptimo, pero que han desarrollado mecanismos para tolerar condiciones temporales de estrés; los cultivos de interés agronómico se encuentran en esta categoría. Las plantas **hidrófitas** son las que viven sumergidas en suelos saturados de agua, o que crecen sobre la superficie del agua. Por otra parte, las plantas **xerófitas** se han adaptado a un suministro casi nulo de agua o a variaciones grandes en su disponibilidad, como de hecho puede ocurrir a lo largo de las diferentes estaciones del año. Este grupo es muy diverso y posee mecanismos que promueven un uso eficiente del agua para mantener un crecimiento mínimo, ya que la supervivencia en estas condiciones extremas se da a cambio de una baja tasa de crecimiento (Nobel, 1994; Dickinson, 2000).

Las plantas son organismos sésiles que han evolucionado mediante la selección de un conjunto de características que les permiten percibir, responder, tolerar y adaptarse a los cambios del ambiente, en donde adaptación se refiere a cualquier carácter o conjunto de

caracteres que permiten que los individuos que los poseen, se ajusten y superen exitosamente uno o más agentes selectivos (Begon *et al.*, 1996). Durante el proceso evolutivo, las modificaciones de un carácter bajo presión selectiva, tienen como consecuencia una ventaja en el ambiente que ejerció la presión de selección, con respecto a otras poblaciones de la misma u otra especie que carecen de ellas. Las adaptaciones pueden ser morfológicas, histológicas, fisiológicas, metabólicas o de comportamiento, y a nivel molecular representan mutaciones heredables que aparecen en una parte de la población (Begon *et al.*, 1996).

Dentro de la respuesta adaptativa a la falta de agua en las plantas, se han seleccionado varios mecanismos para contender con este tipo de estrés. El **escape** al déficit hídrico se presenta en las plantas que se desarrollan rápidamente, completan su ciclo de vida en la estación lluviosa y producen semillas antes del periodo de sequía, como en el caso de algunas plantas anuales que habitan en zonas desérticas y semidesérticas. Las especies que **evaden** la sequía son las que pueden compensar la pérdida de agua a través de adaptaciones como un sistema radicular muy extenso y/o que alcanza grandes profundidades. Por otra parte, las plantas que **resisten** la falta de agua, son capaces de sobrevivir a través de las adaptaciones morfológicas y fisiológicas como: la modificación del ángulo de la hoja, la succulencia en las hojas de las crasuláceas, la reducción de las hojas a espinas en las cactáceas, la formación de cutículas gruesas, el aumento en la cantidad de tricomas, el uso del metabolismo ácido de las crasuláceas, entre otros (Ramanjulu y Bartels, 2002).

La tolerancia a la sequía es la capacidad de algunas plantas de mantener su desarrollo en condiciones de déficit hídrico; es el resultado de un conjunto de alteraciones fisiológicas y bioquímicas a nivel celular y molecular, como cambios en las propiedades físicas de la pared celular, el ajuste osmótico, la síntesis e incremento de la actividad de enzimas y metabolitos antioxidantes, así como la síntesis de proteínas protectoras de la maquinaria celular (Ingram y Bartels, 1996). El ajuste osmótico se refiere a la acumulación, ante situaciones de déficit hídrico, de compuestos orgánicos solubles de bajo peso molecular que no interfieren con el metabolismo, conocidos como solutos compatibles u osmolitos; entre éstos se encuentran aminoácidos (prolina), azúcares/polioles (manitol, pinitol) y aminas cuaternarias (glicina betaína, dimetilsulfoniopropionato). El resultado es un incremento de solutos en la célula para disminuir el potencial hídrico de los tejidos de la planta, principalmente de las hojas, lo cual mantiene un gradiente favorable para la toma de agua, protegiendo así el turgor de

las células (Kishor *et al.*, 1995; Sakamoto *et al.*, 1998; Thomas *et al.*, 1995; Tarczynski *et al.*, 1993; Hayashi *et al.*, 1997; Garg *et al.*, 2002). Además del ajuste osmótico, también se ha propuesto que éstos compuestos podrían estabilizar proteínas y/o membranas, mediante la disminución en la concentración de radicales libres y la eliminación de iones tóxicos (Wang *et al.*, 2003). Una consecuencia de muchos tipos de estrés ambiental, incluyendo la sequía, es el estrés oxidativo, debido a la acumulación de especies reactivas de oxígeno (EROs) (O_2 , H_2O_2 y OH^\cdot) que dañan membranas y macromoléculas (Smirnov, 1998). En condiciones de estrés oxidativo se ha observado la acumulación de enzimas antioxidantes, como catalasa, superóxido dismutasa, ascorbato peroxidasa, glutatión-S-transferasas y glutatión peroxidasa, y de moléculas como el glutatión, ascorbato, carotenoides y antocianinas (McKersie *et al.*, 1996; van Camp *et al.*, 1996; Roxas *et al.*, 1997; Oberschall *et al.*, 2000). Su mecanismo de protección consiste en un incremento en el secuestro de radicales libres para proteger a los cloroplastos, ya que la maquinaria fotosintética localizada en las membranas y el lumen del tilacoide, es muy sensible a la deshidratación (Wang *et al.*, 2003).

En déficit hídrico hay disminución en la tasa de crecimiento, debido a que la expansión celular disminuye para tener una menor superficie de evaporación (relación área-volumen) en la parte aérea, mientras que la parte subterránea continúa su crecimiento al desviar los fotosintatos a la raíz para favorecer la búsqueda de agua en el suelo durante la sequía. Si la planta se encuentra en floración, se da preferencia a la translocación de carbono y nitrógeno al fruto y al desarrollo de la semilla (Bewley y Black, 1994). La modificación de los componentes de la pared puede ser responsable de que se detenga la expansión celular, lo cual se ha verificado en plantas de maíz sometidas a estrés osmótico (Chazen y Neumann, 1994). En plantas de frijol sometidas a tratamientos de deshidratación y ácido abscísico (ABA) se observó la acumulación de proteínas básicas de peso molecular aparente de 22, 33 y 36 kDa (Covarrubias *et al.*, 1995). Las dos últimas proteínas (denominadas p33 y p36) están relacionadas a proteínas ricas en prolina y también se ha reportado su interacción con una proteína de 80 kDa de la pared celular, análoga a las integrinas de animales, lo cual sugiere que la función de p33 y p36 es facilitar la comunicación entre células y/o proveer soporte físico a las células durante el déficit hídrico (García-Gómez *et al.*, 2000).

Las transferasas no específicas de lípidos (LTPs) son proteínas solubles de bajo peso molecular, secretadas a la pared celular y que se piensa están implicadas en la síntesis de cutícula a través de la secreción y deposición de sustancias lipofílicas desde las

células epidérmicas en expansión (Kader, 1996). Algunos ejemplos de LTPs que se inducen por estrés son la TSW12 de tomate, en respuesta a sal (Torres-Schuman *et al.*, 1992) y PvLTP-24 de frijol, en respuesta a sequía y ABA (Colmenero-Flores *et al.*, 1997).

La actividad proteolítica durante condiciones de estrés es elevada y coincide con el proceso de muerte celular programada. La actividad de proteasas y el proteasoma 26S puede estar involucrada en la degradación de proteínas desnaturalizadas por la deshidratación. Por otra parte, se ha reportado la inducción de inhibidores de proteasas específicamente por deshidratación, como el caso de WSCP de coliflor, que puede proteger a ciertas proteínas de la actividad proteolítica que se presenta ante situaciones de estrés y de muerte celular programada (Nishio y Satoh, 1997). Por otra parte, las proteasas inducidas por sequía pueden alterar el metabolismo al incrementar la tasa de recambio de ciertas proteínas, degradando polipéptidos desnaturalizados por el estrés o proteínas de reserva. Otro daño causado por la falta de agua es la modificación química de los residuos de proteínas, como la desaminación y la oxidación, que pueden causar la pérdida de la actividad o la conformación nativa de las proteínas. En plantas de trigo sujetas a déficit hídrico se ha observado un incremento de la síntesis y actividad de la L-isoaspartil metiltransferasa que repara los residuos dañados de L-isoaspartil a residuos funcionales (Mudgett y Clarke, 1994).

La sequía y la alta concentración de sales en el suelo alteran la homeostasis iónica debido al aumento en la concentración de iones de sodio (Na^+) intracelular, lo cual es altamente tóxico para muchas enzimas. En las plantas que están en proceso de aclimatación a condiciones de deshidratación y salinidad, se ha establecido que el mecanismo que utilizan es la excreción de Na^+ de la raíz y la compartimentalización de este catión en la vacuola de las células de la hoja. En *Arabidopsis thaliana* la entrada de Na^+ está controlada por el transportador HKT1, mientras que la expulsión de Na^+ la llevan a cabo los antiportadores Na^+/H^+ , como el de membrana plasmática denominado SOS1 (por Salt Overly Sensitive) (Shi *et al.*, 2000) o el de la membrana de tonoplasto, conocido como AtNHX1 (Gaxiola *et al.*, 1999), que promueve su almacenamiento dentro de la vacuola. La sobre-expresión de SOS1 en *Arabidopsis* (Shi *et al.*, 2003) y el antisentido del transportador de alta afinidad de potasio HKT1 en trigo (Laurie *et al.*, 2002), también confieren un fenotipo de tolerancia a salinidad. En el caso de SOS1 se propone que el mecanismo incluye una menor acumulación de Na^+ en el xilema para prevenir la acumulación de este ión en los tallos (Zhu, 2002). Para HKT1 se ha observado que la disminución del transportador reduce los niveles de acumulación de Na^+ y resulta en una

menor relación Na^+/K^+ en el estele de la raíz, lo cual podría disminuir a translocación de Na^+ al tallo, limitando la toxicidad del ión en los tejidos fotosintéticos (Shi *et al.*, 2003; Zhang y Blumwald, 2001; Zhang *et al.*, 2001).

Los mecanismos de resistencia a la falta de agua no son excluyentes entre sí, sino que las plantas emplean varios de ellos, para desarrollar tolerancia en diferentes grados, dependiendo de la especie y el hábitat al cual se han adaptado (Bray *et al.*, 2000). Debido a que los cultivos de interés agronómico necesitan un aporte constante de agua, el déficit hídrico tiene efectos severos en el crecimiento vegetativo de la planta y en su rendimiento (granos o frutos). Este tipo de estrés es el que causa una de las mayores pérdidas (50-80%) en la productividad de los cultivos cada año. Por ejemplo, se estima que la mitad de los cultivos de trigo que se siembran en el mundo están afectados regularmente por la sequía, por lo que el estudio de los mecanismos de respuesta de las plantas a la falta de agua tendrá un gran impacto en el campo del mejoramiento genético de plantas cultivadas (Boyer, 1982; Trethowan y Reynolds, 2007). Por otra parte, la tendencia a la desertificación y la salinización de los suelos a nivel mundial se ha intensificado debido a las malas prácticas agrícolas, la deforestación de bosques y selvas y el uso de pesticidas y fertilizantes químicos, provocando la erosión de la tierra, y una disminución en la calidad y la productividad de los cultivos que se siembran en estos suelos. Lo anterior, aunado a la demanda de alimentos para la población en constante crecimiento, son algunos de los factores que justifican el estudio de los mecanismos que las plantas utilizan para contender con el déficit hídrico y por consecuencia, con las condiciones ambientales adversas que se predicen para los próximos años (AAAS, 2001).

II. EL AGUA Y SU IMPORTANCIA EN LAS RELACIONES HÍDRICAS DE LA CÉLULA VEGETAL

El agua cumple un papel fundamental en el crecimiento de las plantas. El crecimiento se da por división celular, como ocurre en el caso de los tejidos meristemáticos ó por elongación o expansión, la cual se produce por la entrada de agua a las células, de acuerdo a gradientes de potencial hídrico que favorezcan su entrada por medio de canales de agua. Las acuaporinas son canales protéicos transmembranales que facilitan el transporte de agua, regulan la conductividad hídrica de las membranas e incrementan la permeabilidad al agua de 10 a 20 veces (Maurel y Chrispeels, 2001). La inducción de estos genes se relaciona con procesos de expansión celular y de déficit hídrico, en este

último caso promoviendo una rápida recuperación de la turgencia al momento de la rehidratación. La acuaporina codificada por el gen *rd28* de *Arabidopsis* se induce por deshidratación (Yamaguchi-Shinozaki *et al.*, 1992). Esta acuaporina pertenece a la familia de proteínas MIP (por "Major Intrinsic Proteins"), de las cuales también se han encontrado proteínas relacionadas en la planta halófila *Mesembryanthum crystallinum* (Bray *et al.*, 2000).

El agua es un componente vital para todos los organismos, y constituye el 70 a 90% del peso fresco de muchas especies de plantas herbáceas. Debido a las propiedades fisicoquímicas de esta molécula como su estructura bipolar, su alto punto de ebullición y sus propiedades de tensión-cohesión, resulta ser el solvente por excelencia de todos los compuestos orgánicos polares, iones y moléculas inorgánicas de la célula vegetal (Fitter y Hay, 2002). Además de ser fuente de oxígeno, el agua se requiere para llevar a cabo la fotosíntesis, y regula la temperatura de las hojas, por su elevado calor de vaporización. Sus propiedades de tensión-cohesión permiten su transporte y el de los solutos disueltos a través del sistema vascular de las plantas (Nobel, 2005). El agua se moviliza a lo largo de la planta por diferencia de potenciales del agua. El potencial hídrico de la célula (Ψ) se define como:

$$\Psi_w = \Psi_p + \Psi_{\pi} + \Psi_m + \Psi_g$$

ya que los factores que modifican su actividad son: el potencial de presión (Ψ_p), el potencial osmótico (Ψ_{π}), el potencial mátrico (Ψ_m) y el potencial gravitacional (Ψ_g).

La presencia de la pared celular en una célula vegetal afecta el potencial de presión ya que, al ser una estructura lignificada, es lo suficientemente rígida para impedir la elongación celular. Esto permite una elevada presión hidrostática al interior de la célula. El potencial osmótico se genera por los solutos presentes en la célula, que disminuyen la actividad del agua y por lo tanto el potencial hídrico. El potencial mátrico (Ψ_m) se considera sólo en la superficie líquido-sólido de los coloides, como las proteínas, ribosomas, organelos, cuya superficie tiende a disminuir la actividad del agua. El potencial gravitacional (Ψ_g) es importante para calcular la diferencia de potencial hídrico en regiones localizadas a diferentes alturas de la planta, como la raíces y las hojas en una especie arbórea (Nobel, 2005). La suma de todos los componentes da como resultado el valor del Ψ_w , por ejemplo, en el caso de una hoja turgente, $\Psi_p > \Psi_{\pi}$. Cualquier proceso que provoca la salida de agua desde la célula hacia los espacios intercelulares (apoplasto) genera una disminución del potencial de presión, en tanto que el potencial osmótico se vuelve más negativo, dando como resultado una disminución del Ψ_w y una

pérdida de turgencia celular. Las temperaturas bajo cero provocan la formación de hielo en los espacios del apoplasto, del xilema y de las traqueidas, lo cual provoca la salida de agua desde el interior de las vacuolas al espacio extracelular. Una solución hiperosmótica también causa la salida de agua de la célula por la disminución del potencial hídrico del medio que la rodea. La salida de solutos depende de su tipo y tamaño del soluto; los de tamaño pequeño como las sales o el manitol atraviesan la pared celular y la membrana plasmática. En cambio, los solutos como el polietilenglicol de alto peso molecular (<6000) son impermeables a la pared celular y provocan la salida de agua porque el potencial hídrico disminuye en el exterior de la célula, como sucede cuando hay sequía (Carpita *et al.*, 1979; Oparka, 1994).

III. EVOLUCION DE LA TOLERANCIA A LA DESECACION

La tolerancia a la desecación se puede definir como la capacidad de algunos organismos (anhidrobióticos) de soportar una deshidratación celular y cesar su metabolismo de manera reversible en respuesta a la pérdida de agua, de tal forma que alcanzan el equilibrio con la humedad del ambiente, cuando éste se encuentra en niveles de humedad relativa iguales o menores a 50%; pueden mantenerse en estado quiescente durante largos periodos de tiempo y al ser rehidratadas se recuperan en pocas horas sin sufrir ningún daño en sus células (Alpert y Oliver, 2002). El metabolismo no se puede llevar a cabo si el contenido de agua de una célula es inferior a 0.1 g H₂O / g de biomasa, ya que no existirían moléculas de agua suficientes para formar una monocapa alrededor de las proteínas y membranas intracelulares para llevar a cabo reacciones químicas (Billi y Potts, 2002). Los organismos que toleran la desecación pueden tener cierto nivel de actividad química en niveles de deshidratación celular muy elevado, y a los organismos que soportan estas condiciones se les conoce como *anhidrobióticos*; por ejemplo, los granos de polen pueden incorporar vapor de agua en compuestos orgánicos. En 1702, Anthonie van Leeuwenhoek sugirió por primera vez que podía existir vida de “animálculos” (que se clasificaron posteriormente como rotíferos), los cuales podían crecer después de rehidratar una porción de sedimento que había permanecido seco en una canaleta de su techo por mucho tiempo. Tomó más de 150 años para que la tolerancia a la desecación se aceptara como un término en la comunidad científica, como el proceso donde se detiene la actividad metabólica y de donde se puede “regresar a la vida” cuando se da un proceso de rehidratación. Más tarde se confirmó que varios *phyla* pueden incluir

organismos anhidrobióticos, como algunos representantes de los tardígrados, rotíferos, nemátodos, así como el caso de semillas ortodoxas, esporas, algas, líquenes y musgos, los cuales pueden tolerar una temperatura hasta 0.5°C arriba de la temperatura del cero absoluto (0°K = temperatura a la cual se detiene todo tipo de movimiento molecular) y que al rehidratarse y regresar a la temperatura ambiente, sobreviven (Becquerel 1951; Schierbeek, 1959, Wilson *et al.*, 1979, citados en Alpert y Oliver, 2002). El costo metabólico de la protección y reparación constitutiva que se tiene que mantener en este tipo de organismos anhidrobióticos se ve reflejado en el pequeño tamaño, nivel de complejidad o baja tasa de crecimiento que estos organismos pueden alcanzar (Alpert, 2006). Desde el punto de vista evolutivo, la tolerancia a la desecación en las plantas se puede analizar utilizando como modelo a las plantas con caracteres más primitivos o plantas sin semilla, que fueron capaces de colonizar el ambiente terrestre.

Las plantas son organismos que en su ciclo de vida presentan una alternancia de fases dentro de una misma generación (fase gametofítica y fase esporofítica); dependiendo de la especie vegetal, una u otra de las fases puede ser dominante y difieren en su capacidad de tolerancia a la desecación. En el caso de las plantas sin semilla como los musgos, la fase más conspicua es el gametofito, en tanto que el esporofito depende por completo del gametofito; mientras que en las plantas con semilla, como las plantas con flores, la fase esporofítica (o también llamada fase vegetativa) es la que predomina, ya que la fase gametofítica se ha reducido a las células germinales: el megagametofito (o saco embrionario) y el microgametofito que contiene a los granos de polen. El microgametofito o gametofito masculino es independiente del esporofito cuando se liberan los granos de polen maduros y éstos son tolerantes a la desecación (Hoekstra, 2002); una vez que germina el tubo polínico y ocurre la doble fecundación, se desarrolla una semilla, que también es tolerante a la desecación (semillas ortodoxas) por medio de mecanismos programados en su desarrollo durante la etapa de maduración y secado de la semilla (Bewley, 1997).

Las plantas terrestres o embriofitas, evolucionaron hace más de 400 millones de años a partir de un ancestro común, que probablemente perteneció al grupo de las algas verdes cariofitas (Graham *et al.*, 2000). La colonización de ambientes hostiles cuando pasaron del ambiente acuático a la tierra, requirió estrategias nuevas que les permitieron adaptarse a estas condiciones, por ejemplo, a través de la formación de una cutícula cerosa que evita la pérdida de agua de la superficie de la planta debido a la constante desecación a la que estaban sujetos sus tejidos (Alpert y Oliver, 2002). Las briofitas

(musgos, hepáticas y antoceros) son plantas que se derivaron de estas plantas ancestrales, y, hasta la fecha, las briofitas modernas conservan características primitivas como la falta de un sistema vascular verdadero para conducir agua y nutrientes. Además no poseen estomas y algunas (como el gametofito de las hepáticas) tienen poros que se mantienen abiertos todo el tiempo, lo cual limita su hábitat a lugares con un constante aporte de humedad, que es crucial para la reproducción de estos organismos (Stern *et al.*, 2003). En el curso de la evolución, la tolerancia a la desecación se fue perdiendo, pero con la aparición de los estomas y de un sistema vascular bien desarrollado (compuesto por células bien diferenciadas como xilema y floema) fue posible un control más preciso del intercambio gaseoso y la regulación del movimiento del agua, para evitar su pérdida por evapo-transpiración. Esto permitió que las plantas pudieran tener un mayor tamaño, lo cual tuvo un gran impacto en su capacidad para colonizar exitosamente diversos ambientes terrestres (Grime, 1994).

En la actualidad, sólo algunas especies vegetales mantienen la capacidad de tolerar la desecación. Entre ellas abundan las plantas sin semilla, como las algas y los líquenes, mientras que en el grupo de las gimnospermas no se tiene reportada ninguna especie tolerante a la desecación; por su parte, sólo 60-70 especies de helechos y alrededor de 60 especies de angiospermas (equivalente al 0.1% del total de especies estimado para este último grupo) son tolerantes a la desecación. Sin embargo, este carácter sí se ha mantenido a lo largo de la evolución de las plantas en estructuras reproductoras que son comunes a las plantas: los granos de polen y las semillas (Alpert y Oliver, 2002).

Para entender los cambios que ocurren en la maduración de la semilla, es necesario recapitular los eventos que suceden desde la formación del embrión, a partir de una sola célula cigótica para originar un embrión multicelular por medio de división y diferenciación celular. Esta primera etapa de *histodiferenciación*, junto con la siguiente etapa de *expansión celular*, están caracterizadas por un rápido incremento en el peso fresco y un aumento en el contenido de agua de la semilla. Posteriormente, un incremento en el peso seco ocurre por la deposición de reservas. Aunque el peso fresco de la semilla se mantiene estable, está perdiendo agua al desplazarla por reservas insolubles dentro de las células que las almacenan. Al alcanzar un máximo de incremento en el peso seco, el contenido de agua disminuye hasta que se alcanza la tercera etapa, cuando ocurre la *maduración y desecación* del embrión, lo cual le permite adquirir un estado quiescente y en este momento la semilla puede ser liberada de la planta. La duración de cada una de estas fases de desarrollo puede durar varios días o meses, según la especie y las

condiciones ambientales que se presentaron en el proceso de desarrollo de las semillas (Bewley y Black, 1994). Los cambios metabólicos que ocurren en la semilla antes o durante la desecación involucran la producción de dos tipos de moléculas que tienen un papel en la protección de los tejidos de la semilla contra los efectos de la desecación: oligosacáridos y azúcares como rafinosa, estaquiosa y sacarosa; así como proteínas específicas tipo chaperonas y proteínas LEA o proteínas abundantes de la embriogénesis tardía (por sus siglas en inglés “Late Embryogenesis Abundant”). La adquisición de la tolerancia a la desecación en la semilla madura permite que el embrión se mantenga en un estado quiescente hasta que se presenten las condiciones óptimas para que ocurra el proceso de germinación. Si el embrión no presenta una dormancia endógena, la germinación inicia con la imbibición (rehidratación de la semilla) y culmina con el alargamiento del eje embrionario (Bewley y Black, 1994). En este proceso ocurren numerosos eventos como la hidratación de proteínas, cambios estructurales subcelulares, aumento en la tasa metabólica y respiratoria, en la síntesis de macromoléculas, en el alargamiento y la activación de la división celular; así como en la translocación y asimilación de nutrientes (Bewley, 1997).

IV. REGULACIÓN DE LA EXPRESIÓN GENÉTICA EN RESPUESTA A DÉFICIT HÍDRICO.

La respuesta al déficit hídrico inicia con su percepción a nivel extracelular, lo cual puede darse por cambios en el volumen celular, que son percibidos por canales mecanosensores que pueden cambiar su conformación dependiendo de la presión intracelular, algunos de ellos promueven la entrada de Ca^{2+} . Adicionalmente podrían activarse otros canales iónicos, que promueven la entrada de iones Ca^{2+} y K^+ para recuperar el volumen celular (Zonia y Munnik, 2007). Se ha demostrado que la sequía, la salinidad y la alta salinidad inducen flujos transitorios de Ca^{2+} hacia el citoplasma, ya sea por un influjo del espacio apoplástico o por su liberación desde un almacén interno, por lo que los canales responsables de este flujo representan un tipo de sensor para estas señales de estrés. La cinética, magnitud y fuente celular de las elevaciones en el Ca^{2+} citoplasmático (conocido como firma de calcio) pueden codificar la información del estímulo particular y determinar la respuesta final específica (Knight y Knight, 2001). En plantas sujetas a una rápida disminución de la temperatura se ha encontrado que ocurre un influjo de Ca^{2+} debido a alteraciones en la fluidez de la membrana y reorganizaciones del citoesqueleto (Xiong et

al., 2002). También se activan vías de transducción de señales que transmiten la información dentro de la célula para después extenderse hacia toda la planta. Los cambios en el metabolismo y en el desarrollo inducidos por estrés se pueden atribuir generalmente a una alteración de la expresión genética, algunos genes se reprimen, mientras que otros se inducen.

Los productos de la expresión genética que responden al déficit hídrico pueden estar influenciados por mecanismos a nivel transcripcional, postranscripcional, traduccional y postraduccional, que incrementan la cantidad de RNA mensajeros específicos, promueven la traducción y estabilización de proteínas, alteran la actividad de enzimas o una combinación de éstos (Bray *et al.*, 2000). Su actividad puede estar involucrada en la respuesta primaria de la célula, en las proteínas que llevan a cabo una función reguladora como las de tipo cinasas, fosfatasas, proteínas G, ubiquitina ligasas y factores transcripcionales (Oono *et al.*, 2003; Seki *et al.*, 2002; Yamaguchi-Shinozaki y Shinozaki, 2006), o pueden actuar en la respuesta secundaria de la célula, al producir proteínas protectoras y reparadoras de la maquinaria celular, como las chaperonas (Vierling, 1991) y las proteínas LEA (Battaglia *et al.*, 2008); proteínas de unión a RNA, enzimas que participan en la vía de síntesis de osmorreguladores y en la desintoxicación de especies reactivas de oxígeno; proteínas que modifican la pared celular, entre otros (Hong *et al.*, 2000; Chen y Murata, 2008).

IV. 1 El ácido abscísico (ABA) y su importancia como señalizador en plantas.

El ácido abscísico (ABA) es un isoprenoide de 15 carbonos derivado de la ruptura oxidativa de un precursor epoxi-carotenoide (9'-cis-neoxantina) en la vía no mevalónica o vía MEP (por "Methyl Erythritol Phosphate"), que se lleva a cabo en los cloroplastos (Milborrow, 2001). Es una de las fitohormonas que están muy íntimamente involucradas en la transducción de señales de la respuesta a déficit hídrico. Su participación se ha relacionado en respuesta a diversos tipos de estrés. Inicialmente se describió como una hormona relacionada a la abscisión de las hojas, aunque ahora se sabe que su acción no es un factor común para este proceso en todas las plantas. Se ha detectado su participación en varios procesos fisiológicos y de desarrollo, como en el cerrado de estomas por cambios en el flujo de iones en las células guarda (Vahisalu *et al.*, 2008; Siegel *et al.*, 2009); además es fundamental en el proceso de maduración y adquisición de la tolerancia a la desecación de las semillas. Se ha demostrado su papel en la

inhibición del crecimiento, así como su mecanismo de acción como sensor de déficit hídrico de la raíz hacia la parte aérea (Korneef *et al.*, 1998; Finkelstein *et al.*, 2002; Christmann *et al.*, 2007). Durante el proceso de la embriogénesis hay dos vías principales de acumulación de ABA, la primera de origen materno y la segunda por síntesis en el embrión. Este último es el responsable de la inducción de los genes necesarios para la acumulación de proteínas de reserva en el desarrollo del embrión, la adquisición de la tolerancia a la desecación y la dormancia de la semilla (Blackman *et al.*, 1995; Finkelstein *et al.*, 2008).

En *Arabidopsis* se han identificado sensores de calcio, inducibles por ABA y varias señales de estrés, que son proteínas similares a calcineurina (CBL), que funcionan como reguladores negativos de la sensibilidad a estos estímulos, además de participar en el cerrado de estomas y la toma de K^+ en las raíces (Pandey *et al.*, 2004; Lee *et al.*, 2007, 2009). El análisis de mutantes de plantas afectadas en la sensibilidad (mutantes *abi* o insensibles a ABA) o en la producción de ABA (mutantes *aba* o deficientes en la producción de ABA), ha permitido explorar el papel de esta hormona en diferentes vías de transducción de señales en respuesta a estrés, sus efectos en el desarrollo, así como la regulación de la expresión genética y su conexión con las vías de señalización de otras hormonas (Finkelstein y Lynch, 2000; Hoth *et al.*, 2002). En *Arabidopsis* se han aislado y caracterizado cinco mutantes insensibles a la aplicación de ABA exógeno. Las mutantes en los genes *ABI1* y *ABI2* (*abi1* y *abi2*) están alteradas en su respuesta a ABA en semillas y tejidos vegetativos, ya que presentan una sensibilidad reducida en la germinación (germinan en presencia de ABA) y mantienen los estomas abiertos en presencia de ABA y deshidratación, por lo que se marchitan rápidamente. Ambos genes codifican para fosfatasa de proteínas tipo 2C (PP2C) y tienen un dominio de unión a calcio en el extremo amino (Xiong *et al.*, 2002). La mutante en el gen *ABI3* (*abi3*) también germina en presencia de ABA, pero se asemeja al tipo silvestre en sus relaciones hídricas. El producto del gen *ABI3* codifica para un factor transcripcional y las semillas de estas mutantes no muestran dormancia. Por su parte *ABI4* y *ABI5* son factores transcripcionales con dominios AP2 y bZIP, y las plantas mutantes en estos genes muestran insensibilidad a ABA en la germinación y también presentan un fenotipo de crecimiento resistente a azúcar (Finkelstein *et al.*, 2002). Se conocen mutantes deficientes en ABA en tomate, chícharo, maíz, tabaco, cebada, papa y *Arabidopsis*. Entre las alteraciones fenotípicas de las mutantes *aba1*, *aba2* y *aba3* están la falta de dormancia en la semilla, la alteración en las relaciones hídricas y la mala aclimatación a una variedad de tipos de estrés (Koorneef

et al., 1998). Las mutantes de maíz que carecen de la enzima NCED (9-cis-epoxi-carotenoide dioxigenasa), y las mutantes del factor transcripcional ABI3 llamadas *viviparus 14* o *vp14*, están afectadas en las vías de síntesis y señalización de ABA, respectivamente, y muestran un fenotipo vivíparo, ya que pierden la dormancia y germinan sobre la mazorca (Finkelstein *et al.*, 2002).

Los tejidos vegetativos sujetos a estrés osmótico muestran un incremento en los niveles endógenos de ABA; por ejemplo, las hojas de frijol cortadas y deshidratadas aumentan 45 veces la concentración de ABA en sólo 5 horas de deshidratación, y al rehidratarlas se revierte a la concentración basal en sólo 1 hora (Yang y Zeevaart, 2006); por otra parte, la aplicación de ABA a tejidos no estresados promueve respuestas fisiológicas y moleculares semejantes a las inducidas por déficit hídrico (Campalans *et al.*, 1999). Se ha postulado que existe una señal hidráulica que desencadena la síntesis de ABA en tejidos vasculares y que se transporta desde el sitio de su biosíntesis por medio de transportadores dependientes de ATP para ser distribuída a largas distancias en toda la planta (Christmann *et al.*, 2005; 2007; Kuromori *et al.*, 2010). El aumento en la concentración de ABA ocurre por la activación del ABA almacenado en la célula, por medio de β -glucosidasas (Lee *et al.*, 2006) y por la síntesis *de novo* que ocurre en las células de parénquima de los haces vasculares de las hojas (Endo *et al.*, 2008).

El ABA también participa en la inhibición del crecimiento cuando las condiciones ambientales son adversas, por ejemplo, promueve la dormancia de los primordios foliares de algunos árboles caducifolios durante el invierno (Horvath *et al.*, 2003). En el proceso de imbibición y germinación de *Arabidopsis* se puede revertir el proceso de establecimiento de la plántula, si las condiciones ambientales son adversas, y en este proceso intervienen los factores transcripcionales ABI3 y ABI5 (López-Molina *et al.*, 2002; Kinoshita *et al.*, 2010).

Las proteínas que han sido reportadas como receptores intracelulares de ABA son: una quelatasa de magnesio, que participa en la síntesis de clorofila y puede estar implicada en la señalización entre núcleo y cloroplasto (Shen *et al.*, 2006), mientras que la proteína GCR2 (G-protein coupled receptor 2), que interactúa con la proteína $G\alpha$ GPA1, aparentemente participa en la percepción de ABA en los estomas, regulando su apertura o cerrado (Liu *et al.*, 2007). Recientemente se ha caracterizado una vía central de señalización en respuesta a ABA en tejidos vegetativos, a través de la identificación de la familia RCAR/PYR1/PYL de la que se conocen 14 proteínas a la fecha (Park *et al.*, 2009). Se ha reportado la cristalización y estructura de PYR1(RCAR11) con ABA (Nishimura *et*

al., 2009; Santiago *et al.*, 2009), así como los complejos triméricos de ABA/ABI1/RCAR12(PYL1) y ABA/HAB1/RCAR14(PYL2) (Melcher *et al.*, 2009; Miyazono *et al.*, 2009; Yin *et al.*, 2009). La constante de disociación entre ABA y los diferentes complejos varía, por lo que es posible que estas interacciones dependan de la cantidad de ABA presente en las diferentes partes de la planta, lo cual a su vez podría generar cierta diversidad en las respuestas. Posteriormente se promueve la inactivación de fosfatasa de tipo 2C, como ABI1, ABI2, HAB1 y, esto provoca la activación de canales iónicos y la expresión de genes dependiente de ABA, por medio de la acción de cinasas de tipo SNF1OST1/SnRK2.6/SnRK2E, SnRK2.2/SnRK2D y SnRK2.3/SnRK2I (Cutler *et al.*, 2010; Raghavendra *et al.*, 2010). En la Figura 1 se esquematiza el modelo de señalización por ABA y su acción en diferentes compartimentos intracelulares, el cual se propone en base a los últimos datos provenientes de la identificación y caracterización de los diferentes receptores intracelulares de esta hormona.

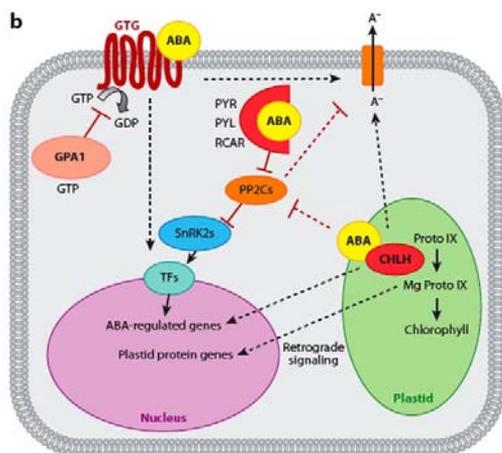


Figura 1. Modelo de las interacciones de ABA con los receptores de diferentes clases descritos a la fecha. Entre ellos se encuentran una quelatasa de magnesio (CHLH) localizada en la membrana de cloroplasto, una proteína G de tipo GPCR, localizada en la membrana plasmática, y las proteínas nucleocitoplásmicas PYR/PYL/RCARs. Estas proteínas intervienen en la expresión genética y la electrofisiología de las células guarda (representadas como un canal aniónico, A⁻). Las interconexiones de estos factores no se han descrito hasta el momento. Las líneas sólidas indican interacciones directas y las líneas punteadas indican interacciones no descritas. Las interacciones positivas se indican con una flecha; las barras indican represión. Reproducido de Cutler *et al.* (2010).

IV.2 Activación transcripcional de genes en respuesta a déficit hídrico.

Se han identificado varios elementos que actúan en *cis* en los promotores de genes que responden a déficit hídrico, entre ellos, el elemento de respuesta a ABA (ABRE) y el elemento de respuesta a deshidratación DRE/CRT (Bray, 1997). El elemento ABRE es una secuencia de 8 a 10 pares de bases presente en los promotores de genes inducibles por ABA que comparten la secuencia consenso (C/T)ACGTGGC, la cual contiene el motivo palindrómico ACGT. Se ha demostrado que este motivo palindrómico está presente también en otros elementos regulatorios, como la caja G, implicada en la regulación de genes regulados por luz, auxinas, ácido jasmónico y ácido salicílico (Ingram & Bartels, 1996). Aunque en muchos genes la presencia de este elemento es suficiente para su expresión mediada por ABA, en algunos genes se han caracterizado elementos adicionales en *cis* que complementan al elemento ABRE, denominados elementos acopladores (CE), que dan una mayor especificidad de la respuesta a través de la unión a diferentes factores transcripcionales. Se han aislado varias proteínas de unión a ABRE, entre ellas, las que poseen una región básica adyacente a un motivo de cierre de leucinas (bZIP), característico de los factores de transcripción como EMBP1 de trigo, TAF-1 de tabaco, TRAB de arroz (Ingram y Bartels, 1996), así como los factores ABF/AREB y ABI5 aislados en *Arabidopsis*, cuya transcripción se induce también en respuesta a estrés osmótico y al ABA (Choi *et al.*, 2000, Uno *et al.*, 2000, Finkelstein y Lynch, 2000). Otro tipo de factores transcripcionales involucrados en la respuesta a ABA son los de tipo MYB y MYC, que se sintetizan después de la acumulación de ABA endógeno. El gen *RD22* de *Arabidopsis* se induce por ABA y su promotor incluye secuencias para la unión de proteínas tipo MYC y MYB (Iwasaki *et al.*, 1995). Otros homólogos que se han identificado también en *Arabidopsis* se inducen por sequía y salinidad (Abe *et al.*, 1997). AtMYC2 y AtMYB2 actúan como activadores transcripcionales en la inducción de la expresión genética en respuesta a ABA y su sobreexpresión incrementa la tolerancia a estrés osmótico en *Arabidopsis* (Abe *et al.*, 2003). Así mismo, la sobreexpresión constitutiva de ABF3 y ABF4 en *Arabidopsis*, incrementa la tolerancia a sequía (Kang *et al.*, 2002).

En el caso de genes cuya inducción en respuesta a estrés es independiente de ABA, se ha encontrado el elemento DRE/CRT (por "Dehydration-Responsive Element/C-Repeat") en sus promotores, cuya secuencia consenso es TA/GCCGACAT. Se han identificado varias proteínas de unión a DRE como las CBF/DREB (por CRT-Binding Factor/DRE Binding proteins) en *Arabidopsis*. Estas proteínas pertenecen a una familia de

factores de transcripción exclusiva de plantas, denominada AP2/EREBP (por “APetala2/Ethylene-Responsive Element-Binding Protein”), de la que varios de sus miembros están involucrados en el desarrollo y la expresión genética en respuesta a hormonas (Kizis *et al.*, 2001). Estos factores transcripcionales comparten una región conservada de 60 a 70 aminoácidos, el dominio AP2, el cual está compuesto de dos módulos, uno de 20 aminoácidos básicos e hidrofílicos en la región amino (elemento YRG) y una secuencia en el extremo carboxilo que puede formar una α -hélice anfipática (elemento RAYD) (Shinozaki y Yamaguchi-Shinozaki, 2000). La expresión de todos los genes CBF/DREB1 en *Arabidopsis* es baja en condiciones normales de crecimiento, pero incrementa en respuesta a frío, como en el caso de los factores CBF1-3 (Gilmour *et al.*, 1998) y en el caso de CBF4, en respuesta a sequía, de manera dependiente de ABA. La sobreexpresión de CBF4 en *Arabidopsis* promueve la inducción constitutiva de genes de respuesta a estrés que contienen el elemento DRE/CRT e incrementa la tolerancia a sequía y frío (Haake *et al.*, 2002). La expresión constitutiva de los genes CBF de *Arabidopsis* en plantas de *Brassica napus* resulta en un incremento en la tolerancia a frío (Jaglo *et al.*, 2001). De igual manera, la sobreexpresión de DREB1A, otorga una marcada tolerancia a deshidratación, sal y frío en plantas de *Arabidopsis* (Kasuga *et al.*, 1999). Un gen ortólogo de arroz (OsDREB1) sobreexpresado en *Arabidopsis*, otorga un fenotipo de tolerancia a sal, sequía y frío (Dubouzet *et al.*, 2003).

Aún cuando en cada respuesta existen atributos específicos, las bajas temperaturas, la sequía y las altas concentraciones de sales son estímulos que promueven el déficit hídrico intracelular, por lo que pueden compartir ciertos elementos que representarían nodos o puntos de interacción entre diferentes vías de señalización; de ahí que es de esperarse que existan múltiples sensores primarios para diferenciar las señales de estrés y diferentes puntos de cruce entre vías de señalización. Algunos de los elementos que se conocen hasta el momento que están involucrados en las vías de señalización en respuesta a temperaturas bajas y a deshidratación se esquematizan en la Figura 2.

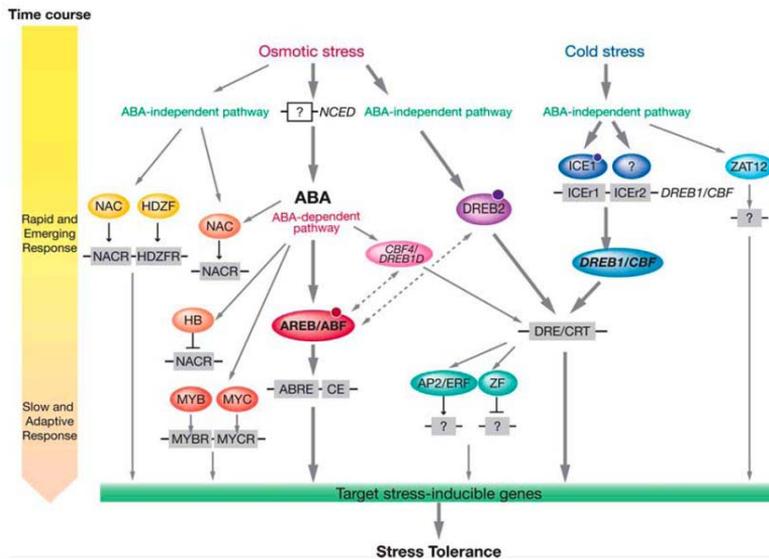


Figura 2. Redes transcripcionales reguladoras de los elementos que actúan *in cis* y los factores transcripcionales de las vías de señalización involucrados en la expresión génica en respuesta a estrés osmótico y frío en *Arabidopsis*. Los factores transcripcionales están mostrados en elipses de color, y los elementos *in cis* en cajas. Los círculos pequeños indican modificaciones postranscripcionales, como fosforilación, que se necesitan para activar a los factores transcripcionales. Las respuestas tempranas de la expresión génica se muestran en la parte superior, mientras que las respuestas tardías y adaptativas se muestran en la parte inferior. Las flechas grises indican las principales vías de señalización; las líneas punteadas indican interacciones proteína-proteína. Reproducido de Shinozaki y Yamaguchi-Shinozaki (2006).

IV.3 Regulación a nivel postranscripcional en la respuesta a déficit hídrico.

La transcripción de un gen se encuentra regulada por eventos que pueden modular la producción, la localización, la estabilidad o la eficiencia de traducción de un RNA mensajero (RNAm) (Kuersten y Goodwin, 2003). El procesamiento del extremo 5' involucra la adición de una G formando un enlace trifosfato (GpppN) que se modifica por metilación, y el extremo 3' contiene secuencias que guían el sitio donde se corta el RNA mensajero y se le adiciona la cola de poli(A), que le confiere estabilidad y permite que se recluten los factores de inicio de la traducción, una vez que el mensajero está maduro y es exportado al citoplasma (Graber *et al.*, 1999). Allí se puede encontrar en tres estados que en condiciones óptimas se encuentran en un equilibrio dinámico: en traducción activa, no traducido o almacenado en complejos de ribonucleoproteínas, y en proceso de degradación. Cuando hay estrés se pueden formar gránulos donde se degradan o se

mantiene inhibida la traducción de una población de RNA mensajeros en particular : los cuerpos de procesamiento y los gránulos de estrés, donde se almacenan diversos elementos que participan en la traducción. Estos gránulos desaparecen cuando se recuperan las condiciones óptimas (Weber *et al.*, 2008). Las regiones no traducidas pueden regular la estabilidad o vida media del RNA mensajero, como se ha encontrado para la región 3' del RNAm que codifica para α -amilasa 3 en plantas de arroz en respuesta a la disponibilidad de carbono (Chan y Yu, 1998a, 1998b) o en el caso de la región 5' a través de la cual se puede regular la traducción del RNAm al promover cierta estructura secundaria y el reclutamiento de factores de inicio de la traducción (Hulzink *et al.*, 2002; Mardanova *et al.*, 2007).

Entre los mecanismos que participan en la regulación postranscripcional de la expresión genética están la degradación de RNAm o la inhibición de la traducción de los mismos debido a la acción de los microRNAs (miRNAs), que son moléculas pequeñas de RNA de 21 a 23 nucleótidos y están involucrados en procesos del desarrollo y respuesta a estrés en eucariotes (Lai, 2003). En general, se ha visto que los miRNAs de plantas tienen una tendencia a regular factores de transcripción tipo MYB, ARF, CBF y AP2 (implicados en el desarrollo y en respuesta a auxinas y frío, entre otras), proteínas de unión a ácidos nucleicos y transductores de señales (Jones-Rhoades y Bartel, 2004; Wang *et al.*, 2004; Arenas-Huertero *et al.*, 2009). Este mecanismo de silenciamiento postranscripcional de genes involucra la producción de un precursor de RNA que se conforma en una estructura particular de tallo-asa (Bartel, 2004). Se ha encontrado que existen otras moléculas pequeñas de RNA cuya biogénesis es distinta a la de los miRNAs, pero su mecanismo de acción es similar, es decir, son moléculas de RNA de 21-24 nt que se generan a partir de un duplex de RNA, como los siRNA (por "small interfering RNA") que pueden generarse por la replicación de virus de RNA de doble cadena que infectan las células o por el sobrelape de transcritos en antisentido (tasiRNA), etc. Una vez que son cargados al complejo proteínico de silenciamiento inducido por RNA (RISC), se aparean con el RNAm que posee secuencias homólogas al RNA pequeño, desencadenando la degradación del mensajero. En animales es más común que los miRNAs dirijan la represión traduccional, la desestabilización del RNAm o una combinación de ambos, al igual que en plantas (Brodersen y Voinnet, 2006; Filipowicz *et al.*, 2008; Covarrubias y Reyes, 2009).

La traducción o síntesis de proteínas consta de tres etapas: la iniciación, la elongación y la terminación. Cada etapa está modulada por distintos complejos proteínicos, aunque

el principal punto de regulación es la iniciación. Es importante considerar que la traducción es un proceso energéticamente costoso, por lo que el estrés provoca una inhibición generalizada de la síntesis de proteínas para dar preferencia a las que se requieren para tolerar el estrés (Bailey-Serres, 1999). Existen algunos ejemplos donde se ha observado que las regiones no traducidas (3') pueden ser importantes para la expresión específica en órganos o tejidos, utilizando fusiones transcripcionales de un gen reportero con la región 3' del gen de interés (Chen *et al.*, 1998; Patel *et al.*, 2004).

