



Universidad Nacional Autónoma de México

---

---

## Posgrado en Ciencias Biológicas

Instituto de Ecología

Análisis de expresión de una familia de proteínas de función desconocida,  
DUF642, en *Arabidopsis thaliana* y su posible papel en las vías de  
transducción de señales durante el desarrollo.

T E S I S

QUE PARA OBTENER EL GRADO DE

Maestra en Ciencias Biológicas

(Biología Experimental)

P R E S E N T A :

Esther Zúñiga Sánchez

Directora de tesis: Dra. Alicia Gamboa de Buen

México D.F.

Enero 2007



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.


Ing. Leopoldo Silva Gutiérrez  
Director General de Administración Escolar, UNAM  
Presente

Por medio de la presente me permito informar a usted que en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día 9 de octubre del 2006, se acordó poner a su consideración el siguiente jurado para el examen de grado de Maestría en Ciencias Biológicas (Biología Experimental) de la alumna **Zuñiga Sánchez Esther** con número de cuenta **96375235** con la tesis titulada: **"Análisis de expresión de una familia de proteínas de función desconocida, DUF642, en *Arabidopsis thaliana* y su posible papel en las vías de transducción de señales durante el desarrollo"** bajo la dirección de la Dra. **Alicia Gamboa de Buen**.

Presidente: Dra. Gladys Cassab López  
Vocal: Dra. Rocío Cruz Ortega  
Secretario: Dra. Alicia Gamboa de Buen  
Suplente: Dra. Tzvetanka Dimitrova Dinkova  
Suplente: Dr. Juan Rafael Riesgo Escovar

Sin otro particular, quedo de usted.

Atentamente  
"POR MI RAZA HABLARA EL ESPIRITU"  
Cd. Universitaria, D.F. a 18 de enero del 2006

  
Dr. Juan Núñez Farfán  
Coordinador del Programa

## ***Reconocimientos***

Agradezco tanto a nuestra gran casa de estudios la UNAM, específicamente a la Dirección General de Estudios de Posgrado (DEGEP), así como al Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo económico que me otorgaron durante el período que abarcó esta maestría.

Este trabajo fue financiado por el PAPIIT (UNAM) a través del proyecto IX239204.

A los miembros de mi Jurado:

**Dra. Gladys Cassab López**

**Dra. Rocío Cruz Ortega**

**Dra. Alicia Gamboa de Buen**

**Dra. Tzvetanca Dimitrova Dinkova**

**Dr. Juan Riesgo Escovar**

## ***Agradecimientos***

Es un hecho, este trabajo no hubiese sido posible sin la dirección de mi asesora la Dra. Alicia Gamboa de Buen. Alicia entiendo que el destino de esta tesis como escrito será el destino de las otras (porque como idea tengo mejores expectativas), con suerte le servirán a unos cuantos para los protocolos o para ver los antecedentes del proyecto (no me corto las venas en lo absoluto), eso me dejará más que satisfecha no pretendo que sea más ni menos, eso sí ojalá que cuando la abras pasees los ojos por los agradecimientos y releas esto con gusto con el gusto con el que lo hice. Gracias por todo el tiempo que invertiste en este proyecto, por el conocimiento transmitido, por tu esfuerzo en momentos críticos, por las pláticas interminables en las que siempre aprendía algo nuevo, gracias por tu sentido del humor (oscuro brillante) y por exponer la realidad sin congoja.

A la Dra. Rocío, quiero agradecerle el apoyo que me brindó como su casi alumna. Dra. Rocío fue un gran apoyo para mí durante la realización de mi proyecto gracias por aceptar haber sido miembro de mi comité tutorial durante estos dos años, por su tiempo sin miramientos ni restricciones y por tener siempre abierta la puerta de su cubículo cuando buscaba respuestas a mis mil dudas.

Al Dr. Jun Riesgo le agradezco las ideas positivas y sugerencias interesantes durante la participación como miembro de mi comité tutorial. Gracias Dr. Riesgo por las correcciones y su interés al revisar el escrito de tesis a pasar de estar a un reino de distancia.

Les agradezco también tanto a la Dra. Gladys Cassab como a la Dra. Tzvetanca por sus valiosas correcciones y aportaciones para el mejoramiento del escrito de tesis.

Caro muchas gracias por tu apoyo en el laboratorio y tus tips que funcionan de maravilla de no ser por esos consejos prácticos, que no aparecen en los protocolos, (dícese quichis quichis y otros) me hubiera estado más horas infructuosas en el laboratorio. Gracias por la clona CBI-26 tomó un lugar importante de mi tesis. Mil gracias también por tu disposición en todo incluyendo la de leer borradores y escritos silvestres.

A la Dra. Alma Orozco quiero agradecerle enormemente el apoyo que siempre me brindó en todos los aspectos y por el cariño transmitido. A Esthercita le agradezco la disposición cuando necesitaba algo, su presencia y lindura.

A mis compañeros del laboratorio a Julio Paéz por la crítica constructiva, por sus protocolos combinados, por la música extrapolada (de Penderecky a Paquita) y por la charla amena; a Saraí a José Luis, a Erendira y a Liber, por su grata compañía al hacer que el laboratorio fuese más cálido en todas las épocas del año.

A mis añorados amigos de todos los días les agradezco los ratos felices a Ale por los buenos tiempos que vienen, a Norma por su risa, a Julián por sus historias de intrepidez, a Fabiola por estar siempre en la tierra, a Elisa por nuestra comunicación suprahumana, a Oyuki por su memoria, a Rocío y Maria Elena que están ausentes mucho tiempo pero imborrables para todas. A mis amigas de toda la vida Ruth, Anabel y Socorro.

A Paula por su amistad y sus visitas en el lab que siempre me caían del cielo como de perlas, a Mitzi por los buenos ratos en las clases que tomamos juntas y la locura del primer semestre de la maestría.

A Isra por el apoyo moral y el entusiasmo de cada día. Gracias por tu amor yo te correspondo!

A mis hermanas les agradezco su cariño y complicidad sana, aunque algunos lo duden!!

***Para mis papás Claudio y Martha  
por su amor incondicional***

***Y he visto que la sabiduría sobrepasa a la necesidad  
como la luz a las tinieblas***

**Ecle. 1:13**

## ÍNDICE GENERAL

	<b>Pag.</b>
Abreviaturas	8
Índice de figuras	9-10
Resumen	11
Summary	12
<b>I. Introducción</b>	<b>13</b>
1.1 La importancia de los dominios de interacción	13
1.2 Los dominios LRR y el establecimiento de interacciones proteína-proteína	16
<b>II. Antecedentes</b>	<b>19</b>
2.1 El genoma de <i>Arabidopsis</i> y las proteínas DUF	19
2.1.2 FLOR1, VSP1 y una proteína DUF642 interactúan <i>in vitro</i> con un decapeptido del dominio MADS de AGAMOUS	20
2.1.3 Nuevos hallazgos de las proteínas DUF642	23
2.1.4 El Proyecto del Genoma Floral y las proteínas DUF642	25
<b>III. Hipótesis</b>	<b>28</b>
<b>IV. Objetivos</b>	<b>29</b>
4.1 Objetivo general	29
4.2