En plantas sometidas a déficit hídrico o hipoxia, la síntesis de proteínas disminuye con respecto al control, la fracción polisomal se reduce a partir de 30 min de inicio del estrés mientras que la fracción monosomal aumenta, y este proceso es reversible al poco tiempo que se recuperan las plantas (Hsiao, 1970; Dhinsa y Cleland, 1975; Branco-Prince *et al.*, 2005; Battaglia, 2008, Tesis Doctoral). En granos de polen maduros, donde ocurre una desecación programada, hay una acumulación de RNAs que se requieren para el inicio de la traducción durante la germinación del polen, de manera semejante a lo que ocurre en la imbibición de la semilla. El gen *Ntp303* se regula a nivel traduccional, ya que sólo se expresa en el momento de la germinación y depende de la región 5' no traducida (Hulzink *et al.*, 2002; Nakabayashi *et al.*, 2005).

V. PROTEÍNAS PROTECTORAS DE LA MAQUINARIA CELULAR

Las proteínas de choque térmico (HSPs, por sus siglas en inglés "Heat Shock Proteins") y las proteínas de embriogénesis tardía (LEAs) son dos tipos de proteínas muy importantes dentro de la respuesta adaptativa ante déficit hídrico, salinidad y temperaturas extremas (Ingram y Bartels, 1996).

V.1 Proteínas de choque térmico (HSP).

Las HSP se agrupan en varias familias, están conservadas desde bacterias hasta mamíferos y se inducen de manera abundante en estrés por altas temperaturas. Su función primordial es la de chaperonas moleculares, que entre otras funciones, ayudan a prevenir el plegamiento incorrecto de proteínas inmaduras o que están siendo translocadas a través de una membrana, aumentan la estabilidad de proteínas en su conformación nativa, facilitan el plegamiento correcto de proteínas desnaturalizadas, reactivándolas, y previenen la agregación de proteínas desnaturalizadas o deshacen los agregados ya formados. Dentro de las cinco familias de HSPs (HSP100, HSP90, HSP70,

HSP60 y sHSP) las sHSP, o proteínas de choque térmico de bajo peso molecular, tienen un dominio conservado relacionado a la α -cristalina de los vertebrados y su actividad es independiente de ATP; se inducen también en respuesta a deshidratación, a diferencia de las otras HSPs que se inducen por otros estímulos, como estrés calórico, etanol y metales pesados (Vierling, 1991; Wang *et al.*, 2004). Las sHSPs son muy abundantes en plantas, son proteínas de bajo peso molecular (17-28 kDa) que forman oligómeros de 9 a más de 30 subunidades. Se han agrupado en cinco familias, dos de ellas se localizan en el citosol y las restantes se localizan en el cloroplasto, retículo endoplásmico y mitocondrias, respectivamente. Su localización diferencial sugiere una divergencia funcional a nivel subcelular y durante el desarrollo, así como durante condiciones de estrés (Waters *et al.*, 1996). Por ejemplo, una sHSP de frijol (PvHSP17/ RSP19) se induce por déficit hídrico principalmente en raíces, pero también responde a ABA, choque térmico y se acumula durante la embriogénesis tardía, expresándose aún en plantas no estresadas (Colmenero-Flores *et al.*, 1997). La sobreexpresión de AtHSP17.6A en *Arabidopsis* incrementa la tolerancia a estrés osmótico y calor (Sun *et al.*, 2001). Una sHSP de chícharo (PsHsp18.1) se utilizó como modelo para identificar los sitios de interacción con sustratos como la malato deshidrogenasa y la luciferasa, encontrando que una región flexible (intrínsecamente no estructurada) hacia el amino terminal de la sHSP es importante para interactuar con sus sustratos y ésta región puede presentar diversas geometrías para la interacción con ellos (Jaya *et al.*, 2009).

V.2 Proteínas de embriogénesis tardía (LEA).

Entre las proteínas que se sintetizan en respuesta a déficit hídrico se encuentran las proteínas LEA, que originalmente fueron descritas por su elevada acumulación en la fase de maduración y desecación de la semilla (Dure, 1993). Por ejemplo, una proteína LEA4 de algodón está distribuida uniformemente en todas las células del embrión en una concentración aproximada de 0.3 mM (Roberts *et al.*, 1993). A pesar de que no existe una similitud significativa entre los miembros de las diferentes familias, las características comunes de estas proteínas son su índice de hidrofobicidad igual o mayor a 1 y un elevado contenido de glicinas y otros aminoácidos pequeños y cargados (Garay-Arroyo *et al.*, 2000). Se ha encontrado que otros organismos como bacterias, levaduras (Garay-Arroyo *et al.*, 2000), nemátodos (Solomon *et al.*, 2000), rotíferos (Tunnacliffe *et al.*, 2005), artrópodos (Kikawada *et al.*, 2006) y crustáceos (Hand *et al.*, 2007; Menze *et al.*, 2009) poseen proteínas con características fisicoquímicas similares. Otra característica en

común es su inducción en situaciones de déficit hídrico, ya sea impuestas en el desarrollo o por el ambiente, por lo que se ha propuesto que las proteínas LEA de plantas formen un subgrupo dentro de las “hidrofilinas”, un grupo de proteínas con características fisicoquímicas similares que incluye organismos eucariontes y procariotes (Garay-Arroyo *et al.*, 2000). La conservación de dominios en los miembros de las diferentes proteínas LEA, la correlación entre los patrones de acumulación de transcritos y proteínas en condiciones de déficit hídrico, impuestas durante la embriogénesis tardía o durante el crecimiento vegetativo, sugieren su papel fundamental en la tolerancia a la desecación en plantas. Las características fisicoquímicas de las hidrofilinas permiten predecir que en solución estas proteínas se encuentran en conformaciones flexibles (“random coils”) por lo que pertenecen al grupo de las proteínas intrínsecamente no-estructuradas (Dyson y Wright, 2005). Aunque no se tiene evidencia experimental directa de la función biológica de estas proteínas, se ha sugerido que su conformación no globular y su gran hidrofilicidad, permitiría su interacción con el agua, reteniéndola en tejidos sujetos a déficit hídrico (McCubin y Kay, 1985). También podrían actuar como solutos compatibles al prevenir la pérdida excesiva de agua en los embriones y crear ambientes de “solvatación” o hidratación preferencial para proteger estructuras citosólicas y macromoléculas (Garay-Arroyo *et al.*, 2000; Hoekstra *et al.*, 2001). En nuestro grupo se ha demostrado que estas proteínas también pueden proteger ciertas proteínas de los efectos dañinos provocados por la deshidratación previniendo cambios conformacionales en las proteínas que afectan su función (ver abajo).

El papel protector de los azúcares y las proteínas en la etapa de desecación de la semilla, se debe a su participación en la vitrificación del agua, la estabilización de lípidos y proteínas de las membranas. La capacidad de interactuar y atraer moléculas de agua por parte de las proteínas LEA mantiene un ambiente localmente enriquecido en agua o en algunos casos puede actuar como reemplazo de las moléculas de agua, por lo que su efecto protector a nivel subcelular es muy importante para la tolerancia a la desecación para proteger las estructuras subcelulares conforme se va perdiendo el agua en la semilla (Hoekstra *et al.*, 2001).

Algunas proteínas LEA pueden inducirse por la aplicación de ABA en embriones aislados de la semilla en etapas inmaduras, que normalmente no son tolerantes a la desecación y la aplicación de la hormona las vuelve resistentes a la deshidratación, así como la deshidratación lenta, en un proceso que puede simular la desecación de la semilla en la planta (Bewley y Black, 1994).

V.3 Clasificación de las proteínas LEA.

Se han propuesto varias clasificaciones para las proteínas LEA, basadas en la conservación de la secuencia primaria en los miembros de una misma familia o grupo (Dure, 1993; Ingram y Bartels, 1996; Cumming, 1999). Recientemente se desarrolló un algoritmo informático denominado POPP (por “Protein or Oligonucleotide Probability Profile”, Wise, 2002) que compara las secuencias de las proteínas basándose en la similitud de sus composiciones peptídicas y las clasifica, con base en la sub- o sobre-representación de aminoácidos particulares dentro de su secuencia primaria en 4 familias y 8 superfamilias (Wise y Tunnacliffe, 2004). Esta última clasificación, al no tomar en cuenta las características estructurales de las familias y la homología de secuencias dentro de cada familia, reagrupa de manera totalmente diferente a las proteínas LEA y podría dificultar el análisis de su función biológica.

La clasificación que se seguirá en este trabajo es la propuesta por Battaglia *et al.* (2008), que las divide en siete grupos, basándose en la conservación de firmas o motivos que distinguen a cada uno de los grupos, considerando las proteínas ASR (no presentes en *Arabidopsis*) y las descritas en nuestro laboratorio cuyo primer representante fue la PvLEA-18 de frijol. Por otra parte, las proteínas LEA atípicas o hidrofóbicas, que no cumplen con las características de las hidrofílicas, se proponen como un grupo separado de las LEAs hidrofílicas (grupo LEA5) y se pueden subclasificar en por lo menos tres subgrupos independientes (ver artículo de revisión en el apéndice).

Los miembros del grupo 1 son proteínas de bajo peso molecular y ricas en glicinas (16-21% de la composición total) por lo que adoptan estructuras secundarias flexibles y poseen al menos un motivo conservado de 20 residuos que se repite en tándem a lo largo de la proteína de 1 a 4 veces. La proteína EM de trigo representa a este grupo, del que se han encontrado homólogos en diversas especies vegetales (Ried y Walker-Simmons, 1993). El grupo 2 (también conocido como dehidrinas) es uno de los más grandes en cuanto al número de proteínas reportadas y se acumulan en altos niveles en respuesta a diferentes tipos de estrés como sequía, alta salinidad, bajas temperaturas, déficit de fosfato o a la aplicación exógena de ABA. Son de tamaño variable (14-150 KDa) y poseen al menos tres motivos representados diferencialmente entre las proteínas de la familia y que están altamente conservados. En el extremo carboxilo el dominio que es característico de todas las dehidrinas, denominado segmento K, está compuesto de una o varias repeticiones en tándem de 15 residuos, rico en lisinas, cuya secuencia consenso es

EKKGIMDKIKEKLPG (Dure, 1993; Close, 1996). Se predice que este segmento forma α -hélices anfipáticas y que puede tener interacciones hidrofóbicas con proteínas parcialmente desnaturalizadas y membranas. Algunas dehidrinas también pueden poseer un motivo en la región amino terminal denominado segmento Y, cuya secuencia consenso es (V/T)DEYGNP, así como un motivo rico en serinas, denominado segmento S, de 6 a 9 residuos fosforilables contiguos, como se ha demostrado en el caso de RAB17 de maíz (Goday *et al.*, 1994). Se han encontrado dehidrinas en el citoplasma, la vacuola y el núcleo (Heyen *et al.*, 2002; Karlson *et al.*, 2003). Las proteínas del grupo 3 se subclasifican en dos subgrupos (D-7/D29), aunque comparten un dominio de 11 residuos, repetido en tándem a lo largo de la proteína de 5 a 13 veces, lo cual hace que tengan tamaños muy variables. Estos onceámeros están conservados en la posición de algunos de los residuos a lo largo del onceámero, lo cual permite que formen α -hélices anfipáticas con la posibilidad de interacciones intra- e intermoleculares (Battaglia *et al.*, 2008).

Las proteínas del grupo 4 son de tamaño variable, de 88 a 175 aminoácidos, son ricas en glicinas, treoninas y, a diferencia de otras familias, en alaninas. Según las predicciones basadas en su secuencia de aminoácidos poseen dos dominios, uno en la región aminoterminal (60-80 aminoácidos) que presenta homología de secuencia entre los miembros de la familia y forma una α -hélice anfipática y otro, con menor homología de secuencia, que forma una estructura aleatoria y varía mucho en tamaño (Dure, 1993; Ingram y Bartels, 1996). El grupo 5 lo integran las proteínas LEA atípicas, ya que no presentan el grado de hidrofobicidad ni la composición general de aminoácidos que caracteriza a las demás proteínas LEA (Galau *et al.*, 1993). El grupo 6 fue descrito más recientemente, está compuesto por proteínas altamente hidrofílicas que se localizan en el citoplasma y en el núcleo (Colmenero-Flores *et al.*, 1997). En el caso de la PvLEA18, que fue la primera proteína descrita para este grupo, se ha propuesto que responde al estatus hídrico de la planta durante el desarrollo, ya que se expresa no sólo en plantas expuestas a estrés hídrico sino también en las regiones de crecimiento de plantas bien irrigadas (Colmenero-Flores *et al.*, 1999).

Las proteínas ASR forman el grupo 7, son proteínas hidrofílicas que se localizan en núcleo y citoplasma y fueron descritas inicialmente en plantas de tomate, y posteriormente se encontraron genes ortólogos en otras especies como papa, maíz, arroz, pino, uva, aunque no hay representantes en *Arabidopsis* (Silhavy *et al.*, 1995; Rossi *et al.*, 1996; Vaidyanathan *et al.*, 1999; Maskin *et al.*, 2008). Estas proteínas también se expresan en la embriogénesis tardía y en respuesta a déficit hídrico, y en algunos casos se ha visto que

tienen una función en la señalización por azúcares y ABA (Carrari *et al.*, 2004). Para cada una de las familias LEA, se llevó a cabo una búsqueda de motivos conservados que se muestran en la Figura 3 y se discuten con detalle en el artículo de revisión publicado por Battaglia *et al.* (2008, ver artículo en el anexo 2 de este manuscrito).

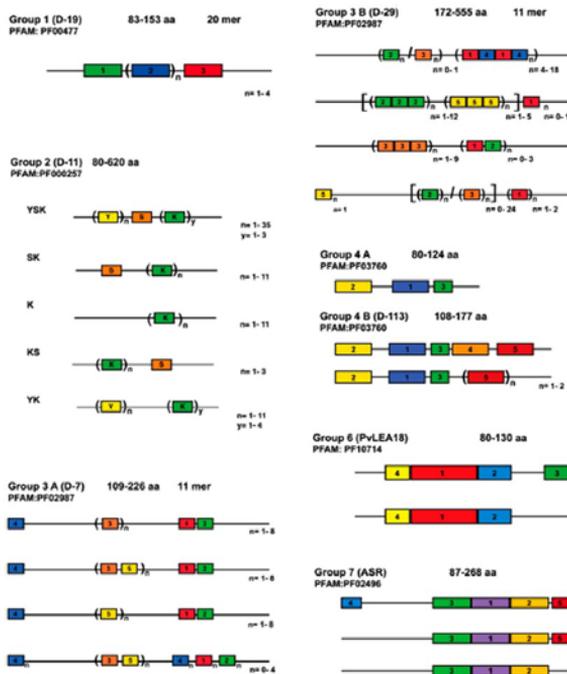


Figura 3. Motivos conservados para las principales familias de proteínas LEA hidrofílicas. Para cada grupo se indica la longitud promedio de las proteínas que lo conforman, así como la correspondencia con la nomenclatura asignada por la clasificación de PFAM, y los subgrupos que se encontraron para cada grupo. Las cajas de colores indican los motivos y el número asignado a cada motivo, pero no implica ninguna relación entre las secuencias de cada grupo. Reproducido de Battaglia *et al.* (2008).

V. 4. Sobreexpresión de genes que codifican para proteínas tipo LEA.

Existe una elevada correlación de la acumulación de estas proteínas ante situaciones de déficit hídrico, ya sea impuestas durante el desarrollo (embriogénesis) o durante el crecimiento vegetativo. Su posible papel en la respuesta adaptativa de las plantas a la limitación de agua se ve apoyado por experimentos de sobreexpresión de algunos genes *LEA* en plantas y en levaduras. La sobre-expresión del gen *HVA1* de cebada, que codifica para una proteína LEA3, aumenta la tolerancia de las plantas de arroz a deshidratación y salinidad (Xu *et al.*, 1996; Babu *et al.*, 2004), mientras que su sobreexpresión en trigo

incrementa la productividad en biomasa y eficiencia en el uso de agua bajo condiciones de estrés hídrico (Sivamani *et al.*, 2000). La proteína LE25 de tomate, de la familia 4, aumenta la tolerancia a salinidad y congelamiento cuando el gen que la codifica se sobreexpresa en levadura (Imai *et al.*, 1996). Se ha reportado la presencia de la proteína EM en plántulas de trigo resistentes a desecación (Ried y Walker-Simmons, 1993) y se ha observado que la expresión heteróloga del gen *EM* en levadura, promueve su crecimiento en presencia de una variedad de compuestos osmóticamente activos (NaCl 1M, KCl 1M y sorbitol 1.5 M) por lo que se propone un papel de osmoprotector (Swire-Clark y Marcotte, 1999). La sobreexpresión de los genes que codifican para las proteínas PMA80 y PMA1959 provenientes de plantas de trigo en arroz, otorgan una mayor integridad a membranas en tratamientos de estrés por alta salinidad, y las plantas muestran una mayor acumulación de biomasa y recuperación después de estar expuestas a tratamientos de sequía y alta salinidad (Cheng *et al.*, 2002). La expresión constitutiva de genes *COR* incrementa la tolerancia de cloroplastos al congelamiento *in vivo*, como el caso del gen *COR15A* (Steponkus *et al.*, 1998), pero el efecto de este mismo gen no es significativo en la sobrevivencia de plantas completas durante congelamiento (Jaglo-Ottosen *et al.*, 1998). La expresión constitutiva de múltiples genes de respuesta a estrés, entre ellos los relacionados a genes *LEA*, como *RD29A*, *COR6.6*, *COR15A*, al ser activados por la sobreexpresión de los genes que codifican para factores de transcripción (CBF1/DREB1A, CBF3), correlaciona con la tolerancia a sequía, salinidad y congelamiento en *Arabidopsis* (Jaglo-Ottosen *et al.*, 1998; Kasuga *et al.*, 1999; Gilmour *et al.*, 2000); así como con la tolerancia a sequía en trigo (Pellegrineschi *et al.*, 2004). Lo anterior sugiere que los productos de estos genes, al igual que las diferentes familias de proteínas *LEA*, podrían actuar de manera sinérgica para incrementar la tolerancia al déficit hídrico.

Sin embargo, la expresión ectópica de una proteína puede dar resultados que no necesariamente indican su papel en la respuesta adaptativa de la planta o en un proceso específico en su desarrollo, ya que se está expresando de manera constitutiva en prácticamente todos los tejidos. Es por ello que en este trabajo decidimos utilizar el análisis por pérdida de función, en donde la ausencia de los productos finales de la expresión genética esté relacionada a fenotipos de sensibilidad ante el déficit hídrico en diferentes etapas del desarrollo de la planta.

VI. LA GENÉTICA REVERSA: UNA HERRAMIENTA EN EL ANALISIS FUNCIONAL DE GENES DE RESPUESTA A DÉFICIT HÍDRICO

Una de las metodologías de genética reversa que ha tenido un gran auge en estos últimos años para conocer la función de genes es el mecanismo de interferencia de RNA (RNAi), es decir, el silenciamiento de genes por medio de la degradación postranscripcional del RNA mensajero que codifican (mecanismo conocido también como silenciamiento postranscripcional de genes, PTGS o interferencia de RNA de doble cadena). El PTGS es un mecanismo que tiene como resultado la degradación específica de una población de RNAs homólogos al RNA interferente (Vaucheret *et al.*, 2001). El mecanismo fue descubierto por primera vez en plantas transgénicas que expresaban copias extra de un gen endógeno y se le llamó cosupresión, debido a que se observó la degradación de los mensajeros codificados por el gen endógeno y el transgen, dando como resultado un fenotipo de pérdida de función (Napoli *et al.*, 1990). Posteriormente, se encontró que este proceso también ocurre de manera similar en otros eucariotes como ciliados, hongos y en diferentes clases de animales como nemátodos y mamíferos (Tijsterman *et al.*, 2002). En los sistemas de silenciamiento de estos organismos, un componente clave involucra moléculas de RNA de doble cadena (dsRNA) que pueden ser generadas por RNA polimerasas dependientes de RNA producidas por el hospedero, por virus (Dalmay *et al.*, 2000; Mourrain *et al.*, 2000; Vance y Vaucheret, 2001) o por la transcripción de secuencias invertidas repetidas (Waterhouse *et al.*, 1998; Chuang y Meyerowitz, 2000; Escobar *et al.*, 2001). Este dsRNA es degradado a fragmentos de 21-26 nucleótidos (denominados siRNA), los cuales se postula actúan como guías para la maquinaria de degradación, de la que también forma parte una ribonucleasa de varias subunidades (RISC, por RNA Induced Silencing Complex) que degrada RNAs con homología de secuencia al dsRNA, la especificidad se logra por medio del apareamiento de bases entre el siRNA y la secuencia blanco (Hamilton y Baulcombe, 1999; Zamore *et al.*, 2000; Zamore, 2001). El mecanismo de silenciamiento en plantas actúa también a nivel transcripcional y está relacionado con metilación y otras modificaciones en el DNA. Por otra parte, el silenciamiento no es autónomo: puede ser transmitido de manera sistémica a toda la planta y debe involucrar una señal, desconocida hasta el momento, que se mueve a través de los plasmodesmos y el sistema vascular, y que podría incluir un componente de ácido nucleico para asegurar la especificidad de secuencia (Mlotshwa *et al.*, 2002). Una señal similar puede existir en *Caenorhabditis elegans*, en el cual el

mecanismo de silenciamiento de RNA también se puede inducir localmente para extenderse posteriormente a otras partes del organismo (Fire *et al.*, 1998). Recientemente se ha clasificado a los siRNA en dos clases: una de 21-22 nucleótidos que dirige la degradación secuencia-específica de RNA mensajero y otra de 24-26 nucleótidos que correlaciona con el silenciamiento sistémico y la metilación de DNA (Hamilton *et al.*, 2002). Se han asignado varias funciones para el silenciamiento de RNA, como parte de la protección adaptativa en contra del ataque por virus, defensa en contra de elementos móviles de DNA, degradación de RNAs aberrantes y en la regulación de expresión genética durante el desarrollo (Tijsterman *et al.*, 2002).

Según lo reportado en la literatura, la eficiencia de silenciamiento de genes utilizando esta metodología puede alcanzar niveles cercanos al 100% y con una elevada especificidad de secuencia, de tal manera que genes homólogos hasta en un 80% no son afectados (Schweizer, *et al.*, 2000; Smith *et al.*, 2000), lo cual permitiría anular de manera específica la producción de proteínas de una sola familia.

Otro mecanismo de PTGS está basado en el uso de microRNAs artificiales (a-miRNA), que también se basan en el uso de la maquinaria de silenciamiento endógeno de la planta, pero que pueden utilizarse para silenciar con una sola construcción a más de un mensajero que comparta homología de secuencia con el a-miRNA y se puede emplear en condiciones en las cuales no es efectivo el silenciamiento por siRNAs; por ejemplo, a temperaturas bajas donde la maquinaria de silenciamiento inducido por RNAi no es funcional. Esta metodología novedosa se ha utilizado para conferir resistencia a virus en plantas de *Arabidopsis* y se ha aplicado con éxito para silenciar postranscripcionalmente la expresión de uno o varios genes de interés en una sola construcción, en lugar del blanco original que es un factor transcripcional tipo MYB, para el cual está dirigido el microRNA que se utiliza como templado (ath-mir159a) (Niu *et al.*, 2006).

Aunque el análisis genético de mutaciones que afecten la función de alguna(s) hidrofili(na)s es muy limitado, la inactivación de genes que codifican para este tipo de proteínas, algunas de ellas homólogas a las proteínas LEA, se ha llevado a cabo en bacterias. Este es el caso de una proteína homóloga a proteínas LEA del grupo 1 en la bacteria *Deinococcus radiodurans*, cuya deficiencia provoca un fenotipo de sensibilidad a la desecación (Battista *et al.*, 2001). De manera similar, la remoción del gen *rmf* de *Escherichia coli*, el cual codifica para una hidrofili(na), resulta en un fenotipo osmosensible (Garay-Arroyo *et al.*, 2000).

En plantas existen algunos reportes de mutantes sencillas para genes tipo LEA: una en el musgo *Physcomitrella patens* que codifica para una dehidrina y otra en *Arabidopsis*, que codifica para uno de los dos miembros que conforman al grupo 1 en esta planta modelo. En el musgo se encontró que la falta de esta dehidrina provoca un fenotipo osmosensible, ya que no son capaces de recuperarse de un estrés osmótico o salino (Saavedra *et al.*, 2006). En *Arabidopsis* se detectó un ligero fenotipo en el desarrollo de las silicuas, ya que algunas de las semillas que se encuentran en el extremo distal de la silicua mostraban una maduración prematura (Manfre *et al.*, 2006). El análisis de los efectos provocados por el silenciamiento de un gen para una proteína LEA del grupo 4 de cacahuete se llevó a cabo en plantas de tabaco, en donde se expresó transitoriamente una construcción que desencadenaba el silenciamiento del gen endógeno por medio de un virus, lo cual conllevó aparentemente a que los tejidos vegetativos de las plantas de tabaco fueran más sensibles a ciertos tratamientos de sequía (Sentil-Kumar y Udayakumar, 2006). Debido a que existen varias familias de proteínas LEA en una misma planta, podría haber redundancia funcional entre los miembros de las diferentes familias; sin embargo, a la fecha no existen reportes de análisis con mutantes de toda una familia de estas proteínas, por lo que sería conveniente abordar un estudio mediante este enfoque (Wang y Waterhouse, 2001).

ANTECEDENTES

Se han abordado muy pocos estudios acerca del mecanismo de acción de las proteínas LEA, en particular, acerca de los blancos con los que cada familia puede interactuar *in vivo*. Los experimentos *in vitro* que se han llevado a cabo de manera independiente en nuestro laboratorio y en otros, indican que algunas proteínas LEA de varias familias pueden interactuar con enzimas como LDH o MDH, protegiendo su actividad en ciclos de congelamiento/descongelamiento o en condiciones de deshidratación controlada (Hara *et al.*, 2001; Bravo *et al.*, 2003; Reyes *et al.*, 2005). La cantidad de agua que queda disponible en estos ensayos (<1%) está dentro del intervalo en el cual los organismos sensibles a la desecación responden a un déficit hídrico severo (Hoekstra *et al.*, 2001). En nuestro laboratorio se han ensayado las proteínas DSP16, ERD10 (LEA2), AtLEA76 (LEA 3) y AtD-113 (AtLEA4-5), así como las hidrofílicas YCIG de *Escherichia coli* y la SIP18 de *Saccharomyces cerevisiae*. Todas ellas mostraron una protección de la actividad enzimática de LDH o MDH después de la rehidratación, a

diferencia de la PvLEA-18 (LEA 6) y otras proteínas no relacionadas a estrés, como la RNasa A y la α -cristalina (Reyes *et al.*, 2005). Es importante destacar que aunque el efecto protector se observa a diferentes relaciones molares entre la enzima y la hidrofiliina (1:5–1:1), una relación molar 1:1 es suficiente para prevenir la pérdida de actividad enzimática bajo condiciones de deshidratación parcial. La capacidad de protección de cada proteína, puede o no depender de la presencia de los dominios conservados y de la severidad del déficit hídrico, pero supera el efecto protector que posee la albúmina (BSA), y es más efectiva que la relación molar necesaria de moléculas como trealosa (2 órdenes de magnitud mayor) para proteger en estas condiciones severas de deshidratación (Reyes *et al.*, 2008). Sin embargo, no todas las proteínas de otras familias son capaces de proteger la actividad de estas enzimas bajo las condiciones ensayadas (ej. proteínas LEA del grupo 6), lo cual sugiere que no todas las hidrofiliinas son capaces de interactuar con proteínas, sino que sus blancos naturales pudieran ser otro tipo de macromoléculas como los ácidos nucleicos (Rivera, L, datos no publicados). Por otra parte, se determinó que la inactivación de estas enzimas correlaciona con la exposición de superficies hidrofóbicas durante la remoción parcial de agua. Esto se midió por ensayos de fluorescencia con ANS (1-anilino-naftaleno-8-sulfonato), cuya emisión a 460 nm incrementa cuando se une a grupos hidrofóbicos. Cuando la deshidratación ocurre en presencia de una hidrofiliina (DSP16 o SIP18) los niveles de fluorescencia son muy similares a los de muestras que no estuvieron expuestas a deshidratación. El uso de un agente entrecruzador permitió detectar la interacción de una hidrofiliina (SIP18) con ambas enzimas, indicando que la interacción proteína-proteína forma parte del mecanismo de protección. Los resultados sugieren que el mecanismo de protección involucra la organización de moléculas de agua y la interacción directa entre proteínas para prevenir cambios estructurales en la enzima que modifican su estructura terciaria y su actividad (Reyes *et al.*, 2005; 2008). En el caso de proteínas LEA con señales de localización a mitocondria o cloroplasto, su localización y su capacidad de protección *in vitro* a enzimas presentes en esos compartimentos subcelulares, han sido verificadas. Tal es el caso de una proteína LEA del grupo 3 de chícharo, que se localiza en la matriz mitocondrial y para la que se comprobó que otorga protección *in vitro* a enzimas que se encuentran en ese compartimento específico, como fumarasa, rodanasa y citrato sintasa (Grelet *et al.*, 2005).

Es importante destacar que la presencia ubicua de estas proteínas en organismos procaríotes y eucaríotes sugiere que su papel es muy importante en la respuesta ante déficit hídrico, sin embargo hasta la fecha no se sabe exactamente cuál es su mecanismo

de acción. Los estudios *in vitro* sugieren que su acumulación en respuesta a deshidratación puede tener un papel similar al de las chaperonas que se inducen en estrés por calor. La función propuesta para las LEA como “chaperonas del estrés osmótico”, sería la de proteger estructuras celulares y macromoléculas, para evitar su desnaturalización y pérdida de actividad, antes de que se lleguen a formar agregados de proteínas. En los ensayos *in vitro*, observamos que las hidrofílicas de plantas no son capaces de proteger la actividad enzimática cuando se adicionan a la enzima deshidratada en niveles tan severos que ya se formaron agregados de proteínas (Reyes *et al.*, 2005). Su actividad protectora previene que se formen estos agregados, pero no pueden desagregar o renaturalizar proteínas.

Debido a que existen varias familias de proteínas LEA en una misma planta, podría haber redundancia funcional entre los miembros de una familia y el análisis *in vivo* por pérdida de función resulta difícil en familias con muchos miembros; sin embargo a la fecha no existen reportes de análisis con mutantes de toda una familia de estas proteínas. Este tipo de enfoque se puede hacer en *Arabidopsis*, ya que se dispone de bancos de mutantes por inserción en la mayoría de los genes, así como una versatilidad de metodologías de silenciamiento post-transcripcional (PTGS, RNAi, microRNAs artificiales). En nuestro laboratorio se llevó a cabo un análisis de secuencia en el genoma de *Arabidopsis* para determinar el número de miembros que posee cada una de las diferentes familias de proteínas LEA, encontrando que las familias más pequeñas son la LEA 1, LEA 4 y LEA 6, que están compuestas por 2 a 3 miembros.

Elegimos como modelo de estudio a la familia AtLEA4, que posee sólo tres miembros en *Arabidopsis*, los cuales denominamos AtLEA4-1, AtLEA4-2 y AtLEA4-5, en donde los números indican el cromosoma en donde se encuentran (*loci* At1g32560, At2g35300 y At5g06760, respectivamente). La mayor identidad de aminoácidos ocurre entre la AtLEA4-1 y AtLEA4-2, comparadas con la AtLEA4-5, y se ha sugerido que son el resultado de un evento de duplicación (Biès-Etheve *et al.*, 2008). El hecho de que esta familia esté formada por tres miembros la convierte en un buen candidato para llevar a cabo análisis por pérdida de función, minimizando la posible redundancia funcional que se ha observado en familias multigénicas. Las proteínas LEA de la familia 4 están muy conservadas en la región N-terminal y se predice que puede formar α -hélices anfipáticas y “coiled-coils”, mientras que en la región C-terminal, la longitud varía de 20 a 80 residuos y no está conservada en secuencia ni en longitud (Dure, 1993; Lupas *et al.*, 1991; Battaglia *et al.*, 2008).

Durante mis estudios de Maestría, en el Programa de Posgrado en Ciencias Bioquímicas, inicié el análisis funcional de uno de los miembros de la familia LEA 4 (AtLEA4-5), cuya expresión a nivel de transcrito y proteína se demostró durante la embriogénesis en condiciones óptimas de desarrollo, así como en respuesta a déficit hídrico en tejidos vegetativos y reproductivos. Se demostró que esta proteína se acumula en la embriogénesis tardía y en respuesta a tratamientos de déficit hídrico, particularmente, en tratamientos de sequía. Se llevó a cabo un análisis fenotípico de plantas transgénicas que sobreexpresaban el gen que codifica para esta proteína. Las plantas mostraron una mayor tolerancia a déficit hídrico, cuantificado como un mayor contenido relativo de agua durante la sequía, así como una mejor recuperación después de un período prolongado de sequía. Esta recuperación fue cuantificada como una mayor acumulación de biomasa, mayor porcentaje de sobrevivencia y producción de un mayor número de botones florales con respecto a los obtenidos con plantas silvestres.

Aunque el estudio a la fecha de estas proteínas ha permitido avanzar en nuestro conocimiento en cuanto a su estructura y función, aún resta por contestar muchas preguntas importantes, algunas de ellas relativas a su participación en la adaptación o ajuste de la planta a una condición de déficit hídrico; ya que, a pesar de que se ha propuesto desde hace varios años su posible contribución en la tolerancia a la sequía, y de que existe cierta evidencia circunstancial en este sentido, a la fecha no existe evidencia directa de ello.

Con estos antecedentes y preguntas en mente se abordó el enfoque de pérdida de función utilizando plantas con niveles disminuidos o nulos para los miembros de la familia LEA4 en plantas de *Arabidopsis*, para determinar si había un fenotipo de sensibilidad en las plantas como resultado de la falta de expresión de uno, dos o los tres miembros de esta familia durante el proceso de deshidratación y la posterior rehidratación de las plantas.

DESCRIPCION DE LA ESPECIE EN ESTUDIO

Arabidopsis thaliana es una planta herbácea anual de la familia Brassicaceae, que tiene hojas en forma de roseta y presenta una clara transición entre las fases de crecimiento vegetativa y la reproductiva. Después de la germinación, el meristemo apical produce hojas con poca elongación entre hojas sucesivas formando una roseta. El número de hojas que se forman depende del genotipo y de las condiciones de crecimiento, y una vez que se induce la transición floral el meristemo apical empieza la producción de flores. Después de la formación de los primordios florales la planta forma un eje caulinar debido a un incremento en la elongación de los internodos entre las hojas más jóvenes y las primeras flores. Las posiciones basales en el eje caulinar de la inflorescencia primaria están ocupadas por un número reducido de hojas caulinares (del tallo), mientras que las posiciones apicales pueden estar ocupadas por un número indeterminado de flores. Los meristemos de las inflorescencias secundarias reiteran el patrón de desarrollo de la inflorescencia primaria y también surgen ramificaciones para formar inflorescencias terciarias en las axilas de las hojas caulinares de los ejes secundarios. También se forman meristemos de inflorescencias de forma basipétala en las axilas de las hojas de la roseta. La inflorescencia es un racimo abierto y el desarrollo de las flores ocurre de manera acropétala (de posiciones basales a apicales). Las flores son perfectas, sin brácteas, y se autopolinizan de manera natural.

La transición a la floración en las plantas anuales es un paso decisivo para dejar de formar hojas en la roseta y empezar la formación de flores. De la respuesta apropiada de las plantas de *Arabidopsis* a las condiciones ambientales en el tiempo adecuado, depende que se logre una producción máxima de semillas. Los factores que afectan de manera muy compleja esta transición son externos o ambientales (como la luz, la temperatura, los nutrientes disponibles) e internos (edad de la planta). El incremento en un fotoperiodo de días largos promueve la floración pero no es un requerimiento absoluto para todos los genotipos, por lo que *Arabidopsis* se considera una planta facultativa de día largo (Bowman, 1994).

MATERIALES Y MÉTODOS

I. Análisis bioinformático de la familia de proteínas LEA4: clasificación, motivos conservados y predicción de estructura secundaria.

Para determinar el número de miembros de la familia LEA4 en *Arabidopsis* se realizó un alineamiento de secuencias con el algoritmo BLAST (tigrblast.tigr.org/) usando como templatado una proteína representante de la familia LEA 4 de algodón (*Gossypium hirsutum*) reportada en los bancos de datos (LEA D-113, número de acceso NCBI), con lo cual se obtuvo el locus At5g06760, que a su vez fue utilizado como templatado para una nueva búsqueda en *Arabidopsis*. De esta segunda búsqueda se encontraron dos secuencias con un porcentaje de similitud significativo correspondientes a los *loci* At1g32560 y At2g35300. En la Tabla 1 se muestran los porcentajes de similitud e identidad a nivel de aminoácidos de las proteínas codificadas por estos genes, que en adelante se referirán como AtLEA4-1, AtLEA4-2 y AtLEA4-5, porque están localizadas en los cromosomas 1, 2 y 5, respectivamente.

PROTEÍNA	AtLEA4-1 vs. AtLEA4-2	AtLEA4-1 vs. AtLEA4-5	AtLEA4-2 vs. AtLEA4-5
% SIMILITUD	66.7	44.3	40.5
% IDENTIDAD	56.3	32.9	29.7

Para llevar a cabo el análisis filogenético de la familia LEA4 se utilizaron proteínas previamente reportadas en la base de datos de SwissProt (expasy.org/cgi-bin/blast.pl), así como proteínas que se encuentran en la base de datos de PFAM clasificadas erróneamente como LEA_1 (pfam.sanger.ac.uk/family?entry=PF03760). También se rastrearon secuencias de cDNAs correspondientes a proteínas ortólogas a la AtLEA4-5 reportadas en los bancos de ESTs para las especies vegetales disponibles utilizando el algoritmo *tblastn* (NCBI BLAST, Protein query vs. Translated database, <http://www.ncbi.nlm.nih.gov/BLAST/>). Adicionalmente se encontraron 5 secuencias en una briofita, en la base de datos COSMOSS (www.cosmoss.org, computational biology resources for the moss *Physcomitrella patens*, Lang *et al.*, 2005). Se seleccionaron únicamente las secuencias que al traducirlas *in silico* en el servidor de Proteómica del Instituto Suizo de Bioinformática (expasy.org/tools/dna.html) correspondieran a proteínas

con un marco de lectura abierto (desde el posible codón de inicio al codón de término). También se verificó que cada proteína tuviera características fisicoquímicas y una predicción de estructura secundaria semejante a los miembros reportados de la familia LEA4; su alta hidrofiliidad, masa molecular y punto isoeléctrico se verificaron utilizando el algoritmo de Kyte y Doolittle (1982) y con el software de análisis de secuencias de proteínas Protean (DNASTar, Inc.). La estructura secundaria predicha para esta familia de proteínas se verificó utilizando el programa PSIPRED y se confirmó que está conformada por una región de α -hélice de aproximadamente 70 residuos, seguida de una región de estructura aleatoria (McGuffin *et al.*, 2000).

Para el análisis estructural de la familia LEA4 de *Arabidopsis* se utilizó el programa COILS (Prediction of coiled-coil regions in proteins, Lupas *et al.*, 1991) para determinar la posible formación de estructuras tipo “coiled-coil” en la región conservada de esta familia de proteínas.

Con un total de 77 secuencias que cumplían con las características que distinguen a la familia LEA 4 se llevó a cabo el análisis de los motivos conservados y de las relaciones filogenéticas dentro de la familia. Los motivos conservados se buscaron con el programa MEME (Multiple Em for Motif Elicitation; Bailey y Elkan, 1994; Bailey *et al.*, 2006) utilizando los parámetros “default” siguiendo el modelo zoops (zero or one per sequence). De este análisis se obtuvieron 10 motivos, los cuales se identificaron a lo largo de la secuencia de la proteína con la herramienta MAST (Motif Alignment y Search Tool; Bailey y Gribskov, 1998).

Para resaltar gráficamente las similitudes en las secuencias de proteínas de las diferentes especies se hizo un alineamiento de 50 especies representativas con el programa T-Coffee (Notredame *et al.*, 2000) utilizando la región conservada de las proteínas y el alineamiento se introdujo al servidor web BOXShade (www.ch.embnet.org/software/BOX_form.html).

Todos los análisis filogenéticos se obtuvieron con el paquete de análisis Phylip (Felsenstein, 2005). Los alineamientos múltiples de las 77 secuencias se realizaron con ClustalW (Larkin *et al.*, 2007) utilizando las proteínas completas o con el “bloque principal conservado”, el cual corresponde a la concatenación de los motivos 4-1-2-6/7 de acuerdo a los resultados del análisis con el programa MEME. Se construyeron las filogenias con Seqboot usando 100 repeticiones para obtener el valor de “Bootstrap” o la confiabilidad en la predicción de cada clado. Las matrices de distancia se obtuvieron con Protdist. Las filogenias basadas en métodos de distancia se realizaron con Neighbor y Fitch. Las

filogenias basadas en el método de “Máxima probabilidad” se construyeron con Proml usando el modelo de sustitución de aminoácidos de Jones-Taylor-Thornton. Los árboles consenso se obtuvieron con Consense. Los árboles filogenéticos se visualizaron y editaron con ATV (A Tree Viewer, Zmasek y Eddy, 2001).

II. Material vegetal y condiciones de crecimiento

Las semillas silvestres de *Arabidopsis thaliana* (ecotipo Columbia) fueron germinadas en cajas de Petri con MS 1X pH 5.7: 4.3 g/L de sales Murashige and Skoog (Caisson laboratories, North Logan, UT, USA), 1% de sacarosa (Research Organics, Cleveland, OH, USA), 0.5 g/L de MES [2-(N-morfolino) ácido etano-sulfónico, Research Organics] y 0.7-0.9% de agar (Research Organics). Previamente a la siembra, las semillas fueron esterilizadas en campana de flujo laminar con etanol absoluto durante 1 minuto y después con una solución de hipoclorito de sodio al 40% y Tritón X-100 al 0.02% (Sigma, St. Louis, MO, USA) durante 8 minutos, se enjuagaron 5 veces con agua destilada estéril. Las semillas se estratificaron a 4°C por 4 días en oscuridad y después se transfirieron a cuartos de crecimiento a 25 °C con un fotoperiodo de 16h/8h, iluminadas con luz blanca a una intensidad de 60-80 $\mu\text{E m}^{-2} \text{s}^{-1}$ y una humedad relativa de 60-65%, hasta que se utilizaron como se indica posteriormente.

II.1 Crecimiento de plantas silvestres y colecta de silicuas para el análisis de los patrones de acumulación de transcritos y proteínas de la familia *AtLEA4* durante la embriogénesis y el establecimiento de la plántula en condiciones óptimas de desarrollo.

Se germinaron semillas de *Arabidopsis* ecotipo Columbia en medio MS 1X y se transplantaron 4 plántulas de 10 días de edad a macetas de 3"x3"x3" con sustrato Metromix 200 (Hummert Int., St. Louis, Mo, USA) y se mantuvieron en condiciones de irrigación óptima con solución nutritiva (5mM KNO_3 , 2.5mM KH_2PO_4 pH 5.6, 2mM MgSO_4 , 2mM $\text{Ca}(\text{NO}_3)_2$, 50 μM Fe-EDTA, micronutrientes {70 μM H_3BO_3 , 14 μM MnCl_2 , 0.5 μM CuSO_4 , 1 μM ZnSO_4 , 0.2 μM $\square\text{NaMoO}_4$, 10 μM NaCl , 0.01 μM CoCl_2 }) durante el período de duración del experimento. Se mantuvieron en un cuarto de crecimiento a 21°C, con una humedad relativa de 70%, un fotoperiodo de 16h/8h con luz blanca a una

intensidad de $60 \mu\text{E m}^2 \text{s}^{-1}$. Las plantas se dejaron crecer hasta la etapa de floración, se marcaron las flores en el día de la antesis y posteriormente se colectaron silicuas de diferentes edades para cubrir las etapas de formación del embrión hasta llegar a la maduración y desecación de las semillas. Los intervalos seleccionados fueron: inflorescencias con botones florales y flores en antesis, silicuas en desarrollo de 1-5 DPA (días post-antesis), 6-10 DPA, 11-15 DPA, 16-20 DPA y semillas secas (post-abscisión). También se sembraron semillas en MS 1X y se colectaron en diferentes intervalos desde la estratificación (3 días a 4°C), semillas germinadas (al emerger la radícula), así como plántulas en desarrollo de 2 días después de la germinación (DAG), 5 DAG, 8 DAG y 11 DAG. El material vegetal se congeló en nitrógeno líquido y se almacenó a -70°C hasta su uso para la extracción de RNA y proteína totales.

III. Tratamientos de estrés abiótico durante la germinación, en plántulas de dos semanas, y en plantas adultas de 6 a 7 semanas de edad.

El análisis de los patrones de acumulación de los transcritos y proteínas de la familia AtLEA4 en respuesta a déficit hídrico se llevó a cabo en dos etapas del desarrollo de las plantas en las que las plantas muestran una alta susceptibilidad a la sequía. La primera de ellas fue en la etapa de plántula de dos semanas de edad (con hojas cotiledonarias y primeros pares de hojas de la roseta), debido a que es una etapa muy sensible a la desecación, donde la sobrevivencia es determinante para el establecimiento y desarrollo de la planta. La segunda fue en plantas adultas de 5 semanas de edad que fueron sometidas a tratamientos de sequía (con los primeros botones florales), en donde las plantas de interés agronómico han mostrado ser muy sensibles a la sequía y, por tanto se considera una etapa determinante para la producción de descendencia de la especie.

Se aplicaron tratamientos de estrés abiótico que se han reportado para la inducción de la expresión de otras proteínas LEA (Horvath *et al.*, 1993; Yamaguchi-Shinozaki y Shinozaki, 1994), así como ensayos en los que las plantas de *Arabidopsis* fueron deshidratadas (van der Weele *et al.*, 2000; Verslues *et al.*, 2006). Las plántulas de dos semanas de edad se cosecharon de las cajas Petri para transferirlas a medio líquido en donde fueron sometidas a tratamientos de estrés. En campana de flujo laminar se trasplantaron de 20 a 25 plántulas por tratamiento a cajas Petri de vidrio estériles, de 6 cm de diámetro, con 15 ml de MS 1X líquido (sin agar) adicionando polietilenglicol 25% (peso

molecular 8000, PEG-8000 Research Organics) o cloruro de sodio (NaCl, 100, 150 y 200 mM, J.T. Baker, México). También se incluyó un tratamiento con ácido abscísico 100 μM (ABA, Sigma) para simular el estrés abiótico y como control se colocaron plántulas en el medio líquido sin agregar ningún inductor de estrés. Las cajas tenían un soporte de tela plástica para que únicamente quedara en contacto la raíz con el medio líquido y se evitara la hipoxia en los tejidos aéreos. Se incubaron tapadas a 25°C en el cuarto de crecimiento y con agitación constante. También se aplicaron tratamientos de temperatura baja, transfiriendo las plantas crecidas en cajas Petri a 25° C a un cuarto frío a 4°C .

Para el caso de los experimentos de sequía con plantas adultas, en nuestro laboratorio se estableció un método en el cual se utiliza un sustrato inerte de muy baja retención de agua que consiste en gránulos de arcilla calcinada (Turface, Hummert). Las plántulas crecidas en cajas Petri desde la germinación y hasta 10-12 días de edad, se trasplantaron a macetas de 3"x3"x3" con 250 g de este sustrato. El diseño experimental fue de 3 réplicas (macetas) con 4 plantas cada una. Las charolas con macetas se colocaron en un cuarto de crecimiento a 21°C, con una humedad relativa de 70%, un fotoperiodo de 16h/8h con luz blanca a una intensidad de 60 $\mu\text{E m}^2 \text{ s}^{-1}$. Debido a que el sustrato es inerte y posee muy baja retención de agua, las plantas se mantuvieron hidratadas a capacidad de campo ($[\Psi_{\text{sustrato}} = - 0.455 \text{ MPa } (\pm 0.70)]$) hasta alcanzar la edad en la que se aplicó el tratamiento de estrés, regándolas siempre a saturación con solución nutritiva hasta el inicio de la sequía, o hasta el final del experimento en el caso de las plantas control. A las cuatro semanas de edad de las plantas, las charolas se transfirieron del cuarto de crecimiento al invernadero para llevar a la cabo las cinéticas de deshidratación en este sustrato con baja retención de agua, lo cual nos permitió realizar los experimentos en condiciones donde la velocidad de deshidratación es más severa que en los cuartos de crecimiento debido a la menor humedad relativa del invernadero (35-50%). Hubo una variación en la tasa de pérdida de agua en el sustrato dependiendo de la estación (seca-Marzo a Mayo, lluviosa- Julio-Septiembre), por lo cual se seleccionó la estación seca del año donde la humedad relativa se encuentra en niveles del 35-50% y la temperatura controlada a 25°C. La baja humedad relativa y la poca retención de agua en el sustrato, permitió imponer una sequía rápida y progresiva para analizar la contribución de los transgenes en los fenotipos cuantificados.

Una vez que se había desarrollado el primer eje caulinar y antes de haberse formado silicuas se iniciaron los tratamientos de sequía (aproximadamente 5 semanas de edad), para lo cual se absorbió previamente el exceso de agua de cada maceta con

toallas de papel y se registró el peso inicial de cada una. Las macetas se transfirieron a una charola sin agua y se evaluó la pérdida de agua de cada maceta pesándolas cada tercer día y rotándolas dentro de la charola para obtener una deshidratación homogénea en todas las macetas. Para analizar la acumulación de transcritos y/o proteínas, se tomaron muestras de los diferentes órganos de las plantas al inicio del experimento, así como a los 2-13 días después de suspender el riego, donde se alcanzaron porcentajes de pérdida de agua en el sustrato (PAS) entre 20 a 83%. Una vez concluidos los tratamientos de sequía, se congeló el material vegetal en nitrógeno líquido y se almacenaron a -70°C hasta el momento de utilizarlos para la extracción de RNA y proteína total. En el caso de plantas adultas, se separaron las flores, raíces y hojas (hojas de la roseta y tallos de los ejes caulinares).

II. 1 Tratamientos para cuantificar fenotipos en condiciones de déficit hídrico en las sobreexpresoras y en las mutantes de la familia AtLEA4

Para determinar si la alteración de los niveles de expresión de la familia AtLEA4 confería un fenotipo comparado con el de plantas silvestres en condiciones de déficit hídrico, se llevaron a cabo ensayos de germinación y experimentos de sequía con plantas adultas. En el caso de la germinación, se sembraron las semillas en medio MS 1X con NaCl 250 mM, o manitol 350 mM, se estratificaron y se transfirieron al cuarto de cultivo a 25°C donde se evaluó la germinación *senso stricto* (emergencia de la radícula) entre 48 h y 2 semanas, dependiendo del tratamiento, hasta que la cinética de germinación llegaba a la fase final o fase de meseta.

Los experimentos de sequía consistieron en dejar deshidratar las plantas por un período de 10-14 días, seguido de un período de rehidratación de 6-10 días. En estos experimentos se utilizaron plantas adultas, en las que se determinaron diferentes parámetros de importancia fisiológica y agronómica. Por una parte se analizó el contenido relativo de agua (CRA) durante el proceso de deshidratación y, en experimentos independientes, se analizaron la acumulación de biomasa, el porcentaje de sobrevivencia y la producción de yemas florales..

Los experimentos con plantas adultas en sequía que se aplicaron fueron semejantes a los descritos para la caracterización de la expresión de la familia AtLEA4. Para analizar el CRA se usaron plantas completas o discos de hojas de la roseta, tanto al inicio del experimento, como a los 4, 8, 10 y 13 días después de suspender el riego, donde se

alcanzaron porcentajes de pérdida de agua en el sustrato (PAS) entre 30 y 80%. El CRA se definió por la fórmula $(P_f - P_s)/(P_t - P_s)$, en donde P_f es el peso fresco de las plantas, P_s el peso seco después de aplicar calor seco en un horno a 80 °C por 3 días y P_t es el peso turgente de las plantas después de flotarlas de 3-24 horas en agua destilada estéril a temperatura ambiente (Tezara *et al.*, 2002).

Para los experimentos de sequía/rehidratación se siguió el mismo procedimiento antes mencionado hasta la deshidratación. Para los experimentos que involucraban la caracterización de mutantes por pérdida de función, las macetas se dejaron deshidratar durante 10 a 12 días, cuando el sustrato alcanzó un potencial hídrico (Ψ_{sustrato}) de $\Psi_{\text{sustrato}} = -4.62$ MPa (± 0.62). En el caso del análisis de la sobre-expresión de la AtLEA4-5 se aplicó un estrés más severo, sin aplicar riego hasta por 14 días, equivalente a $\Psi_{\text{sustrato}} = -6.45$ MPa (± 0.57). Una vez que llegaban al nivel de deshidratación deseado, se rehidrataban con solución nutritiva y se dejaban recuperar de 6 a 10 días, dependiendo de la severidad del estrés, y al cabo de este periodo, se cuantificaba la sobrevivencia, como la capacidad de recuperación (reverdecimiento del tejido fotosintético) después de rehidratar las plantas con solución nutritiva. También se cuantificó el número de yemas florales y axilares por planta, y, posteriormente, se cuantificó la acumulación de biomasa total de cada planta después de la recuperación, para lo cual, se lavaron las raíces, se desecharon las partes senescentes de las hojas, y se secaron en horno a 80°C en bolsas de papel por separado para determinar la biomasa seca total de cada planta.

El potencial hídrico del sustrato (Ψ) se determinó por el método psicrométrico *in situ* ($n=10$, \pm SD), con termopares que se insertaban de 4 a 5 cm de profundidad en el centro de las macetas y se dejaban estabilizar 30 min, para tomar las lecturas por duplicado con ayuda de un microvoltímetro Dew Point (modelo HR-33T, Wescor, Logan, UT) siguiendo el manual de instrucciones del proveedor.

III. 2 Análisis estadístico

Los experimentos de germinación incluían tres réplicas (cajas Petri) con 100 semillas por réplica. El porcentaje de germinación acumulada en el tiempo se ajustó a una curva exponencial sigmoide (curvas dosis-respuesta) de pendiente variable ($Y = \text{Bottom} + \{(\text{Top} - \text{Bottom}) / (1 + 10^{-(\text{LogEC50} - X) * \text{Hillslope}})\}$), también conocida como ecuación logística de 4 parámetros. Bottom es el valor del eje Y que se encuentra en el inicio de la curva (el cual fue restringido a cero); Top es el valor de Y al final de la curva de germinación; LogEC50 es el valor de X cuando la respuesta está en un intermedio entre Bottom y Top; Hillslope

describe la pendiente de la curva. La hipótesis nula era que no había diferencias en dos o tres parámetros del ajuste (Hillslope, LogEC50, Top) de cada curva.

Para los tratamientos de déficit hídrico, el diseño experimental incluyó tres réplicas biológicas (con 4 a 6 plantas por réplica). Para verificar si las diferencias obtenidas entre los tratamientos de sequía y rehidratación en las diferentes líneas transgénicas con respecto a plantas silvestres eran significativas, se realizaron análisis de varianza (ANOVA) de una vía. Se realizó la prueba de homogeneidad de varianzas de Bartlett para cumplir con uno de los supuestos de la prueba de ANOVA y no se encontraron diferencias significativas ($p < 0.05$) en la varianza de los datos dentro de cada grupo. Una vez comprobado que la varianza dentro de cada grupo no era significativa se realizaron comparaciones múltiples entre grupos con la prueba de Dunnett y de Tukey. La hipótesis nula era que no existían diferencias significativas entre los grupos (tratamientos o genotipos).

Para realizar los ajustes de las curvas de germinación, los análisis de varianza y la representación gráfica de las figuras, se utilizó el programa Prism5 (GraphPad Software, La Jolla, CA).

IV. Construcciones, vectores, cepas bacterianas y transformación de plantas.

IV.1 Construcciones con el marco abierto de lectura del gen *AtLEA4-5*.

El marco de lectura abierto (ORF, 477 pb) para el gen *AtLEA4-5* se clonó por una reacción de PCR con la enzima *Taq* DNA polimerasa (Invitrogen) en el vector pBluescriptKS+ (Amp^R; Stratagene) a partir de una biblioteca de cDNA de plantas de *Arabidopsis* (hojas, flores y silicuas) utilizando oligonucleótidos específicos que agregaron los sitios de restricción *NcoI-SaII* (5'-AAA CCA TGG AGT CGA TGA AAG AAA C-3' y 5'-GCG GTC GAC CCG TTT ATC CAG TAT ATC C-3'), utilizando una T_m de 55°C. La inserción del fragmento de interés se verificó por secuencia (Unidad de Secuenciación, Instituto de Biotecnología, UNAM).

Este cDNA se utilizó como templado para hacer PCR y marcar radioactivamente fragmentos internos generados al azar, que se usaron como sondas específicas para hibridar en experimentos tipo northern (ver sección más adelante). El vector se cortó con las enzimas *NcoI-SaII* para subclonarlo al vector binario pBin35S-NOS (Kan^R; Bevan,

1984) para sobre-expresar la proteína correspondiente en plantas y bacterias (4-5 OE). Este cDNA se subclonó también en el vector pHannibal (Amp^R; Commonwealth Scientific and Industrial Research Organization vectors) insertando el ORF del gen *AtLEA4-5* en forma de invertidos repetidos separados por un intrón (*XhoI-KpnI*, para la orientación sentido y *BamHI-XbaI* para la orientación antisentido). La finalidad de esta construcción es que, al transcribirse el precursor, se forme un transcrito que sea autocomplementario y forme una estructura de tallo (gen) y asa (espaciador), el cual será procesado por la maquinaria endógena de RNAi para silenciar el mRNA correspondiente a este gen. Se ha visto que en plantas este tipo de estructuras de RNA de doble cadena son las que inducen en mayores porcentajes el silenciamiento del gen endógeno (con identidad de secuencia al transgen) (Smith *et al.*, 2000). Un fragmento de restricción *NotI* de 4Kb que contiene la construcción descrita de RNAi para el gen *AtLEA4-5*, se subclonó al vector binario pArt27 (RNAi 4-5; Kan^R, Spect^R; Gleave, 1992). La construcción transformada en plantas de *Arabidopsis* en la generación T₃ no fue efectiva para producir el silenciamiento en semillas maduras ni en tejidos vegetativos durante deshidratación severa (ver Discusión). Por esta razón, se seleccionaron líneas homocigotas T₂, en donde el silenciamiento fue evidente, con los mayores niveles de silenciamiento comparado con plantas silvestres, sometiendo plántulas de 2 semanas a estrés osmótico utilizando PEG 25%. Se seleccionaron tres líneas independientes que fueron utilizadas para hacer los análisis fenotípicos.

Para la producción de anticuerpos dirigidos contra la *AtLEA4-5*, el ORF se subclonó en los sitios *NcoI-SalI* del vector de expresión para bacterias pTrc99A (pTrc99A, Amp^R; Amann *et al.*, 1988), en el cual el gen queda bajo el control del promotor del gen *lacZ* de tal manera que se induce con IPTG la sobreproducción de la proteína recombinante de forma nativa en *Escherichia coli* (ver más adelante la sección de producción de anticuerpos).

IV.2 Construcciones con los genes *AtLEA4-1* y *AtLEA4-2*.

El cDNA del gen *AtLEA4-1* se amplificó por RT-PCR con la SuperScriptIII Reverse Transcriptase (Invitrogen) y *Taq* DNA polimerasa utilizando una preparación de RNA total de plantas de *Arabidopsis*, ecotipo Columbia, de dos semanas de edad, sometidas a un tratamiento de estrés por deshidratación en cámara húmeda (8 h, 25°C). Los oligonucleótidos 5'-AAT CCA TGG AAT CGG CGA AAC AGA TAA GCG ATA TGGC-3' y

5'-GCC GTC GAC CGG ATT AGT AGT GAT GAT GAT TAT GAT-3' permitieron amplificar el ORF completo, utilizando una T_m de 70°C. El gen *AtLEA4-2* se amplificó a partir de DNA genómico de plantas de *Arabidopsis* del mismo ecotipo, ya que el gen no posee intrones. Los oligonucleótidos 5'-CGA ATC GAT GTC ATG CGG AGA AGA CGA TGG CAAG3-' y 5'-TCG AAG CTT GAT CTG TCC CGG CGG GTA ATT AGC-3' amplificaron un fragmento de 200 pb, de los 297 pb que constituyen al gen correspondiente, utilizando una T_m de 70°C. Ambos productos de PCR se clonaron en el vector PCRII-TOPO (Amp^R, Kan^R; Invitrogen) y se secuenciaron los insertos para verificar las secuencias (Unidad de Secuenciación, Instituto de Biotecnología, UNAM). Se utilizaron como templados en reacciones de PCR para utilizarlos como sondas en experimentos de tipo "Northern" (ver sección adelante). En todos los casos se realizaron 35 ciclos de amplificación en las reacciones de PCR. También se clonaron en el vector de entrada pENTR/SD/D-TOPO (Kan^R, Invitrogen) usando el sistema de Gateway (Invitrogen) para hacer fusiones traduccionales al extremo amino de la GST para las proteínas *AtLEA4-1* y *AtLEA4-2*, una vez recombinadas con el vector destino pDEST24 (Amp^R, Invitrogen), estas fusiones se utilizaron para la producción de anticuerpos.

Para compensar la falta de mutantes que mostraran una reducción significativa de los transcritos para los genes *AtLEA4-1* y *AtLEA4-2*, reportados hasta ese momento, se diseñó un microRNA artificial (a-miR) dirigido contra los genes de interés (a-miR 4-1/2) para silenciar post-transcripcionalmente sus transcritos utilizando una sola construcción que permite desencadenar el silenciamiento a través de los microRNAs (Niu *et al.*, 2006). Se diseñaron cebadores específicos para introducir a un precursor templado de un microRNA (*ath-miR159a*) mutaciones específicas, por medio de reacciones de PCR, lo cual permitiría la producción del a-miR 4-1/2 para desencadenar el silenciamiento de los genes de interés y no el de su blanco original. Las secuencias de los oligonucleótidos usados fueron: 5' -ATAGATCTTGATCTGACGATGGAAGGACATGGCCAGATCGG-TCAAACATG-3' y 5'-TTGACCCGGGATGGACATGGCCAGTACAGCC-AAAGAAG -3', utilizando una T_m de 65°C. El precursor mutado se clonó en el vector pENTR/SD/D-TOPO y se recombinó con el vector binario pK2GW7 (Spect^R, Kan^R; Karimi *et al.*, 2002). El análisis de las plantas transgénicas que llevaban la construcción del precursor del a-miR 4-1/2 mostraron la expresión del microRNA maduro (5'-UUUGGCUGUACUGGCCAUGUC-3') de manera constitutiva hasta la generación T_3 . El silenciamiento de los transcritos para *AtLEA4-1* y *AtLEA4-2* se determinó mediante la detección de los transcritos y proteínas correspondientes; aquéllas líneas que mostraron

los mayores niveles de silenciamiento se seleccionaron para su caracterización posterior (a-miR 4 – 1/2).

IV.4. Transformación con *Agrobacterium tumefaciens*.

Se transformaron células electrocompetentes de *Agrobacterium tumefaciens* (cepa C58/pGV2260, Rif^R, Amp^R) con los vectores binarios para cada construcción y, en las colonias resistentes a los marcadores correspondientes se verificó la presencia del gen en ambas orientaciones mediante PCR de colonias aisladas. Las bacterias se utilizaron para transformar al vacío 25 plantas adultas silvestres de *Arabidopsis* ecotipo Columbia por el método de infiltración floral (Clough y Bent, 1998). Se cosecharon semillas de la T₀ y se seleccionaron las plantas transgénicas plaqueando aproximadamente 2000 semillas en un medio con 50 µg/ml de sulfato de kanamicina (Sigma) y/o 15-50 µg/ml glufosinato de amonio (BASTA; Finale, Aventis). Las plantas resistentes al marcador se trasplantaron a tierra y se dejaron autofecundar para cosechar las semillas de la T₁. Se sembraron 200 semillas T₁ para estimar el número de inserciones del transgen por la segregación del marcador y se seleccionaron aquellas líneas T₁ que mostraron una segregación en proporción 3:1, indicativo de una sola inserción. Se extrajo DNA genómico de una minipreparación utilizando hojas y flores de las plantas T₁ para verificar la expresión de los transgenes. Se autofecundaron 15 plantas de 5 líneas independientes con esta segregación para obtener cuando menos una planta homocigota para el transgen en la generación T₂ por cada línea, verificándolo también al sembrarlas en medio con los agentes selectivos apropiados. En el caso de las sobre-expresoras de la AtLEA4-5 (OE 4-5) y las mutantes que llevan el a-miR 4- 1/2, se autofecundaron líneas seleccionadas de la generación T₂ para generar plantas homocigotas en la generación T₃. Las líneas T₂, T₃ y F₂ que fueron seleccionadas eran 100% resistentes al marcador de selección: kanamicina (OE 4-5, RNAi 4-5), BASTA (dSpm 4-5, a-miR 4- 1/2) o kanamicina + BASTA (triple silenciada, complementación de la mutante dSpm 4-5).

La expresión de los transgenes se confirmó en las plantas homocigotas por experimentos tipo RT-PCR y northern, utilizando oligonucleótidos específicos para cada transcrito, los niveles de proteína también se confirmaron por experimentos tipo western, utilizando anticuerpos específicos para cada uno de los miembros del grupo. La caracterización fenotípica de las líneas transgénicas se llevó a cabo sembrando las semillas en medio MS 1X (sin kanamicina y/o BASTA) y transplantándolas al sustrato correspondiente en

aquellas líneas que previamente se había confirmado que eran homocigotas para cada construcción. En estas líneas se comprobó que mostraran una expresión constitutiva de la proteína o, en otros casos, que mostraran una disminución o ausencia en la acumulación de la proteína en condiciones de sequía debido al silenciamiento con respecto al control, el cual consistió en plantas silvestres que recibieron el mismo tratamiento.

V. Extracción de DNA genómico y experimentos tipo Southern (hibridación DNA-DNA)

Una vez que se tenían clonadas las tres sondas de los genes que conforman a la familia *AtLEA4*, se confirmó por experimentos tipo Southern (hibridación DNA-DNA) la especificidad de reconocimiento de cada una con el cDNA correspondiente. Para ello se hizo una digestión con *EcoRI*, en el caso de las sondas para los genes *AtLEA4-1* y *AtLEA4-2*, y con *NcoI/SalI* en el caso del gen *AtLEA4-5*. Los productos de la digestión se cargaron por triplicado y se separaron por electroforesis en un gel de agarosa al 1.5% y, posteriormente, se transfirieron a una membrana de nylon (Hybond N+, Amersham). Los productos de las digestiones presentes en esta membrana se hibridaron con cada una de las sondas por separado.

Para obtener como templados algunas regiones genómicas se llevaron a cabo reacciones de PCR utilizando minipreparaciones de DNA genómico obtenidas por el método descrito por Edwards *et al.* (1991). Cuando se utilizó DNA genómico a gran escala éste se extrajo con el método descrito por Taylor *et al.* (1993). Para verificar la mutante en el gen *AtLEA4-5* obtenida del Banco Europeo de Arabidopsis se realizó un experimento tipo Southern utilizando DNA genómico (40 µg) de plantas silvestres y plantas mutantes homocigotas (N122943) digeridas con *Clal*, *HindIII* y *BamHI*, se separaron en gel de agarosa al 1.2 % y se transfirieron por capilaridad a una membrana de nitrocelulosa (N+; Amersham). Se entrecruzó la membrana con luz UV y se hibridó usando protocolos estándar en condiciones de alta severidad, una sonda correspondiente al gen *AtLEA4-5*, que fue obtenida por el método de marcado radiactivo de cebadores al azar (3 millones de cpm incorporadas en la sonda, partiendo de 3,000 Ci/mmol α -³²P-dCTP), a partir de un producto de PCR producto de la amplificación de una región genómica de 800 pb obtenido con oligonucleótidos específicos para el gen *AtLEA4-5* que abarcan la región promotora cercana a la región esperada según lo reportado en el banco de mutantes.

Después de lavar la membrana en condiciones de alta severidad, se expusieron películas Kodak a -70°C en cassettes con pantalla intensificadora (Amersham).

VI. Extracción de RNA y experimentos tipo northern (hibridación DNA-RNA) y RT-PCR.

Todo el material utilizado en la extracción, cuantificación y separación en geles fue esterilizado en la autoclave (120°C , 20 min) y horneado a 80°C durante 24-48 h. Se extrajo RNA total de flores, silicuas y semillas maduras utilizando el protocolo descrito por Vicient y Delseny (1999). Para la extracción de RNA total de plantas de dos semanas y los órganos de plantas adultas se utilizó Trizol (Invitrogen), siguiendo las instrucciones del proveedor. Para la extracción de RNAs pequeños se omitió el pase de lavado con etanol al 75% y se resuspendieron los extractos en formamida.

Se cuantificó la absorbancia a 260 y 280 nm en el espectrofotómetro y se tomaron alícuotas de 15-20 μg por muestra para separarlas en geles de formaldehído-agarosa al 1.5%, se lavaron los geles con SSC 10x (1.5 M NaCl, 150 mM citrato de sodio) durante 30 minutos y se transfirieron por capilaridad a membranas de nylon (Hybond N+, Amersham) de 18 a 24 h en presencia de SSC 20X. Las membranas se entrecruzaron con luz UV y se tiñeron con azul de metileno al 0.03% disuelto en acetato de sodio 0.5M pH 5.2 con la finalidad de verificar que se hubieran transferido correctamente. Se prehibridaron con 15 ml de PSE 1x (7% SDS, 0.3 M NaH_2PO_4 pH 7.2, 1 mM EDTA) durante 30 min a 65°C , se hibridaron con las sondas de DNA marcadas con α - ^{32}P -dCTP por el método de cebadores al azar (2.5-3 millones de cpm incorporadas en la sonda, partiendo de 3,000 Ci/mmol), se purificaron las sondas en columnas Ilustra Microspin G-25 (GE Healthcare). Previamente a la hibridación las sondas se desnaturalizaron por ebullición o con 250 μl de 0.1 M NaOH y diluidas en 10 ml PSE 1x. Las membranas se hibridaron en condiciones de alta severidad a 65°C durante 14-18 h y, posteriormente, se lavaron una vez con 0.5 x SSC-O.1% SDS y dos veces en condiciones de alta severidad (0.1 x SSC-O.1% SDS). Las membranas se emplearon para exponer películas Kodak X-Omat durante 1 a 3 días dentro de cassettes con pantalla intensificadora (Kodak) a -70°C . Para la detección de microRNAs, las muestras de RNA total se resolvieron en geles de acrilamida 15%-urea 8M-Tris-borato/EDTA, se transfirieron en semiseco a membranas de nitrocelulosa (N+; Amersham). Se prehibridó la membrana con Ultra-Hyb Oligo Solution (Ambion). Las hibridaciones se dejaron toda la noche a 42°C , se lavaron dos veces 30 minutos con 2X

SSC, 0.1% SDS, se expusieron en películas Kodak o en pantalla Phosphorimager (GE). La sonda fue marcada con 15 pmol de [γ -³²P]ATP (3,000 Ci/mmol) y se purificó con columnas mini Quick Spin (Roche). La sonda corresponde al antisentido del a-mir 4 - 1/2.

Para los análisis tipo RT-PCR se trató con DNAsa (Invitrogen) 1U/ml y se sintetizó el cDNA a partir de 2 μ g de RNA total con un oligonucleótido dT (1 μ g/ml) de secuencia 24 (T)TVN en donde V= A/G/C y N= A/G/C/T, se incubó a 72°C/10' y se amplificaron las secuencias poliadeniladas con la enzima SuperScriptIII (Invitrogen) a 50°C por 60 minutos, se inactivó la enzima incubando la reacción a 72°C durante 10 minutos, posteriormente se tomó una alícuota de 1/20 de la reacción y se agregó al coctel de reacción de PCR con la enzima Platinum *Taq* polimerasa (Invitrogen) y los oligonucleótidos específicos para cada cDNA. Para la *AtLEA4-1* los iniciadores corresponden a los mismos que se utilizaron para la obtención de la sonda correspondiente, mientras que para *AtLEA4-2* la secuencia corresponde a 5'-GCG GGA TCC CTC GAG ATG CAG TCG GCG AAG GAA AAG -3' y 5'-CGA TCT AGA GGT ACC GTT TTA GAT CTG TCC CGG CGG-3' con una T_m de 65°C. Para *AtLEA4-5* la secuencia de los iniciadores fue 5'- GTA GGA TCC CTC GAG ATG CAG TCG ATG AAA GAA ACA GC3-' y 5'-TCG TCT AGA GGT ACC CCG TTT ATC CAG TAT ATC CCC C3-' con una T_m de 55°C. Todos los iniciadores amplifican el ORF completo del cDNA correspondiente, se aplicaron 20 ciclos de amplificación y se cargó en un gel de agarosa al 1.5% lo correspondiente a 1/6 de la reacción de PCR. El gen *ACT2* se utilizó para normalizar la carga y se amplificó con los oligonucleótidos (5'-GTATGTTGCCATTCAGGCCGTTCTTTCTCT- 3' y 5' - CGACCCGCAAGATCAAGAC-GAAGGA -3'.

VII. Extracción de proteínas y experimentos tipo western (Interacción proteína-anticuerpo)

Con el objetivo de optimizar la obtención de proteínas tipo LEA se utilizaron extractos proteicos totales de plantas silvestres de dos semanas de edad, ensayando tres métodos de extracción de proteína total de tejido vegetativo y dos buffers de solubilización reportados para otras proteínas, incluyendo proteínas LEA (Hurkman y Tanaka, 1986; Colmenero-Flores *et al.*, 1999). El método que dio mejores resultados fue el reportado por Hurkman y Tanaka (1986) con algunas modificaciones. El tejido congelado con nitrógeno líquido se homogenizó en presencia de buffer (sacarosa 0.7 M, Tris Base 0.5 M, HCl 30

mM, β -mercaptoetanol 2% y 12 mg/ml PVPP, polivinilpirrolidona) y fenol saturado en agua, en proporción de 1/3 buffer: 2/3 fenol. Se centrifugó y se tomó la fase orgánica, extrayendo una vez más con la misma proporción de buffer (sin PVPP) y fenol, la fase orgánica se centrifugó y se precipitó toda la noche a $-20\text{ }^{\circ}\text{C}$ utilizando 3 volúmenes de acetato de amonio 0.1 M en metanol. Se centrifugó 15 minutos y la pastilla resultante se lavó con acetona al 80% y se resuspendió en buffer Tris 20 mM, NaCl 50 mM y SDS 0.2%. La cuantificación de proteína se realizó por el método de Bradford (Bio-Rad) utilizando como referencia una curva estándar de albúmina de suero bovino.

Debido al tamaño de las proteínas (10-16.2 KDa), se decidió utilizar geles de poliacrilamida al 16.5% en presencia de un buffer Tris 0.1 M, Tricina 0.1 M y SDS 0.1%, que permiten resolver proteínas en un intervalo de 1 a 100 KDa (Schägger y von Jagow, 1987). Se cargaron alícuotas de 10-20 μg de proteína total y se separaron por electroforesis a 30 mA. Se transfirieron las proteínas en líquido (buffer Tris 25 mM, glicina 0.2 M, metanol 20% v/v) a membranas de nitrocelulosa (Hybond C⁺, Amersham) durante 2 h a 100 V ($4\text{ }^{\circ}\text{C}$). Se verificó la transferencia tiñendo la membrana con rojo de Ponceau, se lavó el colorante con PBS 1x (NaCl 0.27 M, KCl 5.4 mM, Na_2HPO_4 20 mM, KH_2PO_4 3.5 mM) y se bloqueó la membrana 2 h a temperatura ambiente en 10 ml de solución PBSL (PBS 1 X adicionada con 5% de leche), se incubó toda la noche en 10 ml de PBSL con los anticuerpos (inmunopurificados o directamente del suero, ver más adelante) anti-AtLEA4-5 (1:2000), anti-AtLEA4-2 (1:2000) y anti-AtLEA4-1 (1:2000) a 4°C y al día siguiente se lavó 3 veces con PBS 1x durante 10 minutos a temperatura ambiente con agitación constante antes de agregar el segundo anticuerpo (1:5,000-20,000 peroxidasa de rábano anti-conejo, Zymed) incubando 1 h a temperatura ambiente. Se lavó 3 veces más con PBS 1x y se incubó con el sustrato de la peroxidasa (SuperSignal West Pico, Pierce) siguiendo las instrucciones del proveedor. Se utilizó película Kodak X-Omat para detectar la señal de peroxidasa, exponiendo las películas de 15 segundos a 5 minutos dentro de cassettes con pantalla intensificadora (Amersham).

VIII. Obtención de anticuerpos dirigidos contra las proteínas AtLEA4

Los anticuerpos dirigidos a la AtLEA4-5 se produjeron utilizando proteínas recombinantes nativas, mientras que los anticuerpos dirigidos contra la AtLEA4-1 y AtLEA4-2 son proteínas de fusión a GST. Se indujo la expresión de la proteína AtLEA4-5 con 1mM de IPTG (isopropiltio- β -galactopiranosido) aplicado a un cultivo de bacterias (*E. coli*)

transformadas con la construcción que sobre-produce la proteína nativa en un buffer frío (4°C) de Tris 20 mM, NaCl 50mM y 1mM PMSF (fenilmetilsulfonil fluoruro). Se enriqueció la fracción no coagulable sonicando las bacterias concentradas a 50 D.O. por ml, durante 3 minutos con ciclos de 30 segundos cada uno, se centrifugó el sobrenadante y se hirvió durante 10 minutos y después se colocó de inmediato en hielo. La mayoría de las proteínas solubles pertenecen a la proteína recombinante (Jepson y Close, 1995). Para obtener la proteína AtLEA4-1-GST y AtLEA4-2-GST, se resuspendieron las células en PBS frío y se sonicaron tres veces en frío. Se le adicionó TritónX-100 (concentración final 1%), se descartó la pastilla resultante y el sobrenadante se incubó con perlas de glutatión agarosa (Sigma) pre-equilibradas en PBS. Se lavó tres veces con PBS frío y se obtuvo la proteína hirviendo las perlas con SDS 10%. Para la producción de los anticuerpos policlonales dirigidos contra estos dos miembros del grupo AtLEA4, cada proteína de fusión se separó en geles de acrilamida, después de teñir con azul de coomasie, la banda correspondiente a la proteína de interés se separó del gel y se equilibró con agua destilada para emulsificarla con el adyuvante completo de Freund's (Gibco) en una proporción 1:1, se homogenizó exhaustivamente y se usó para inmunizar conejos hembra New Zealand, a través de múltiples inyecciones intradérmicas (Bioterio, IBT).

El título y la especificidad de los anticuerpos se determinó por el método de múltiples inoculaciones utilizando diluciones de los anticuerpos y extractos de semillas de *Arabidopsis* y plántulas tratadas con estrés osmótico. Se titularon los anticuerpos mediante un experimento tipo western incubando diluciones del suero (1:500, 1:1000 y 1:2000) con diferentes concentraciones del extracto de proteína enriquecido por calor (10, 20, 50 y 100 ng). Para determinar si los anticuerpos reconocen específicamente a las proteínas de interés, se extrajo proteína total de semillas maduras y de raíces en sequía de plantas de *A. thaliana*. Dado que hay ESTs reportados para los tres genes obtenidos a partir de RNA de semillas, y plantas en deshidratación, se espera que las tres proteínas se acumulen en estos órganos y, por tanto, ésta fue la fuente para obtener el extracto de proteína total para este análisis. En un experimento tipo western se observó una sola banda con una masa molecular aparente de 20 KDa correspondiente a la proteína AtLEA4-5.

Debido a que en experimentos iniciales tipo western utilizando extractos de proteína total, correspondientes a cinéticas de inducción con plantas de dos semanas de edad, se detectaron bandas de peso molecular superior al correspondiente a las proteínas AtLEA4 y, quisimos descartar que éstas fuesen producto de una detección inespecífica,

por lo que los anticuerpos se inmunopurificaron siguiendo el método descrito por Lillie y Brown (1987). Antes de usarse se titularon por el mismo procedimiento que en la titulación de anticuerpos sin inmunopurificar, pero en este caso las diluciones utilizadas fueron 1:250, 1:125 y 1:60. Como control de la inmunopurificación se utilizaron extractos vegetales de plantas tratadas con una solución de PEG 25% durante 8 horas, en los que ya se habían detectado bandas de alto peso molecular con los anticuerpos sin inmunopurificar.

Para confirmar la especificidad de la detección se realizó un experimento tipo western en el que se utilizaron los anticuerpos inmunopurificados previamente competidos proteína total de semillas secas y de plantas de dos semanas de edad tratadas con PEG 25% durante 8 horas. Se eligieron estas dos etapas del desarrollo, ya que se había determinado en los estudios anteriores que eran condiciones en las cuales se acumulaban las proteínas abundantemente. Con estos experimentos se demuestra que las proteínas AtLEA4-1, AtLEA4-2 y AtLEA4-5 son reconocidas específicamente por anticuerpos generados contra el antígeno utilizado. En el ensayo de competencia los anticuerpos (en 1 ml de buffer PBSL) fueron incubados con diferentes cantidades de proteína (0, 50 ng, 500 ng, 1, 2 y 5 μ g) durante toda la noche con agitación suave y posteriormente la mezcla antígeno/anticuerpo se agregó a 9 ml de buffer PBSL y se incubaron con las membranas de nitrocelulosa que tenían los extractos de semillas y plantas tratadas con PEG 25%.

HIPÓTESIS

La disminución en los niveles de acumulación de las hidrofílicas del grupo LEA4 podría provocar fenotipos de sensibilidad ante déficit hídrico. La función protectora de esta familia de proteínas puede ser mediante la interacción con proteínas durante el proceso de deshidratación y/o en el proceso de recuperación del estrés.

OBJETIVO GENERAL

Evaluar la función de la familia de hidrofílicas AtLEA4 en la respuesta al déficit hídrico.

OBJETIVOS PARTICULARES

- Analizar los patrones de acumulación de transcritos y proteínas de la familia AtLEA4 en varias etapas de la embriogénesis, la germinación y el establecimiento de la plántula bajo riego óptimo; así como su acumulación en tejidos vegetativos y reproductivos bajo déficit hídrico.
- Obtener plantas transgénicas de *Arabidopsis thaliana* que no expresen uno, dos o ninguno de los miembros de la familia AtLEA4 y evaluar su repercusión en el fenotipo de estas plantas durante diferentes etapas del desarrollo, crecidas en condiciones óptimas y en situaciones de estrés hídrico.
- Detectar los blancos de interacción de la proteína AtLEA4-5 en situaciones de déficit hídrico mediante el uso de técnicas *in vitro* e *in vivo*.

RESULTADOS

I. Análisis estructural y filogenético del grupo LEA4: estructura secundaria, motivos conservados y sub-clasificación.

Las proteínas de este grupo están conservadas hacia la región amino terminal (AtLEA4-1, 1-78; AtLEA4-2, 1-74; AtLEA4-5 1-76; Figura 4A), donde se predice la formación de estructuras tipo α -hélice y “coiled-coils” por análisis *in silico* (Lupas *et al.*, 1991; McGuffin *et al.*, 2000; Figura 4B). En contraste, la región carboxilo terminal (AtLEA4-1, 79-134; AtLEA4-2, 75-97; AtLEA4-5 77-158, con una longitud de 56, 23 y 82 residuos, respectivamente) se predice como una región tipo “random coil” (Figura 4B).

Las proteínas de la familia LEA4 se pueden definir con base en la similitud de sus secuencias (Battaglia *et al.*, 2008). Se detectaron 5 motivos conservados, tres de los cuales son comunes a todas las secuencias, mientras que los otros están presentes sólo en las proteínas de mayor longitud. En este trabajo, se llevó a cabo un análisis más detallado del grupo LEA4 partiendo de una colección de 74 proteínas obtenidas de angiospermas, gimnospermas y una briofita. Se encontró una región conservada en el extremo amino (Figura 5) y una región carboxilo mucho más variable, en donde se encuentran motivos repetidos y con rearreglos frecuentes (Figura 6).

Para evitar errores en el alineamiento de la región variable, se utilizó MEME (Multiple em for Motif Elicitation, Bailey y Elkan, 1994), un algoritmo que encuentra motivos con número y longitud determinados y que no se basa en alineamientos. Se encontraron 10 motivos que no corresponden en secuencia exactamente con el reporte anterior, ya que el algoritmo es sensible al número de secuencias, por ellos nos referiremos a la nomenclatura anterior. Los motivos 1, 2 y 4 están distribuidos casi universalmente en todas las secuencias, ya que están en 70, 74 y 66 de las 77 proteínas, respectivamente (Figura 6). La aparición del motivo 4 puede ser mayor a la observada, ya que al ser el motivo que se localiza en el extremo izquierdo, puede ser el resultado de su pérdida en la construcción de las bibliotecas de cDNA. El motivo 6 y 7 aunque son diferentes en secuencia, pueden ser variantes divergentes de un mismo motivo, ya que ocupan posiciones equivalentes en la proteína y son mutuamente exclusivos, 35 tienen el motivo 6 y 32 contienen el motivo 7. El motivo 5 también es común hacia el extremo carboxilo y puede repetirse. Los motivos 3, 9 y 10 aparecen asociados al motivo 7 y no están presentes en las proteínas que llevan el motivo 6 (Tabla 2, Figura 6).

A

```

AtLEA4-1 MQSAKQKISDMASTAKEKRWICQAKADEKAFAMARTKEKSIHQRRRAKEAEANMCMH
AtLEA4-2 MQSAKEKISDMASTAKEKLNIGGARAQGHAEKTMARKKEKLAQEREKSEKAEAKADLH
AtLEA4-5 MQSMKETASNIASAASGMDKTKATLEEKAEKMKTRDPVQKMATQVKEDKINQASMKR
*** *.: *1:1:1**.: | *.: | 1**1 1* 1*1:1* | .:1* 1*1* .: |

AtLEA4-1 MAKAAHAEDKLMA-----KQSHYHVTDHGPHVPPQAPVPAFAPVMGHGYGHNPT
AtLEA4-2 QSKAEHAADAQV-----HGHLPGHSTYPTRA-----T
AtLEA4-5 ETRQHNAAMKEAAGAGTGLGLOTGTHSTIGQVGHGTGTHQMSALP-----GHGTGQLTD
:| :* ** :.

AtLEA4-1 GV-----TSVFPQTYHPTY---PPTGHHNHHHY--
AtLEA4-2 GA-----NYPPGQI-----
AtLEA4-5 RVVEGTAVTDPIGRNTGTGRTHAHTHVGGGGATGYGTGGGYTG

```

B

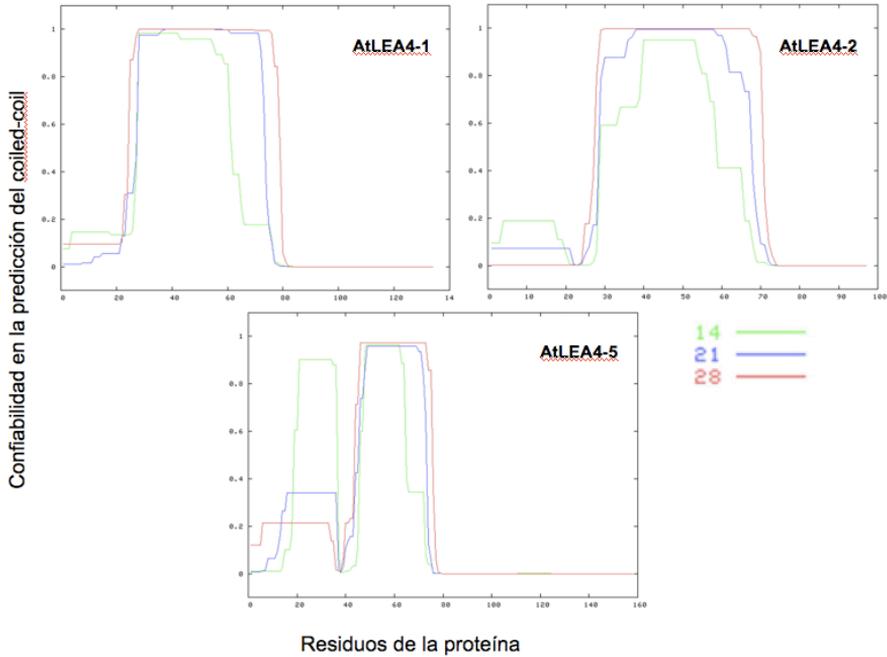


Figura 4. Similitud de secuencias del grupo AtLEA4 y predicción de estructuras tipo “coiled-coil”. A) Alineamiento de las proteínas AtLEA4. (*) = residuos idénticos, (.)(:) = residuos similares. B) Predicción de estructuras tipo “coiled-coil” en la región amino terminal. El eje “y” representa la confiabilidad de la predicción, utilizando tres longitudes del coiled-coil: 14, 21 ó 28 residuos (representados por la línea verde, azul o roja, respectivamente) y el eje “x” indica los residuos de la proteína.

```

Agrostis      1  -----KQHGTTKMDGSAATKATITQVKAIVAAATASHDRLRERGKAVAAAEELDLPVTVREAMEHHIHKHG--
Lolium       1  -----KDDGSAATKATITQVKSIRGAAATASHDRLRERGKAVAAAEELDLPVTVREAMEHHIHKH--
Hordeum2    1  -----MHLAIEVVDASAAATKAKIIEHVAITVAATASHDRLRERGKAVAAAEELDLPVTVREAMEHRLHKRH--
Triticum    1  -----MENAIEVVDGSAATKSTIIEKVAIEVAATASHDRLRERGAAVAAAEALQDQDVAEREDAMSHIHK---
Festuca     1  -----MCAAIEVVDGVSAAVTKETIKKAEIEVAATASHAERERGKAVAAAEALQDQDVAEREDAMSHIHK---
Os06g0110200 1  -----MCAVIEVVDVSAVTKKQVETKADIEVAATASHAERERGKAVAAAEALQDQDVAEREDAMSHIHK---
Saccharum   1  -----MCAVIEVVDVSAATKAREQEKAETVAATASSDIEVVRKAKVAAAKDLQDQDVAEREDAMSHIHK---
Panicum     1  -----MCAVIEVVDVSAATKSKKEQKAEIEVAATASRAEIEVVRGKRVAAAEALQDQDVAEREDAMSHIHK---
Ipomoea     1  -----MCSAIEVVDAAIAALELDILAKTDIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Nicotiana   1  -----MSVIEVVDVSAAGQEVVDILTAAEIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Lactuca     1  -----MSVIEVVDVSAATINVESTNNEIEVAATATQEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Pinus2      1  -----MOPVKNATAAAIEVVDVSAATSEKVDATSAEIEVVSATKGGIIEVVRGAAIEEATIEVAATAEKLAQKQA---
Pseudotsuga 1  -----MOPVKNATAAAIEVVDVSAATSEKVDATSAEIEVVSATKGGIIEVVRGAAIEEATIEVAATAEKLAQKQA---
AtLEA4-1    1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
AtLEA4-2    1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Theilungia  1  -----MISAIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Craterostigma 1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
GlycinePM29 1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Antirrhinum 1  -----KPAIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Sesamum     1  -----CSMIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Brassica    1  -----CSMIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Medicago   1  -----QGAAGKAGETIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Sisymbrium  1  -----CSMIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Phaseolus   1  -----MGGKKAESVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Glycine     1  -----MGGKKAESVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Atrachis    1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Vitis       1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Citrus      1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
AtLEA4-5    1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Fisum       1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Lycopersicon 1  -----MGTGKDAIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Solanum     1  -----MGTGKDAIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Os06g0127700 1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Petunia     1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Gossypium   1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Helianthus  1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Physcomitrella 1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Cryptomeria 1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Pinus1      1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Picea       1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Os06g0124400 1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Triticum2   1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Hordeum1    1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Os04g0589800 1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Sorghum     1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---
Zea         1  -----MCAVIEVVDVSAATSEKLVICQKADIEVAATASKEEIEIVVRKATGEGKQNLIEVAATAEKLAQKQA---

```

Figura 5. Alineamiento múltiple de secuencias de algunas proteínas LEA4 de plantas. El alineamiento se hizo con la región amino de la proteína el programa T-Coffee (Notredame *et al.*, 2000) y se introdujo al servidor Boxshade 3.21 para resaltar los residuos idénticos (negro) y los similares (gris).

Encontramos un región que denominamos Bloque Principal Conservado (BPC), la cual coincide con el alineamiento de PFAM, un programa que utiliza otro algoritmo para encontrar motivos conservados. En PFAM el grupo LEA4 está clasificado erróneamente como LEA_1 PF03760; sin embargo, contiene a los motivos 4, 1, 2 y el 6 o 7, en este orden, lo cual lo identifica como grupo 4 de acuerdo a nuestros criterios de clasificación. No hay espacios en el alineamiento de estos motivos y es fácilmente reconocible puesto que tiene la misma longitud en todas las secuencias. Se utilizó el BPC para reconstruir la filogenia de esta familia. Aunque ninguno de los métodos logró una alta confiabilidad en las ramificaciones -lo que se atribuye a una gran cantidad de secuencias con pocas columnas alineables-, los árboles filogenéticos resultantes son muy similares entre sí, y son consistentes con la filogenia de las especies, aunque hay una ramificación basal que sugiere una duplicación muy antigua (se muestra un árbol representativo en la Figura 9 del artículo anexo). Si se omite la secuencia de trigo (*Triticum aestivum*) que tiene el valor esperado (e) más bajo de $MAST=0.00037$, esta ramificación basal coincide con la presencia de los motivos 6 o 7 con un valor de bootstrap 62/100 (método Neighbor Joining), 80/100 (Fitch) o 99/100 (Fitch) (ver Tabla 2). El método de Máxima Probabilidad

(Proml) también otorga una alta confiabilidad a esta clasificación (con bootstrap 96/100, si se elimina la secuencia de trigo).

Esta clasificación no es un artefacto de la variación entre el motivo 6 y el 7, ya que se obtienen árboles similares utilizando el motivo 2 en lugar del BPC. El motivo 2 es informativo de la historia evolutiva de las proteínas, por lo que representa el núcleo que define a las proteínas LEA del grupo 4, y la adquisición de motivos no ha sido un proceso al azar. La identificación de los motivos 6 y 7 como motivos excluyentes permitieron identificar motivos accesorios o asociados a ellos; así como establecer una ramificación más basal lo cual permitió clasificar a las proteínas LEA4 en dos sub-grupos 4A y 4B.

Es posible deducir que la duplicación de los genes en esta familia es anterior a la evolución de las plantas con semilla, ya que existen proteínas 4A en gimnospermas, en tanto que en una briofita se detectaron secuencias que corresponden a proteínas LEA4B.

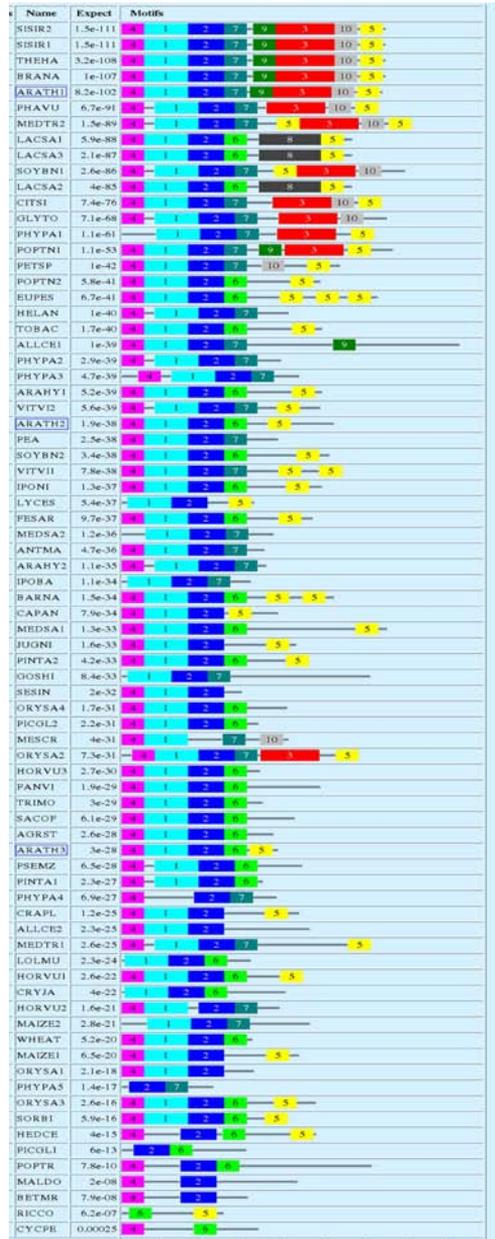


Figura 6. Motivos conservados y su arreglo dentro de las secuencias de proteínas de las LEA4 de plantas. Los motivos se encontraron utilizando MEME (Bailey y Gribskov, 1998) a partir de 77 secuencias obtenidas de bases de datos de proteínas y ESTs. Basado en el bajo valor de e, así como la ausencia del motivo 2, las últimas dos secuencias (RICCO y CYCPE) no se utilizaron en análisis posteriores.

Tabla 2. Clasificación taxonómica de las secuencias de proteínas utilizadas para la búsqueda de motivos. Los motivos están descritos en orden ascendente pero este orden no representa su arreglo dentro de las secuencias (ver Figura 6). Se incluye el número de acceso de las secuencias de proteínas reportadas previamente o los marcos de lectura de las bases de datos de ESTs en la siguiente página. Se excluyeron del análisis filogenético las secuencias que mostraron el valor más bajo con el algoritmo MEME (Seq_25, Seq_46, y Seq_47).

shortName	ncbiTax	CLASS	ORDER	FAMILY	GENUS	1	2	3	4	5	6	7	8	9	10
AGRST	15297	monocots	Poales	Poaceae	Agrostis	X	X		X	X					
ALLCE1	4678	monocots	Asparagales	Alliaceae	Allium	X	X	X	X	X		X	X		
ALLCE2	4678	monocots	Asparagales	Alliaceae	Allium	X	X	X	X	X					
ANTMA	4150	dicots	Lamiales	Plantaginaceae	Antirrhinum	X	X	X				X			
ARAHY1	3817	dicots	Fabales	Fabaceae	Arachis	X	X	X	X	X					
ARAHY2	3817	dicots	Fabales	Fabaceae	Arachis	X	X	X	X	X		X			
ARATH1	3701	dicots	Brassicales	Brassicaceae	Arabidopsis	X	X	X	X	X		X	X	X	
ARATH2	3701	dicots	Brassicales	Brassicaceae	Arabidopsis	X	X	X	X	X					
ARATH3	3701	dicots	Brassicales	Brassicaceae	Arabidopsis	X	X	X	X	X					
BETMR	3554	dicots	Caryophyllales	Amaranthaceae	Beta	X	X	X							
BRANA	3705	dicots	Brassicales	Brassicaceae	Brassica	X	X	X	X	X		X	X	X	
BRANA	3705	dicots	Brassicales	Brassicaceae	Brassica	X	X	X	X	X					
CAPAN	4071	dicots	Solanales	Solanaceae	Capsicum	X	X	X	X						
CITS1	2706	dicots	Sapindales	Rutaceae	Citrus	X	X	X	X	X		X			X
CRAFL	4152	dicots	Lamiales	Linderniaceae	Craterostigma	X	X	X	X						
CRYJA	3368	Coniferopsida	Coniferales	Cupressaceae	Cryptomeria	X	X					X			
CYCP2	59973	dicots	Ericales	Myrsinaceae	Cydamen	X	X								
EUPES	3990	dicots	Malpighiales	Euphorbiaceae	Euphorbia	X	X	X	X	X					
FESAR	4605	monocots	Poales	Poaceae	Festuca	X	X	X	X	X					
GLYTO	3846	dicots	Fabales	Fabaceae	Glycine	X	X	X	X			X			X
GOSHI	3633	dicots	Malvales	Malvaceae	Gossypium	X	X	X	X	X					
HEDCE	58416	dicots	Gerraniales	Rubiaceae	Hedyotis	X	X	X	X	X					
HELAN	4231	dicots	Asterales	Asteraceae	Helianthus	X	X	X	X	X		X			
HORVU1	4512	monocots	Poales	Poaceae	Hordeum	X	X	X	X	X					
HORVU2	4512	monocots	Poales	Poaceae	Hordeum	X	X	X	X	X					
HORVU3	4512	monocots	Poales	Poaceae	Hordeum	X	X	X	X	X					
IPOBA	4119	dicots	Solanales	Convolvulaceae	Ipomoea	X	X	X	X			X			
IPONI	4119	dicots	Solanales	Convolvulaceae	Ipomoea	X	X	X	X						
JUGNI	16718	dicots	Fagales	Juglandaceae	Juglans	X	X	X	X						
LACSA1	4235	dicots	Asterales	Asteraceae	Lactuca	X	X	X	X	X			X		
LACSA2	4235	dicots	Asterales	Asteraceae	Lactuca	X	X	X	X	X		X			
LACSA3	4235	dicots	Asterales	Asteraceae	Lactuca	X	X	X	X	X		X			
LOLMU	4520	monocots	Poales	Poaceae	Lolium	X	X	X	X	X					
LYCES	4107	dicots	Solanales	Solanaceae	Solanum	X	X	X	X						
MAIZE1	4575	monocots	Poales	Poaceae	Zea	X	X	X	X						
MAIZE2	4575	monocots	Poales	Poaceae	Zea	X	X	X	X			X			
MALDO	3749	dicots	Rosales	Rosaceae	Malus	X	X	X	X						
MEDSA1	3877	dicots	Fabales	Fabaceae	Medicago	X	X	X	X	X					
MEDSA2	3877	dicots	Fabales	Fabaceae	Medicago	X	X	X	X	X		X			
MEDTR1	3877	dicots	Fabales	Fabaceae	Medicago	X	X	X	X	X					
MEDTR2	3877	dicots	Fabales	Fabaceae	Medicago	X	X	X	X	X					X
MESCR	3543	dicots	Caryophyllales	Alzooaceae	Mesembryanthemum	X	X	X	X	X					
ORYSA1	4527	monocots	Poales	Poaceae	Oryza	X	X	X	X			X			
ORYSA2	4527	monocots	Poales	Poaceae	Oryza	X	X	X	X			X			
ORYSA3	4527	monocots	Poales	Poaceae	Oryza	X	X	X	X	X					
ORYSA4	4527	monocots	Poales	Poaceae	Oryza	X	X	X	X	X					
PANV1	4539	monocots	Poales	Poaceae	Panicum	X	X	X	X	X					
PEA	3887	dicots	Fabales	Fabaceae	Pisum	X	X	X	X	X		X			
PETSP	4101	dicots	Solanales	Solanaceae	Petunia	X	X	X	X	X					X
PHAVU	3883	dicots	Fabales	Fabaceae	Phaseolus	X	X	X	X	X		X			X
PHYPA1	3217	Bryopsida	Funariales	Funariaceae	Physcomitrella	X	X	X	X	X					
PHYPA2	3217	Bryopsida	Funariales	Funariaceae	Physcomitrella	X	X	X	X	X					
PHYPA3	3217	Bryopsida	Funariales	Funariaceae	Physcomitrella	X	X	X	X	X					
PHYPA4	3217	Bryopsida	Funariales	Funariaceae	Physcomitrella	X	X	X	X	X					
PHYPA5	3217	Bryopsida	Funariales	Funariaceae	Physcomitrella	X	X	X	X	X					
PICGL1	3328	Coniferopsida	Coniferales	Pinaceae	Picea	X	X	X	X	X					
PICGL2	3328	Coniferopsida	Coniferales	Pinaceae	Picea	X	X	X	X	X					
PINTA1	3337	Coniferopsida	Coniferales	Pinaceae	Pinus	X	X	X	X	X					
PINTA2	3337	Coniferopsida	Coniferales	Pinaceae	Pinus	X	X	X	X	X					
POPTN1	3689	dicots	Malpighiales	Salicaceae	Populus	X	X	X	X	X		X	X		
POPTN2	3689	dicots	Malpighiales	Salicaceae	Populus	X	X	X	X	X					
POPTR	3689	dicots	Malpighiales	Salicaceae	Populus	X	X	X	X	X					
PSEMZ	3356	Coniferopsida	Coniferales	Pinaceae	Pseudotsuga	X	X	X	X	X					
RICCO	4107	dicots	Solanales	Solanaceae	Solanum	X	X	X	X	X					
SACOF	4546	monocots	Poales	Poaceae	Saccharum	X	X	X	X	X					
SESI1	4181	dicots	Lamiales	Pedaliaceae	Sesamum	X	X	X	X	X					
SISIR1	3729	dicots	Brassicales	Brassicaceae	Sisymbrium	X	X	X	X	X		X	X	X	
SISIR2	3729	dicots	Brassicales	Brassicaceae	Sisymbrium	X	X	X	X	X		X	X	X	
SORBI	4557	monocots	Poales	Poaceae	Sorghum	X	X	X	X	X					
SOYBN1	3846	dicots	Fabales	Fabaceae	Glycine	X	X	X	X	X		X			X
SOYBN2	3846	dicots	Fabales	Fabaceae	Glycine	X	X	X	X	X					
THEHA	72663	dicots	Brassicales	Brassicaceae	Thellungiella	X	X	X	X	X		X	X	X	
TOSAC	4085	dicots	Solanales	Solanaceae	Nicotiana	X	X	X	X	X					
TRIMO	4564	monocots	Poales	Poaceae	Triticum	X	X	X	X	X					
VITV1	3603	dicots	Vitales	Vitaceae	Vitis	X	X	X	X	X					
VITV2	3603	dicots	Vitales	Vitaceae	Vitis	X	X	X	X	X					
WHEAT	4564	monocots	Poales	Poaceae	Triticum	X	X	X	X	X					

Tabla 2. Continuación...

shortName	sequenceHeading
AGRST	>gi82773414 gb DV966976.1 CRP6709 Creeping bentgrass EST Agrostis stolonifera cDNA clone CRP315T_G03 3', mRNA sequence Frame = +3
ALLCE1	>EST555884 B1935995 cTOD2L13 onion frame +1
ALLCE2	>EST686129 CF449784 ACACW17 onion frame +1
ANTMA	>Antirrhinum majus AJ560281 Frame = +3
ARAHY1	>Arachis hypogaea AA220279.1 GI:7104066 DQ097719.1
ARAHY2	>Gsi-22 Arachis hypogaea AAC03589.1 GI:28194655
ARATH1	>O96272 O96272_ARATH LEA D113 homologue type1 At5g06760
ARATH2	>Q39138 Q39138_ARATH LEA protein Le25 homolog At1g32560
ARATH3	>O96273 O96273_ARATH LEA D113 homologue type2 At2g35300
BETMR	>gi26114695 gb BQ585113.1 E011827-024-002-P05-SP6 MPIZ-ADIS-024-inflorescence Beta vulgaris cDNA clone 024-002-P05 5-PRIME, Frame = +3
BRANA	>AAT77223.1 GI:50429308 Brassica napus Group 1 late embryogenesis abundant protein
BRANA	>83831777 CX280000.1 Brassica napus 16ACDHDS Frame = +3
CAPAN	>TC6258 ORF 39_331 frame +3 pepper Capsicum annuum
CITSI	>gi38052736 gb CF837084.1 Orange Shoot Meristem cDNA Library frame +2 Citrus sinensis
CRAPL	>Q3YMR0 Q3YMR0_CRAPL Group 4 LEA protein
CRYJA	>gi9962207 db AU084224.1 AU084224 Cryptomeria japonica inner bark cDNA clone CC1106 5', mRNA sequence Frame = +3
CYCEP	>gi81488050 emb AJ886946.1 AJ886946 Cyclamen persicum induced calli and somatic embryos clone 33T7.E05, mRNA sequence Frame = +2
EUPES	>gi76862853 gb DV153846.1 CV03093A1H12.F1 CV03-normalized library Euphorbia esula cDNA clone CV03093A1H12.F1 5, mRNA sequence Frame = +2
FESAR	>gi74424198 gb DT680373.1 s130FA13B03DS029_78397 Tall fescue, Festuca arundinacea Schreb, Drought Stress frame = +2
GLYTO	>Q9FNW4 Q9FNW4_GLYTO Glycine tomentella Seed maturation protein LEA 4
GOSHI	>P09441 LEA13_GOSHI Late embryogenesis abundant protein D-113
HEDEC	>gi27911444 gb CB087252.1 h98b07.g1 Hedyotis cernianthoides flower cDNA clone h98b07, mRNA sequence Frame = +2
HELAN	>P46515 LEA11_HELAN Helianthus annuus 11 kDa
HORVU1	>Hordeum vulgare BQ740127 var vulgare frame +3
HORVU2	>TC141523 ORF 65_360 frame +2 Hordeum vulgare barley
HORVU3	>gi13084331 gb BF624252.2 HVSMEa0013C05f Hordeum vulgare seedling shoot EST library HVCdNA0001 (Cold stress)HVSMEa0013C05f, frame = +2
IPOBA	>TC118882 ORF 422_163 frame -2 Potato
IPOBI	>27260544 BJ578716 mixture of flower and flower bud Ipomoea nil cDNA clone jm39c17 3', mRNA sequence Frame = -2
JUGNI	>gi52125110 gb CV196273.1 CGF1003499_A05 Seed coat from mid-season walnut embryos Juglans regia cDNA clone WSC0008_IF_A05 5' Frame = +3
LACSA1	>gi83876902 gb DW078982.1 CLPX14950.b1_L18.ab1 CLP(XYZ) Lactuca penninis cDNA clone CLPX14950, mRNA sequence Frame = +1
LACSA2	>gi84018004 gb DW155274.1 CLVX731.b1_F16.ab1 CLV(XYZ) Lactuca virosa cDNA clone CLVX731, mRNA sequence Frame = +3
LACSA3	>gi83989100 gb DW135209.1 CLSY2723.b1_E09.ab1 CLS(XYZ) Lactuca sativa cDNA clone CLSY2723, mRNA sequence Frame = +2
LOLMU	>gi46504283 db AU247014.1 AU247014 CL Lolium multiflorum cDNA clone CL013G06-5, mRNA sequence Frame = +2
LYCES	>Q00747 LE25_LYCES Protein LE25 tomato
MAIZE1	>Zea mays CD445945 Endosperm frame +1
MAIZE2	>zmwrs48_0B20-011-e12.s3 CF633097_maiz frame -1
MALDO	>gi44614169 gb CN493697.1 Mdfw200724.y1 Mdfw Malus x domestica cDNA clone Mdfw200724 5', mRNA sequence Frame = +3
MEDSA1	>BF633170 NFD46G12DT Alfalfa frame +2 similar a pm29
MEDSA2	>TC94509 ORF 83_387 Alfalfa frame +2
MEDTR1	>TC94296 ORF 122_597 7 frame +2
MEDTR2	>TC96465 ORF 91_623 frame +1 Medicago truncatula alfalfa
MESCR	>TC7047 Mesembryanthemum ORF 512_220 frame -3
ORYSA1	>O52771 Q52771_ORYSA Os06g0324400 Hypothetical protein P0421H01.27
ORYSA2	>O62000 O62000_ORYSA Os06g0327700 Putative seed maturation protein
ORYSA3	>O7XLZ7 O7XLZ7_ORYSA Os04g0589600 OSJNBa0086006.12 protein
ORYSA4	>O8H683 O8H683_ORYSA Hypothetical protein Os06g0110200 OSJNBa0004120.22
PANVI	>gi59873285 gb DN152434.1 5229_G10_M19 Switchgrass callus cDNA library Panicum virgatum cDNA clone 5229_G10_M19 5' Frame = +3
PEA	>Pisum sativum CD859796 Frame +2
PETSP	>EST888206 CV300863 PetuniaRF-5-C02 frame +3
PHAVU	>Phaseolus vulgaris EC911272 negro jamapa drought root frame +3
PHYPA1	>PP014005026R FrameD:1.504:1
PHYPA2	>PP013003013R FrameD:49.348:1 PP_cl.22605.singlet
PHYPA3	>PP_22646_C1 FrameD:1.336:1 nseq=2
PHYPA4	>PP_22605_C1 FrameD:1.96:486:-2 nseq=3
PHYPA5	>PP014002070R FrameD:208.393:-2 PP_cl.22646.singlet
PICGL1	>Q40841 Q40841_PICGL Lea-like protein EMb14
PICGL2	>TC8884 ORF 141_394 Picea frame +3
PINTA1	>Pinus2 taeda DT629371 frame +2
PINTA2	>TC65044 ORF 134_486 frame +2 Pinus
POPTN1	>TC19988 ORF 54_571 frame +3 Populus
POPTN2	>gi60708624 gb DN498434.1 S026H08.5pR imbibed seed cDNA library Populus tremula cDNA Frame = +2
POPTR	>gi60711848 gb DN501658.1 V013B10.5pR male catkins cDNA library Populus trichocarpa cDNA clone V013B10 5', mRNA sequence Frame = +3
PSEMZ	>Q40928 Q40928_PSEMZ Pseudotsuga menziesii (douglas-fir)
RICCO	>gi688714 gb T15059.1 Ricinus communis cDNA clone pcrs596 similar to Lycopersicon esculentum gene le25 protein, Frame = +2
SACOF	>gi35002515 gb CA125027.1 SCOSLR1089F10.g LR1 Saccharum officinarum cDNA clone SCOSLR1089F10 5', Frame = +3
SEGIN	>Sesamum indicum BU668967 frame+2
SISR1	>Sisymbrium irio AAY25119.1 GI:63020472
SISR2	>Sisymbrium irio AAY25119.1 GI:63020472
SORBI	>Sorghum bicolor BG411045 EM1 frame +1
SOYBN1	>Q541U1 Q541U1_SOYBN Glycine max Seed maturation protein LEA 4
SOYBN2	>Q8XES9 Q8XES9_SOYBN Glycine max Seed maturation protein PMZ9
THEHA	>Thehungella halophila ABY86896.1 GI:166359608
TOBAC	>gi76886494 gb DV157487.1 KG9B.001A12F_050628T7 KG9B Nicotiana tabacum cDNA clone KG9B.001A12 Frame = +3
TRIMO	>gi22016470 gb BQ80150.1 WHE2815_B02_D03Zs Triticum monococcum vernalized apex cDNA library Frame = +2
VITV1	>AN63334.1 GI:147806171 Vitis vinifera
VITV2	>TC45434 ORF 39_415 frame +3 Vitis vinifera
WHEAT	>Triticum2 aestivum CJ841147 dormant embryo frame +1

II. Análisis de los patrones de acumulación de transcritos y proteínas de la familia *AtLEA4* durante la embriogénesis y el establecimiento de la plántula, así como en tejidos vegetativos y reproductivos sujetos a déficit hídrico.

La presencia y la abundancia de un transcrito *LEA* y su correspondiente proteína durante una etapa del desarrollo o en respuesta a estrés en un órgano en particular nos puede ofrecer información acerca de su sensibilidad a diferentes tipos de condiciones ambientales adversas, que pueden ser valiosas en la elucidación de su función. Por ello, analizamos los transcritos y proteínas del grupo *AtLEA4* durante la embriogénesis y el establecimiento de la plántula en condiciones óptimas de crecimiento, así como en plantas sujetas a tratamientos de déficit hídrico.

El análisis de la expresión de transcritos de la familia *AtLEA4* se llevó a cabo mediante experimentos de RT-PCR utilizando oligos específicos para cada transcrito. Se utilizó RNA total de plantas silvestres de *Arabidopsis* (ecotipo Columbia) crecidas en condiciones óptimas y los resultados obtenidos mostraron que los transcritos de la familia *AtLEA4* se encuentran en flores y se acumulan gradualmente desde etapas tempranas durante el desarrollo del embrión, aunque la mayor abundancia se detectó en la etapa de semilla madura (post-abscisión). Durante el proceso de la germinación y el establecimiento de la plántula en condiciones óptimas de crecimiento se observó una disminución gradual en la acumulación de estos transcritos (ver Figura 1 en el artículo anexo).

Con la finalidad de obtener información pertinente a los patrones de acumulación de proteínas de todos los miembros de la familia *AtLEA4*, se obtuvieron anticuerpos policlonales en conejos. Para la *AtLEA4-5* se utilizó como antígeno la proteína recombinante nativa y para las proteínas *AtLEA4-1* y *AtLEA4-2* se utilizaron como antígenos a la proteínas recombinantes fusionadas a la glutatión-S-transferasa (GST, pDEST Invitrogen), ya que se ha visto en el laboratorio que las proteínas pequeñas pueden ser poco antigénicas y la fusión con GST aumenta su título (Colmenero-Flores *et al.*, 1997). Con el uso de proteínas recombinantes purificadas de extractos bacterianos y de extractos de proteínas totales vegetales se confirmó la especificidad y la sensibilidad de ambos anticuerpos en ensayos de tipo western.

En el caso de la *AtLEA4-1*, se encontró un patrón de acumulación diferente al de sus correspondientes transcritos, ya que la proteína se acumula abundantemente en flores y en silicuas inmaduras. De manera inesperada, en la semilla madura, ya sea seca o imbibida y durante la germinación, la banda del monómero de 14.9 KDa no se detectó, en su lugar se detectó una banda de mayor peso molecular (*AtLEA4-1-L*), aún utilizando anticuerpos inmunopurificados en estos mismos extractos. Los niveles de

acumulación de la AtLEA4-1-L son mayores en semilla seca comparados con semillas imbibidas y estratificadas o semillas durante el proceso de germinación. Aunque en el panel B de la Figura 1 (ver el artículo anexo) no se detecta la señal de la proteína AtLEA4-1-L en el extracto de semillas en germinación (24h después de incubar a 25°C), una exposición mayor permite su detección en niveles bajos (dato no mostrado). Sin embargo, una vez que las semillas han germinado la proteína AtLEA4-1 se detecta con el peso molecular esperado (14.9 KDa), en tanto que los niveles de acumulación de la proteína AtLEA4-1-L ya no se detectan en el estado de establecimiento de la plántula, reiniciando la detección del monómero con el peso molecular esperado a partir de 2 días después de la germinación (ver Figura 1B en el artículo anexo).

Las otras dos proteínas del grupo AtLEA4 se acumulan en la embriogénesis tardía con un máximo en la etapa de semilla madura y sus niveles disminuyen gradualmente después de la germinación, de acuerdo a su correspondiente patrón de acumulación de transcritos (ver Figura 1 en el artículo anexo). En el caso de la proteína AtLEA4-2 ésta no se detectó en los experimentos tipo western con el peso molecular esperado (10.5 KDa) sino que se detectaron específicamente una o dos bandas de mayor peso molecular (>30 KDa), cuya detección corresponde con el patrón de acumulación de sus transcritos, alcanzando el máximo de acumulación en la fase de semilla madura. Con la finalidad de descartar que fuera una detección inespecífica, se usaron anticuerpos comerciales (GenScript) manufacturados contra un péptido sintético, que abarca una región poco conservada de la familia y con una alta antigenicidad predicha *in silico*. Los resultados obtenidos fueron los mismos. Por lo tanto, obtuvimos dos evidencias independientes que indican que esta(s) banda(s) de mayor peso molecular son detectadas específicamente por los anticuerpos. Por un lado, la detección con estos anticuerpos se compite con un péptido de la AtLEA4-2 y también con la proteína recombinante AtLEA4-2/GST. Por otra parte, estas bandas se dejan de detectar en extractos de plantas transgénicas que expresan un microRNA artificial (a-miR, Figura 7) dirigido a silenciar los transcritos de los genes *AtLEA4-1* y *AtLEA4-2* en condiciones de sequía.

Los patrones de acumulación de transcritos y proteínas se determinaron también en plántulas de dos semanas de edad, sometidas a tratamientos de estrés por 12 horas en medio líquido, en presencia de MS con o sin ácido abscísico (ABA, 100 μ M), polietilenglicol peso molecular 8000 (PEG-8000 25%) o cloruro de sodio (NaCl, 100, 150 o 200 mM). En ensayos de tipo northern se detectaron con sondas específicas para cada transcrito a los tres miembros de la familia, encontrando que los tres transcritos se inducen en respuesta al tratamiento con PEG, y en el caso de la

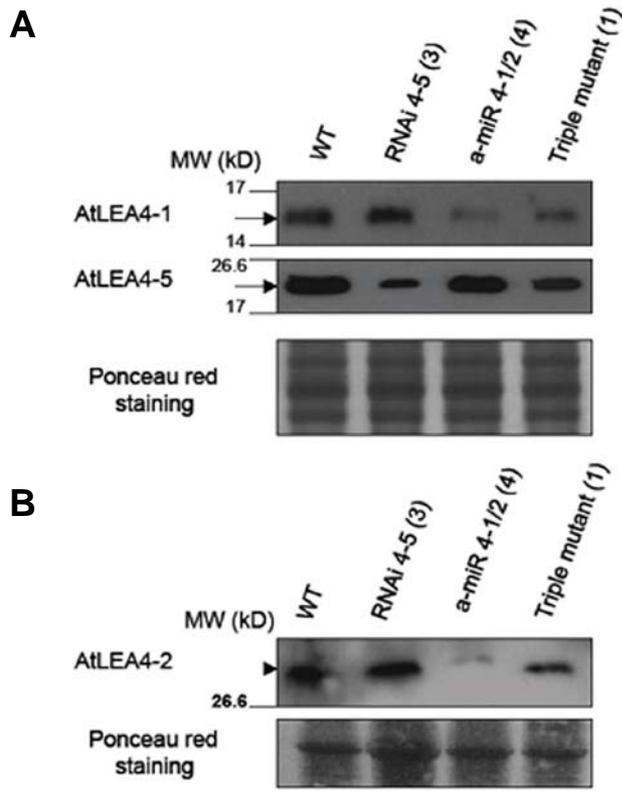


Figura 7. Niveles de reducción de los niveles de proteína en plantas homocigotas sencillas, dobles y triple mutantes por silenciamiento post-transcripcional de la familia AtLEA4 en experimentos de sequía. Análisis tipo western utilizando anticuerpos específicos dirigidos contra cada proteína y extractos proteicos totales (10 μ g) de plantas adultas en tratamiento de deshidratación para mostrar los niveles de silenciamiento en las plantas mutantes comparados con los niveles de las plantas silvestres (WT). Las líneas homocigotas utilizadas en los experimentos fueron: mutantes por el mecanismo de RNAi (RNAi 4-5) en el gen *AtLEA4-5*, mutantes para los genes *AtLEA4-1* y *AtLEA4-2* utilizando un microRNA artificial (a-miR 4 – 1/2), así como las cruzas F₂ de mutantes por RNAi con a-miR (triple mutante). Como control de carga se muestra la tinción reversible con rojo de Ponceau después de la transferencia. La cabeza de flecha muestra la señal de mayor masa molecular a la esperada (aprox. 30 KDa), obtenida específicamente con los anticuerpos dirigidos contra la AtLEA4-2.

AtLEA4-5 se inducen también en respuesta a ABA y NaCl. El aumento en la concentración de NaCl provocó un gradual incremento en la acumulación de los transcritos para el gen *AtLEA4-5*, mientras que para los otros dos no se detectó un incremento significativo (Figura 2A del artículo anexo). La mayor sensibilidad del análisis por RT-PCR mostró que los tres transcritos se acumulan en respuesta a los diferentes tratamientos de estrés, siendo el transcrito para el gen *AtLEA4-5* el más abundante en todos los casos. En el caso de la determinación de los niveles de proteína, encontramos que sus patrones de acumulación son similares a los mostrados por los transcritos en respuesta a PEG, pero no así en respuesta a tratamientos con ABA y NaCl, en donde se encontraron diferencias entre ambos patrones (Figura 2A del artículo anexo). Tal fue el caso en respuesta a estrés por sal, en donde aunque los transcritos para los genes *AtLEA4-1* y *AtLEA4-2* se detectaron en bajas cantidades, los niveles de las proteínas correspondientes en esta condición claramente mostraron un incremento. En el caso del tratamiento con ABA no hubo cambios en la acumulación de las proteínas *AtLEA4-1* y *AtLEA4-2* con respecto a la condición control. El hecho de que la *AtLEA4-1* se observara acumulada a pesar de los bajos niveles de sus transcritos sugiere la participación de mecanismos de control post-transcripcional que modulan los niveles de acumulación de la proteína. En el caso de la *AtLEA4-5*, los resultados obtenidos por ensayos de tipo western indican que se detecta una abundante acumulación en respuesta a los tratamientos con ABA y PEG, (Figura 2 artículo anexo). Sin embargo, se puede observar que los niveles de acumulación de la proteína *AtLEA4-5* en respuesta a las concentraciones de NaCl utilizadas fueron menores comparados con los obtenidos por los tratamientos de ABA, aún cuando los transcritos correspondientes se acumularon en niveles similares en ambos tratamientos (ABA vs 200 mM NaCl, Figura 2 artículo anexo), lo cual sugiere nuevamente la participación de mecanismos de regulación post-transcripcional en la síntesis de esta proteína.

De manera semejante a los análisis anteriormente descritos, para el caso de la proteína *AtLEA4-2* no se detectó una banda con el peso molecular esperado, sino que en respuesta a PEG se observaron dos bandas de alto peso molecular, y un aumento en la acumulación de ambas en respuesta al incremento en la concentración de sal. Estos dobletes de alto peso molecular se observaron en geles de SDS-tricina, con sistema discontinuo de pH, que son los geles que utilizamos de manera rutinaria para separar las proteínas de los extractos vegetales ya que permiten la migración correcta de proteínas pequeñas(1-30 kDa), como en el caso de la *AtLEA4-1* y *AtLEA4-5* que migraron de acuerdo al peso esperado (Schägger y von Jagow, 1987). Aunque con

una menor resolución el doblete de alto peso molecular también se detectó utilizando geles de SDS-glicina. Este fenómeno se repitió con cualquiera de los anticuerpos disponibles dirigidos contra la AtLEA4-2. La incubación de los extractos vegetales y las proteínas recombinantes con el suero preinmune no concuerda con las bandas obtenidas con el suero inmune para la AtLEA4-2, lo que sugiere que los dobletes de alto peso molecular corresponden a homo-oligómeros que la proteína pudiera estar formando por interacciones intramoleculares.

En conclusión, los experimentos que se describen arriba indican que la síntesis de estas proteínas está sujeta a una regulación no sólo a nivel transcripcional sino también a nivel postranscripcional o traduccional. Por otro lado, la acumulación diferencial de los transcritos y proteínas de esta familia en plantas crecidas bajo condiciones óptimas, durante el desarrollo del embrión, durante la germinación y el establecimiento de la plántula; al igual que sus patrones de acumulación en condiciones de déficit hídrico, sugiere (como también lo indica el análisis filogenético) que la función de estas proteínas se ha diversificado después de su duplicación y no necesariamente se sobrelapan sus funciones en diferentes condiciones osmóticas de la célula, etapa de desarrollo o tejidos de la planta.

III. La mutante por inserción en el gen *AtLEA4-5*, las mutantes por silenciamiento post-transcripcional (PTGS) en los genes *AtLEA4-1* y *AtLEA4-2* y las plantas de las cruas por PTGS muestran fenotipos de sensibilidad en respuesta a estrés osmótico durante la germinación y en tratamientos de sequía en plantas adultas.

El análisis por pérdida de función es el enfoque que permite contar con información relevante sobre la participación de las proteínas del grupo AtLEA4 en diferentes procesos, tanto en condiciones óptimas de crecimiento como en situaciones de limitación de agua. Por ejemplo, durante y/o después de una situación de déficit hídrico, durante la germinación en condiciones de estrés osmótico; o bien durante un proceso de deshidratación/rehidratación en plantas adultas o durante el desarrollo normal de la semilla. El análisis fenotípico se llevó a cabo durante la germinación, en condiciones normales o de estrés iónico/osmótico (NaCl 250 mM, manitol 350 mM); así como, de plantas sometidas a tratamientos de deshidratación/rehidratación en estado adulto, utilizando una severidad de deshidratación menor a la que se usó con las plantas sobreexpresoras de la *AtLEA4-5* (ver Figura 3 del artículo anexo), puesto que se esperaba que las mutantes deficientes en la producción de esta familia de proteínas fueran más sensibles a este tratamiento.

Para determinar si los fenotipos observados se podían atribuir en particular a la disminución de la expresión de esta familia de proteínas durante déficit hídrico, se determinó previamente que las líneas transgénicas utilizadas en los experimentos tuvieran una sola inserción del transgen y fueran homocigotas para el o los transgenes. Se verificó que expresaran la construcción que desencadena el silenciamiento específico de los genes blanco en el caso de mutantes por silenciamiento post-transcripcional o que tuvieran la inserción en el sitio esperado del genoma en el caso de la mutante por inserción y, en consecuencia, tuvieran abatidos los niveles de expresión de proteína de los respectivos miembros de la familia LEA4. A continuación se detallan los experimentos para cada tipo de mutantes.

1) Mutante por inserción en el gen *AtLEA4-5* y su complementación con una construcción *35S:AtLEA4-5:NOS* en plantas transgénicas estables

Se obtuvo una mutante por inserción de un transposón no-autónomo (Enhancer/Suppressor-mutator, dSpm) en el promotor del gen *AtLEA4-5*, la cual está disponible a través del Banco Europeo de Mutantes de Arabidopsis (NASC, Número de acceso N122943, Tissier *et al.*, 1999, Figura 8A) que lleva como marcador de selección de la inserción, el gen que otorga resistencia a glufosinato de amonio (BASTA). Se confirmó el sitio de inserción a través de experimentos tipo Southern usando una sonda específica que abarca una parte de la región promotora del gen cerca del sitio esperado de la inserción, hasta el final del marco abierto de lectura del gen (Figura 8C). Se utilizó DNA genómico de plantas silvestres y plantas mutantes para digerir con enzimas que cortan en sitios aledaños a la inserción del transposón (Figura 8B). Para confirmar que la inserción afectara la producción de la proteína se llevaron a cabo ensayos de tipo northern y western. Se encontró una severa reducción en los niveles de transcritos y proteínas comparados con los de plantas silvestres, utilizando extractos de semillas maduras y de raíces de plantas adultas en tratamientos de sequía (ver Figura 4 del artículo anexo). Esta mutante se utilizó para analizar su fenotipo durante la germinación en medio con o sin la adición de NaCl (250 mM) y manitol (350 mM), así como en plantas adultas sujetas a tratamientos de deshidratación-rehidratación. Hubo un retraso en el inicio de la germinación y una menor tasa de germinación en condiciones control ($P < 0.0011$), una significativa reducción en la tasa de germinación y un menor porcentaje final de germinación en condiciones de déficit hídrico ($P < 0.0001$), lo cual sugiere que la proteína *AtLEA4-5* es necesaria para una eficiencia de germinación óptima en estrés osmótico y iónico (ver Figura 5 del artículo anexo).

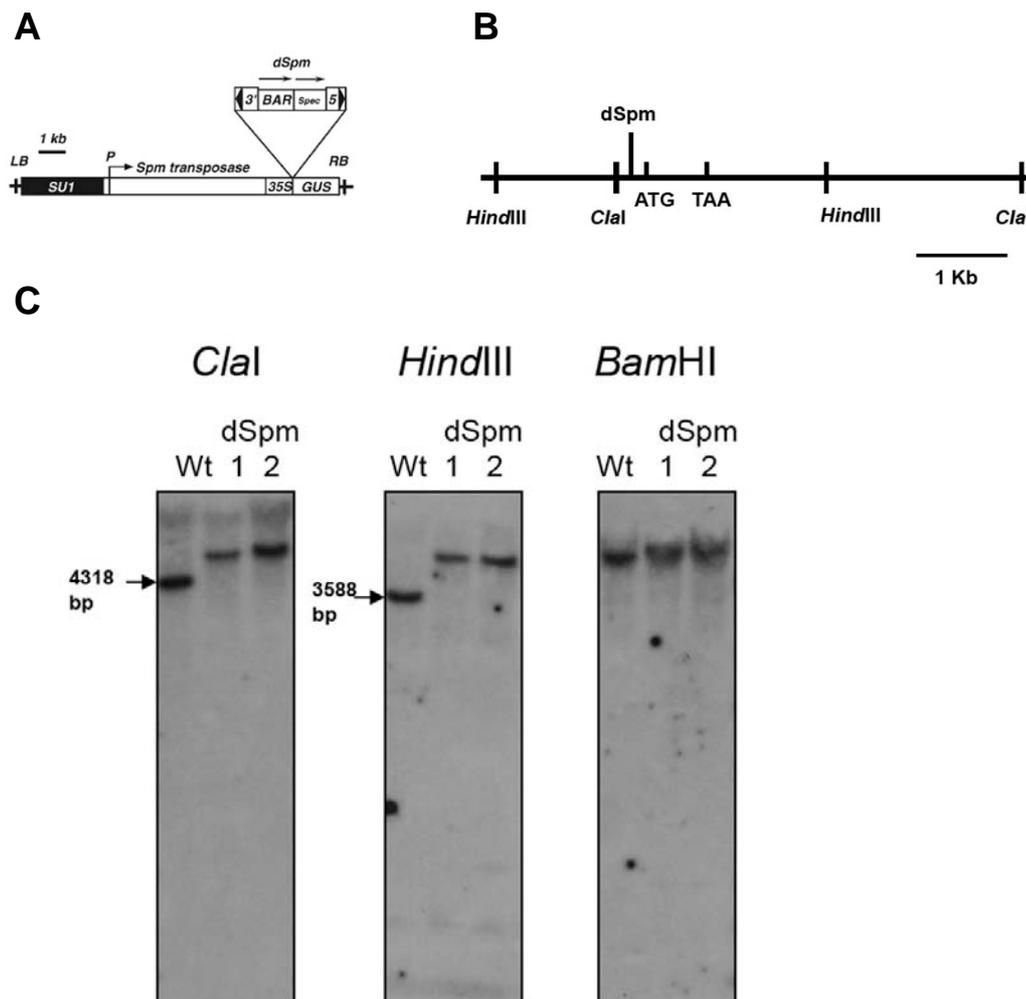


Figura 8. Descripción de la mutante por inserción de un transposón (dSpm 4-5) en el gen *AtLEA4-5*. A) Construcción utilizada para seleccionar las transposiciones estables del transposón dSpm en la colección de mutantes SLAT del John Innes Centre (imagen reproducida de Tissier *et al.*, 1999). Se introdujo un elemento Enhance/Supresor-mutador (En/Spm) en plantas de Arabidopsis en una construcción de T-DNA, la cual lleva un transposón Spm no-autónomo defectuoso (dSpm) con el gen de resistencia al herbicida fosfotricina o BASTA (BAR), un cassette de expresión de una transposasa y un marcador de contra-selección. B) Localización genómica de la inserción estable del dSpm en la región promotora del gen *AtLEA4-5* (208 pb del codon de inicio) y sitios de restricción utilizados para el análisis de tipo Southern. Las semillas mutantes se obtuvieron del Banco Europeo de Mutantes de Arabidopsis (NASC, Número de acceso N122943). Se crecieron plantas de cada semilla y la progenie de cada planta se colectó por separado. C) Análisis tipo Southern utilizando una sonda específica para detectar la inserción del transposón dSpm en la región esperada del promotor del gen *AtLEA4-5* y DNA genómico (40 mg) extraído de plantas silvestres (WT) y mutantes de dos líneas T4 independientes (dSpm 1, 2). El DNA genómico fue digerido con las enzimas *Clal* o *HindIII*, que dieron señal en el tamaño esperado para las plantas silvestres y de un mayor tamaño en las líneas mutantes, de acuerdo con el tamaño y la orientación del transposón. Se utilizó como control negativo la digestión con la enzima *BamHI*.

Para evaluar la participación de la proteína AtLEA4-5 en la capacidad de la planta para mantener la producción de botones florales y axilares bajo restricción de agua, se compararon las plantas mutantes y plantas silvestres, crecidas en la misma maceta, para que la sequía fuera lo más comparable posible. Se crecieron plantas en condiciones óptimas de desarrollo hasta el inicio de la floración, en este punto, se sometieron a un tratamiento de deshidratación hasta que el sustrato alcanzó un potencial osmótico (Ψ_{sustrato}) de $-4.62 (\pm 0.62)$, -cabe recordar que el tratamiento de deshidratación en este experimento es menos severo que el utilizado para los experimentos que se llevaron a cabo con las plantas sobre-expresoras de la *AtLEA4-5* en fondo silvestre-. Las plantas se re-hidrataron con solución nutritiva y se dejaron recuperar por un periodo de 6 días, al cabo de este tiempo se cuantificó el número de botones florales y axilares, antes de cosechar las plantas completas para determinar la acumulación total de biomasa (una vez que las plantas se secaron en un horno a 80°C). Se encontraron diferencias significativas ($P < 0.0001$) comparadas con plantas silvestres en la acumulación de biomasa después de la recuperación del periodo de sequía, así como en la producción de botones florales y axilares en las mutantes bajo estas condiciones ($P < 0.0001$).

Para verificar que los fenotipos detectados en las plantas mutantes se deben a la ausencia del gen funcional *AtLEA4-5*, se llevaron a cabo experimentos de complementación, transformando las plantas mutantes por inserción del transposón dSpm (fondo mutante *AtLEA4-5::dSpm*), con una construcción que expresa el ORF del gen *AtLEA4-5* bajo el promotor constitutivo 35S, cuyo marcador de selección confiere resistencia a kanamicina (*AtLEA4-5::dSpm*, *35S::AtLEA4-5::NOS*).

El análisis fenotípico de las plantas transformadas en estado homocigoto para el transgen (generación T_3) mostraron que la sobre-expresión del gen *AtLEA4-5* puede complementar la mutación del gen endógeno durante la germinación bajo condiciones de baja disponibilidad de agua. Se obtuvo una mayor tasa de germinación en las semillas de plantas silvestres y en las plantas complementadas con la construcción comparación con la obtenida con las semillas de plantas que llevan la inserción del transposón (*AtLEA4-5::dSpm*, ver Figura 5 del artículo anexo). Este experimento es indicativo de que el fenotipo de sensibilidad observado en las mutantes por inserción en el gen *AtLEA4-5* se debe al abatimiento en la expresión de este gen, que en condiciones normales se expresa muy abundantemente en las últimas etapas de la embriogénesis. También se observó la complementación del fenotipo en plantas adultas después de haber estado sometidas a tratamientos de sequía y rehidratación (ver Figura 6 del artículo anexo). Cabe notar que en este tipo de experimento, donde la deshidratación no fue tan severa como la que se utilizó para caracterizar a las líneas

sobre-expresoras del gen *AtLEA4-5* en fondo silvestre (Col, 35S::*AtLEA4-5*::NOS), no hubo diferencias significativas entre la complementación del gen *AtLEA4-5* en el fondo mutante (dSpm::*AtLEA4-5*, 35S::*AtLEA4-5*::NOS) con el desempeño de las plantas silvestres. Se utilizó como control positivo a las plantas que sobre-expresan este mismo gen en el fondo silvestre (Col, 35S::*AtLEA4-5*::NOS).

2) Mutantes por silenciamiento post-transcripcional en los genes *AtLEA4*.

Debido a que no se encontraron mutantes para los genes *AtLEA4-1* y *AtLEA4-2*, con una disminución significativa de los transcritos y proteínas correspondientes, se realizó una construcción de un microRNA artificial (a-miR) para su silenciamiento post-transcripcional (PTGS). La homología entre los transcritos de los genes *AtLEA4-1* y *AtLEA4-2* permitió el diseño de un a-miR (a-miR 4-1/2), que podría tener como blanco de silenciamiento a ambos transcritos, utilizando una región de máxima identidad de nucleótidos entre ellos (Figura 9). El precursor del a-miR 4-1/2 se expresó bajo el promotor 35S para favorecer un silenciamiento eficiente de los genes de interés y las plantas transgénicas llevan el marcador de selección que confiere resistencia a BASTA. Se seleccionaron aquellas líneas transgénicas que mostraron la mayor acumulación del a-miR para validar el silenciamiento esperado. El a-miR 4-1/2 maduro se detectó por medio de ensayos tipo northern para RNAs pequeños utilizando plantas de dos semanas de edad de líneas T₃ crecidas bajo condiciones óptimas (ver Figura 7A del artículo anexo). Los resultados de los ensayos tipo western utilizando extractos de proteína total de plantas adultas de líneas T₃ seleccionadas bajo tratamientos de sequía mostraron que el a-miR es funcional y puede silenciar específicamente los transcritos correspondientes a los genes *AtLEA4-1* y *AtLEA4-2* (ver Figura 7B del artículo anexo). Una vez confirmada su funcionalidad se caracterizaron fenotípicamente estas líneas, aplicando los mismos tratamientos descritos para la mutante por inserción en el gen *AtLEA4-5*. En los ensayos de germinación el mayor efecto se observó en presencia de NaCl (250 mM), ya que las semillas de la mutante silenciada en los genes *AtLEA4-1* y *AtLEA4-2* mostraron un 20% de germinación comparado con un 70% observado en las semillas silvestres. Cuando se germinaron en presencia de manitol (350 mM) no se obtuvieron diferencias en el porcentaje final de germinación pero la doble mutante mostró una menor tasa de germinación (P<0.0001, ver Figura 6 del artículo anexo). Asimismo, se detectó una mayor susceptibilidad a la deshidratación–rehidratación en experimentos con plantas adultas, ya que la mutante mostró una reducción significativa en el número de botones en comparación con las plantas silvestres (ver Figura 6 del artículo anexo 1 y Tabla 3).

Para poder obtener una triple mutante que estuviera afectada en la producción de los tres miembros de la familia de genes *AtLEA4*, y debido a que ambas mutantes, la de inserción y la de PTGS son resistentes a BASTA, generamos una mutante por PTGS para el gen *AtLEA4-5*, que llevara como marcador de selección al gen de resistencia a kanamicina, utilizando como herramienta la interferencia de RNA (RNAi). Para ello, se transformaron plantas que expresaran bajo el promotor 35S a la región codificante del gen *AtLEA4-5* en forma de repetidos invertidos (RNAi 4-5) y que desencadenaran la maquinaria endógena de la planta de silenciamiento por RNAi debido a la generación de siRNAs (Figura 9B).

A

1) Search for a region of 22 nucleotides with maximal identity between *AtLEA4-1* and *AtLEA4-2* transcripts.

```

AtLEA4-1 AUGCAAUCGGCGAAACAGAAGAUAGCGAUAUGGCCUAGUACAGCCAAGGA
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
AtLEA4-2 AUGCAGUCGGCGAAGGAAAAGAUCAUGGACAGUGACAUGGCCAGUACGGCCAAGGA
  
```

2) Search for a region with the least number of unpairings with the microRNA template (*ath-miR159a*).

```

AtLEA4-1/2 mRNA      5' GACAUGGCCAGUACAGCCAAG 3'
                   ||| | | | | | | | | | | | | | | | | | | | | |
ath-miR159a        3' AUCUCGAGGGGAAGUUAGGUUU 5'
  
```

3) Introduce point mutations in the microRNA template (*ath-miR159a*) to eliminate unpaired sites.

```

AtLEA4-1/2 mRNA      5' GACAUGGCCAGUACAGCCAAG 3'
                   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a-miRLEA4-1/2      3' CUGUACCGGUC AUGUCGGUUU5'
  
```

4) From the original stem-loop precursor sequence in *ath-miR159*,

```

ath-miR159a*  5' -- UAGAGCUCCUU AA AGUUCAAA-----
                ||||| ||||| ||||| ||||| |||||
ath-miR159a  3' -- AUCUCGAGGGA AG UUAGGUUU-----
                ||||| ||||| ||||| ||||| |||||
                U G GCU UUA A U
ath-miR159a*  5' -- A A CC A G UCAAA-----
                | | || | |
a-miRLEA4-1/2 3' -- U U GG U U GGUUU-----
                C G ACC UCA G C
  
```

5) restore the secondary structure with the corresponding changes in the sequence of *ath-miR159**.

```

ath-miR159a*  5' -- GACAUGGCCAG AA CCGUCAAA-----
                ||||| ||||| ||||| ||||| |||||
a-miRLEA4-1/2 3' -- CUGUACCGGUC AU GUCGGUUU-----
                ||||| ||||| ||||| ||||| |||||
  
```

B

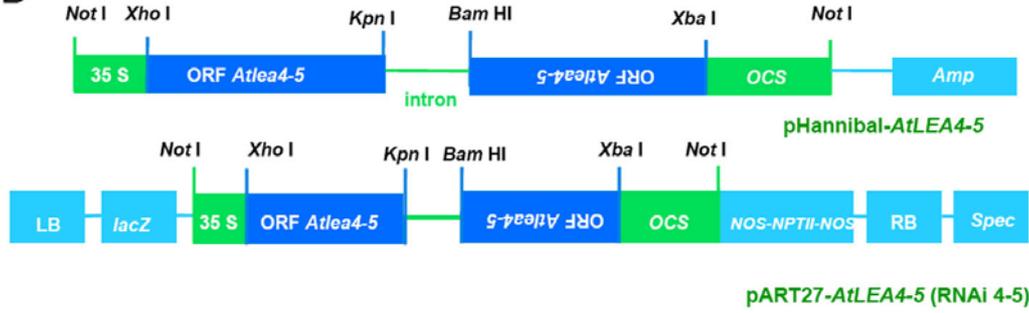


Figura 9. Diseño de un microRNA artificial para silenciar post-transcripcionalmente a los genes *AtLEA4-1* y *AtLEA4-2* (A) y clonación de la construcción que desencadena el silenciamiento del gen *AtLEA4-5* por medio de RNAi (B), con una construcción que expresa el marco abierto de lectura (ORF) en forma de invertidos repetidos separados por un intrón, para desencadenar el silenciamiento por siRNAs que específicamente tendrían como blanco a los transcritos del gen *AtLEA4-5*.

Algunas líneas se seleccionaron por sus mayores niveles de silenciamiento del gen endógeno, al ser comparados con los niveles de acumulación del gen en plantas silvestres (líneas 3-5, ver Figura 7C del artículo anexo 1). Sin embargo, el silenciamiento con esta construcción sólo fue efectivo en las generaciones T_1 y T_2 , por lo que todos los experimentos con esta construcción se llevaron a cabo con plantas homocigotas en la generación T_2 . Estas líneas homocigotas fueron caracterizadas fenotípicamente antes de utilizarlas en la generación de la triple mutante (Figura 10). Estos análisis se llevaron a cabo con plantas adultas en tratamientos de deshidratación similares a los descritos anteriormente, hasta que el sustrato alcanzó un potencial hídrico $\Psi_{\text{sustrato}} = -4.62 (\pm 0.62)$ en condiciones de invernadero. Los resultados de estos análisis mostraron que las líneas seleccionadas acumularon menos biomasa y recuperaron un menor número de botones florales después de la rehidratación. Además se mantuvieron en condiciones óptimas hasta que completaran su ciclo de vida, con la finalidad de determinar la producción de semillas, observando que produjeron una menor cantidad de semillas que las plantas silvestres que también sufrieron el mismo tratamiento de sequía (Figura 10). Las líneas seleccionadas de RNAi (línea 3, ver Figura 7C del artículo anexo 1, T_2 RNAi 4-5) y de a-miR (líneas 3 y 4 de la Figura 7B del artículo anexo, T_2 a-miR 4 – 1/2) se usaron como parentales para hacer las cruzas que permitieran obtener una triple mutante; así se obtuvieron semillas de las cruzas y posteriormente se crecieron las plantas de la F_1 y se autofecundaron para obtener las semillas de la generación F_2 con el fin de seleccionar las líneas homocigotas para ambos transgenes. Se seleccionaron líneas homocigotas con 100% de resistencia a ambos marcadores y que mostraron los menores niveles de acumulación de las proteínas correspondientes en tratamientos de sequía (Figura 7). Estas cruzas no se analizaron en la germinación ya que previamente habíamos observado que las semillas de las plantas que llevan la construcción RNAi *AtLEA4-5* no presentan un abatimiento de la proteína en la semilla madura en la generación T_3 . También se hicieron los experimentos de sequía con plantas silvestres y plantas de las cruzas, crecidas en la misma maceta. Los resultados indicaron que las plantas homocigotas obtenidas de la cruce tienen un fenotipo de sensibilidad a la sequía en plantas adultas, reflejado en una menor acumulación de biomasa y menor número de botones florales después de un periodo de recuperación (Figura 6 artículo anexo, Tabla 3). Este fenotipo es más severo al obtenido con las plantas mutantes para uno de los genes (RNAi *AtLEA4-5*, dSpm::*AtLEA4-5*) o para dos genes (*AtLEA4-1* y *AtLEA4-2*).

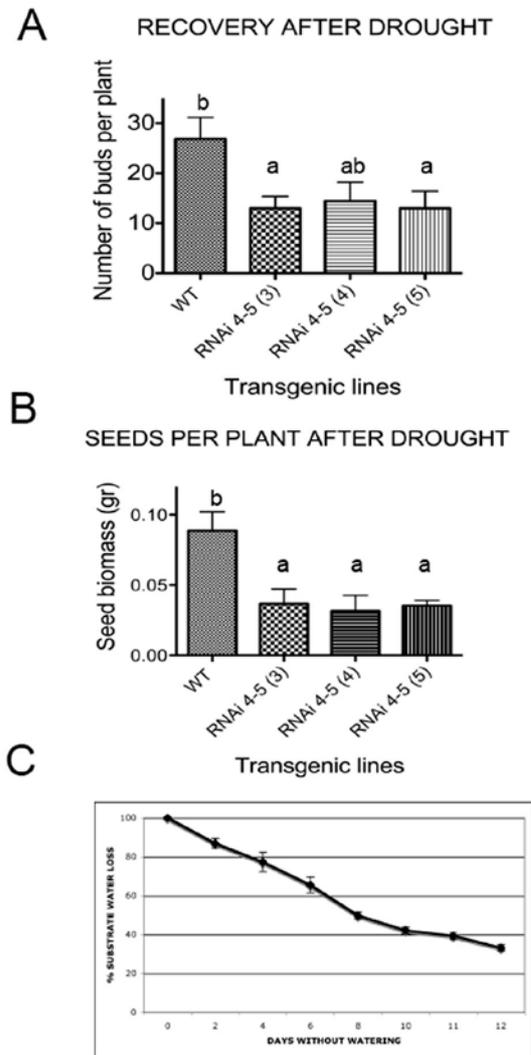


Figura 10. Análisis fenotípico de plantas homocigotas T_2 que expresan una construcción que desencadena el silenciamiento por RNAi de los transcritos *AtLEA4-5* en plantas adultas después de la recuperación del estrés. Se sembraron en la misma maceta plantas silvestres (Wt) y tres líneas transgénicas (ver Figura 15, líneas 3-5), que mostraron una reducción significativa de los niveles de proteína *AtLEA4-5* bajo tratamientos de sequía. Las plantas se crecieron en invernadero y se mantuvieron en condiciones de riego óptimo hasta el inicio de la floración y se sometieron a deshidratación por un periodo de 12 días (potencial hídrico del sustrato $\Psi = -4.62 (\pm 0.62)$ MPa). A) Promedio del número de botones florales y axilares por planta después de 6 días de recuperación del estrés. Las barras indican el promedio (\pm SE, error estándar, $n=7$), encontrando diferencias significativas entre grupos ($P=0.0273$) con análisis de varianza (ANOVA) de una vía. B) Biomasa seca de semillas provenientes de plantas que sufrieron los tratamientos de sequía y recuperación, obtenida mediante la cosecha total de semillas producidas al mantener las plantas en riego óptimo hasta la senescencia. Las barras indican el promedio (\pm SE, $n=3$). Se encontraron diferencias significativas entre grupos ($P=0.0133$) con el ANOVA de una vía. Las letras encima de las barras indican diferencias significativas entre grupos según la prueba de Dunnett ($P<0.05$). C) Porcentaje de pérdida de agua del sustrato (\pm SE) durante la cinética de deshidratación de macetas con plantas silvestres y mutantes con niveles reducidos de los miembros de la familia *AtLEA4*, indicado como el porcentaje del contenido inicial de agua en el sustrato.

Así como se detectó inestabilidad de los transgenes para la construcción de RNAi *AtLEA4-5*, detectamos que el silenciamiento en las cruzas sólo era funcional en las generaciones F_1 y F_2 . Cuando se analizaron las cruzas en la generación F_3 el fenotipo de sensibilidad encontrado anteriormente no se reprodujo en todas las plantas, ya que los transgenes se habían silenciado y los niveles de expresión de las proteínas *AtLEA4* se revirtieron a los de plantas silvestres (datos no mostrados). Debido a que la construcción de RNAi no era efectiva en reducir los niveles de la proteína *AtLEA4-5* en las semillas, estas líneas no se analizaron para la caracterización fenotípica durante germinación.

Debido a que los tres miembros de la familia *AtLEA4* se acumulan durante la embriogénesis, analizamos el efecto de su disminución en la producción de semillas en las diferentes mutantes en condiciones óptimas de desarrollo. Para ello, se crecieron plantas silvestres y plantas de dos líneas transgénicas para cada construcción de manera individual en macetas con Metromix y se mantuvieron en riego óptimo durante todo su ciclo de vida, cosechando las semillas de cada planta por separado hasta la senescencia. Se determinó la biomasa seca de las semillas por planta y se encontró que las plantas que no expresan o que tienen una reducción en los niveles de acumulación de proteínas de los miembros de la familia *AtLEA4* (dSpm 4-5, RNAi 4-5, a-miR 4 1/2) producen una menor cantidad total de semillas durante su ciclo de vida, comparado con las plantas silvestres (ver Figura 8 del artículo anexo).

El incremento en los niveles de expresión de este gen en las plantas que sobreexpresan el gen *AtLEA4-5* en el fondo silvestre (Col, 35S::*AtLEA4-5*::*NOS*) como en el fondo de la mutante por transposición en este mismo gen (dSpm 4-5, 35S::*AtLEA4-5*::*NOS*) no confirió una ventaja en cuanto a la producción de semillas cuando se compara la biomasa de semillas de plantas silvestres (ver Figura 8 del artículo anexo 1).

Tabla 3. Análisis fenotípico de plantas mutantes T₄ por inserción de un transposón en el gen *AtLEA4-5* (dSpm 4-5) y plantas homocigotas T₂ por silenciamiento post-transcripcional en de la familia *AtLEA4* durante sequía y después de la recuperación del estrés. Las plantas fueron crecidas en invernadero y se mantuvieron en riego óptimo hasta la floración, se dejaron de regar hasta que el sustrato alcanzó un potencial ($\Psi_{\text{sustrato}} = -4.62 \pm [0.62]$) y en este punto se rehidrataron y recuperaron durante 6 días. Se contaron los botones florales por planta (Columna 3) y también se obtuvo la biomasa seca total de cada planta bajo riego óptimo (Columna 1) o después de la recuperación (Columna 2). En experimentos independientes se calculó el contenido relativo de agua (RWC) en discos de hojas de plantas control (Columna 4) o después de una sequía en donde el sustrato alcanzó un potencial de $\Psi_{\text{sustrato}} = -5.42 [\pm 0.54]$, Columna 5).

Genotype	Dry biomass of whole plants kept under optimum irrigation (gr)[§]	Dry biomass of whole plants after recovery from drought (gr)[¶]	Number of buds per plant after recovery from drought[¶]	RWC of leaf discs from plants kept under optimum irrigation[§]	RWC of leaf discs from plants after 13 days of withholding irrigation[§]
WT 1 (±SD) (lower 95% CI, upper 95% CI)	0.1021 (±0.02358) (0.07734, 0.1268)	0.03505 (±0.008034) (0.02833, 0.04177)	13.75 (±2.486) (12.42, 15.58)	0.9019 (±0.02323) (0.8775, 0.9263)	0.17 (±0.00674) (0.1593, 0.1808)
dSpm 4-5 (1) (±SD) (lower 95% CI, upper 95% CI)	0.1141 (±0.01072) (0.1028, 0.1253)	0.01826 (±0.00529) (0.01384, 0.02268)	7.58 (±3.059) (5.64, 9.527)	0.8985 (±0.03627) (0.8604, 0.9365)	0.1552 (±0.01283) (0.1347, 0.1756)
WT 2 (±SD) (lower 95% CI, upper 95% CI)	0.1078 (±0.0398) (0.06601, 0.1496)	0.03485 (±0.00798) (0.02818, 0.04152)	16.25 (±3.991) (12.91, 19.59)	0.9147 (±0.03054) (0.8826, 0.9467)	0.2164 (±0.01408) (0.194, 0.2388)
RNAi 4-5 (3) (±SD) (lower 95% CI, upper 95% CI)	0.1224 (±0.03627) (0.08434, 0.1605)	0.01816 (±0.00397) (0.01485, 0.02148)	6.5 (±2.39) (4.502, 8.498)	0.9135 (±0.05177) (0.8591, 0.9678)	0.1337 (±0.0368) (0.07519, 0.1923)
WT 3 (±SD) (lower 95% CI, upper 95% CI)	0.106 (±0.0199) (0.08515, 0.1269)	0.03769 (±0.00398) (0.03436, 0.04102)	15.88 (±5.276) (11.46, 20.29)	0.9255 (±0.02639) (0.8978, 0.9532)	0.1567 (±0.02432) (0.118, 0.1954)
a-miR 4-1/2 (4) (±SD) (lower 95% CI, upper 95% CI)	0.1088 (±0.02774) (0.07968, 0.1379)	0.02015 (±0.00420) (0.01664, 0.02366)	7.5 (±2.507) (5.404, 9.596)	0.9138 (±0.01799) (0.8949, 0.9327)	0.1478 (±0.03539) (0.09152, 0.2041)
WT 4 (±SD) (lower 95% CI, upper 95% CI)	0.1099 (±0.02415) (0.07155, 0.1484)	0.0342 (±0.00649) (0.02877, 0.03963)	16.38 (±4.897) (12.28, 20.47)	0.9203 (±0.02924) (0.8896, 0.951)	0.1883 (±0.00859) (0.1746, 0.202)
Triple mutant (1) (±SD) (lower 95% CI, upper 95% CI)	0.1236 (±0.01129) (0.1056, 0.1415)	0.0207 (±0.00538) (0.0162, 0.0252)	7.875 (±3.523) (4.93, 10.82)	0.9247 (±0.01299) (0.911, 0.9383)	0.1279 (±0.01971) (0.0965, 0.1592)
WT 5 (±SD) (lower 95% CI, upper 95% CI)	0.0967 (±0.02718) (0.06813, 0.1252)	0.03986 (±0.00759) (0.03352, 0.04621)	12.88 (±3.563) (9.896, 15.85)	0.9019 (±0.02323) (0.8775, 0.9263)	0.1932 (±0.01446) (0.1702, 0.2162)
Compl dSpm (1) (±SD) (lower 95% CI, upper 95% CI)	0.1143 (±0.03298) (0.08381, 0.1448)	0.04573 (±0.00547) (0.04115, 0.0503)	11.25 (±5.97) (6.259, 16.24)	0.9083 (±0.02954) (0.8773, 0.9393)	0.1987 (±0.03705) (0.1398, 0.2577)
WT 6 (±SD) (lower 95% CI, upper 95% CI)	0.1244 (±0.01351) (0.08965, 0.1591)	0.03721 (±0.00811) (0.03044, 0.04399)	11.88 (±4.824) (7.842, 15.91)	0.9338 (±0.02097) (0.9117, 0.9558)	0.228 (±0.007428) (0.2162, 0.2398)
OE 4-5 (4) (±SD) (lower 95% CI, upper 95% CI)	0.1244 (±0.01947) (0.104, 0.1448)	0.04755 (±0.00674) (0.04191, 0.05319)	14.38 (±4.438) (10.66, 18.09)	0.9536 (±0.02697) (0.9253, 0.9819)	0.1971 (±0.03261) (0.1452, 0.249)

[§]Equivalent to $\Psi_{\text{sustrato}} = - 0.451 (\pm 0.212)$ MPa

[¶]Equivalent to $\Psi_{\text{sustrato}} = - 4.617 (\pm 0.619)$ MPa

[§]Equivalent to $\Psi_{\text{sustrato}} = - 5.415 (\pm 0.543)$ MPa

DISCUSION

La gran conservación en las secuencias de cada uno de los diferentes grupos de proteínas LEA y la alta correlación entre sus patrones de expresión y el déficit hídrico denotan un papel muy relevante para estas proteínas en este tipo de situaciones. Aunque existe evidencia circunstancial de que estas proteínas participan en la adaptación de plantas superiores a una baja disponibilidad de agua, no existía hasta la fecha una evidencia directa reportada en este aspecto. En este trabajo, analizamos la familia de genes *AtLEA4* demostrando que participan en la respuesta adaptativa ante déficit hídrico en plantas de *Arabidopsis* a lo largo de diferentes etapas de su ciclo de vida.

Para abordar de manera apropiada el análisis fenotípico de las mutantes obtenidas, se determinaron y analizaron los patrones de acumulación de los transcritos y proteínas de esta familia en diferentes órganos y durante diferentes etapas de desarrollo de la planta. Aún cuando la mayor parte de la información con la que se cuenta en relación a los genes *LEA* y sus proteínas en *Arabidopsis* y otras plantas modelo se refiere a los patrones de acumulación de sus transcritos (Delseny *et al.*, 2001; Hoth *et al.*, 2002; Seki *et al.*, 2002; Oono *et al.*, 2003; Schmid *et al.*, 2005; Winter *et al.*, 2007; Hruz *et al.*, 2008); ésta información aunque útil no necesariamente correspondía a las condiciones experimentales ni de crecimiento de nuestro interés, y por otro lado, no se contaba con la información referente a los niveles de acumulación de las proteínas correspondientes, por lo que hasta ese momento no era posible hacer correlaciones entre la presencia y abundancia de un transcrito y la proteína correspondiente. Nuestros resultados y los publicados en las bases de datos disponibles de microarreglos (Schmid *et al.*, 2005; Winter *et al.*, 2007; Hruz *et al.*, 2008), concuerdan en que la mayor acumulación para todos los transcritos del grupo *AtLEA4* correlacionan con la etapa de adquisición de la tolerancia a la desecación durante el desarrollo de la semilla y hasta dos días después de la germinación. También nuestros resultados comprueban que los transcritos para *AtLEA4-5* son los que se acumulan de manera más abundante en comparación con los otros miembros de la familia en la mayoría de las condiciones de estrés ensayadas, en particular, en respuesta a estrés hiper-osmótico, sequía y tratamiento con ABA. Los genes que muestran la mayor similitud a nivel de nucleótidos *AtLEA4-1* y *AtLEA4-2*, también mostraron una correlación en sus patrones de acumulación de transcritos, lo cual sugiere que pueden llevar a cabo funciones similares, y está de acuerdo con nuestro análisis filogenético, el cual localiza a ambas proteínas en el mismo subgrupo (ver adelante). Este resultado es congruente con un análisis de las proteínas LEA de

Arabidopsis en donde se menciona que los genes que codifican para las proteínas AtLEA4-1 y AtLEA4-2 provienen de una duplicación reciente (Hundertmark y Hinch, 2008).

Los niveles de transcritos contrastan con la abundancia de las proteínas correspondientes en flores y durante el desarrollo de la plántula en el caso de la AtLEA4-1, en donde los niveles de proteína son similares a la etapa de semilla madura, aunque se detectan bajos niveles de acumulación de transcrito en esos órganos o etapas del desarrollo. En semillas maduras, tanto secas como imbibidas (durante el tratamiento de estratificación en frío) y durante la germinación, la proteína AtLEA4-1 se detectó con una mayor masa molecular aparente, lo que sugiere que bajo las condiciones de deshidratación severa que se tienen en la semilla seca, la proteína se modifica post-traduccionalmente o sufre cambios conformacionales, y que durante el proceso de la germinación (que inicia con la imbibición de la semilla) estos cambios se pierden conforme avanza el proceso. De tal forma, la proteína se vuelve a detectar en el tamaño esperado del monómero en etapas posteriores a la germinación, y su acumulación incrementa gradualmente hasta los 12 días después de la germinación, por lo que su papel puede ser relevante para el exitoso establecimiento de la plántula.

En el caso de la AtLEA4-2, no pudimos detectar a la proteína con una masa molecular esperada para el monómero (10.5 KDa); en su lugar, detectamos específicamente una proteína con una masa molecular aparente mayor a 30 KDa, cuyo patrón de acumulación correlaciona con el obtenido para los transcritos. Una posible explicación para este resultado, es que a diferencia de la AtLEA4-5 y la AtLEA4-1, las predicciones de estructura secundaria sugieren la menor proporción de estructura aleatoria o random coil (20% de la secuencia de la proteína hacia la región carboxilo) para la AtLEA4-2, indicando que la mayor proporción de estructura secundaria tipo "coiled-coil" en esta proteína puede favorecer la formación de homo u hetero-oligómeros. En algunos casos (ver Figuras 1B y 2B del artículo anexo), se detectaron dos bandas en lugar de una, sugiriendo la formación de dos tipos de oligómeros debido probablemente a interacciones preferenciales, como se ha reportado para la proteína PvLEA18 (Colmenero_Flores *et al.*, 1999). Cabe recordar que para los tres miembros de la familia AtLEA4 se predice que pueden formar estructuras de tipo "coiled-coil", las cuales consisten en dos o más α -hélices que se enrollan unas con otras en un super enrollamiento de orientación izquierda, pudiendo formar homo- o hetero-dímeros, en complejos de dos a cinco hélices y se pueden alinear en orientación paralela o antiparalela. La secuencia de aminoácidos de un "coiled-coil" se caracteriza por una repetición conservada de 7 residuos (a-b-c-d-e-f-g)_n. Los residuos de la posición 1 (a) y 4 (d) forman la interface de dimerización entre

las hélices y generalmente están ocupados por residuos hidrofóbicos. Los demás residuos de la héptada son en su mayoría polares o cargados y pueden formar interacciones intra- o inter-hélices, que contribuyen a la estabilidad y la especificidad de la formación del complejo (Mason *et al.*, 2009). Lo anterior refuerza la hipótesis de que las proteínas LEA que forman “coiled-coils” pueden interactuar con ellas mismas o con otras proteínas “blanco” en la célula. En el caso particular de la AtLEA4-2, se predice que la mayoría de sus residuos (80%) pueden estar involucrados en la formación de “coiled-coils”, apoyando la existencia de interacciones fuertes entre los oligómeros, que no pudieron ser resueltos con las condiciones utilizadas para su separación en la electroforesis con SDS.

El análisis de los patrones de acumulación de proteínas para los diferentes miembros de la familia AtLEA4 sugiere la participación de mecanismos de control postranscripcional que modulan los niveles de proteína, como ha sido sugerido para otras proteínas LEA (Ried y Walker-Simmons, 1993; Colmenero-Flores *et al.*, 1999). Un ejemplo de esta situación es evidente en el caso de los niveles de acumulación de transcritos vs de proteína para la AtLEA4-5 en tratamientos con ABA y NaCl, en donde los niveles de transcritos fueron similares para ambas condiciones mientras que los niveles de proteína en respuesta a ABA son mucho mayores que en respuesta a NaCl, lo cual indica que el ABA puede estar involucrado en algún mecanismo de regulación postranscripcional de los transcritos para este gen.

Para obtener información del papel de la familia de genes *AtLEA4* en la adaptación de las plantas ante condiciones de déficit hídrico, llevamos a cabo un análisis fenotípico de plantas de *Arabidopsis* deficientes en los productos génicos de esta familia. Las plantas mutantes que no pueden sintetizar la proteína AtLEA4-5 (dSpm 4-5) mostraron que ésta tiene un papel relevante en el ajuste de la planta a condiciones de estrés hiper-osmótico (manitol 350 mM) y de alta salinidad (NaCl 250 mM), ya que alcanzaron una menor tasa y eficiencia germinativa. La ausencia de esta proteína también afectó negativamente la capacidad de las plantas de *Arabidopsis* para recuperarse de los tratamientos de sequía y recuperación aplicados en la etapa adulta. Las plantas mutantes mostraron una menor acumulación de biomasa y una menor capacidad para producir botones florales y axilares comparadas con plantas silvestres. De acuerdo con esto, la deficiencia de AtLEA4-5 en las plantas silenciadas por RNAi provocó una baja producción de semillas después de haber sufrido tratamientos de sequía y recuperación. Un efecto similar se observó en la capacidad de recuperación de las plantas silenciadas para los genes AtLEA4-1 y AtLEA4-2 con un a-miR, sugiriendo que no hay una redundancia funcional en su participación en estas condiciones ambientales adversas. Esta interpretación se ve apoyada aún más por los

resultados obtenidos con las plantas en las cuales los tres genes de la familia estaban silenciados (RNAi 4-5 x a-miR 4 -1/2). Estas plantas con niveles indetectables o muy bajos de las tres proteínas mostraron una mayor reducción en la acumulación de biomasa, así como en la producción de botones florales y axilares en respuesta a tratamientos de sequía.

De manera muy interesante, encontramos que aún bajo condiciones óptimas de crecimiento, la biomasa total de semillas producidas en el ciclo de vida completo de las plantas mutantes fue menor que el de las plantas silvestres, probablemente debido a su papel protector en la maquinaria de la planta necesaria para un desarrollo óptimo de la semilla y/o el fruto. También se debe considerar el posible papel protector de estas proteínas en las regiones meristemáticas o los primordios (de acuerdo con su efecto en el número de botones florales y axilares después de sequía), lo cual puede ser relevante para la formación de inflorescencias y consecuentemente en el número total de semillas. Realizamos la determinación del tamaño de las semillas calculando el peso de 1000 semillas, lo cual indicó que las mutantes no están afectadas en este fenotipo, por lo que se puede descartar la posibilidad de que un llenado ineficiente de las semillas sea el responsable de una menor biomasa total de semillas por planta. Las plantas que sobre-expresan la AtLEA4-5 no mostraron una mayor cantidad de biomasa de semillas bajo irrigación óptima, lo cual sugiere que durante el desarrollo normal de las semillas, la presencia de niveles mayores que los que ocurren en las plantas silvestres no son necesarios para una exitosa producción de semillas.

El análisis *in silico* de las proteínas del grupo AtLEA4 confirmó nuestra propuesta anterior (Battaglia *et al.*, 2008) de que las proteínas de este grupo pueden definirse por la conservación de sus secuencias, el cual es un criterio más fuerte que las propiedades de sus aminoácidos (Wise y Tunnacliffe, 2004) y que pueden subdividirse en dos subgrupos, cada uno caracterizado por diferentes motivos conservados. Los motivos 1, 2, 4 y 6/7 están tan ampliamente distribuidos que pueden representar la firma de este grupo de proteínas, ya que se encuentran en 52 de 77 secuencias; sin embargo, el motivo 2 puede ser considerado el representativo de la familia, ya que se encuentra en todas las proteínas y es informativo de la filogenia de las proteínas completas para esta familia. Al sobreponer la taxonomía de los organismos con la filogenia del grupo LEA4 se puede concluir que los dos subgrupos se originaron a partir de una duplicación muy temprana que antecede a la divergencia de las dicotiledóneas y las monocotiledóneas, ya que podemos encontrar los dos tipos de proteínas se pueden encontrar en ambos taxones. El subgrupo 4A se encuentra representado en una briofita, mientras que el subgrupo 4B se encuentra en coníferas. Puede sugerirse que el grupo 4B es más antiguo que el 4A, ya que el 4B está

representado en plantas vasculares y no vasculares, mientras que las proteínas 4A están ausentes en el genoma secuenciado de *Physcomitrella patens*. Sin embargo, las proteínas 4A pueden estar representadas en otras plantas no vasculares conforme se examinen y se secuencien un mayor número de especies.

Por otra parte, es notable que varios géneros de plantas monocotiledóneas y dicotiledóneas tienen representantes de ambos subgrupos. Esta observación sugiere que la duplicación inicial dio lugar a una divergencia funcional o subfuncionalización. Una hipótesis puede ser que al proteger otras proteínas o estructuras subcelulares, ambos subgrupos utilicen mecanismos similares pero tengan blancos diferentes, lo cual se apoya con el hecho de que ambos comparten los motivos 1, 2, 4 y 6/7, pero difieren en la presencia de motivos adicionales que pueden estar involucrados en la especificidad de sus blancos. En el caso de *Arabidopsis*, las dos proteínas del subgrupo 4A (AtLEA4-1 y AtLEA4-2) tienen patrones de expresión similares, que son diferentes a los de la proteína del subgrupo 4B (AtLEA4-5). Nuestros datos experimentales correlacionan con la conclusión del análisis bioinformático, ya que se espera que tengan patrones diferentes al ser proteínas cuyas funciones han divergido.

Desde el primer reporte en donde se reconocen varias familias de estas proteínas hidrofílicas e intrínsecamente no estructuradas por Dure *et al.* (1989), el enfoque principal de la investigación sobre estas proteínas ha sido la expresión constitutiva o inducible en plantas y micro-organismos, para probar su papel protector en la sobrevivencia de éstos (para una revisión reciente de este tema ver el artículo de Battaglia *et al.*, 2008 en el anexo 2 de este manuscrito). En el caso particular del grupo LEA4, se han llevado a cabo pocos estudios de análisis funcional y éstos han sido reportados en tejidos vegetativos únicamente. Tal es el caso de un homólogo del gen *AtLEA4-5* en *Brassica napus* (BnLEA4-1) cuya expresión constitutiva incrementa la tolerancia de plántulas (Dalal *et al.*, 2009). En lo que se refiere al análisis por pérdida de función, se ha reportado sólo un estudio para uno de los genes de esta familia en cacahuate (*Arachis hypogea*), en donde la expresión transitoria (inducida por la infección con un virus) del cDNA causa el silenciamiento del gen endógeno en plantas de tabaco (*Nicotiana tabacum*), provocando una sensibilidad de los tejidos vegetativos a un tratamiento de sequía (Senthil-Kumar y Udayakumar, 2006). En contraste con estos y otros reportes, en el presente estudio, reportamos el papel de una familia completa de proteínas en la adaptación de plantas superiores a la sequía y la desecación utilizando un enfoque genético, en donde correlacionamos la presencia y abundancia de los miembros de esta familia con el ajuste de la planta al estrés por déficit hídrico a lo largo de su ciclo de vida. El impacto biotecnológico que tendría la sobre-expresión de estas proteínas tipo LEA en plantas de cultivo puede ser de gran

importancia debido al constante incremento en los niveles de sequía que se predicen para los años próximos. Se ha reportado para otros grupos de proteínas LEA su eficacia en la protección contra sequía en plantas de interés agronómico como arroz y trigo, lo cual además de que apoya la hipótesis de su papel como protectores de la maquinaria celular, confiere una ventaja adicional al expresar estas proteínas tipo chaperona respecto a otras que modifican el metabolismo de la planta, ya que su sobre-expresión no va en detrimento de la planta en condiciones normales, a diferencia de lo que se ha reportado en la literatura para el caso de la sobre-expresión de factores transcripcionales o enzimas de la ruta de biosíntesis de osmolitos.

CONCLUSIONES

- 1) La persistencia de los sub-grupos LEA4A y LEA4B en grupos taxonómicos de plantas vasculares y no vasculares sugiere que su duplicación dio origen a una especialización funcional de tal forma que una proteína de un subgrupo no puede substituir el papel del otro subgrupo.
- 2) La acumulación diferencial de los transcritos y proteínas de esta familia durante el desarrollo del embrión, la germinación y el establecimiento de la plántula en condiciones óptimas; así como sus patrones de acumulación en condiciones de déficit hídrico, indica una diversificación en sus patrones de expresión después de su duplicación.
- 3) La alteración en los niveles de expresión de los genes *AtLEA4* tiene un efecto negativo en las plantas con abatimiento de uno o dos de los productos de su expresión en el proceso de la germinación en condiciones normales y bajo déficit hídrico, impuesto por altas concentraciones de sales u osmolitos en el medio.
- 4) Hay un defecto en la recuperación de plantas mutantes adultas después de haber pasado por un periodo de sequía, cuantificado como una menor acumulación de biomasa y un menor número de botones florales y axilares por planta.
- 5) En experimentos llevados a cabo en riego óptimo durante todo el ciclo de vida de las plantas se producen menos semillas por planta, ya que requiere de estas proteínas para obtener una producción óptima, lo cual indica que estas proteínas

están implicadas en alguno(s) de los procesos involucrados en la producción de semillas y/o silicuas maduras en condiciones de riego óptimo.

PERSPECTIVAS

Tomando como base los resultados obtenidos con estos experimentos, planteamos realizar otro tipo de experimento de sequía, cuyas principales diferencias son: el uso de un sustrato con mayor retención de agua (mezcla de 40% metromix, 30% agrolita y 30% vermiculita vs. 100% turface) y la aplicación de una sequía menos severa que permitiera continuar el desarrollo de las plantas, ya que el tratamiento de sequía es más lento en un sustrato con mayor retención de agua y la restricción de agua se puede controlar restaurando el agua que van perdiendo las macetas durante dos semanas una vez que el sustrato llega a un 20% de la capacidad de campo. Este experimento involucra una sequía sostenida, a diferencia de una sequía progresiva (cercana al punto de marchitez permanente) y una posterior rehidratación y recuperación de las plantas, usado en los experimentos reportados en esta tesis. Con este tipo de experimento se pueden analizar parámetros relacionados con el crecimiento y la productividad de las plantas silvestres y las diferentes líneas de plantas transgénicas, comparando una situación de irrigación óptima con una condición de sequía sostenida durante un periodo de dos semanas (p.ej. acumulación de biomasa, relación peso fresco:peso seco, coeficiente raíz:parte aérea, fotosíntesis índice de cosecha, eficiencia en el uso de agua, número de semillas por silicua, porcentaje de aborción de semillas por silicua)

Los resultados de este experimento se están analizando (datos no mostrados) para determinar si existen diferencias significativas entre los genotipos debidas al tratamiento de sequía, y, de manera preliminar podemos decir que encontramos un fenotipo de una mayor aborción de semillas y silicuas en la mutante por inserción en el promotor del gen *AtLEA4-5*, lo cual sugiere que la presencia y abundancia de los productos de la expresión del gen *AtLEA4-5* son necesarios para una óptima producción de semillas en condiciones de deshidratación moderada.

CAPITULO 2

INTRODUCCION

I. ANÁLISIS ESTRUCTURA-FUNCIÓN EN LAS PROTEÍNAS LEA.

Aunque los experimentos genéticos pueden indicar los procesos en los que participa el grupo de proteínas LEA4, éstos sugieren poco sobre su mecanismo de acción. Así que con la finalidad de dilucidar las bases de su mecanismo de acción, en nuestro grupo se planteó la hipótesis de que las proteínas LEA, incluyendo a las del grupo 4, protegen a otras proteínas de los efectos deletéreos sobre su actividad biológica causados por condiciones de limitación de agua. Como se menciona en la Introducción del capítulo 1 de este manuscrito, se demostró que varias proteínas, representantes de diferentes grupos de proteínas LEA (AtLEA4, AtLEA2, AtLEA3) son capaces de prevenir la pérdida de la conformación nativa y, por tanto de la actividad, de dos proteínas blanco (LDH, MDH) ante una deshidratación parcial o ante tratamientos de congelamiento-descongelamiento inducidos *in vitro* (Reyes *et al.*, 2005; 2008). También se aportaron datos que indicaban que bajo estas condiciones *in vitro* estas proteínas podrían interactuar con sus proteínas blanco (Reyes *et al.*, 2005). Es por ello que en la segunda parte del proyecto se abordó la pregunta de cuáles son los blancos de la proteína AtLEA4-5, a través de determinar su identidad por análisis *in vivo* e *in vitro*, en condiciones óptimas y/o bajo déficit hídrico.

Las características físico-químicas de la proteínas LEA permiten establecer hipótesis para su mecanismo de acción. Típicamente, la estructura primaria de las proteínas LEA hidrofílicas incluye residuos cargados y polares no cargados (Glu, Lys, Arg, Gly, Gln, Ser, Pro), y muy pocos o ningún residuo hidrofóbico (Ile, Leu, Val, Trp, Phe, Tyr, Cys, Asn), por lo que son proteínas no globulares (desplegadas) y muy flexibles. Es por ello, que se ha sugerido que las proteínas LEA tienen la capacidad de unirse a moléculas de agua y asociarse a la matriz subcelular formando interacciones no covalentes o débiles, funcionando como una superficie de "solvatación" para las moléculas de la célula (Close *et al.*, 1993; Dure, 1993). En nuestro laboratorio propusimos clasificarlas dentro de las "hidrofilinas", un grupo de amplia distribución, que abarca organismos de los tres dominios Eubacteria, Archaea y Eukarya representado por proteínas con un índice de hidrofiliidad mayor a 1 y una alta proporción de aminoácidos pequeños (Gly, Ser, Ala) y que se expresan en condiciones de estrés por déficit hídrico (Garay-Arroyo *et al.*, 2000, Battaglia *et al.*, 2008; Campos, com. pers).

Debido a la importancia de las proteínas LEA en la tolerancia a la sequía en tejidos vegetativos y reproductivos, así como en la adquisición de la tolerancia a la desecación de la semilla, muchos estudios se han enfocado en determinar las estructuras de las proteínas LEA. La estructura secundaria de varias de ellas ha sido caracterizada por dicroísmo circular (CD), resonancia magnética nuclear (NMR) o espectroscopía por transformación de Fourier (FTIR). Además, se ha determinado la estructura tridimensional de algunas proteínas LEA “atípicas” o hidrofóbicas (revisado en Battaglia *et al.*, 2008; Shih *et al.*, 2008).

A pesar de la alta probabilidad que presentan estas proteínas de mostrarse como proteínas desplegadas en solución acuosa, varios algoritmos de predicción de estructuras secundarias (Chou y Fasman, 1978; Muñoz y Serrano, 1994; Rost y Liu, 2003) indican que las proteínas LEA de los grupos 1-4 pueden adoptar estructuras secundarias definidas, como α -hélices o β -plegadas, en algunas regiones. Por ejemplo, se predice que algunos dominios conservados de los grupos de proteínas LEA1 (20-mero), LEA2 (segmento K), LEA3 (11-mero) o la región-amino conservada del grupo LEA4 pueden formar α -hélices anfipáticas (Dure, 1993). Sin embargo, con enfoques experimentales de análisis espectroscópicos se ha encontrado que la mayoría de las proteínas LEA contienen regiones desordenadas en la mayor parte de su extensión cuando se encuentran en solución acuosa. Esto ha llevado a considerar a las proteínas LEA hidrofílicas de plantas como parte del grupo de las proteínas intrínsecamente no estructuradas o desordenadas (IUP o IDPs). Las características estructurales que hacen únicas a este tipo de proteínas son, entre otras: una baja compactación, ausencia de globularidad, bajo contenido de estructuras secundarias y una gran flexibilidad (Tompa, 2002; Uversky, 2002). La mayoría de las proteínas globulares contienen un núcleo hidrofóbico que estabiliza su conformación nativa, mientras que las cadenas laterales cargadas o polares de la superficie pueden interactuar con las moléculas de agua que las rodean; en cambio, en las IUPs por la naturaleza cargada e hidrofílica de los aminoácidos que las componen se favorece la interacción con las moléculas de agua, en competencia con las interacciones intramoleculares las cuales son muy inestables por la repulsión de cargas entre sus aminoácidos en comparación con las interacciones hidrofóbicas que ocurren en las proteínas globulares; de ahí que no puedan adoptar estructuras terciarias o cuaternarias estables (Shih *et al.*, 2008).

A pesar de ello, se ha propuesto por resultados obtenidos con algunas IUPS que éstas son capaces de adoptar estructuras secundarias estables y diversas ante diferentes condiciones ambientales; así mismo, se ha propuesto que pueden adoptar una estructura secundaria estable al interactuar con su(s) blanco(s), lo que permite

un control más preciso en la termodinámica del proceso de unión, logrando así regular varios procesos celulares (Wright y Dyson, 1999). Lo anterior se ha comprobado en proteínas desordenadas que están involucradas en procesos de señalización y regulación de la transcripción como p53 (Bell *et al.*, 2002), p27^{Kip1} (Sivakolundu *et al.*, 2005) o BRCA1 (Mark *et al.*, 2005). En estudios *in vitro* las IUP se han caracterizado por tener una gran rapidez en su velocidad de interacción, en algunos casos se ha determinado que a pesar de su flexibilidad estructural y de no tener una fuerza de unión excesiva no pierden especificidad por sus blancos, aunque son promiscuas pues se adaptan a múltiples blancos de interacción. También se ha propuesto que estas interacciones podrían ser regulables o selectivas a través de modificaciones post-traduccionales (Hegyi y Tompa, 2008). Por todo ello se ha sugerido que su gran flexibilidad pudiera representar una ventaja funcional en condiciones fisiológicas pues les permitiría interactuar eficientemente con diferentes blancos presentes en una condición determinada (Wright y Dyson, 1999). Las IUPs, como el caso de las proteínas LEA hidrofílicas, pueden cambiar su conformación al interactuar con otras moléculas como proteínas, ácidos nucleicos o iones metálicos (Battaglia *et al.*, 2008; Shih *et al.*, 2008). Ante este panorama, las proteínas LEA representan un paradigma para las IUPs que permitiría relacionar aspectos estructurales y funcionales de la proteína.

El proceso de adquisición de la tolerancia a la desecación, involucra una respuesta compleja que incluye alteraciones en los componentes celulares; hay una inducción de sistemas de reparación y desintoxicación de radicales libres, así como un incremento en la acumulación de macromoléculas (Hoekstra *et al.*, 2001). Durante el proceso de deshidratación hay acumulación de proteínas LEA de todos los grupos, los cuales pueden interactuar con oligosacáridos, aunque esta interacción no ocurra en solución. En comparación con los cristales de oligosacáridos puros, la presencia de proteínas LEA incrementa la temperatura de transición vítrea (T_g), sugiriendo que actúan de manera sinérgica incrementando la viscosidad citoplasmática en el proceso de formación de matrices vitrificadas o biocristales (Wolkers *et al.*, 2001; Shih *et al.*, 2004). Se ha sugerido que los biocristales otorgan protección intracelular en contra de la desnaturalización de macromoléculas (Burke, 1986). La trealosa es un oligosacárido que incrementa en un 20% el peso seco de nemátodos durante deshidratación lenta y está presente en esporas de levaduras (Madin y Crowe, 1975; Crowe *et al.*, 1984). En cambio, en las plantas, la sacarosa y algunos otros oligosacáridos como estaquirosa o rafinosa, se acumulan durante la maduración de las semillas y en los granos de polen (Hoekstra *et al.*, 1992; 2001). Las características biofísicas de los biocristales difieren de las características de los cristales de azúcares, ya que se ha sugerido que las

proteínas LEA tienen un papel primordial en la conformación de los biocristales (Bewley *et al.*, 1993; Blackman *et al.*, 1995; Oliver *et al.*, 2001). La deshidratación causa un incremento en la concentración de solutos en las células, incrementando su viscosidad, lo cual promovería un proceso termodinámicamente desfavorable entre la superficie hidrofílica de las proteínas con la superficie de las membranas al desaparecer el agua “libre”, que normalmente se encuentra en niveles del 80% en las células, ocasionando la agregación o desnaturalización de proteínas y la fusión de membranas. La presencia de las proteínas LEA permite establecer una capa de hidratación o de “solvatación” intermolecular, que se establece a través de puentes de hidrógeno entre las proteínas LEA y los oligosacáridos, con otras moléculas de agua y los componentes celulares, para evitar la desnaturalización de los componentes celulares. El hecho de que posean una alta proporción de aminoácidos cargados y polares permite que también actúen como atrapadores de iones (Hoekstra *et al.*, 2001; Garay-Arroyo *et al.*, 2001). A diferencia de las semillas ortodoxas, las semillas recalcitrantes son liberadas de la planta madre con un contenido de agua cercano al 23% en peso fresco. Estas semillas no son tolerantes a la deshidratación o al frío (Finch-Savage *et al.*, 1994; Farrant *et al.*, 1996) y, aunque se ha observado la presencia de algunas proteínas LEA y oligosacáridos, la transición vítrea ocurre en estas semillas en niveles de deshidratación entre 10-15%, cuando ya se rebasó el límite de tolerancia a la deshidratación que pueden tener estas semillas; es por ello que pierden su viabilidad aún antes de que los mecanismos de tolerancia a la deshidratación severa puedan actuar (Berjak y Pammenter, 2008).

A la fecha los análisis espectroscópicos no sólo otorgan una evidencia directa que define a las proteínas LEA como IUPs sino que también aportan evidencia que indica que éstas sufren cambios conformacionales dependiendo de la disponibilidad de agua en el ambiente (Battaglia *et al.*, 2008; Shih *et al.*, 2008). Se ha comprobado por dicroísmo circular que algunas proteínas de la familia LEA 1, LEA 3 y LEA 4 presentan transiciones de estructura desplegada a una conformación de α -hélice en presencia de algunos solventes, como SDS o trifluoroetanol (TFE) o un incremento en la conformación de hélice extendida o tipo poli-L-prolina conforme disminuye la temperatura (revisado en Battaglia *et al.*, 2008); sin embargo, no existen evidencias *in vivo* que demuestren que las proteínas LEA hidrofílicas puedan interactuar con azúcares, iones o membranas de fosfolípidos en solución.

Para el grupo 4 de las proteínas LEA se han reportado secuencias similares provenientes de angiospermas, gimnospermas y briofitas; un análisis global de esta información indica que son proteínas pequeñas, con un promedio de masa molecular de 12.6 kDa, contienen una alta proporción de aminoácidos pequeños (Gly, Ala) y

cargados (Lys, Arg), por lo que son altamente hidrofílicas. Para el total de las proteínas LEA4 reportadas, el intervalo de pI para estas proteínas es alrededor de 6 (22%), entre 6 y 9 (22%) y mayor a 9 (56%), por lo que la mayoría de las proteínas LEA4 son básicas (Shih *et al.*, 2008). Se ha demostrado por CD de UV lejano y espectroscopía FTIR que la proteína LEA4 de soya GmPM16 se encuentra desordenada en solución, y que su estructura no cambia en presencia de oligosacáridos no reductores como sacarosa, rafinosa, estaquiosa o trealosa ni en soluciones con pH 4.8, 7 ó 9.4. La adición de SDS o TFE promueve la formación de estructuras α -hélices en un 90% de la longitud de la proteína (Shih *et al.*, 2004).

En el caso de otra LEA del grupo 4 de soya (GmPM1) se observó por CD que la proteína contiene estructuras extendidas, tipo poli-L-prolina II en temperaturas bajas, pero en solución contienen una alta proporción de estructuras desordenadas o de tipo “random coil”. Después de deshidratar la proteína lenta o rápidamente hay un cambio (de 30 a 49%) en la proporción de estructura secundaria definida (α -hélice y β -plegada). Además en el estado deshidratado pueden interactuar con azúcares no reductores para mejorar la T_g de cristales celulares (o biocristales), así como con fosfolípidos para mantener la fase líquido-cristal en un amplio intervalo de temperaturas (Shi *et al.*, 2010). Al someter la GmPM1 o lisozima en secado lento, solas o en una proporción 1:1 (w/w) con poli-L-lisina, los resultados mostraron que ninguna de las proteínas, excepto la poli-L-lisina, se agregan por el secado lento. La lisozima adopta una estructura de α -hélice en secado lento. El espectro de las mezclas de poli-L-lisina mostraron que la GmPM1 previene el cambio conformacional de las hojas β de la poli-L-lisina en secado lento, mientras que la lisozima no puede prevenir la agregación de la poli-L-lisina. Los resultados sugieren a los autores, que las proteínas LEA4 en su estado funcional pueden afectar la conformación de otras proteínas en el proceso de secado lento (Shih *et al.*, 2010).

Existen otras evidencias experimentales que sugieren que las proteínas LEA pueden interactuar entre sí, además de las interacciones que pueden tener con otras proteínas. Se ha demostrado que la proteína Cor15am de *Arabidopsis* puede formar oligómeros y tiene actividad crioprotectora *in vitro*; por otro lado, la proteína GmPM8 de soya pueden oligomerizarse *in vitro* (Nakayama *et al.*, 2007; Hsing *et al.*, 2010).

Por otra parte, se han caracterizado dehidrininas (LEA2) de *Arabidopsis thaliana* (Cor47, Ltí29, Ltí 30, y Rab18), así como péptidos sintéticos con las secuencias correspondientes a los segmentos conservados Y, S y K (Mouillon *et al.*, 2006). Se encontró que en condiciones fisiológicas los espectros de dicroísmo circular de las proteínas completas y los péptidos indican una estructura desordenada con un

contenido variable de hélices poli-L-prolina. Sin embargo, no cambian su estructura en respuesta a alteraciones en la temperatura, adición de metales o incremento en la concentración de proteína, por lo que los autores concluyen que estas proteínas han evolucionado para mantener una alta flexibilidad en su configuración y que el papel de los segmentos conservados es ejercer su función biológica de manera local al interactuar con su blanco específico (Mouillon *et al.*, 2006). Además, la transición de estructuras es totalmente reversible, ya que al rehidratarlas, vuelven a su estado de estructura aleatoria (Tolleter *et al.*, 2007).

En otros organismos también se han detectado transiciones en la conformación de proteínas en respuesta a deshidratación, ya que una proteína similar a LEA de la familia 3 (AavLEA1) en el nemátodo *Aphelenchus avenae*, cambia de su estado nativo, sin plegamiento, a una estructura ordenada de α -hélice y posiblemente forma estructuras tipo “coiled-coil” en respuesta a la desecación (Goyal *et al.*, 2003). Los cambios conformacionales durante deshidratación ocurren debido a que los grupos superficiales de las proteínas ya no se encuentran formando puentes de hidrógeno con el agua, y en su lugar, forman los puentes entre ellas de manera intra- e intermolecular (Prestrelski *et al.*, 1993). Los estudios hidrodinámicos sugieren que los volúmenes de las proteínas Em de trigo (LEA 1) y AavLEA1 (LEA3) son más grandes que los de otras proteínas globulares de tamaño equivalente, lo cual indica que tienen un mayor número de moléculas de agua asociadas, por su estructura abierta o de tipo “random coil” en solución (McCubin y Kay, 1985; Goyal *et al.*, 2003).

II. LAS INTERACCIONES PROTEÍNA-PROTEÍNA Y EL ESTUDIO DE LAS PROTEÍNAS DESORDENADAS: EL CASO DE LAS PROTEÍNAS LEA.

En las últimas décadas, se han utilizado una variedad de métodos para explorar las interacciones proteína-proteína. La interacción específica de complejos proteicos puede ser estudiada a través de enfoques *in vivo*, tales como despliegue de fagos, ensayos de dos híbridos en levadura y ensayos de fluorescencia por complementación (Smith y Petrenko, 1997; cita, y2h, Walter *et al.*, 2004; Lalonde *et al.*, 2008), así como enfoques *in vitro* como la coimmunoprecipitación, la purificación por afinidad utilizando proteínas fusionadas a etiquetas, o la electroforesis azul nativa (Eubel *et al.*, 2005; Miernyk y Thelen, 2008). El perfeccionamiento en la ultrasensibilidad de la espectrometría de masas (LC-MS/MS) junto con la creciente disponibilidad de bases de datos de proteínas para identificar los péptidos que provienen de este análisis, ha favorecido el desarrollo de métodos de purificación por afinidad de complejos proteicos *in situ*.

Los estudios de experimentos *in vivo* e *in vitro* no son excluyentes entre sí, sino que ambos se complementan y son necesarios para corroborar los resultados obtenidos, ya que se ha observado por ejemplo, que el uso de sistemas heterólogos de expresión (como el análisis de proteína híbridas en levadura) puede arrojar muchos falsos positivos o que la posición en la que se inserta la etiqueta para la purificación de la proteína de fusión (como es el caso de la purificación por TAP) puede alterar la interacción con sus proteínas blanco (Rubio *et al.*, 2005; Lalonde *et al.*, 2008). Las características intrínsecas de las proteínas LEA sugieren que existen dos posibles mecanismos que afectarían la transición de estructura de estas IUPs: podrían interactuar con sus blancos adecuando su estructura para poder unirse a ellos una vez que los reconocen como su compañero de interacción, o alternativamente, las condiciones intracelulares podrían influir en la conformación espacial de las proteínas LEA, lo cual determinaría los blancos con los que podrían interactuar una vez que la condición osmótica causó la modificación en su estructura secundaria. Con estas ideas en mente y ante la oferta de opciones metodológicas para el análisis de interacción proteína-proteína, nos enfocamos a determinar la identidad de sus posibles ligandos proteícos.

MATERIALES Y MÉTODOS

I. Material vegetal y tratamientos de sequía con plantas de tres semanas de edad.

Las semillas mutantes de *Arabidopsis thaliana* (Columbia, mutantes dSpm, número de acceso NASC: N122943; Tissier *et al.*, 1999) fueron germinadas en cajas de Petri con MS 1X, agregando BASTA para verificar la inserción del transposón. Para los experimentos de sequía se cosecharon 12 g en peso fresco de plantas en irrigación óptima y 6 g de peso fresco de plantas en sequía, a partir de plantas transgénicas de la generación T₁ de 3 semanas de edad germinadas y crecidas en un sustrato inerte de baja retención de agua (Turface). Se llenaron charolas de 20 x 55 x 10 cm con Turface y se regaron a saturación con solución nutritiva. Se sembraron en la superficie las semillas transgénicas y se mantuvieron tapadas con un domo hasta 10 días después de la germinación, manteniendo una humedad relativa alta en estas etapas tempranas del establecimiento, debido a que el sustrato retiene muy poco la humedad. Debido a que el sustrato es inerte y posee muy baja retención de agua, las plantas se mantuvieron hidratadas a capacidad de campo hasta alcanzar la edad en la que se aplicó el tratamiento de estrés (16 días después de la germinación), regándolas con solución nutritiva hasta el inicio de la sequía, o hasta el final del experimento en el caso de las plantas control (22 días después de la germinación).

Para aplicar el tratamiento de sequía se absorbió todo el exceso de agua de las charolas y se mantuvieron sin riego durante 6 días, se rotaron de posición para una deshidratación homogénea. Se colectaron las plantas evitando las que se encuentran en los bordes, ya que estaban más deshidratadas que el resto de las plantas de la charola. Para la mitad del tejido colectado, tanto en control como en sequía, se llevó a cabo un entrecruzamiento con formaldehído al vacío durante 30 min, siguiendo el protocolo descrito por Rohila *et al.* (2004). Posteriormente, se lavó el tejido y se congeló en N₂ líquido para su uso en la extracción de complejos protéicos por TAP (Rubio *et al.*, 2005).

II. Obtención de plantas transgénicas que sobre-expresan la proteína de fusión AtLEA4-5-TAPa.

Se transformaron células electrocompetentes de *Agrobacterium tumefaciens* (cepa C58/pGV2260, Rif^R, Amp^R) con el vector binario para la construcción que lleva la etiqueta TAP (Tandem Affinity Purification) $2x35S\Omega::AtLEA4-5-cTAPa::NOS$ (Rubio *et al.*, 2005). En las colonias resistentes a los marcadores correspondientes se verificó la

presencia del gen mediante PCR de colonias aisladas. Las bacterias se utilizaron para transformar al vacío 25 plantas mutantes de *Arabidopsis* ecotipo Columbia que previamente se había comprobado que llevan una inserción del transposón dSpm en el promotor del gen *AtLEA4-5* (Tissier *et al.*, 1999; Olvera-Carrillo *et al.*, 2010). Se utilizó el método de infiltración floral al vacío (Bechtold *et al.*, 1993). Se cosecharon semillas de la T_0 y se seleccionaron las plantas transgénicas plaqueando aproximadamente 2000 semillas en un medio con 50 $\mu\text{g/ml}$ de sulfato de kanamicina (Sigma) y 25 $\mu\text{g/ml}$ glufosinato de amonio (BASTA; Finale, Aventis). Las plantas resistentes al marcador se trasplantaron a tierra y se dejaron autofecundar para cosechar las semillas de la T_1 . Se extrajo DNA genómico de una minipreparación utilizando hojas y flores de las plantas T_1 para verificar la expresión del transgen. Se sembraron 200 semillas T_1 de 12 líneas independientes para estimar el número de inserciones del transgen por la segregación del marcador. Se seleccionaron las líneas T_1 que mostraron una mayor expresión del transgen por experimentos tipo western, para llevar a cabo un experimento piloto con plantas en condiciones de irrigación óptima, comparando la muestra obtenida con plantas transformadas con el vector vacío como control negativo. Se verificaron los niveles de la proteína de fusión utilizando anticuerpos específicos para la *AtLEA4-5* y extractos totales de proteínas de plantas T_1 crecidas en condiciones de irrigación óptima. Las líneas T_1 que fueron seleccionadas para este primer experimento eran 100% resistentes al marcador de selección: kanamicina + BASTA (complementación de la mutante dSpm 4-5), por lo que contienen más de una inserción del transgen. Las demás líneas T_1 que mostraron una segregación de los marcadores de resistencia en una proporción 3:1, indicativo de una sola inserción, mostraron también una menor acumulación de proteína, por lo que estas cuatro líneas independientes fueron utilizadas para autofecundar nueve plantas de cada línea con segregación 3:1 para obtener cuando menos una planta homocigota para el transgen en la generación T_2 por cada línea.

Para verificar la funcionalidad de la proteína de fusión con la etiqueta TAP se tiene como perspectiva llevar a cabo experimentos de complementación, a través de analizar el fenotipo bajo condiciones de sequía severa y rehidratación en experimentos similares a los descritos en Olvera-Carrillo *et al.* (2010). Para ello se utilizarán plantas adultas de las líneas T_2 homocigotas en las que el transgen se encuentra en el fondo mutante dSpm::*AtLEA4-5*.

III. Espectrometría de masas.

Después del segundo paso de purificación de los complejos proteicos por afinidad de

con las perlas de Níquel-NTA, se lavaron las perlas dos veces con 5 ml de buffer de lavado (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% glycerol, 0.1% Nonidet P-40) durante 5 min y, posteriormente, se lavaron con 5 ml de 0.1 M CH₃CN durante 5 minutos. Las perlas que provenían de las purificaciones de extractos (en control y sequía) de tejidos con tratamiento de entrecruzamiento previo a la extracción de proteína se hirvieron durante 5 min en presencia de 1x buffer de Laemli sin colorante. La digestión con tripsina se llevó a cabo después de evaporar el β-mercaptoetanol del buffer Laemli, incubando las muestras a 37°C durante 24 hr. Las muestras con los péptidos se analizaron por espectrometría de masas en un ESI QUAD-TOF (nano spray and nano flow 0.2-1 μl/min-HPLC, Electro-spray ionization (ESI) MS/MS analysis, data interpretation, and database search, Waters/Micromass, Milford, MA, USA). Se inyectaron en la columna 10 μl de cada muestra y los resultados fueron procesados usando el algoritmo de búsqueda de la base de datos MASCOT (Hirosawa *et al.*, 1993). Las proteínas se consideraron identificadas confiablemente cuando MASCOT las enlistó como significativas (con un “score” mayor a 35) y uno o dos péptidos correspondían con la proteína. La base de datos buscada fue la base no redundante (NCBI nr 2009; 9617243 secuencias; 3288814006 residuos) del National Center for Biotechnology Information (NCBI, *Arabidopsis thaliana*; thale cress; 62336 secuencias). Los parámetros utilizados para la búsqueda fueron:

Modificaciones Variables: Acetyl (K), Acetyl (N-term), Formyl (N-term), Gln->pyro-Glu (N-term Q), Glu->pyro-Glu (N-term E), Oxidation (M), Phospho (ST)

Valores de la masas : Monoisotoópicos

Masa de las proteínas : Sin restricción

Masa de los péptidos : $\hat{A} \pm 80$ ppm

Masa de los fragmentos : $\hat{A} \pm 0.18$ Da

IV. Ensayos de dos híbridos en levadura.

La región codificante del gen (ORF) *AtLEA4-5* se clonó en fase con el vector pGBKT7 (Clontech) utilizando oligonucleótidos específicos para agregar los sitios de restricción *NdeI* en el extremo 5' y *BamHI* en el extremo 3'. Se hicieron dobles digestiones *NdeI/BamHI* (New England Biolabs) con los productos de PCR y, para linearizar el vector vacío se dejaron toda la noche a 16°C. Las bandas correspondientes a los productos de digestión se purificaron después de separarlas por electroforesis (Qiagen). Los productos de PCR se ligaron direccionalmente con el vector y, se verificó por secuenciación que quedara en fase con el dominio de unión a DNA de GAL4 presente en el vector. La cepa Mat α (MaV203) se transformó en presencia de

PEG (Yeast Protocols Handbook, Clontech) con el vector que lleva la proteína híbrida usada como anzuelo. La biblioteca de cDNAs de *Arabidopsis* con la que se contaba (obsequiada por el Dr John Cushman, obtenida de plántulas de *Arabidopsis* en estrés abiótico) estaba transformada en una cepa Mat a (AH109). Se verificó la expresión de la proteína de fusión AtLEA4-5-en experimentos tipo western y se procedió con el apareamiento, incubando ambas cepas en cajas con medio rico. Posteriormente, se sembraron en cajas con medio mínimo, utilizando como criterio de selección las auxotrofías de ambos plásmidos (leucina, triptofano) y la del gen reportero (histidina), así como el inhibidor 3-amino-triazol (3AT), en la concentración previamente determinada para evitar el escape del gen *HIS3*. La selección de las proteínas híbridas que interaccionan y activan el gen reportero se llevó a cabo en presencia de estrés osmótico (1.5 y 2 M de sorbitol), estrés iónico (1 M NaCl), así como en condiciones control (SD-L-T-H, 3AT 50mM). Se sembraron de tres a cinco réplicas de cada condición, y se repitió dos veces el tamizado utilizando una menor proporción de la cepa que lleva la biblioteca transformada. Se llevaron a cabo ensayos de actividad de β -galactosidasa utilizando las colonias positivas del ensayo de apareamiento replicando las colonias a filtros de nitrocelulosa ("colony-lift filter assay", Yeast Protocols Handbook, Clontech).

RESULTADOS Y DISCUSION

I. El análisis de dos híbridos en levadura sugiere que la AtLEA4-5 tiene interacciones electrostáticas con otras proteínas y los dominios de interacción pueden estar relacionados a la estructura secundaria de las proteínas interaccionantes.

Se utilizó el sistema de dos híbridos en levadura para tener un enfoque *in vivo* de la interacción de la proteína AtLEA4-5 en un sistema de expresión heterólogo y porque nos permitía aplicar una selección bajo condiciones de déficit hídrico. El tamizado de colonias se llevó a cabo en una situación de estrés, ya que una de las dificultades del estudio de este tipo de proteínas hidrofílicas es la flexibilidad en su estructura, por lo que pensamos que si la selección se llevaba a cabo en las condiciones intracelulares que se imponen durante un estrés osmótico, se podrían determinar interacciones que no podrían detectarse bajo una condición en donde el agua estuviera disponible en abundancia. Se utilizó una biblioteca de cDNAs de plantas de Arabidopsis de plántulas de dos semanas de edad, sometidas a tratamientos de deshidratación, salinidad, alta luz y frío. La biblioteca se clonó como una fusión al dominio de activación de GAL4 que se usó como “presa” para el ensayo de dos híbridos, usando la auxotrofia a leucina (L) como marcador de selección para este plásmido. Esta biblioteca fue obtenida de plántulas de Arabidopsis en condiciones de estrés por deshidratación, alta salinidad, temperaturas bajas y alta iluminación, y fue donada por el laboratorio del Dr. John Cushman, de la Universidad de Nevada. Esta sección del trabajo se llevó a cabo como parte de las actividades de una estancia de investigación en la Universidad de California, Riverside, en el laboratorio del Dr. Jian-kan Zhu, dentro del Programa de Becas Mixtas de Investigación otorgada por CONACYT. El plásmido pGBKT7 (Clontech) en forma de DNA de doble cadena y la biblioteca clonada en el vector pGADT7-Rec ya se encontraba transformada en una cepa de levadura con un tipo de apareamiento “a” (AH109), ambos fueron enviados a la Universidad de California. Se clonó el cDNA de la AtLEA4-5 como una fusión traduccional al dominio de unión a DNA de GAL4 del vector pGBKT7 y se transformó en una cepa con el tipo de apareamiento “ α ” de tal forma que pudiera ser usado como “anzuelo” en el ensayo de dos híbridos en levadura, usando la auxotrofia a triptófano (T) como marcador de selección. Se determinó la concentración del inhibidor 3-amino triazol (3-AT), que permite eliminar el escape en el gen *HIS3* (un gen que codifica para una enzima de la ruta de biosíntesis de histidina, H), que es uno de los genes reporteros cuya transcripción se activa al interaccionar las dos proteínas híbridas. En los ensayos de

apareamiento entre la cepa que contiene el anzuelo y la cepa que contiene la biblioteca de cDNAs se utilizó 3-AT 50 mM. Después de cruzar (aparear) las dos cepas durante 4 h en cajas con medio rico, se procedió a la selección de las colonias diploides capaces de crecer con las auxotrofías correspondientes. Para ello se utilizó medio mínimo (placas con medio SD-T-L-H + 50 mM 3AT), o bien medio mínimo adicionado con sorbitol para imponer condiciones de estrés osmótico (placas con medio SD-T-L-H + 50 mM 3AT + 1.5 ó 2 M sorbitol), o bien con sal para inducir estrés iónico (placas con medio SD-T-L-H + 50 mM 3AT + 1M NaCl). Adicionalmente, se sembró una dilución 1:30,000 de cada apareamiento y se sembraron en cajas sin histidina (SD-T-L) para cuantificar el número de colonias diploides en ausencia de la selección de la auxotrofia para la activación del gen reportero. En promedio se estimaron un total de 2'210,000 colonias diploides en el tamizado, (entre 1'740,000 y 2'460,000 colonias), a partir de las cuales se hizo la selección en cada una de las condiciones descritas. Las cajas se incubaron por cuatro días y dos semanas hasta que se observó un crecimiento diferencial entre colonias grandes y pequeñas, considerando a las primeras como positivas por su capacidad de reproducirse a mayor velocidad en el medio de selección (Figura 11).

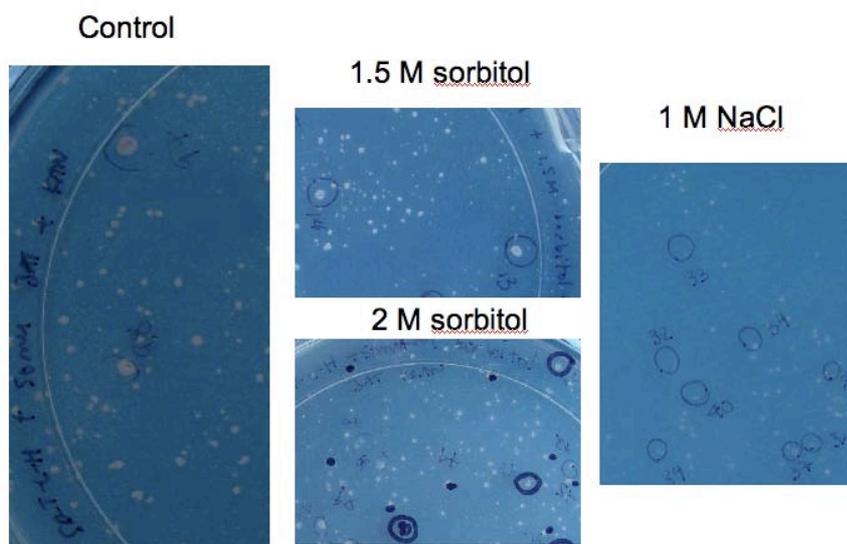


Figura 11. Tamizado de colonias diploides en ensayos de dos híbridos de levadura en condiciones control o en presencia de estrés osmótico (sorbitol 1.5 y 2M) o estrés salino (NaCl 1M). La interacción de las proteínas híbridas se llevó a cabo por apareamiento de las cepas MaV203 que llevan la proteína AtLEA4-5 fusionada al dominio de unión a DNA de GAL4, la cual fue usada como anzuelo, y la cepa AH109 transformada con una biblioteca de cDNAs de plántulas de Arabidopsis en tratamientos de estrés abiótico, fusionadas al dominio de activación de GAL4.

Adicionalmente, en colaboración con el estudiante de maestría César Cuevas, se llevó a cabo el ensayo de activación de otro gen reportero, *lacZ* que codifica para la enzima β -galactosidasa utilizando 120 colonias diploides que se obtuvieron como

candidatos en el ensayo de activación de *HIS3*. Se crecieron en medio SD-T-L para posteriormente transferirse a filtros de nitrocelulosa, en donde se incubaron con el sustrato X-gal, el cual genera una coloración azul en las colonias si éstas contienen células en donde ocurrió la activación del gen reportero. Se revisaron a partir de las 4 h después de la incubación, y se descartaron las colonias que no dieron coloración hasta después de 24 h de incubación con el sustrato.

Se extrajo el plásmido de los candidatos de interacción que presentaron coloración azul, después de crecer los cultivos de levadura en medio líquido (SD-L) y se transformaron células competentes de *E. coli* con estos plásmidos para su análisis por secuenciación, a partir de cultivos realizados en placas de 96 pozos. El análisis de las secuencias obtenidas indicó que la construcción de la biblioteca de cDNAs utilizada fue deficiente, ya que la longitud promedio de los insertos en las clonas fue inferior a lo esperado y, en muchas clonas se detectaron cambios de fase y codones de término prematuros.

De manera preliminar se puede decir que se detectaron entre las colonias positivas a varios péptidos ricos en aminoácidos pequeños y cargados, y que este tipo de interacción con la proteína híbrida AtLEA4-5 se observó tanto en condiciones óptimas como en condiciones de estrés. A pesar de los problemas encontrados, fue posible identificar de este primer análisis una clona en fase al dominio de activación de GAL4 para cada tipo de estrés y para la condición control.

Tabla 4. Secuencias clonadas en fase al dominio de activación de GAL4, obtenidas del ensayo de activación del gen reportero de histidina bajo condiciones óptimas (SD-L-T-H, 3AT 50 mM). Las clonas marcadas con negritas, son las que dieron coloración azul cuando se probaron en el ensayo de activación del gen reportero de β -galactosidasa.

No.	Secuencia clonada.
2	CG RERRT KRR LHHE SQVEE AYEET KEE APKDET AIQVH SIGRV SK FLS YV IPF
3	GAEKRER KQN KKWL PE KL
5	GNRRR LS VR FSTN GEG
43	GERRNT KDADLREN LDRQDH SRGREQRRHQQ GQDP RQRNP SGS AEIDLRR KAA RR WPY FG
52	GSLG HYT LK TINSLQ
53	GVLELRRRQDRVR KTC GEVC FIQEV SPL EE WAQT SS TSSL TQT LPP STIG YT ATT LG ASTVEPR G
63	GI APITIYSL
68	GERRR WRRR IE
73	LGKKKKKKKKK XCRPPR LEGG HRY GI HR ARAA XE SXIL KNP ASSL QLCIXHH LNF HLYIVLPSFM
95	GVLELRRRQDRVR KTC GEVC FIQEV SPL EE WAQT SS TSSL TQT LPP STIG YT ATT LG ASTVE PR G
96	At 5g 5000 protein kinase, putative AMPNN ICCVWEY LPGA ALKS YLIKNRRR KL TFKI WQ LAL DLARGLS YLHSQKI VHRDV KTIENM LL DK TRIV KI ADFGVARVE ASNNDMI GET GTLG YM APEV LNGN PYNR KQVY SFGI CLW EIYC DMPYD LTFSEVT SAVRNQ LRPD IPRCP SALAA VM KRC WDANPD KRPEDEVPM LESIDIT KCGM PD QQ GCLCFRR KRGP
97	KPIIW HFF QV FLDLYLRLES L FRR ASKI RLS RY SK TWG RY
99	GS RGRX GHM IHH DIL XH XXX FWLGG ERIL YQ
103	GSL VGG

(Tablas 4-6). En el caso del tamizado de colonias diploides en condiciones control, se obtuvieron colonias de diferentes tamaños, al cabo de una semana se tomaron las

más grandes para replicarlas en otra caja con las auxotrofías correspondientes. Se encontró en fase una putativa proteína cinasa (At5g50000, Tabla 4). En el caso del tamizado de colonias diploides en presencia de estrés salino (1 M NaCl) se observó un severo retraso en el crecimiento de las colonias, debido a la toxicidad iónica; al cabo de dos semanas se eligieron las colonias aisladas que estaban ligeramente más grandes y después de analizar varias colonias que también dieron positivo en el ensayo de activación del gen *lacZ*, se encontró una clona en fase que corresponde a una chaperona de bajo peso molecular 23.6 localizada en la mitocondria (At4g25200, Tabla 5).

Tabla 5. Secuencias clonadas en fase al dominio de activación de GAL4, obtenidas del ensayo de activación del gen reportero de histidina bajo condiciones de estrés osmótico impuesto con NaCl 1M (SD-L-T-H, 50 mM 3AT, 1 M NaCl). Las clonas marcadas con negritas, son las que dieron coloración azul cuando se probaron en el ensayo de activación del gen reportero de β -galactosidasa.

No.	Secuencia clonada
4	GRVRPCLCPYIPLSPRT
5	GAGSTRGKDLIGWRITRCSRIHRRWMKEILITRNR
6	GKAKRICFYEQRASQRE
7	GVNSRK
10	GEGEDQTANSRS YRSTVESHDTLPST QIQALRGTS
12	GALWESCNYNFYLSTDFSQIQLTIKPLQNP
22	ATHSP23.6-MITO (MITOCHONDRION-LOCALIZED SMALL HEAT SHOCK PROTEIN 23.6) AT4G25200 GAHSSQINTKLTVKTFSTKLPVSLQMASALALKRLLSSS IAPRSRSLRPAVSSRLFNNAV RSYDDDDGENGDGVDLYRRSVPRRRGDFFSGNKPIVSLFLQIYVSLLLT
30	AIGRGRGLALPRWQLRPIGFGYRPRGIR
32	GQDGRYGSEIRLRKVLRV
34	PEQSRTRMVISRPPCSSILKNTEVSKGLMV QIRSIQTPYWLRIVISSSLQHLVVSSTG
35	GEEIERRIV
36	WGGRIPIRRRTIPRRRWATRRRGKLLPPLLQRLP WLLKVLFCRRSGSDSVWVSLNLHNIS TTMSVRFHLYIHLRVYVI

En el caso del tamizado de colonias en presencia de dos concentraciones de sorbitol (1.5 M y 2 M) se observó que las colonias presentaban una clara diferencia de tamaño de manera semejante a lo que se observó en condiciones control. Esta diferencia en el tamaño fue evidente al cabo de dos semanas de incubación, y no se encontraron diferencias en la cantidad de colonias de tamaño grande obtenidas ante una concentración mayor de sorbitol, aunque sí presentaron una apariencia cristalina y color más claro, lo cual es diferente a la de colonias crecidas en medio sin estrés osmótico (Figura 18). En este caso, el análisis por secuenciación indicó que también hay variedad de péptidos que se produjeron por el método de clonación utilizado en el sistema Clontech, produciendo muchas clonas con cambios de fase que produjeron péptidos aberrantes o con codones de término prematuros. Interesantemente, en algunas clonas positivas, se observó que algunos péptidos que se fusionaron con el dominio de activación de GAL4 correspondían a las regiones 3' no traducidas de genes que tenían secuencias poliadeniladas, por lo que al traducir estas secuencias

en el sistema híbrido, se produjeron péptidos ricos en lisinas, los cuales pudieron activar el reportero, a diferencia de lo que se observa transformando con el vector anzuelo que lleva la AtLEA4-5 con el vector presa vacío, es decir, sin la biblioteca de cDNAs clonada en fase con el dominio de activación de GAL4. En este tamizado de colonias se encontró una clona en fase que codifica para un fragmento de la proteína ribosomal L17, que forma parte de la subunidad 60S del ribosoma (Tabla 6). Este tipo de proteínas que se encuentran en el citoplasma son un candidato de interacción que se puede comprobar tomando como antecedente otros experimentos preliminares de localización celular por fluorescencia, donde se ha observado que la fluorescencia tiene una distribución núcleo-citoplasmática (Cuevas, C., com, pers).

Tabla 6. Secuencias clonadas en fase al dominio de activación de GAL4, obtenidas del ensayo de activación del gen reportero de histidina bajo condiciones de estrés osmótico impuesto con sorbitol 1.5 M (SD-L-T-H, 50 mM 3AT, 1.5 M sorbitol). Las clonas marcadas con negritas, son las que dieron coloración azul cuando se probaron en el ensayo de activación del gen reportero de β -galactosidasa.

No.	Secuencia clonada
9	GGAKRRASSGAEVEHK
49	GEKRSREE
A	GREAKEKKAV
51	GALXPSLTR
64	GRRDREKNSIIL
86	GALRPSLTR
88	GRKRRKGIWRS
90	GVLLHKCRGSGTTLRKRERKSHLISLKNLEHPDRLLSSFWG
91	At1g67430 60S ribosomal protein L17 (RPL17B) GQPILLLAAMVKYSQEPDNQTKSCKARGSDLRVHFKNTRTAHAIRKLPLIAKARYL EDVIAHKQAIPFTRFCRGVGRTA QAKNRHSNGQGRVPAKSAQFVLDLLKNAESNAEVKGLD DALFISHIQVNQAAKQ RRRTYRAHGRINPYMSNPCHIELIL SEKEEPVKKEPETQLAAKSKKSAA

Este enfoque ha sido continuado por otro miembro del laboratorio inscrito en el programa de Maestría (Cuevas-Velázquez, C.) para indagar si hay cambios en la activación del gen reportero *lacZ* en presencia o ausencia de estrés en ensayos cuantitativos en medio líquido. Con base en los resultados obtenidos seleccionará aquéllos que se consideren positivos, los cuales retransformará para confirmar la interacción.

Idealmente, estas interacciones deberán ser probadas *in vivo* en un sistema vegetal, aprovechando las técnicas de fluorescencia por complementación (BiFC), que se basan en la expresión transitoria de proteínas híbridas, fusionadas al extremo amino o al extremo carboxilo de una proteína fluorescente. En este caso se construyen dos tipos de proteínas híbridas una que contiene a la proteína anzuelo (e.g. LEA4-5) fusionada a uno de los extremos de una proteína fluorescente (e.g. N-

YFP), y la otra que contendría al ligando candidato fusionado al extremo complementario de la misma proteína fluorescente (e.g. YFP-C); estas construcciones se expresan de forma transitoria o estable en plantas y se determina si se genera fluorescencia o no. Con ello se podría determinar si dicha interacción ocurre *in vivo* con la proteína AtLEA4-5 y, de ser el caso, si tal interacción ocurre de manera diferencial en presencia de estrés osmótico o iónico. Esta estrategia se usaría para determinar la funcionalidad de las interacciones que hayan sido seleccionados de los diferentes enfoques analizados, tanto del sistema de doble híbrido en levadura, como del sistema *in vitro* conocido como TAP (Tandem Affinity Purification), a través de la purificación de complejos proteicos obtenidos de extractos vegetales provenientes de plantas transgénicas que expresan la proteína de interés etiquetada y que serían crecidas bajo condiciones de déficit hídrico, en las que se ha demostrado que la expresión de la proteína de interés es muy abundante y sostenida.

II. La purificación de complejos proteicos por TAP (“Tandem Affinity Purification”) mostró una variedad de proteínas localizadas en diferentes compartimentos celulares que pueden interactuar con la AtLEA4-5.

Para complementar y comparar los resultados obtenidos con el ensayo de dos híbridos en levadura, se usó un enfoque *in vitro* utilizando métodos bioquímicos de purificación de proteínas. Originalmente se había pensado utilizar una columna de afinidad en donde se inmovilizara a la AtLEA4-5 y se enfrentara con extractos de proteínas vegetales. Sin embargo, debido a las características intrínsecas de estas proteínas flexibles relacionadas con los cambios estructurales inducidos por cambios en las condiciones del medio que podrían afectar sus interacciones, preferimos usar otro sistema que permitiera purificar complejos de la proteína de interés con sus blancos de interacción, directamente de extractos que provinieran de plantas en sequía. El sistema denominado TAP-tag (por Tandem Affinity Purification) consiste en una doble purificación a través de columnas de afinidad en condiciones de extracción relajadas, para evitar que los complejos que llevan la proteína etiquetada y las proteínas que interactúan con ella, no se pierdan en el proceso de purificación; entre ambos pasos de purificación hay un corte con una proteasa que, en esta versión modificada del vector TAP (TAPa, Rubio *et al.*, 2005), corresponde a una proteasa de rinovirus que funciona a 4°C, lo cual favorece que los complejos se mantengan estables durante este paso. Elegimos hacer la fusión hacia el extremo carboxilo de la proteína de interés, debido a que el extremo que está más conservado se encuentra en la región amino de la proteína y, bajo la hipótesis de que esta región conservada es

la que permitiría interactuar con sus blancos por su potencial para formar estructuras tipo “coiled-coil”. La secuencia de aminoácidos y la disposición del “cassette” de expresión en la fusión hacia el extremo carboxilo de la proteína de interés se muestran en la Figura 12.

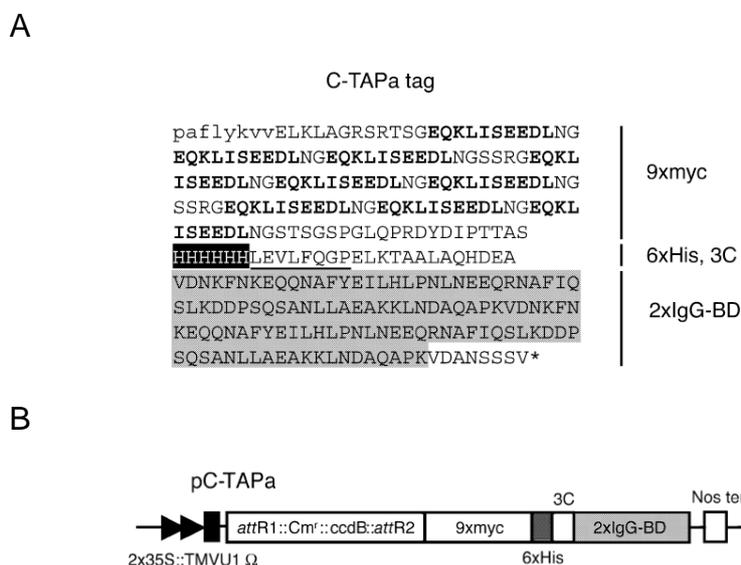


Figura 12. Vector utilizado para la obtención de la proteína de fusión AtLEA4-5-TAPa (reproducido de Rubio *et al.*, 2005). A) Secuencia de aminoácidos de la etiqueta TAP. La secuencia corresponde a dos copias del dominio de unión a IgG de la proteína A, marcada en gris. El repetido de seis histidinas se resalta con negro. La secuencia de corte que reconoce la proteasa 3C está subrayada. Las nueve repeticiones del epítipo Myc se muestran en negritas. La secuencia del “cassette” de recombinación GATEWAY *attB2* se muestran en minúsculas. El asterisco corresponde al codón de término de la fusión C-terminal. La etiqueta C-TAPa contiene 307 residuos. B) Representación diagramática del “cassette” de expresión en el vector C-TAPa, que permite su fusión traduccional en el extremo carboxilo de la proteína de interés. La expresión está regulada por dos copias del promotor 35S del virus de mosaico de coliflor (2x35S) y un “enhancer” traduccional del virus de mosaico del tabaco (TMVU1 Ω). La etiqueta TAPa consiste en dos copias del dominio de unión a IgG de la proteína (2xIgG-BD), una secuencia de ocho aminoácidos que corresponde al sitio de corte de la proteasa 3C (3C), seis histidinas en tandem (6xHis), y nueve repeticiones del epítipo Myc (9xMyc). El vector pC-TAPa contiene los sitios de recombinación GATEWAY (*attR1::Cmr::ccdB::attR2*). La secuencia del terminador NOS (Nos ter) está localizada después del cassette de expresión.

En primer lugar, se generó la construcción por recombinación del vector pENTR/D/TOPO-LEA4-5, que lleva el ORF de la AtLEA4-5 y utilizando el sistema Gateway (Invitrogen) se recombinó con el vector C-TAPa (Rubio *et al.*, 2005), se transformaron bacterias con esta construcción y se comprobó por secuenciación que la proteína estuviera en fase con la etiqueta, dando lugar a una proteína de 55 kDa. La construcción obtenida se introdujo en plantas de Arabidopsis con un fondo genético mutante para la proteína de interés, minimizando así la competencia de la proteína

endógena con la proteína de fusión, ya que sabemos que los niveles de acumulación de la proteína endógena son muy elevados en condiciones de déficit hídrico. Se analizaron 8 líneas transgénicas independientes T₁ y se verificó por análisis de tipo western de extractos totales con anticuerpos específicos dirigidos contra la proteína AtLEA4-5 que expresaran de manera constitutiva el transgen que codifica para la proteína de fusión (AtLEA4-5-cTAP). De las plantas analizadas se seleccionaron tres líneas que expresaron en mayor cantidad dicha proteína (Líneas 5, 17, 23, Figura 13).

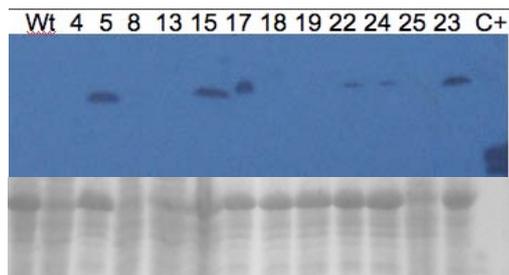


Figura 13. Experimento tipo Western utilizando anticuerpos específicos contra la AtLEA4-5 y extractos totales de plantas silvestres y plantas transgénicas mutantes que llevan la proteína de fusión (indicada con una flecha). Como control positivo se utilizó un extracto de la proteína recombinante, y los anticuerpos reconocen a la proteína en el tamaño esperado para el monómero (indicada con un asterisco). Como control de carga se muestra la tinción reversible con rojo de Ponceau después de la transferencia.

Para la purificación de los complejos proteicos por el método de TAP se siguió el protocolo descrito por Rubio *et al.* (2005), que se encuentra esquematizado la Figura 13.

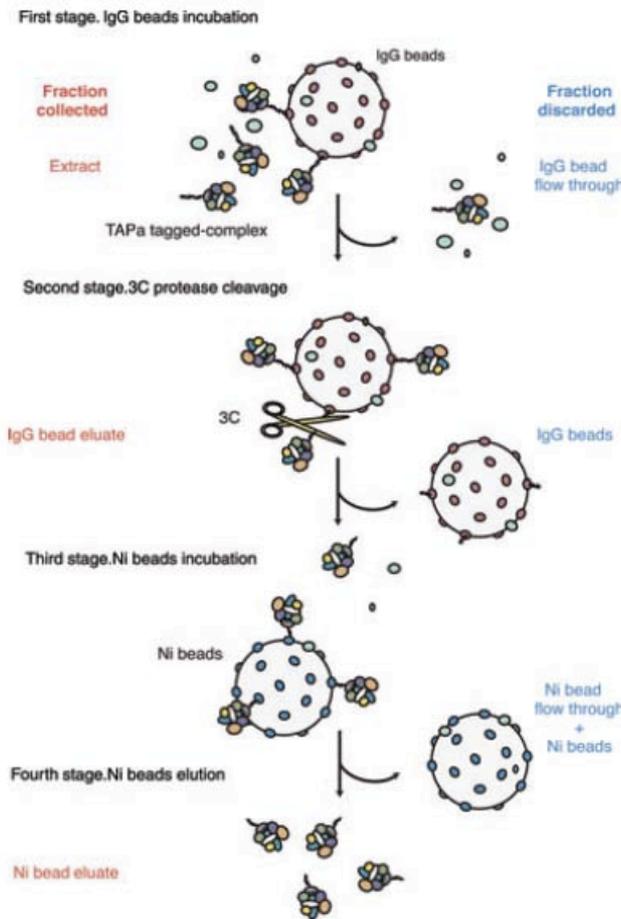


Figura 14. Procedimiento para la extracción de complejos proteicos con el método de TAPa (reproducido de Rubio *et al.*, 2005). 1) Los extractos totales de proteína se incuban con perlas de IgG en el primer paso de purificación por afinidad. 2) La elución de las proteínas etiquetadas con TAPa requiere la digestión específica de las perlas de IgG, con una proteasa de rinovirus, (3C protease) que es activa a temperaturas bajas. 3) El siguiente paso de purificación por afinidad consiste en la incubación de los sobrenadantes (recuperados después de la digestión de perlas de IgG) con perlas de Ni o, alternativamente, con perlas α -Myc. 4) La elución de las proteínas no se llevó a cabo en el presente trabajo, sino que se lavaron 3 veces con el buffer de lavado, y después con el buffer de digestión de tripsina, y, para las perlas que provenían de los tratamientos de entrecruzamiento, se hirvieron en buffer de Laemmli antes de lavar con el buffer de tripsina. Las fracciones colectadas durante cada etapa, se muestran en color rojo, y las fracciones que se descartaron se muestran en color azul.

Los complejos de proteínas que quedan inmovilizados en la segunda purificación en una matriz de agarosa acoplada con níquel se digieren con tripsina y se identifican los péptidos, por espectrometría de masas tanto de la proteína de interés como los correspondientes a las proteínas interactoras. Como experimento inicial se llevó a cabo la purificación de complejos de proteínas en condiciones sin estrés, para determinar si la proteína de interés se detectaba entre los péptidos presentes en la muestra y no en el control negativo. Se utilizaron plantas transgénicas que expresaban el vector sin la proteína de interés como control negativo. La proteína de interés se encontró en la muestra de plantas obtenidas en condiciones óptimas de crecimiento,

junto con proteínas de unión a RNA, proteínas ribosomales, y proteínas similares a cinasas. (resaltadas en amarillo en la Figura 22). La proteína de interés no se expresa en condiciones de riego óptimo y debido a que las plantas transformadas con la construcción contienen la inserción estable de un transposón en el promotor del gen AtLEA4-5, no se detecta la proteína endógena en la muestra del control negativo. Al comparar los péptidos identificados en cada una de las condiciones, se pudieron descartar los péptidos que se copurifican como fondo, ya que se obtienen de los extractos proteicos de plantas en donde no se está expresando la proteína de fusión (marcados en color azul en la Figura 15).

gi 15237059	ATRABE1B (ARABIDOPSIS RAB GTPASE HOMOLOG E1B); GTP binding / GTPase/ translation elongation f
2. gi 15240144	late embryogenesis abundant group 1 domain-containing protein / LEA group 1 domain-containing protein [A
gi 30692067	NIT1; indole-3-acetonitrile nitrilase/ indole-3-acetonitrile nitrile hydratase/ nitrilase [Arabidopsis thaliana]
gi 15222741	FBR12 (FUMONISIN B1-RESISTANT12); translation initiation factor [Arabidopsis thaliana]
gi 166706	cystolic glyceraldehyde-3-phosphate dehydrogenase
gi 15230291	ATRBP47B (RNA-binding protein 47B); RNA binding [Arabidopsis thaliana]
gi 30695318	transducin family protein / WD-40 repeat family protein [Arabidopsis thaliana]
gi 15236849	hydroxyproline-rich glycoprotein family protein [Arabidopsis thaliana]
gi 15220146	PTAC17 (PLASTID TRANSCRIPTIONALLY ACTIVE17) [Arabidopsis thaliana]
gi 30696487	50S ribosomal protein L24, chloroplast (CL24) [Arabidopsis thaliana]
gi 15229936	NIT3 (NITRILASE 3); indole-3-acetonitrile nitrilase/ indole-3-acetonitrile nitrile hydratase/ nitrilase [Arabidopsis
12. gi 468773	50 S ribosomal protein L12 [Arabidopsis thaliana]
gi 10740175	Rubisco activase [Arabidopsis thaliana]
14. gi 22329848	kinase-related [Arabidopsis thaliana] Atlg29350/F15D2_27
gi 222423868	AT1G07930 [Arabidopsis thaliana]
gi 227202852	AT5G60390 [Arabidopsis thaliana]
17. gi 186511104	P5CS2 (DELTA 1-PYRROLINE-5-CARBOXYLATE SYNTHASE 2); catalytic/ glutamate 5-kinase/ oxidoreduct
18. gi 42571135	P5CS1 (DELTA1-PYRROLINE-5-CARBOXYLATE SYNTHASE 1); delta1-pyrroline-5-carboxylate synthe
	AT2G39800
19. gi 12323530	RNA-binding protein MEI2, putative; 36123-32976 [Arabidopsis thaliana] AT1G29400
gi 15234126	CIP7 (COPI-INTERACTING PROTEIN 7); transcription activator [Arabidopsis thaliana]
21. gi 21537296	ribosomal protein L4 [Arabidopsis thaliana] RPL4; poly(U) binding
gi 22329999	polypyrimidine tract-binding protein, putative / heterogeneous nuclear ribonucleoprotein, putative [Arabidopsis tha
gi 16934	catalase [Arabidopsis thaliana]

Figura 15. Comparación de péptidos obtenidos en el análisis por espectrometría de masas, a partir de extractos proteicos de plantas mutantes dSpm::AtLEA4-5 transformadas con la proteína de fusión AtLEA4-5- C-TAPa o a partir de extractos de plantas mutantes transformadas con el vector vacío.

Los péptidos identificados que se consideraron provenientes de una interacción específica con la proteína de fusión y que tienen una puntuación confiable (mayor a 35) se muestran en la Figura 16.

2.	gi 15240144	Mass: 16169	Score: 365	Queries matched: 10	emPAI: 0.76				
	late embryogenesis abundant group 1 domain-containing protein / LEA group 1 domain-containing								
Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
387	657.3525	1312.6904	1312.6987	-6.29	0	32	12	1	R.VVEGTAVTDPGR.N
567	855.4108	1708.8070	1708.8123	-3.10	1	117	5.2e-008	1	-.MQSMKETASNIAASAK.S 566
577	863.4157	1724.8168	1724.8073	5.56	1	(113)	1.5e-007	1	-.MQSMKETASNIAASAK.S 576
582	871.4110	1740.8074	1740.8022	3.02	1	(75)	0.0009	1	-.MQSMKETASNIAASAK.S
753	707.3584	2119.0534	2119.0401	6.25	2	11	2.4e+003	7	K.QMATQVKEDKINQAEQK.R

14. [gi|22329848](#) Mass: 90833 Score: 72 Queries matched: 4 emPAI: 0.04
kinase-related [Arabidopsis thaliana]

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
541	819.9428	1637.8710	1637.8009	42.8	0	13	8.6e+002	1	R.LENIIHIAESGPFSESR.K
544	829.4235	1656.8324	1656.7757	34.2	0	17	5.3e+002	1	Q.SASGHQTAWFGAPGQR.S
762	1074.9867	2147.9588	2147.9582	0.30	0	72	0.002	1	R.ELDFQYSPFSAQQSMQSR.T
900	862.4778	2584.4116	2584.2186	74.7	2	9	2.3e+003	8	K.GIQDIPSGSRKIVQSLTEIVNS.P

Proteins matching the same set of peptides:
[gi|18087627](#) Mass: 91016 Score: 72 Queries matched: 4
 Atlg29350/F15D2_27 [Arabidopsis thaliana]
[gi|22329844](#) Mass: 90824 Score: 72 Queries matched: 4
 unknown protein [Arabidopsis thaliana]

18. [gi|42571135](#) Mass: 66718 Score: 63 Queries matched: 3 emPAI: 0.10
P5CS1 (DELTA1-PYRROLINE-5-CARBOXYLATE SYNTHASE 1); delta1-pyrroline-5-carboxylate synthetase [Arabidopsis]

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
335	599.8396	1197.6646	1197.6718	-5.93	0	28		19	R.GPVGVEGLLTR.W
781	735.7693	2204.2861	2204.2304	25.3	1	66	0.0016	1	R.EEIPDLLKLDVLDVIPR.C
807	756.0582	2265.1528	2265.0119	62.2	2	6	7.1e+003	3	T.GPPSPNSKLIHTFVKEK.H

Proteins matching the same set of peptides:
[gi|938021](#) Mass: 77771 Score: 63 Queries matched: 3
 delta1-pyrroline-5-carboxylate synthetase [Arabidopsis thaliana]
[gi|15225571](#) Mass: 77654 Score: 63 Queries matched: 3
 P5CS1 (DELTA1-PYRROLINE-5-CARBOXYLATE SYNTHASE 1); delta1-pyrroline-5-carboxylate synthetase [Arabidopsis]
[gi|19423916](#) Mass: 77684 Score: 63 Queries matched: 3
 putative delta-1-pyrroline 5-carboxylase synthetase P5C1 [Arabidopsis thaliana]
[gi|222424578](#) Mass: 72320 Score: 63 Queries matched: 3
 AT2G39800 [Arabidopsis thaliana]

19. [gi|12323530](#) Mass: 85759 Score: 60 Queries matched: 2 emPAI: 0.08
RNA-binding protein MEI2, putative; 36123-32976 [Arabidopsis thaliana]

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
771	1083.5349	2165.0552	2165.0364	8.71	0	50	0.29	1	-_MDIPHEAEAGANGILPGFGR.H 772

Proteins matching the same set of peptides:
[gi|30690712](#) Mass: 88157 Score: 60 Queries matched: 2
 AML5 (ARABIDOPSIS MEI2-LIKE PROTEIN 5); RNA binding / nucleic acid binding / nucleotide binding [Arabidopsis thaliana]

21. [gi|21537296](#) Mass: 30566 Score: 56 Queries matched: 2 emPAI: 0.11
ribosomal protein L4 [Arabidopsis thaliana]

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
108	365.2148	728.4150	728.3970	24.8	0	6	3.2e+003	1	V_FGPRPK.D
960	957.8753	2870.6041	2870.5905	4.71	1	56	0.033	1	R.TLNLFDILNADKLVLTAAVEFLNAR.Y

Proteins matching the same set of peptides:
[gi|2791998](#) Mass: 30506 Score: 56 Queries matched: 2
 ribosomal protein L4 [Arabidopsis thaliana]
[gi|18390754](#) Mass: 30540 Score: 56 Queries matched: 2
 RPL4; poly(U) binding / structural constituent of ribosome [Arabidopsis thaliana]
[gi|13660078](#) Mass: 30386 Score: 56 Queries matched: 2

Figura 16. Péptidos obtenidos en el análisis por espectrometría de masas, a partir de extractos proteicos de plantas mutantes dSpm::AtLEA4-5 transformadas con la proteína de fusión AtLEA4-5-C-TAPa. Se muestran los péptidos después de su identificación con el algoritmo MASCOT.

En otro experimento independiente, se procedió a extraer proteína de muestras de plantas que sufrieron un tratamiento de sequía antes del proceso de purificación; se aplicó entrecruzamiento con formaldehído a una fracción del material vegetal de ambos tratamientos (A, sequía, con entrecruzamiento; B, riego óptimo, con

entrecruzamiento), usando las plantas completas una vez cosechadas y previo a la homogenización en presencia de nitrógeno (ver Materiales y Métodos). Se analizaron cuatro muestras en una misma corrida de MS/MS, comparando los péptidos identificados en cada condición. A continuación se muestran los perfiles MASCOT correspondientes a algunos péptidos identificados en el tratamiento de sequía con la aplicación de entrecruzamiento con formaldehído, previo a la homogenización del tejido (Figura 17). Es interesante notar que las proteínas encontradas aparentemente se localizan en diferentes compartimientos celulares, tales como núcleo y citoplasma, o membranas por lo que resulta relevante el conocer la distribución de esta proteína en la célula bajo condiciones de déficit hídrico. Para ello se llevarán a cabo estudios de la localización intracelular de la proteína por medio de inmunolocalización con anticuerpos dirigidos contra la proteína AtLEA4-5, lo que permitiría definir si estas proteínas son candidatos reales de interacción en situaciones de déficit hídrico.

A

1. [gi|1592679](#) Mass: 16164 Score: 200 Queries matched: 3 emPAI: 0.46
LEA D113 homologue type1 [Arabidopsis thaliana]
 Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
<input checked="" type="checkbox"/> 521	855.4077	1708.8008	1708.8123	-6.73	1	(74)	3.7e-005	1	-_MQSMKETASNIAASAK.S + Acetyl (N-term)
<input checked="" type="checkbox"/> 522	855.4100	1708.8054	1708.8123	-4.04	1	117	1.9e-009	1	-_MQSMKETASNIAASAK.S + Acetyl (N-term)
<input checked="" type="checkbox"/> 532	863.4063	1724.7980	1724.8073	-5.34	1	(94)	3.6e-007	1	-_MQSMKETASNIAASAK.S + Acetyl (N-term); Oxidation (M)

Proteins matching the same set of peptides:
[gi|15240144](#) Mass: 16169 Score: 200 Queries matched: 3
late embryogenesis abundant group 1 domain-containing protein / LEA group 1 domain-containing protein [Arabidopsis thaliana]

2. [gi|15236401](#) Mass: 17305 Score: 72 Queries matched: 1 emPAI: 0.20
major latex protein-related / MLP-related [Arabidopsis thaliana]
 Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
<input checked="" type="checkbox"/> 607	940.4816	1878.9486	1878.9575	-4.69	1	72	5.3e-005	1	M_ATSGTYVTEVPLKGSAEK.H + Acetyl (N-term)

Proteins matching the same set of peptides:
[gi|15236566](#) Mass: 17507 Score: 72 Queries matched: 1
major latex protein-related / MLP-related [Arabidopsis thaliana]
[gi|15236568](#) Mass: 17463 Score: 72 Queries matched: 1
major latex protein-related / MLP-related [Arabidopsis thaliana]
[gi|18379240](#) Mass: 17502 Score: 72 Queries matched: 1
MLP328 (MLP-LIKE PROTEIN 328); copper ion binding [Arabidopsis thaliana]
[gi|227202622](#) Mass: 11616 Score: 72 Queries matched: 1
AT4G23670 [Arabidopsis thaliana]

3. [gi|18087513](#) Mass: 60143 Score: 34 Queries matched: 3 emPAI: 0.05
AT5g18480/F20L16_200 [Arabidopsis thaliana]
 Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
<input checked="" type="checkbox"/> 328	577.2871	1152.5596	1152.6139	-47.05	1	(22)	5.2	1	K_VKTLSSYTGR.D + Acetyl (N-term)
<input checked="" type="checkbox"/> 329	577.2893	1152.5640	1152.6139	-43.23	1	34	0.29	1	K_VKTLSSYTGR.D + Acetyl (N-term)
<input checked="" type="checkbox"/> 330	577.2917	1152.5688	1152.6139	-39.07	1	(20)	7.3	1	K_VKTLSSYTGR.D + Acetyl (N-term)

4. [gi|9955540](#) Mass: 75537 Score: 26 Queries matched: 2 emPAI: 0.04
putative protein [Arabidopsis thaliana]
 Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
786	1094.0309	2186.0472	2186.0327	6.68	2	(0)	8.1e+002	8	K_YWQNNMNNVKANKTAHK.E + 2 Acetyl (K); Acetyl (N-term)
<input checked="" type="checkbox"/> 790	1095.0364	2188.0582	2188.0119	21.2	2	26	2.2	1	K_YWQNNMNNVKANKTAHK.E + 2 Acetyl (K); Formyl (N-term); Oxidati

Proteins matching the same set of peptides:
[gi|79512802](#) Mass: 72195 Score: 26 Queries matched: 2
FY; protein binding [Arabidopsis thaliana]

5. [gi|15219730](#) Mass: 39176 Score: 23 Queries matched: 4 emPAI: 0.08
CAK4 (CDK-ACTIVATING KINASE 4); kinase/ protein binding / protein serine/threonine kinase [Arabidopsis thaliana]
 Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
<input checked="" type="checkbox"/> 778	1082.0211	2162.0276	2162.0895	-28.62	1	(22)	5.5	1	R_YFSSSPSPTEPGKLPASK.G + Acetyl (K)
<input checked="" type="checkbox"/> 779	1082.0265	2162.0384	2162.0895	-23.62	1	23	4.1	1	R_YFSSSPSPTEPGKLPASK.G + Acetyl (K)
780	1082.0298	2162.0450	2162.0895	-20.57	1	(4)	3.4e+002	2	R_YFSSSPSPTEPGKLPASK.G + Acetyl (K)
793	1103.0390	2204.0634	2204.1001	-16.62	1	(2)	5.8e+002	3	R_YFSSSPSPTEPGKLPASK.G + Acetyl (K); Acetyl (N-term)

Figura 17. Péptidos obtenidos en el análisis por espectrometría de masas, a partir de extractos proteicos de plantas mutantes dSpm::AtLEA4-5 de tres semanas de edad, transformadas establemente con la proteína de fusión AtLEA4-5-C-TAPa que sufrieron un periodo de sequía y se entrecruzaron los tejidos con formaldehído antes de la extracción de los complejos proteicos. Se muestran los péptidos después de su identificación con el algoritmo MASCOT.

En el caso de los complejos provenientes de plantas crecidas bajo riego óptimo y a las que se les aplicó el tratamiento con formaldehído (B) se encontraron pocos péptidos con una puntuación significativa. A diferencia del experimento anterior, los péptidos que se encontraron corresponden a una proteína de función desconocida (Figura 17). Cuando se extrajeron los complejos proteicos sin el tratamiento previo de entrecruzamiento (tratamientos C y D) se obtuvieron algunos péptidos que fueron idénticos a los encontrados en las muestras provenientes del tratamiento de sequía tratadas con formaldehído; tal fue el caso de una cinasa putativa (Figura 18, Tratamiento C). Cuando se crecieron las plantas en condiciones de riego óptimo (Figura 19, Tratamiento D) se encontraron péptidos que concuerdan con los encontrados en el experimento anterior (que se consideraron como proteínas inespecíficas) y otros péptidos correspondientes a proteínas con localización celular en núcleo o citoplasma. Se identificó una proteína similar a las proteínas Se volvió a identificar la proteína cinasa que se obtuvo también en los tratamientos A y C, aunque en esta ocasión con una puntuación menor a 30. Se detectó otra cinasa que no se había obtenido en los análisis anteriores, que corresponde a una cinasa de serinas y treoninas. También se detectó un péptido que corresponde a una histona y una proteína de respuesta temprana a la deshidratación (ERD10, por “Early Responsive to Dehydration”).

De esta manera se pudieron comparar los candidatos de interacción provenientes de plantas crecidas en condiciones óptimas con plantas crecidas bajo sequía y

también se compararon con los resultados del sistema de doble híbrido en levadura, con el fin de tener información que nos permitiera determinar si los candidatos obtenidos corresponden a blancos de interacción específicos de la proteína AtLEA4-5 y, si éstos son diferentes dependiendo de las condiciones osmóticas de la célula.

B

1. [gi|1592679](#) Mass: 16164 Score: 90 Queries matched: 1 emPAI: 0.21
LEA D113 homologue type1 [Arabidopsis thaliana]

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
563	855.4076	1708.8006	1708.8123	-6.85	1	90	8.3e-007	1	-_MQSMKETASNIAASAK.S

Proteins matching the same set of peptides:
[gi|15240144](#) Mass: 16169 Score: 90 Queries matched: 1
late embryogenesis abundant group 1 domain-containing protein / LEA group 1 domain-containing protein [Arabidopsis thaliana]

3. [gi|15240611](#) Mass: 17429 Score: 17 Queries matched: 3 emPAI: 0.19
unknown protein [Arabidopsis thaliana]

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
283	526.3024	1050.5902	1050.5202	66.7	2	17	13	1	K_GLMAKKCSK.L 281 282

Proteins matching the same set of peptides:
[gi|42573646](#) Mass: 16673 Score: 17 Queries matched: 3
unknown protein [Arabidopsis thaliana]

Figura 18. Péptidos obtenidos en el análisis por espectrometría de masas, a partir de extractos proteicos de plantas mutantes dSpm::AtLEA4-5 de tres semanas de edad, transformadas establemente con la proteína de fusión AtLEA4-5-C-TAPa que se mantuvieron en condiciones de riego óptimo y se entrecruzaron los tejidos con formaldehído antes de la extracción de los complejos proteicos. Se muestran los péptidos después de su identificación con el algoritmo MASCOT.

C

3. [gi|20197434](#) Mass: 40207 Score: 50 Queries matched: 3 emPAI: 0.08
putative calmodulin-binding protein [Arabidopsis thaliana]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
<input checked="" type="checkbox"/> 298	679.3258	1356.6370	1356.7249	-64.73	1	(18)	15	1	R_VVSEEVERALAK.L + Formyl (N-term)
<input checked="" type="checkbox"/> 299	679.3258	1356.6370	1356.7249	-64.73	1	50	0.0085	1	R_VVSEEVERALAK.L + Formyl (N-term)
<input checked="" type="checkbox"/> 301	679.3273	1356.6400	1356.7249	-62.52	1	(13)	47	1	R_VVSEEVERALAK.L + Formyl (N-term)

Proteins matching the same set of peptides:
[gi|30687005](#) Mass: 67549 Score: 50 Queries matched: 3
calmodulin-binding protein [Arabidopsis thaliana]
[gi|42569137](#) Mass: 70156 Score: 50 Queries matched: 3
calmodulin-binding protein [Arabidopsis thaliana]

5. [gi|1592679](#) Mass: 16164 Score: 40 Queries matched: 1 emPAI: 0.21
LEA D113 homologue type1 [Arabidopsis thaliana]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
<input checked="" type="checkbox"/> 346	657.3642	1312.7138	1312.6987	11.5	0	40	0.091	1	R.VVEGTAVTDPIGR.N

Proteins matching the same set of peptides:
[gi|15240144](#) Mass: 16169 Score: 40 Queries matched: 1
late embryogenesis abundant group 1 domain-containing protein / LEA group 1 domain-containing

8. [gi|15219730](#) Mass: 39176 Score: 24 Queries matched: 5 emPAI: 0.08
 CAK4 (CDK-ACTIVATING KINASE 4); kinase/ protein binding / protein serine/threonine kinase [Arabidopsis thaliana]
 Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
<input checked="" type="checkbox"/> 678	1082.0353	2162.0560	2162.0895	-15.48	1	(17)	19	1	R.YFSSSPSPTEPGKIQIPASK.G + Acetyl (K)
<input checked="" type="checkbox"/> 679	721.6929	2162.0569	2162.0895	-15.10	1	(14)	32	1	R.YFSSSPSPTEPGKIQIPASK.G + Acetyl (K)
<input checked="" type="checkbox"/> 680	1082.0369	2162.0592	2162.0895	-14.00	1	(16)	20	1	R.YFSSSPSPTEPGKIQIPASK.G + Acetyl (K)
<input checked="" type="checkbox"/> 681	1082.0425	2162.0704	2162.0895	-8.82	1	24	3.2	1	R.YFSSSPSPTEPGKIQIPASK.G + Acetyl (K)
682	1082.0449	2162.0752	2162.0895	-6.60	1	(4)	3.3e+002	2	R.YFSSSPSPTEPGKIQIPASK.G + Acetyl (K)

Figura 19. Péptidos obtenidos en el análisis por espectrometría de masas, a partir de extractos proteicos de plantas mutantes dSpm::AtLEA4-5 de tres semanas de edad, transformadas establemente con la proteína de fusión AtLEA4-5-C-TAPa que sufrieron un periodo de sequía y no se les aplicó entrecruzamiento con formaldehído. Se muestran los péptidos después de su identificación con el algoritmo MASCOT.

Un estudio reciente utilizando como modelo una chaperona de bajo peso molecular de chícharo (PsHSP 18.1) indicó que esta chaperona puede presentar alrededor de 300 estequiometrías diferentes de interacción, lo cual permite que pueda interactuar con varios sustratos, y es imposible detectar estas variaciones en la estructura de los complejos chaperona-cliente utilizando los enfoques estructurales tradicionales (Stengel *et al.*, 2010).

En el caso de las proteínas LEA el estudio de los blancos de interacción se dificulta por su naturaleza intrínsecamente desordenada, lo cual supone que la naturaleza de las macromoléculas con las que puede interactuar no dependan de su estructura primaria, sino de la conformación espacial particular que pueden adoptar, tanto la proteína LEA, como el blanco de interacción, en una condición determinada en la célula. Como se mencionó anteriormente, los resultados descritos en este segundo capítulo son preliminares, y debido a las dificultades técnicas que se tuvieron tanto en el análisis de dos híbridos, como en la purificación de complejos proteicos por TAP, es necesario llevar a cabo más experimentos para corroborar los datos obtenidos. Es fundamental conocer la localización intracelular de la proteína AtLEA4-5 para determinar si los candidatos son realmente proteínas que se encuentran en el mismo contexto celular o, como se ha visto para otras proteínas LEA, que las condiciones osmóticas intracelulares puedan causar una relocalización de la proteína y de esta forma, una vez que ha adoptado la estructura secundaria necesaria para interactuar con su blanco, puedan otorgar protección en esa condición particular de la célula. Cabe destacar que se están iniciando experimentos en colaboración con el Dr Fernando Lledías, quien tiene experiencia en la purificación de complejos proteicos, utilizando una técnica conocida como geles azules nativos (Eubel, *et al.*, 2005), para lo cual, se utilizarán anticuerpos inmunopurificados de la AtLEA4-5 y extractos vegetales de plantas silvestres, en diferentes órganos y etapas de desarrollo de la planta, tanto en condiciones óptimas de irrigación como en tratamientos de sequía. Este método se basa en la purificación en dos dimensiones de los componentes presentes en

D

2.	gi 3269293	Mass: 27902	Score: 58	Queries matched: 1	emPAI: 0.12						
	putative protein [Arabidopsis thaliana]										
	<input type="checkbox"/> Check to include this hit in error tolerant search or archive report										
	Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide	
	✓ 403	849.8894	1697.7642	1697.7719	-4.50	0	58	0.0011	1	-_MNPDSYPHLPNIK.I + Acetyl (N-term); Oxidation (M)	
	Proteins matching the same set of peptides:										
	gi 18416966	Mass: 28089	Score: 58	Queries matched: 1							
	universal stress protein (USP) family protein [Arabidopsis thaliana]										
	gi 51968678	Mass: 28161	Score: 58	Queries matched: 1							
	unknown protein [Arabidopsis thaliana]										
5.	gi 15235366	Mass: 110214	Score: 31	Queries matched: 5	emPAI: 0.03						
	BAM3 (BARELY ANY MERISTEM 3); ATP binding / protein binding / protein kinase/ protein serine/threonine kinase [Arabidopsi:										
	<input type="checkbox"/> Check to include this hit in error tolerant search or archive report										
	Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide	
	✓ 398	566.2876	1695.8410	1695.9056	-38.13	2	(6)	2.2e+002	1	K_ENNHVIGKGGRGIVYK.G + Acetyl (K); Formyl (N-term)	
	✓ 399	566.2916	1695.8530	1695.9056	-31.06	2	(17)		17	K_ENNHVIGKGGRGIVYK.G + Acetyl (K); Formyl (N-term)	
	✓ 400	566.2926	1695.8560	1695.9056	-29.29	2	(17)		19	K_ENNHVIGKGGRGIVYK.G + Acetyl (K); Formyl (N-term)	
	✓ 401	566.2957	1695.8653	1695.9056	-23.80	2	(19)		12	K_ENNHVIGKGGRGIVYK.G + Acetyl (K); Formyl (N-term)	
	✓ 402	566.2999	1695.8779	1695.9056	-16.37	2	31		0.73	1	K_ENNHVIGKGGRGIVYK.G + Acetyl (K); Formyl (N-term)
7.	gi 227206250	Mass: 21749	Score: 28	Queries matched: 1	emPAI: 0.15						
	AT2G30620 [Arabidopsis thaliana]										
	<input type="checkbox"/> Check to include this hit in error tolerant search or archive report										
	Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide	
	✓ 149	421.2638	840.5130	840.5069	7.32	0	28	0.8	1	K_ATVVAKPK.G + Formyl (N-term)	
	Proteins matching the same set of peptides:										
	gi 15224536	Mass: 28470	Score: 28	Queries matched: 1							
	histone H1.2 [Arabidopsis thaliana]										
8.	gi 15233100	Mass: 86631	Score: 27	Queries matched: 3	emPAI: 0.04						
	early-responsive to dehydration protein-related / ERD protein-related [Arabidopsis thaliana]										
	<input type="checkbox"/> Check to include this hit in error tolerant search or archive report										
	Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide	
	✓ 374	827.8701	1653.7256	1653.8131	-52.87	2	(6)	1.4e+002	1	R.KISEEKETVMSSTK.S + Acetyl (K); Oxidation (M)	
	✓ 375	827.8723	1653.7300	1653.8131	-50.21	2	(22)		4.4	1	R.KISEEKETVMSSTK.S + Acetyl (K); Oxidation (M)
	✓ 376	827.8739	1653.7332	1653.8131	-48.27	2	27		1.4	1	R.KISEEKETVMSSTK.S + Acetyl (K); Oxidation (M)
	Proteins matching the same set of peptides:										
	gi 62319233	Mass: 86659	Score: 27	Queries matched: 3							
	hypothetical protein [Arabidopsis thaliana]										
9.	gi 15219730	Mass: 39176	Score: 24	Queries matched: 4	emPAI: 0.08						
	CAK4 (CDK-ACTIVATING KINASE 4); kinase/ protein binding / protein serine/threonine kinase [Arabidopsis thaliana]										
	<input type="checkbox"/> Check to include this hit in error tolerant search or archive report										
	Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide	
	✓ 576	1082.0259	2162.0372	2162.0895	-24.18	1	(15)		27	1	R.YFSSSPSPTEPGKLQIPASK.G + Acetyl (K)
	✓ 577	1082.0266	2162.0386	2162.0895	-23.53	1	(15)		25	1	R.YFSSSPSPTEPGKLQIPASK.G + Acetyl (K)
	✓ 578	1082.0273	2162.0400	2162.0895	-22.88	1	24		3.7	1	R.YFSSSPSPTEPGKLQIPASK.G + Acetyl (K)
	✓ 579	1082.0326	2162.0506	2162.0895	-17.98	1	(10)		96	1	R.YFSSSPSPTEPGKLQIPASK.G + Acetyl (K)

Figura 20. Péptidos obtenidos en el análisis por espectrometría de masas, a partir de extractos proteicos de plantas mutantes dSpm::AtLEA4-5 de tres semanas de edad, transformadas establemente con la proteína de fusión AtLEA4-5-C-TAPa que se mantuvieron en condiciones de riego óptimo y no se les aplicó entrecruzamiento con formaldehído. Se muestran los péptidos después de su identificación con el algoritmo MASCOT.

complejos proteicos en condiciones nativas, lo cual implica una ventaja para el caso de la proteínas desordenadas como la AtLEA4-5, ya que no se estaría alterando la

proteína al adicionarle alguna etiqueta, como es el caso de la purificación con TAP u otras etiquetas. Estos experimentos representarían una herramienta adicional muy importante en el estudio de las proteínas que pueden interactuar y, de acuerdo a nuestra hipótesis, ser protegidas por la AtLEA4-5.

Debido a que cada uno de los métodos para buscar interacciones proteína-proteína tiene ventajas y desventajas, es necesario comparar los resultados obtenidos con cada uno de los métodos y corroborar las interacciones *in vivo* en un sistema vegetal. Los resultados del análisis de dos híbridos indicó que las clonas positivas muestran interacciones débiles, ya que ninguno de los candidatos dio una intensidad del color azul, semejante a la que se obtiene con un control de interacción fuerte (Cuevas-Velázquez, C., datos no publicados) y la mayoría de las clonas positivas corresponden a péptidos ricos en lisina y otros aminoácidos cargados, que son capaces de activar la transcripción del gen reportero, lo cual sugiere que algunos motivos de interacción con la proteína anzuelo (AtLEA4-5) pueden ser péptidos pequeños que sean propensos a formar α -hélices y que al encontrarse la célula en las condiciones osmóticas que promueven la formación de estas estructuras, la proteína AtLEA4-5 se conforme y pueda haber una interacción con sus blancos.

Por otra parte, los resultados del análisis por TAP arrojó péptidos correspondientes a proteínas con diversas localizaciones intracelulares y solamente uno de los péptidos, que corresponden a una posible cinasa de proteínas, se detectó en tres de los cuatro tratamientos explorados en el segundo experimento. Es necesario comprobar la funcionalidad de esta proteína de fusión en los experimentos de sequía y rehidratación, como ya se había mencionado anteriormente. Se deben llevar a cabo más experimentos de purificación por TAP para ver si algunos de los candidatos que se han encontrado en los experimentos anteriores, se vuelven a identificar en la espectrometría de masas. Esto permitiría que se puedan hacer análisis más detallados en cuanto a la conformación de ambas proteínas en una situación de estrés, una vez que se conozca la localización intracelular de las proteínas de interés.

ANEXO 1. ARTICULO DE INVESTIGACIÓN PUBLICADO DEL TRABAJO DEL DOCTORADO.

ANEXO 2. ARTICULO DE REVISIÓN PUBLICADO PARA CUMPLIR CON EL REQUISITO DE ACTIVIDAD DOCENTE EN EL PROGRAMA DE DOCTORADO.

LITERATURA CITADA

- AAAS: Atlas of Population and environment (2001) American Association for the advancement of science. Lyons J, Hawkins C (eds) The University of California Press.
- Abe H, Yamaguchi-Shinozaki K, Urao T, Iwasaki T, Hosokawa D, Shinozaki K (1997) Role of *Arabidopsis* MYC and MYB homologues in drought- and abscisic acid- regulated gene expression. *Plant Cell* **9**: 1859-1868.
- Abe H, Urao T, Ito T, Seki M, Shinozaki Y, Yamaguchi-Shinozaki K (2003) *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell* **15**: 63-78.
- Alpert P (2006) Constraints of tolerance: why are desiccation-tolerant organisms so small or rare? *J Exp Biol* **209**:1575-1584.
- Alpert P, Oliver M (2002) Desiccation and Survival in Plants, pp. 3-43, En: Drying Without Dying. M Black, H Pritchard (eds) CAB Intl Publishing, London.
- Allan AC, Trewavas AJ (1994) Abscisic acid and gibberellin perception: inside or out? *Plant Phys* **104**: 1107-1108.
- Amann E, Ochs B, Abel KJ (1988) Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* **69**: 301-315.
- Apse MP, Aharon GS, Snedden WA, Blumwald E (1999) Salt tolerance conferred by overexpression of a vacuolar Na⁺/H⁺ antiport in *Arabidopsis*. *Science* **285**: 1256-1258.
- Arenas-Huertero C, Pérez B, Rabanal F, Blanco-Melo D, De la Rosa C, Estrada-Navarrete G, Sánchez F, Covarrubias AA, Reyes JL (2009) Conserved and novel miRNAs in the legume *Phaseolus vulgaris* in response to stress. *Plant Mol Biol* **70**: 385-401.
- Artus NN, Uemura M, Steponkus PL, Gilmour SJ, Lin C, Thomashow MF (1996) Constitutive expression of the cold-regulated *Arabidopsis thaliana* COR15a gene affects both chloroplast and protoplast freezing tolerance. *Proc Natl Acad Sci USA* **93**: 13404-13409.
- Babu RC, Zhang J, Ho D, Wu R, Nguyen HT (2004) *HVA1*, a LEA gene from barley confers dehydration tolerance in transgenic rice (*Oryza sativa* L.) via cell membrane protection. *Plant Sci* **166**: 855-862.
- Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, pp. 28-36, AAAI Press, Menlo Park, California.
- Bailey TL, Gribskov M (1998) Combining evidence using P-values: application to sequence homology searches. *Bioinformatics* **14**: 48-54
- Bailey-Serres J (1999) Selective translation of cytoplasmic mRNAs in plants. *Trends Plant Sci.* **4**:142-148.
- Bajaj S, Targolli J, Liu L, Ho D, Wu R (1999) Transgenic approaches to increase dehydration-stress tolerance in plants. *Mol Breed* **5**: 493-503.
- Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**: 281-297.
- Battaglia M, Olvera-Carrillo Y, Garcarrubio A, Campos F, Covarrubias A (2008) The enigmatic LEA proteins and other hydrophilins. *Plant Physiol* **148**: 6-24.
- Battista JR, Mie-Jung P, McLemore A (2001) Inactivation of two homologues of proteins presumed to be involved in the desiccation tolerance of plants sensitizes *Deinococcus radiodurans* R1 to desiccation. *Cryobiol* **43**: 133-139.
- Bechtold, N.; Ellis, J. & Pelletier, G. (1993) *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *CR Academy of Sciences Paris Sci. Vie/Life Sci.* **316**: 1194-1199.
- Bell S, Klein C, Muller L, Hansen S, Buchner J (2002) p53 contains large unstructured regions in its native state. *J Mol Biol* **322**: 917-927.
- Begon M, Townsend CR, Harper JL (1996) *Ecology: from individuals to ecosystems*. Wiley-Blackwell Publishing, London.
- Berjak P, Pammenter N W (2008) From *Avicennia* to *Zizania*: Seed recalcitrance in perspective. *Annals Bot* **101**: 213-228.
- Bevan M (1984) Binary *Agrobacterium* vectors for plant transformation. *Nucleic Acids Res* **12**: 8711-8721.
- Bewley JD, Black M (1994) *Seeds. Physiology of development and germination*. Plenum Press, New York.
- Bewley JD (1997) Seed germination and dormancy. *Plant Cell* **9**:1055-1066.
- Bewley JD, Reynolds TL, Oliver MJ (1993) Evolving strategies in the adaptation to desiccation. En: *Plant Responses to Cellular Dehydration During Environmental Stress*, TJ Close, EA Bray (eds.), pp. 193-201. American Society of Plant Physiologists Series, USA.
- Bies N, Aspart L, Carles C, Gallois P, Delseny M (1998) Accumulation and degradation of Em proteins in *Arabidopsis thaliana*, evidence for post-transcriptional controls. *J Exp Bot* **49**: 1925-1933.
- Bies-Ethève N, Gaubier-Comella P, Debures A, Lasserre E, Jobet E, Raynal M, Cooke R, Delseny M (2008) Inventory, evolution and expression profiling diversity of the LEA (late embryogenesis abundant) protein gene family in *Arabidopsis thaliana*. *Plant Mol Biol* **67**:107-124.
- Billi D, Potts M (2002) Life and death of dried prokaryotes. *Res Microbiol* **153**:7-12.
- Blackman SA, Obendorf RL, Leopold AC (1995) Desiccation tolerance in developing soybean seeds: The role of stress proteins. *Physiol Plantarum* **93**: 630-638.
- Bordas M, Montesinos C, Debauza M, Salvador A, Roig LA, Serrano R, Moreno V (1997) Transfer of the yeast salt tolerance gene HAL1 to *Cucumis melo* L. cultivars and in vitro evaluation of salt tolerance. *Transgenic Res* **6**: 41-50.

- Bowman J (1994) *Arabidopsis: an atlas of morphology and development*. Springer-Verlag, New York.
- Boyer JS (1982) Plant productivity and environment. *Science* **218**: 443-448.
- Bravo LA, Gallardo J, Navarrete A, Olave N, Martinez J, Alberdi M, Close TJ, Corcuera LJ (2003) Cryoprotective activity of a cold-induced dehydrin purified from barley. *Physiol Plant* **118**: 262-269.
- Bray EA (1997) Plant responses to water deficit. *Trends Plant Sci* **2**: 48-54.
- Brodersen P, Voinnet O (2009) Revisiting the principles of microRNA target recognition and mode of action. *Nat Rev Mol Cell Biol* **10**:141-148.
- Browne J, Tunnacliffe A, Burnell A (2002) Plant desiccation gene found in a nematode. *Nature* **416**: 38.
- Browne JA, Dolan KM, Tyson T, Goyal K, Tunnacliffe A, Burnell AM (2004) Dehydration-specific induction of hydrophilic protein genes in the anhydrobiotic nematode *Aphelenchus avenae*. *Eukaryot Cell* **3**: 966-975.
- Buchanan B, Grissem W, Jones R (eds), *Biochemistry and Molecular Biology of Plants*, pp. 1158-1203, American Society of Plant Physiologists, Rockville, MD
- Campalans A, Messegueur R, Goday A, Pagès M (1999) Plant responses to drought, from ABA signal transduction events to the action of the induced proteins. *Plant Physiol Biochem* **37**: 327-340.
- Carpita N, Sabulase D, Montezinos D, Delmer DP (1979) Determination of the pore size of cell walls of living plant cells. *Science* **205**: 1144-1147.
- Carrari F, Fernie AR, Iusem ND (2004) Heard it through the grapevine? ABA and sugar cross-talk: the ASR story. *Trends Plant Sci* **9**: 57-59.
- Chan MT, Yu SM (1998) The 3' untranslated region of a rice alpha-amylase gene mediates sugar-dependent abundance of mRNA. *Plant J* **15**: 685-695.
- Chan MT, Yu SM (1998) The 3' untranslated region of a rice alpha-amylase gene functions as a sugar-dependent mRNA stability determinant. *Proc Natl Acad Sci USA* **95**: 6543-6547.
- Chandra Babu R, Zhang JS, Blum A, Ho T, Wu R, Nguyen HT (2004) *HVA1*, a LEA gene from barley confers dehydration tolerance in transgenic rice (*Oryza sativa* L.) via cell membrane protection. *Plant Sci* **166**: 855-862.
- Chazen O, Neumann PM (1994) Hydraulic signals from the roots and rapid cell wall hardening in growing maize (*Zea mays* L.) leaves are primary responses to polyethylene glycol-induced water deficit. *Plant Physiol* **104**: 1385-1392.
- Chen RD, Yu LX, Greer AF, Cheriti H, Tabaeizadeh Z (1994) Isolation of an osmotic stress- and abscisic acid-induced gene encoding an acidic endochitinase from *Lycopersicon chilense*. *Mol Gen Genet* **245**:195-202.
- Chen CN, Chiang YC, Ho D, Schaal BA, Chiang TY (2004) Coalescent processes and relaxation of selective constraints leading to contrasting diversity at paralogs *AtHVA22d* and *AtHVA22e* in *Arabidopsis thaliana*. *Mol Phylogenet Evol* **32**: 616-626.
- Chen TH, Murata N. (2008) Glycinebetaine: an effective protectant against abiotic stress in plants. *Trends Plant Sci.***13**: 499-505.
- Cheng Z, Targolli J, Huang X, Wu R (2002) Wheat LEA genes, PMA80 and PMA1959, enhance dehydration tolerance of transgenic rice (*Oryza sativa* L.). *Mol Breed* **10**: 71-82.
- Chinnusamy V, Ohta M, Kanrar S, Lee BH, Hong X, Agarwal M, Zhu JK (2003) ICE1: a regulator of cold-induced transcriptome and freezing tolerance in Arabidopsis. *Genes Dev* **17**: 1043-1054.
- Christmann A, Hoffmann T, Teplova I, Grill E, Müller A (2005) Generation of active pools of abscisic acid revealed by in vivo imaging of water-stressed Arabidopsis. *Plant Physiol* **137**: 209-219.
- Christmann A, Weiler EW, Steudle E, Grill E (2007) A hydraulic signal in root-to-shoot signalling of water shortage. *Plant J* **52**:167-174.
- Choi HI, Hong JH, Ha J, Kang JY, Kim SY (2000) ABFs, a family of ABA-responsive element binding factors. *J Biol Chem* **275**: 1723-1730.
- Chou J, Fasman GD (1978) Prediction of the secondary structure of proteins from their amino acid sequence. *Advances Enzymol* **47**: 45-148.
- Chuang C, Meyerowitz E (2000) Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* **97**: 4985-4990.
- Close TJ (1996) Dehydrins: Emergence of a biochemical role of a family of plant dehydration proteins. *Physiol Plantarum* **97**: 795-803.
- Close TJ, Fenton RD, Yang A, Asghar R, DeMason DA, Crone DE, Meyer NC, Moonan F (1993) Dehydrin: The protein. In TJ Close, EA Bray (eds), *Plant Responses to Cellular Dehydration During Environmental Stress*, pp. 104-118. American Society of Plant Physiologists Series, USA.
- Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium* mediated transformation of *Arabidopsis thaliana*. *Plant J* **16**: 735-743.
- Colmenero-Flores JM, Campos F, Garcarrubio A, Covarrubias AA (1997) Characterization of *Phaseolus vulgaris* cDNA clones responsive to water deficit: identification of a novel late embryogenesis abundant-like protein. *Plant Mol Biol* **35**: 393-405.
- Colmenero-Flores JM, Moreno LP, Smith C, Covarrubias AA (1999) *Pvlea-18*, a member of a new late-embryogenesis abundant protein family that accumulates during water stress and in the growing regions of well irrigated bean seedlings. *Plant Physiol* **120**: 93-103.
- Covarrubias AA, Ayala JW, Reyes JL, Hernández M, Garcarrubio A (1995) Cell wall proteins induced by water deficit in bean (*Phaseolus vulgaris* L.) seedlings. *Plant Physiol* **107**: 1119-1128.

- Covarrubias AA, Reyes JL. (2010) Post-transcriptional gene regulation of salinity and drought responses by plant microRNAs. *Plant Cell Environ* **33**: 481-489.
- Crowe JH, Crowe LM, Chapman D (1984) Preservation of membrane in anhydrobiotic organisms: The role of trehalose. *Science* **223**: 701-703.
- Cuming AC (1999) LEA proteins. En PR Shewry, R Casey (eds), *Seed Proteins*, pp. 753-780. Kluwer Academic Publishers, Dordrecht.
- Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR (2010) Abscisic acid: emergence of a core signalling network. *Annu Rev Plant Biol.* **61**: 651-679.
- Dalal M, Tayal D, Chinnusamy V, Bansal KC (2009) Abiotic stress and ABA-inducible Group 4 LEA from *Brassica napus* plays a key role in salt and drought tolerance. *J Biotech* **139**: 137-145
- Dalmay T, Hamilton A, Rudd S, Angell S Baulcombe D (2000) An RNA-dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell* **101**: 543-553.
- Davison PA, Hunter CN, Horton P (2002) Overexpression of β -carotene hydroxylase enhances stress tolerance in *Arabidopsis*. *Nature* **418**: 203-206.
- Delseny M, Biès-Etheve N, Carles C, Hull G, Vicient C, Raynal M, Grellet F, Aspart L (2001) Late embryogenesis abundant (LEA) protein regulation during *Arabidopsis* seed maturation. *J Plant Physiol* **158**: 419-427.
- Dickinson, W (2000) Integrative plant anatomy. *Ecological Anatomy*, pp 295-337, Academic Press, San Diego.
- Dubouzet JG, Sakuman Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold responsive gene expression. *Plant J* **33**: 751-763.
- Dure L III (1993) Structural motifs in LEA proteins. En: Close T, Bray E (eds), *Plant responses to cellular dehydration during environmental stress*. The American Society of Plant Physiologists, Rockville, MD. Pp. 91-103.
- Dure L, Crouch M, Harada JJ, Ho T, Mundy J, Quatrano RS, Thomas TL, Sung ZR (1989) Common amino acid sequence domains among the LEA proteins of higher plants. *Plant Mol Biol* **12**: 475-486
- Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* **6**:197-208.
- Edwards K, Johstone C, Thompson C (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nuc Acid Res* **19**:1349
- Endo A, Sawada Y, Takahashi H, Okamoto M, Ikegami K, Koiwai H, Seo M, Toyomasu T, Mitsuhashi W, Shinozaki K, Nakazono M, Kamiya Y, Koshiha T, Nambara E. (2008) Drought induction of Arabidopsis 9-cis-epoxycarotenoid dioxygenase occurs in vascular parenchyma cells. *Plant Physiol* **147**: 1984-1993.
- Escobar M, Civerolo E, Summerfelt K, Dandekar A (2001) RNAi-mediated oncogene silencing confers resistance to crown gall tumorigenesis. *Proc Natl Acad Sci USA* **98**: 13437-13442.
- Eom J, Baker W, Kintanar A, Wurtele ES (1996) The embryo-specific EMB-1 protein of *Daucus carota* is flexible and unstructured. *Plant Sci* **115**: 17-24
- Eubel H, Braun HP, Millar AH (2005) Blue-native PAGE in plants: a tool in analysis of protein-protein interactions *Plant Methods* **1**:1-11.
- Farrant JM, Pammenter NW, Berjak P, Farnsworth EJ, Vertucci CW (1996) Presence of dehydrin-like proteins and levels of abscisic acid in recalcitrant (desiccation sensitive) seeds may be related to habitat. *Seed Sci Res* **6**: 175-182.
- Felsenstein J (2005) PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.
- Filipowicz W, Bhattacharyya SN, Sonenberg N (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* **9**:102-114.
- Finch-Savage WE, Blake PS (1994) Indeterminate development in desiccation-sensitive seeds of *Quercus robur* L. *Seed Sci Res* **4**: 127-133.
- Finkelstein RR, Lynch TJ (2000) The *Arabidopsis* abscisic acid response gene *ABI5* encodes a basic leucine zipper transcription factor. *Plant Cell* **12**: 599-609.
- Finkelstein RR, Gampala SS, Rock CD (2002) Abscisic acid signaling in seeds and seedlings. *Plant Cell Supplement* **14**: S15-S45.
- Finkelstein R, Reeves W, Ariizumi T, Steber C (2008) Molecular aspects of seed dormancy. *Annu Rev Plant Biol* **59**: 387-415.
- Fitter A, Hay R (2002) *Environmental physiology of plants*. Academic Press, San Diego.
- Fire A, Xu S, Montgomery M, Kostas S, Driver S, Mello C (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**: 806-811.
- Fu D, Huang B, Xiao Y, Muthukrishnan S, Liang GH (2007) Overexpression of barley *HVA1* gene in creeping bentgrass for improving drought tolerance. *Plant Cell Rep* **26**: 467-477.
- Galau GA, Wang HYC, Hughes DW (1993) Cotton *Lea5* and *Lea14* encode atypical late embryogenesis-abundant proteins. *Plant Physiol* **101**: 695-696.
- Garay-Arroyo A, Colmenero-Flores JM, Garciarubio A, Covarrubias AA (2000) Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. *J Biol Chem* **275**: 5668-5674.

- García-Gómez BI, Campos F, Hernández M, Covarrubias AA (2000) Two bean cell wall proteins more abundant during water deficit are high in proline and interact with plasma membrane protein. *Plant J* **22**: 277-288.
- Garg AK, Kim JK, Owens TG, Ranwala AP, Choi YD, Kochian LV, Wu RJ (2002) Trehalose accumulation in rice plants confers high tolerance levels to different abiotic stresses. *Proc Natl Acad Sci USA* **99**: 15898–15903.
- Gaxiola RA, Rao R, Sherman A, Grisafi P, Alper SL, Fink GR (1999) The *Arabidopsis thaliana* proton transporters, AtNhx1 and Avp1, can function in cation detoxification in yeast. *Proc Natl Acad Sci USA* **96**: 1480-1485.
- Gilmour SJ, Zarka DG, Stockinger EJ, Salazar MP, Houghton JM, Thomashow MF (1998) Low temperature regulation of the Arabidopsis CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression. *Plant J* **16**: 433-442.
- Gilmour SJ, Sebolt AM, Salazar MP, Everard JD, Thomashow MF (2000) Overexpression of the *Arabidopsis* CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. *Plant Physiol* **124**:1854-1865.
- Gleave AP (1992) A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Mol Biol* **20**: 1203-1207.
- Goday A, Jensen A, Culiáñez-Macià FA, Albà MM, Figueras M, Serratosa J, Torrent M, Pagés M (1994) The maize abscisic acid responsive protein RAb17 is located in the nucleus and interacts with nuclear localization signals. *Plant Cell* **6**: 351-360.
- Goyal K, Tisi L, Basran A, Browne J, Burnell A, Zurdo J, Tunnacliffe A (2003) Transition from natively unfolded to folded state induced by desiccation in an anhydrobiotic nematode protein. *J Biol Chem* **278**: 12977-12984.
- Goyal K, Walton LJ, Tunnacliffe A (2005) LEA proteins prevent protein aggregation due to water stress. *Biochem J* **388**: 151-157.
- Graber JH, Cantor CR, Mohr SC, Smith TF (1999) In silico detection of control signals: mRNA 3'-end-processing sequences in diverse species. *Proc Natl Acad Sci USA* **96**:14055-14060.
- Graham LE, Cook ME, Busse JS (2000) The origin of plants: body plan changes contributing to a major evolutionary radiation. *Proc Natl Acad Sci USA* **97**: 4535-4540.
- Grelet J, Benamar A, Teyssier, E, Avelange-Macherel MH, Grunwald D, Macherel D (2005) Identification in pea seed mitochondria of a late-embryogenesis abundant protein able to protect enzymes from drying. *Plant Physiol* **137**: 157-167
- Grime JP (1994) The role of plasticity in exploiting environmental resources. En: Cladwell MM, Pearcy RW (eds), *Exploitation of environmental Heterogeneity by plants: ecophysiological process above and belowground*. Academia Press, San Diego.
- Gupta AS, Heinen JL, Holaday AS, Burke JJ, Allen RD (1993) Increased resistance to oxidative stress in transgenic plants that overexpress chloroplastic Cu/Zn superoxide dismutase. *Proc Natl Acad Sci USA* **90**: 1629–1633.
- Haake V, Cook D, Riechmann JL, Pineda O, Thomashow MF, Zhang JZ (2002) Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis. *Plant Physiol* **130**: 639-648.
- Hall A (1993) Is dehydration tolerance relevant to genotypic differences in leaf senescence and crop adaptation to dry environments? En: *Plant responses to cellular dehydration during environmental stress*. Close T, Bray E (eds) The American Society of Plant Physiologists. Rockville, MD. Pp. 1-10.
- Hamilton A, Baulcombe D (1999) A novel species of small antisense RNA in post-transcriptional gene silencing. *Science* **286**: 950-952.
- Hamilton A, Voinnet O, Chappell L, Baulcombe D (2002) Two classes of short interfering RNA in RNA silencing. *EMBO J* **21**: 4671-4679.
- Hand SC, Jones D, Menze MA, Witt TL. (2007) Life without water: expression of plant LEA genes by an anhydrobiotic arthropod. *J Exp Zool A Ecol Genet Physiol* **307**:62-6.
- Hara M, Terashima S, Kuboi T (2001) Characterization and cryoprotective activity of cold responsive dehydrin from *Citrus unshiu*. *J Plant Physiol* **58**: 1333-1339.
- Hayashi HA, Mustardy L, Deshniem P, Ida M, Murata N (1997) Transformation of *Arabidopsis thaliana* with the *codA* gene for choline oxidase: accumulation of glycinebetaine and enhanced tolerance to salt and cold stress. *Plant J* **12**: 133-142.
- Hayashi, H.A.; Chen, T.H.H. & Murata, N. (1998). Transformation with a gene for choline oxidase enhances the cold tolerance of *Arabidopsis* during germination and early growth. *Plant Cell Environ* **21**: 232-239.
- Hegyí H, Tompa P. (2008) Intrinsically disordered proteins display no preference for chaperone binding in vivo. *PLoS Comput Biol* **4**(3):e1000017.
- Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* **89**: 10915-10919.
- Heyen BJ, Alsheikh MK, Smith EA, Torvik CF, Seals DF, Randall SK (2002) The calcium-binding activity of a vacuole-associated, dehydrin-like protein is regulated by phosphorylation. *Plant Physiol* **130**: 675-687.
- Hoekstra F, Golovina E, Buitink J (2001) Mechanisms of plant desiccation tolerance. *Trends Plant Sci* **6**: 431-438.

- Hoekstra F (2002) Pollen and spores: desiccation tolerance in pollen and the spores of lower plants and fungi. En: Drying without dying. Black, M, Pritchard, H (eds) CABI Publishing, London.
- Hong Z, Lakkineni K, Zhang K, Verma DPS (2000) Removal of feedback inhibition of Δ^1 -pyrroline-5-carboxylate synthetase results in increased proline accumulation and protection of plants from osmotic stress. *Plant Physiol* **122**: 1129-1136.
- Honjoh K, Yoshimoto M, Joh T, Kajiwara T, Miyamoto T, Hatano S (1995) Isolation and characterization of hardening-induced proteins in *Chlorella vulgaris* C-27: identification of late embryogenesis abundant proteins. *Plant Cell Physiol* **36**: 1421-1430.
- Honjoh K, Matsumoto H, Shimizu H, Ooyama K, Tanaka K, Oda Y, Takata R, Joh T, Suga K, Miyamoto T, Lio M, Hatano S (2000) Cryoprotective activities of group 3 late embryogenesis abundant proteins from *Chlorella vulgaris* C-27. *Biosci Biotechnol Biochem* **64**: 1656-1663.
- Horvath DP, McLarney BK, Thomashow MF (1993) Regulation of *Arabidopsis thaliana* L. (Heyn) cor78 in response to low temperature. *Plant Physiol* **103**: 1047-1053.
- Horvath DP, Anderson JV, Chao WS, Foley ME (2003) Knowing when to grow: signals regulating bud dormancy. *Trends Plant Sci* **8**:534-540.
- Hoth S, Morgante M, Sanchez JP, Hanafey MK, Tingey SV, Chua NH (2002) Genome-wide gene expression profiling in *Arabidopsis thaliana* reveals new targets of abscisic acid and largely impaired gene regulation in the *abi1-1* mutant. *J Cell Sci* **115**: 4891-4900.
- Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P (2008) Genevestigator V3: a reference expression database for the meta-analysis of transcriptomes. *Advances in Bioinformatics* 420747 doi:10.1155/2008/420747
- Hsieh TH, Lee JT, Yang PT, Chiu LH, Charng YY, Wang YC, Chan MT (2002a) Heterology expression of the *Arabidopsis* C-repeat/dehydration response element binding factor 1 gene confers elevated tolerance to chilling and oxidative stresses in transgenic tomato. *Plant Physiol* **129**: 1086-1094.
- Hsieh TH, Lee JT, Charng YY, Chan MT (2002b) Tomato plants ectopically expressing *Arabidopsis* CBF1 show enhanced resistance to water deficit stress. *Plant Physiol* **130**: 618-626.
- Hulzink RJ, de Groot PF, Croes AF, Quaedvlieg W, Twell D, Wullems GJ, Van Herpen MM (2002) The 5'-untranslated region of the *ntp303* gene strongly enhances translation during pollen tube growth, but not during pollen maturation. *Plant Physiol* **129**: 342-353.
- Hundertmark M, Hinch DK (2008) LEA (Late Embryogenesis Abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics* **9**: 118-139.
- Hurkman W, Tanaka C (1986) Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. *Plant Physiol* **81**: 802-806.
- Imai R, Chang L, Ohta A, Bray E, Takagi M (1996) A lea-class gene of tomato confers salt and freezing tolerance when expressed in *Saccharomyces cerevisiae*. *Gene* **170**: 243-248.
- Ingram J, Bartels D (1996) The molecular basis of dehydration tolerance in plants. *Annu Rev Plant Physiol Plant Mol Biol* **47**: 377-403.
- Ismail AM, Hall AE, Close TJ (1999) Purification and partial characterization of a dehydrin involved in chilling tolerance during seedling emergence of cowpea. *Plant Physiol* **120**: 237-244.
- Iwasaki T, Yamaguchi-Shinozaki K, Shinozaki K (1995) Identification of a *cis* regulatory region of a gene in *Arabidopsis thaliana* whose induction by dehydration is mediated by abscisic acid and requires protein synthesis. *Mol Gen Genet* **247**: 391-398.
- Jaglo KR, Kleff S, Amundsen KL, Zhang X, Haake V, Zhang JZ, Deits T, Thomashow MF (2001) Components of the *Arabidopsis* C-repeat/dehydration-responsive element binding factor cold-response pathway are conserved in *Brassica napus* and other plant species. *Plant Physiol* **217**: 910-917.
- Jaglo-Ottosen KR, Gilmour SJ, Zarka DG, Schabenberger O, Thomashow MF (1998) *Arabidopsis* CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science* **280**: 104-106.
- Jaya N, Garcia V, Vierling E (2009) Substrate binding site flexibility of the small heat shock protein molecular chaperones. *Proc Natl Acad Sci USA* **106**:15604-15609.
- Jepson SG, Close T (1995) Purification of a maize dehydrin protein expressed in *Escherichia coli*. *Protein Expr Purific* **6**: 632-636.
- Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Molec Cell* **14**: 787-799.
- Kader JC (1996) Lipid-transfer proteins in plants. *Ann Revi Plant Physiol Plant Mol Biol* **47**: 627-654.
- Kang JY, Choi HI, Im MY, Kim SY (2002) *Arabidopsis* basic leucine zipper proteins that mediate stress-responsive abscisic acid signaling. *Plant Cell* **14**: 343-357.
- Karimi M, Inzé D, Depicker A (2002) GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci* **7**:193-195.
- Karlson DT, Fujino T, Kimura S, Baba K, Itoh T, Ashworth EN (2003) Novel plasmodesmata association of dehydrin-like proteins in cold-acclimated Red-osier dogwood (*Cornus sericea*). *Tree Physiol* **23**:759-67.
- Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1999) Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nat Biotechnol* **17**: 287-291.

- Kinoshita N, Berr A, Belin C, Chappuis R, Nishizawa NK, Lopez-Molina L (2010) Identification of growth insensitive to ABA3 (*gia3*), a recessive mutation affecting ABA Signaling for the control of early post-germination growth in *Arabidopsis thaliana*. *Plant Cell Physiol* **51**: 239-251
- Kikawada T, Nakahara Y, Kanamori Y, Iwata K, Watanabe M, McGee B, Tunnacliffe A, Okuda T (2006) Dehydration-induced expression of LEA proteins in an anhydrobiotic chironomid. *Biochem Biophys Res Commun* **348**:56-61.
- Kishor KPB, Hong Z, Miao GH, Hu CAA, Verma DPS (1995) Overexpression of Δ^1 -pyrroline-5-carboxylate synthetase increase proline production and confers osmotolerance in transgenic plants. *Plant Physiol* **108**: 1387–1394.
- Kizis D, Lumbreras V, Pagès M (2001) Role of AP2/EREBP transcription factors in gene regulation during abiotic stress. *FEBS Lett* **498**: 187-189.
- Knight H, Knight MR (2001) Abiotic stress signaling pathways: specificity and cross-talk. *Trends Plant Sci* **6**: 262-267.
- Koag MC, Fenton RD, Wilkens S, Close T (2003) The binding of maize DHN1 to lipid vesicles. Gain of structure and lipid specificity. *Plant Physiol* **131**: 309-316.
- Koornneef M, Leon-Kloosterziel KM, Schuartz SH, Zeevaart JAD (1998) The genetic and molecular dissection of abscisic acid biosynthesis and signal transduction in *Arabidopsis*. *Plant Physiol Biochem* **36**: 83-89.
- Kuersten S, Goodwin EB. (2003) The power of the 3' UTR: translational control and development. *Nat Rev Genet* **4**: 626-637.
- Kuromori T, Miyaji T, Yabuuchi H, Shimizu H, Sugimoto E, Kamiya A, Moriyama Y, Shinozaki K (2010) ABC transporter AtABCG25 is involved in abscisic acid transport and responses. *Proc Natl Acad Sci USA* **107**: 2361-2366.
- Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**: 105-132.
- Lai EC (2003) microRNAs: runts of the genome assert themselves. *Curr Biol* **13**: R925-R936.
- Lal S, Gulyani V, Khurana P (2007) Overexpression of HVA1 gene from barley generates tolerance to salinity and water stress in transgenic mulberry (*Morus indica*). *Transgenic Res* **17**: 651-663
- Lalonde S, Ehrhardt DW, Loqué D, Chen J, Rhee SY, Frommer WB (2008) Molecular and cellular approaches for the detection of protein-protein interactions: latest techniques and current limitations. *Plant J* **53**: 610-635.
- Lang D, Eisinger J, Reski R, Rensing S (2005) Representation and high-quality annotation of the *Physcomitrella patens* transcriptome demonstrates a high proportion of proteins involved in metabolism among mosses. *Plant Biol* **7**: 238-250.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2. *Bioinformatics* **23**: 2947-2948.
- Laurie S, Feeney KA, Maathuis FJ, Heard PJ, Brown SJ, Leigh RA (2002) A role for HKT1 in sodium uptake by wheat roots. *Plant J* **32**: 139-149.
- Lee JH, Hübel A, Schöffl F (1995) Derepression of the activity of genetically engineered heat shock factor causes constitutive synthesis of heat shock proteins and increased thermotolerance in transgenic *Arabidopsis*. *Plant J* **8**: 603-612.
- Lee KH, Piao HL, Kim HY, Choi SM, Jiang F, Hartung W, Hwang I, Kwak JM, Lee IJ, Hwang I (2006) Activation of glucosidase via stress-induced polymerization rapidly increases active pools of abscisic acid. *Cell* **126**:1109-1120.
- Lee SC, Lan WZ, Kim BG, Li L, Cheong YH, Pandey GK, Lu G, Buchanan BB, Luan S (2007) A protein phosphorylation/dephosphorylation network regulates a plant potassium channel. *Proc Natl Acad Sci USA* **104**:15959-15964.
- Lee SC, Lan W, Buchanan BB, Luan S (2009) A protein kinase-phosphatase pair interacts with an ion channel to regulate ABA signaling in plant guard cells. *Proc Natl Acad Sci USA* **106**: 21419-21424.
- Lian HL, Yu X, Ye Q, Ding XS, Kitagawa Y, Kwak SS, Su WA, Tang Z (2004) The role of aquaporin RWC3 in drought avoidance in rice. *Plant Cell Physiol* **45**: 481-489.
- Lillie SH, Brown SS (1987) Artfactual immunofluorescent labeling in yeast, demonstrated by affinity purification of antibody. *Yeast* **3**: 63-70.
- Lisse T, Bartels D, Kalbitzer HR, Jaenicke R (1996) The recombinant dehydrin-like desiccation stress protein from the resurrection plant *Craterostigma plantagineum* displays no defined three-dimensional structure in its native state. *Biol Chem* **377**: 555-567.
- Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, DREB1 and DREB2, with an EREB/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell* **10**: 1391-1406.
- Liu Y, Zheng Y (2005) PM2, a group 3 LEA protein from soybean, and its 22-mer repeating region confer salt tolerance in *Escherichia coli*. *Biochem Biophys Res Commun* **331**: 325-332.
- Liu X, Yue Y, Li B, Nie Y, Li W, Wu WH, Ma L (2007) A G protein-coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid. *Science* **315**: 1712-1716.
- López-Molina L, Mongrand S, McLachlin DT, Chait BT, Chua NH. (2002) ABI5 acts downstream of ABI3 to execute an ABA-dependent growth arrest during germination. *Plant J* **32**:317-328.

- Lupas A, Van Dyke M, Stock J (1991) Predicting coiled coils from protein sequences. *Science* **252**: 1162-1164
- Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. *Science* **290**: 1151-1154.
- Madin KAC, Crowe JH (1975) Anhydrobiosis in nematodes: Carbohydrate and lipid metabolism during dehydration. *J Exp Zool* **193**: 335-337.
- Manfre AJ, Lanni LM, Marcotte WR (2006) The *Arabidopsis* group 1 late embryogenesis abundant protein ATEM6 is required for normal seed development. *Plant Physiol* **140**: 140-149.
- Mardanov ES, Zamchuk LA, Ravin NV (2007) The 5' untranslated region of the maize alcohol dehydrogenase gene provides efficient translation of mRNA in plants under stress conditions. *Mol Biol (Mosk)* **41**: 1002-1008.
- Mark WY, Liao JC, Lu Y, Ayed A, Laister R, Szymczynska B, Chakrabarty A, Arrowsmith CH (2005) Characterization of segments from the central region of BRCA1: An intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions? *J Mol Biol* **345**: 275-287.
- Mason JM, Hagemann UB, Arndt KM (2009) Role of hydrophobic and electrostatic interactions in coiled coil stability and specificity. *Biochemistry* **48**:10380-10388.
- Maurel C, Chrispeels MJ (2001) Aquaporins: a molecular entry into plant water relations. *Plant Physiol* **125**: 135-138.
- Maskin L, Maldonado S, Iusem ND (2008) Tomato leaf spatial expression of stress-induced Asr genes. *Mol Biol Rep* **35**: 501-505.
- McCubin WD, Kay CM (1985) Hydrodynamics and optical properties of the wheat Em protein. *J Biochem* **63**: 803-810.
- McGuffin LJ, Bryson K, Jones DT (2000) The PSIPRED protein structure prediction server. *Bioinformatics* **16**: 404-405.
- McKersie BD, Bowley SR, Harjanto E, Leprince O (1996) Water-deficit tolerance and field performance of transgenic alfalfa overexpressing superoxide dismutase. *Plant Physiol* **111**: 1177-1181.
- McKersie BD, Bowley SR, Jones KS (1999) Winter survival of transgenic alfalfa overexpressing superoxide dismutase. *Plant Physiol* **119**: 839-848.
- McKersie BD, Murnaghan J, Jones KS, Bowley SR (2000) Iron-superoxide dismutase expression in transgenic alfalfa increases winter survival without a detectable increase in photosynthetic oxidative stress tolerance. *Plant Physiol* **122**: 1427-1438.
- Melcher K, Xu Y, Ng LM, Zhou XE, Soon FF, Chinnusamy V, Suino-Powell KM, Kovach A, Tham FS, Cutler SR, Li J, Yong EL, Zhu JK, Xu HE (2010) Identification and mechanism of ABA receptor antagonism. *Nat Struct Mol Biol* **17**:1102-1108.
- Menze MA, Boswell L, Toner M, Hand SC (2009) Occurrence of mitochondria-targeted Late Embryogenesis Abundant (LEA) gene in animals increases organelle resistance to water stress. *J Biol Chem* **284**: 10714-10719.
- Miernyk JA, Thelen JJ (2008) Biochemical approaches for discovering protein-protein interactions. *Plant J* **53**:597-609.
- Milborrow BV (2001) The pathway of biosynthesis of abscisic acid in vascular plants: a review of the present state of knowledge of ABA biosynthesis. *J Exp Bot* **52**: 1145-1164.
- Mishra SK, Tripp J, Winkelhaus S, Tschiersch B, Theres K, Nover L, Scharf KD (2002) In the complex family of heat stress transcription factors, HsfA1 has a unique role as master regulator of thermotolerance in tomato. *Genes Dev* **16**: 1555-1567.
- Miyazono K, Miyakawa T, Sawano Y, Kubota K, Kang HJ, Asano A, Miyauchi Y, Takahashi M, Zhi Y, Fujita Y, Yoshida T, Kodaira KS, Yamaguchi-Shinozaki K, Tanokura M (2009) Structural basis of abscisic acid signalling. *Nature* **462**:609-614.
- Mlotshwa S, Voynet, O, Mette MF, Matzke M, Vaucheret H, Ding S, Pruss G, Vance V (2002) RNA silencing and the mobile silencing signal. *Plant Cell Supplement* **14**: S289-S301.
- Mouillon JM, Gustafsson P, Harryson P (2006) Structural investigation of disordered stress proteins. Comparison of full-length dehydrins with isolated peptides of their conserved segments. *Plant Physiol* **141**: 638-650
- Mourrain P, Béclin C, Elmayan T, Feuerbach F, Godon C, Morel J, Jouette D, Lacombe A, Nikic S, Picault N, Remoue K, Sanial M, Vo T, Vaucheret H (2000) *Arabidopsis* SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell* **101**: 533-542.
- Moreno-Fonseca LP, Covarrubias AA (2001) Downstream DNA sequences are required to modulate *Pvlea-18* gene expression in response to dehydration. *Plant Mol Biol* **45**: 501-515.
- Morison JI, Baker NR, Mullineaux PM, Davies WJ (2008) Improving water use in crop production. *Philos Trans R Soc Lond B Biol Sci* **363**: 639-658
- Mudgett MB, Clarke S (1994) Hormonal end environmental responsiveness of a developmentally regulated protein repair L-isoaspartyl methyltransferase in wheat. *J Biol Chem* **269**: 25605-25612.
- Muñoz V, Serrano L (1994) Elucidating the folding problem of helical peptides using empirical parameters. *Nat Struct Biol* **1**: 399-409.
- Nakayama K, Okawa K, Kakizaki T, Honma T, Itoh H, Inaba T (2007) *Arabidopsis* Cor15am is a chloroplast stromal protein that has cryoprotective activity and forms oligomers. *Plant Physiol* **144**: 513-523

- Napoli C, Lemieu, C, Jorgensen R (1990) Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in *trans*. *Plant Cell* **2**: 279-289.
- Nakayama K, Okawa K, Kakizaki T, Honma T, Itoh H, Inaba T (2007) Arabidopsis Cor15am is a chloroplast stromal protein that has cryoprotective activity and forms oligomers. *Plant Physiol* **144**: 513-523
- Ndong C, Danyluk J, Wilson KE, Pocock T, Huner NPA, Sarhan F (2002) Cold-regulated cereal chloroplast late embryogenesis abundant-like proteins. Molecular characterization and functional analyses. *Plant Physiol* **129**: 1368-1381.
- Nishimura N, Sarkeshik A, Nito K, Park SY, Wang A, Carvalho PC, Lee S, Caddell DF, Cutler SR, Chory J, Yates JR, Schroeder JI (2010) PYR/PYL/RCAR family members are major in-vivo ABI1 protein phosphatase 2C-interacting proteins in Arabidopsis. *Plant J* **61**: 290-299.
- Nishio N, Satoh H (1997) A water-soluble chlorophyll protein in cauliflower may be identical to BnD22, a drought-induced 22-kilodalton protein in rapeseed. *Plant Physiol* **115**: 841-846.
- Niu QW, Lin SS, Reyes JL, Chen KC, Wu HW, Yeh SD, Chua NH (2006) Expression of artificial microRNAs in transgenic *Arabidopsis thaliana* confers virus resistance. *Nat Biotechnol* **24**: 1420-1428
- Nobel PS (1994) Remarkable agaves and cacti. Oxford University Press, New York.
- Nobel PS (2005) Physicochemical and environmental plant physiology. Academic Press, San Diego.
- Notredame C, Higgins D, Heringa J (2000) T-Coffee: A novel method for multiple sequence alignments. *J Mol Biol* **302**: 205-217.
- Nylander M, Svensson J, Palva ET, Wellin BV (2001) Stress-induced accumulation and tissue-specific localization of dehydrins in *Arabidopsis thaliana*. *Plant Mol Biol* **45**: 263-279.
- Oberschall A, Deák M, Török K, Sass L, Vass I, Kovács I, Fehér A, Dudits D, Horvath GV (2000) A novel aldose/aldehyde reductase protects transgenic plants against lipid peroxidation under chemical and drought stresses. *Plant J* **24**: 437-446.
- Oliver AE, LePrince O, Wolkers WF, Hinch DK, Heyer AG, Crowe JH (2001) Non-disaccharide-based mechanisms of protection during drying. *Cryobiology* **43**: 151-167.
- Oono Y, Seki M, Nanjo T, Narusaka M, Fujita M, Satoh R, Satou M, Sakurai T, Ishida J, Akiyama K, Iida K, Maruyama K, Satoh S, Yamaguchi-Shinozaki K, Shinozaki K (2003) Monitoring expression profiles of *Arabidopsis* gene expression during rehydration process after dehydration using ca. 7000 full-length cDNA microarray. *Plant J* **34**: 868-887.
- Oparka K (1994) Plasmolysis. New insights into an old process. *New Phytol* **126**: 109-121.
- Owens CL, Thomashow MF, Hancock JF, Lezzoni AF (2002) CBF1 orthologs in sour cherry and strawberry and the heterologous expression of CBF1 in strawberry. *J Amer Soc Hortic Sci* **127**: 489-494.
- Pandey GK, Cheong YH, Kim KN, Grant JJ, Li L, Hung W, D'Angelo C, Weinl S, Kudla J, Luan S (2004) The calcium sensor calcineurin B-like 9 modulates abscisic acid sensitivity and biosynthesis in Arabidopsis. *Plant Cell* **16**: 1912-1924.
- Park SY, Fung P, Nishimura N, Jensen DR, Fujii H, Zhao Y, Lumba S, Santiago J, Rodrigues A, Chow TF, Alfred SE, Bonetta D, Finkelstein R, Provart NJ, Desveaux D, Rodriguez PL, McCourt P, Zhu JK, Schroeder JI, Volkman BF, Cutler SR (2009) Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. *Science* **324**: 1068-1071.
- Pellegrineschi A, Reynolds M, Pacheco M, Brito RM, Almeraya R, Yamaguchi-Shinozaki K, Hoisington D (2004) Stress-induced expression in wheat of the *Arabidopsis thaliana* DREB1A gene delays water stress symptoms under greenhouse conditions. *Genome* **47**: 493-500.
- Perruc E, Charpentreau M, Ramirez BC, Jauneau A, Galaud JP, Ranjeva R, Ranty B (2004) A novel calmodulin-binding protein functions as a negative regulator of osmotic stress tolerance in *Arabidopsis thaliana* seedlings. *Plant J* **38**: 410-420.
- Pilon-Smiths EAH, Ebskamp MJM, Paul MJ, Jeuken MJW, Weisbeek PJ, Smeekens SCM (1995) Improved performance of transgenic fructan-accumulating tobacco under drought stress. *Plant Physiol* **107**: 125-130.
- Pinyopich A, Ditta GS, Savidge B, Liljegren SJ, Baumann E, Wisman E, Yanofsky MF (2003) Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature* **424**: 85-88.
- Prändl R, Hinderhofer K, Eggers-Schumacher G, Schöffl F (1998) HSF3, a new heat shock factor from *Arabidopsis thaliana*, derepresses the heat shock response and confers thermotolerance when overexpressed in transgenic plants. *Mol Gen Genet* **258**: 269-278.
- Prestrelski SJ, Tedeschi N, Arakawa T, Carpenter JF (1993) Dehydration-induced conformational transitions in proteins and their inhibition by stabilizers. *Biophys J* **65**: 661-671.
- Pruvot G, Cuiné S, Peltier G, Rey P (1996) Characterization of a novel drought-induced 34kDa protein located in the thylakoids of *Solanum tuberosum* L. plants. *Planta* **198**: 471-479.
- Puhakainen T, Hess MW, Makela P, Svensson J, Heino P, Palva ET (2004) Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in Arabidopsis. *Plant Mol Biol* **54**: 743-753
- Raghavendra AS, Gonugunta VK, Christmann A, Grill E (2010) ABA perception and signalling. *Trends Plant Sci* **15**:395-3401.
- Ramanjulu S, Bartels D (2002) Drought- and desiccation-induced modulation of gene expression in plants. *Plant Cell Environ* **25**: 141-151.

- Reichman JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR (2000) *Arabidopsis* transcription factors: genome-wide comparative analysis among eucaryotes. *Science* **290**: 2105-2110.
- Rey P, Pruvot G, Becuwe N, Eymery F, Rumeau D, Peltier GA (1998) A novel thioredoxin-like protein located in the chloroplast is induced by water deficit in *Solanum tuberosum* L. plants. *Plant J* **13**:97-107.
- Reyes JL, Rodrigo MJ, Colmenero-Flores JM, Gil JV, Salamini F, Bartels D, Covarrubias AA (2005) Hydrophilins from distant organisms can protect enzymatic activities from water limitation effects *in vitro*. *Plant Cell Environ* **28**: 709-718.
- Reyes JL, Campos F, Wei H, Arora R, Yang Y, Karlson DT, Covarrubias AA (2008) Functional dissection of hydrophilins during *in vitro* freeze protection. *Plant Cell Environ* **31**: 1781-1790.
- Rhizhsky L, Liang H, Shuman J, Shulaev V, Davletova S, Mittler R (2004) When defense pathways collide. The response of *Arabidopsis* to a combination of drought and heat stress. *Plant Physiol* **134**: 1683-1696.
- Ried JL, Walker-Simmons MK (1993) Group 3 late embryogenesis abundant proteins in desiccation-tolerant seedlings of wheat (*Triticum aestivum* L.). *Plant Physiol* **102**: 125-131.
- Roberts J, DeSimone N, Lingle W, Dure L III (1993) Cellular concentrations and uniformity of cell-type accumulation of two LEA proteins in cotton embryos. *Plant Cell* **5**: 769-780.
- Rohila JS, Chen M, Cerny R, Fromm ME (2004) Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *Plant J* **38**:172-181.
- Romero C, Bellés JM, Vayá JL, Serrano R, Culiáñez-Macià FA (1997) Expression of the yeast *trehalose-6-phosphate synthase* gene in transgenic tobacco plants: pleiotropic phenotypes include drought tolerance. *Planta* **201**: 293-297.
- Rossi M, Lijavetzky D, Bernacchi D, Hopp HE, Iusem N (1996) Asr genes belong to a gene family comprising at least three closely linked loci on chromosome 4 in tomato. *Mol Gen Genet* **252**: 489-492.
- Rost B, Liu J (2003) The PredictProtein server. *Nucleic Acids Research* **31**: 3300-3304.
- Rost B, Sander C (1993) Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol* **232**: 584-599.
- Roxas VP, Smith RK Jr, Allen ER, Allen RD (1997) Overexpression of glutathione S-transferase/glutathione peroxidase enhances the growth of transgenic tobacco seedlings during stress. *Nat Biotech* **15**: 988-991.
- Rubio V, Shen Y, Saijo Y, Liu Y, Gusmaroli G, Dinesh-Kumar SP, Deng XW (2005) An alternative tandem affinity purification strategy applied to *Arabidopsis* protein complex isolation. *Plant J* **41**:767-778.
- Rus AM, Estaño MT, Gisbert C, Garcia-Sogo B, Serrano R, Caro M, Moreno V, Bolarín MC (2001) Expressing the yeast HAL1 gene in tomato increases fruit yield and enhances K⁺/Na⁺ selectivity under salt stress. *Plant Cell Environ* **24**: 875-880.
- Russouw PS, Farrant J, Brandt W, Maeder D, Lindsey GG (1995) Isolation and characterization of a heat-soluble protein from pea (*Pisum sativum*) embryos. *Seed Sci Res* **5**: 137-144.
- Saavedra L, Svensson J, Carballo V, Izemendi D, Welin B, Vidal S (2006) A dehydrin gene in *Physcomitrella patens* is required for salt and osmotic stress tolerance. *Plant J* **45**: 237-249.
- Santiago J, Rodrigues A, Saez A, Rubio S, Antoni R, Dupeux F, Park SY, Márquez JA, Cutler SR, Rodriguez PL. (2009) Modulation of drought resistance by the abscisic acid receptor PYL5 through inhibition of clade A PP2Cs. *Plant J* **60**:575-588.
- Sakamoto A, Murata A, Murata N (1998) Metabolic engineering of rice leading to biosynthesis of glycine betaine and tolerance to salt and cold. *Plant Mol Biol* **38**: 1011-1019.
- Schägger H, Von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analyt Biochem* **166**: 368-379.
- Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Schölkopf B, Weigel D, Lohmann JU (2005) A gene expression map of *Arabidopsis thaliana* development. *Nat Genet* **37**: 501-506
- Schneider K, Wells B, Schmelzer E, Salamini F, Bartels D (1993) Desiccation leads to the rapid accumulation of both cytosolic and chloroplastic proteins in the resurrection plant *Craterostigma plantagineum* Hochst. *Planta* **189**: 120-131.
- Schweizer P, Pokorny J, Schulze-Lefert P, Dudler P (2000) Double-stranded RNA interferes with gene function at the single-cell level in cereals. *Plant J* **24**: 895-903.
- Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, Oono Y, Kamiya A, Nakajima M, Enju A, Sakurai T, Satou M, Akiyama K, Taji T, Yamaguchi-Shinozaki K, Carninci P, Kawai J, Hayashizaki Y, Shinozaki K (2002) Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. *Plant J* **31**: 279-292.
- Senthil-Kumar M, Udayakumar M (2006) High-throughput virus-induced gene-silencing approach to assess the functional relevance of a moisture stress-induced cDNA homologous to *Lea4*. *J Exp Bot* **57**: 2291-2302.
- Shen YY, Wang XF, Wu FQ, Du SY, Cao Z, Shang Y, Wang XL, Peng CC, Yu XC, Zhu SY, Fan RC, Xu YH, Zhang DP (2006) The Mg-chelatase H subunit is an abscisic acid receptor. *Nature* **443**: 823-826.
- Shi H, Ishitani M, Kim C, Zhu JK (2000) The *Arabidopsis thaliana* salt tolerance gene *SOS1* encodes a putative Na⁺/H⁺ antiporter. *Proc Natl Acad Sci USA* **97**: 6896-6901.

- Shi H, Lee BH, Wu SJ, Zhu JK (2003) Overexpression of a plasma membrane Na⁺/H⁺ antiporter gene improves salt tolerance in *Arabidopsis thaliana*. *Nat Biotech* **21**: 81-85.
- Shih MD, Lin SD, Hsieh JS, Tsou CH, Chow TY, Lin TP, Hsing YIC (2004) Gene cloning and characterization of a soybean (*Glycine max* L.) LEA protein, GmPM16. *Plant Mol Biol* **56**: 689-703.
- Shih, M, Hoekstra A, Hsing Y (2008) Late Embryogenesis Abundant Proteins. *Advanc Bot Res* **48**: 211-255.
- Shih M, Hsieh T, Lin T, Hsing Y, Hoekstra F (2010) Characterization of Two Soybean (*Glycine max* L.) LEA IV Proteins by Circular Dichroism and Fourier Transform Infrared Spectrometry. *Plant Cell Physiol* **51**: 395-407.
- Shinozaki K, Yamaguchi-Shinozaki K (2000) Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Curr Op Plant Biol* **3**: 217-223.
- Silhavy D, Hutvágner G, Barta E, Bánfalvi Z (1995) Isolation and characterization of a water-stress-inducible cDNA clone from *Solanum chacoense*. *Plant Mol Biol* **27**: 587-95.
- Siegel RS, Xue S, Murata Y, Yang Y, Nishimura N, Wang A, Schroeder JI (2009) Calcium elevation-dependent and attenuated resting calcium-dependent abscisic acid induction of stomatal closure and abscisic acid-induced enhancement of calcium sensitivities of S-type anion and inward-rectifying K channels in *Arabidopsis* guard cells. *Plant J* **59**: 207-20.
- Singh S, Cornilescu CC, Tyler RC, Cornilescu G, Tonelli M, Lee MS, Markley JL (2005) Solution structure of a late embryogenesis abundant protein (LEA14) from *Arabidopsis thaliana*, a cellular stress-related protein. *Protein Sci* **14**: 2601-2609.
- Sivakolundu SG, Bashford D, Kriwacki RW (2005) Disordered p27Kip1 exhibits intrinsic structure resembling the Cdk2/cyclin A-bound conformation. *J Mol Biol* **353**: 1118-1128.
- Sivamani E, Bahieldin A, Wraith JM, Al-Niemia T, Dyer WE, Hod THD, Wu R (2000) Improved biomass productivity and water use efficiency under water deficit conditions in transgenic rice constitutively expressing the barley *HVA1* gene. *Plant Sci* **155**: 1-9.
- Smirnoff N (1998) Plant resistance to environmental stress. *Curr Op Plant Biotechnol* **9**: 214-219.
- Smith GP, Petrenko VA (1997) Phage display. *Chem Rev* **97**: 391-410.
- Smith N, Singh S, Wang M, Stoutjesdijk P, Green A, Waterhouse P (2000) Total silencing by intron-spliced hairpin RNAs. *Nature* **407**: 319-320.
- Solomon A, Solomon R, Paperna I, Glazer I (2000) Desiccation stress of entomopathogenic nematodes induces the accumulation of a novel heat-stable protein. *Parasitology* **121**: 409-416.
- Soulages JL, Kim K, Walters C, Cushman JC (2002) Temperature-induced extended helix/random coil transitions in a group 1 late embryogenesis-abundant protein from soybean. *Plant Physiol* **128**: 822-832.
- Soulages JL, Kim K, Arrese EL, Walters C, Cushman JC (2003) Conformation of a group 2 late embryogenesis abundant protein from soybean. Evidence of poly (L-proline)-type II structure. *Plant Physiol* **131**: 963-975.
- Stengel F, Baldwin AJ, Painter AJ, Jaya N, Basha E, Kay LE, Vierling E, Robinson CV, Benesch JL (2010) Quaternary dynamics and plasticity underlie small heat shock protein chaperone function. *Proc Natl Acad Sci USA* **107**: 2007-2012.
- Stacy RAP, Aalen RB (1998) Identification of sequence homology between the internal hydrophilic repeated motifs of Group 1 late-embryogenesis-abundant proteins in plants and hydrophilic repeats of the general stress protein GsiB of *Bacillus subtilis*. *Planta* **206**: 476-478.
- Steponkus PL, Uemura M, Joseph RA, Gilmour S, Thomashow MF (1998) Mode of action of the COR15a gene on the freezing tolerance of *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* **95**: 14570-14575.
- Stern G, Jansky S, Bidlack J (2003) *Introductory Plant Biology*. McGrawHill, Boston, MA.
- Sun W, Bernard C, van der Cotte B, van Montagu M, Verbruggen N (2001) At-HSP17.6A, encoding a small heat-shock protein in *Arabidopsis*, can enhance osmotolerance upon overexpression. *Plant J* **27**: 407-415.
- Swire-Clark GA, Marcotte WR (1999) The wheat LEA protein Em functions as an osmoprotective molecule in *Saccharomyces cerevisiae*. *Plant Mol Biol* **39**: 117-128.
- Taiz L, Zeiger E (1991) *Plant Physiology*. The Benjamin/Cumming Publishing Company, Inc.
- Tanaka S, Ikeda K, Miyasaka H (2004) Isolation of a new member of group 3 late embryogenesis abundant protein gene from a halotolerant green alga by a functional expression screening with cyanobacterial cells. *FEMS Microbiol Lett* **236**: 41-45.
- Tarczynski MC, Jensen RG, Bohnert HJ (1993) Stress protection of transgenic tobacco by production of the osmolyte mannitol. *Science* **259**: 508-510.
- Tezara W, Mitchell V, Driscoll SP, Lawlor DW (2002) Effects of water deficit and its interaction with CO₂ supply on the biochemistry and physiology of photosynthesis in sunflower. *J Exp Bot* **53**: 1781-1791.
- Thomas JC, Sepahi M, Arendall B, Bonhert HJ (1995) Enhancement of seed germination in high salinity by engineering mannitol expression in *Arabidopsis thaliana*. *Plant Cell Environ* **18**: 801-806.
- Tijsterman M, Ketting R, Plasterk R (2002) The genetics of RNA silencing. *Annu Rev Genet* **36**: 489-519.
- Tissier A, Marillonnet S, Klimyuka V, Patela K, Torres MA, Murphy G, Jones J (1999) Multiple independent defective Suppressor-mutator transposon insertions in *Arabidopsis*: A Tool for Functional Genomics. *Plant Cell* **11**: 1841-1852.
- Tompa P (2002) Intrinsically unstructured proteins. *Trends Biochem Sci* **27**: 527-533.
- Tolletier D, Jaquinod M, Mangavel C, Passirani C, Saulnier P, Manon S, Teyssier E, Payet N, Avelange-

- Macherel MH, Macherel D (2007) Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation. *Plant Cell* **19**: 1580-1589.
- Torres-Schuman S, Godoy JA, Pintor-Toro JA (1992) A probable lipid transfer protein gene is induced by NaCl in stems of tomato plants. *Plant Mol Biol* **18**: 749-757.
- Tolletier D, Jaquinod M, Mangavel C, Passirani C, Saulnier P, Manon S, Teyssier E, Payet N, Avelange-Macherel MH, Macherel D (2007) Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation. *Plant Cell* **19**: 1580-1589.
- Trethowan, RM, Reynolds, M (2007) Drought resistance: genetic approaches for improving productivity under stress. En: *Wheat production in stressed environments*. Back HT, Nisi JE, Salomon N (eds), Springer-Verlag, Dordrecht, Netherlands. Pp 289-299.
- Tunnacliffe A, Lapinski J, McGee B (2005) A putative LEA protein, but no trehalose, is present in anhydrobiotic bdelloid rotifers. *Hydrobiologia* **546**: 315-321.
- Uno Y, Furihata T, Abe H, Yoshida R, Shinozaki K, Yamaguchi-Shinozaki K (2000) Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *Proc Natl Acad Sci USA* **97**: 11632-11637.
- Urao T, Yakubov B, Satoh R, Yamaguchi-Shinozaki K, Seki M, Hirayama T, Shinozaki K (1999) A transmembrane hybrid-type histidine kinase in Arabidopsis functions as an osmosensor. *Plant Cell* **11**: 1743-1754.
- Uversky VN (2002) What does it mean to be natively unfolded? *Europ J Biochem* **269**: 2-12.
- Vahisalu T, Kollist H, Wang YF, Nishimura N, Chan WY, Valerio G, Lamminmäki A, Brosché M, Moldau H, Desikan R, Schroeder JI, Kangasjärvi J (2008) SLAC1 is required for plant guard cell S-type anion channel function in stomatal signalling. *Nature* **452**: 487-91.
- Vaidyanathan R, Kuruvilla S, Thomas G (1999) Characterization and expression patterns of an abscisic acid and osmotic stress responsive gene from rice. *Plant Sci* **140**: 21-30.
- van der Wee C, Spollen W, Sharp R, Baskin T (2000) Growth of *Arabidopsis thaliana* seedlings under water deficit studied by control of water potential in nutrient-agar media. *J Exp Bot* **51**: 1555-1562.
- van Camp W, Capiou K, Van Montagu M, Inzé D, Slooten L (1996) Enhancement of oxidative stress tolerance in transgenic tobacco plants overproducing Fe-superoxide dismutase in chloroplasts. *Plant Physiol* **112**: 1703-1714.
- Vance V, Vaucheret H (2001) RNA silencing in plants-defense and counterdefense. *Science* **292**: 2277-2280.
- Vaucheret H, Béclin C, Fagard M (2001) Post-transcriptional gene silencing in plants. *J Cell Sci* **114**: 3083-3091.
- Verslues PE, Agarwal M, Katiyar-Agarwal S, Zhu J, Zhu JK (2006) Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. *Plant J* **45**: 523-39. Erratum in: *Plant J* **46**: 1092.
- Vicient CM, Delseny M (1999) Isolation of total RNA from *Arabidopsis thaliana* seeds. *Analyt Biochem* **268**: 412-413.
- Vierling E (1991) The roles of heat-shock proteins in plants. *Ann Rev Plant Biol* **42**: 579-620.
- Wang M, Waterhouse P (2001) Application of gene silencing in plants. *Curr Op Plant Biol* **5**: 146-150.
- Wang W, Vinocur V, Altman A (2003) Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. *Planta* **218**: 1-14.
- Wang XJ, Reyes JL, Chua NH, Gaasterland T (2004) Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets. *Genome Biol* **5**: R65.
- Waterhouse P, Graham M, Wang M (1998) Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc Natl Aca Sci USA* **95**: 13959-13964.
- Waters ER, Lee GJ, Vierling E (1996) Evolution, structure and function of the small heat shock proteins in plants. *J Exp Bot* **47**: 325-338.
- Weber C, Nover L, Fauth M (2008) Plant stress granules and mRNA processing bodies are distinct from heat stress granules. *Plant J* **56**: 517-530.
- Wesley V, Helliwell C, Smith N, Wang M, Rouse D, Liu Q, Gooding P, Singh S, Abbott D, Stoutjesdijk P, Robinson S, Gleave A, Green A, Waterhouse P (2001) Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant J* **27**: 581-590.
- Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ (2007) An "Electronic Fluorescent Pictograph" Browser for Exploring and Analyzing Large-Scale Biological Data Sets. *PLoS ONE* **2**: e718.
- Wise MJ (2002) The POPPs: clustering and searching using peptide probability profiles. *Bioinformatics* **18 Supp 1**: S38-S45.
- Wise MJ (2003) LEAping to conclusions: a computational reanalysis of late embryogenesis abundant proteins and their possible roles. *BMC Bioinformatics* **4**: 52
- Wise MJ, Tunnacliffe A (2004) POPP the question: what do LEA proteins do? *Trends Plant Sci* **9**: 13-17.
- Wolkers WF, McCready S, Brandt WF, Lindsey GG, Hoekstra FA (2001) Isolation and characterization of a D-7 LEA protein from pollen that stabilizes glasses *in vitro*. *Biochim Biophys Acta* **1544**: 196-206.
- Wright PE, Dyson JH (1999) Intrinsically unstructured proteins: re-assessing the protein-structure function paradigm. *J Mol Biol* **293**: 321-331.
- Xiao B, Huang Y, Tang N, Xiong L (2007) Over-expression of a LEA gene in rice improves drought resistance under the field conditions. *Theor Appl Genet* **115**: 35-46.

- Xiong LX, Schumaker KS, Zhu JK (2002) Cell signaling during cold, drought, and salt stress. *Plant Cell Supplement* **14**: S165-S183.
- Xu D, Duan X, Wang B, Hong B, Ho TD, Wu R (1996) Expression of a late embryogenesis abundant protein gene, *HVA1*, from barley confers tolerance to water deficit and salt stress in transgenic rice. *Plant Physiol* **110**: 249-257.
- Yamaguchi-Shinozaki K, Koizumi M, Urao S, Shinozaki K (1992) Molecular cloning and characterization of 9 cDNAs for genes that are responsive to desiccation in *Arabidopsis thaliana*: sequence analysis of one clone that encodes a putative transmembrane channel protein. *Plant Cell Phys* **33**:217-224.
- Yamaguchi-Shinozaki K, Shinozaki, K (1994) A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *Plant Cell* **6**: 251-264.
- Yamaguchi-Shinozaki K, Shinozaki K (2006) Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu Rev Plant Biol* **57**: 781-803.
- Yang SH, Zeevaart JA (2006) Expression of ABA 8'-hydroxylases in relation to leaf water relations and seed development in bean. *Plant J* **47**:675-86.
- Yin P, Fan H, Hao Q, Yuan X, Wu D, Pang Y, Yan C, Li W, Wang J, Yan N. (2009) Structural insights into the mechanism of abscisic acid signaling by PYL proteins. *Nat Struct Mol Biol* **16**:1230-1236.
- Zamore P, Tuschl T, Sharp P, Bartel D (2000) RNAi: double stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**:25-33.
- Zamore P (2001) RNA interference: listening to the sound of silence. *Nature Struct Biol* **8**: 746-750.
- Zhang HX, Blumwald E (2001) Transgenic salt-tolerant tomato plants accumulate salt in foliage but not in fruit. *Nat Biotechnol* **19**: 765-768.
- Zhang HX, Hodson JN, Williams JP, Blumwald E (2001) Engineering salt-tolerant *Brassica napus* plants: characterization of yield and seed oil quality in transgenic plants with increased vacuolar sodium accumulation. *Proc Natl Acad Sci USA* **98**: 12832-12836.
- Zhang L, Ohta A, Takagi M, Imai R (2000) Expression of plant group 2 and group 3 lea genes in *Saccharomyces cerevisiae* revealed functional divergence among LEA proteins. *J Biochem* **127**: 611-616.
- Zhu B, Chang MC, Verma DPS, Fan YL, Wu R (1998) Overexpression of a Δ^1 -pyrroline-5-carboxylate synthetase gene and analysis of tolerance to water-and salt-stress in transgenic rice. *Plant Sci* **139**: 41-48.
- Zhu JK (2002) Salt and drought stress signal transduction in plants. *Annu Rev Plant Biol* **53**: 247-273.
- Zmasek CM, Eddy SR (2001) ATV: Display and manipulation of annotated phylogenetic trees. *Bioinformatics* **17**: 383-384.
- Zonia L, Munnik T (2007) Life under pressure: hydrostatic pressure in cell growth and function. *Trends Plant Sci* **12**: 90-97.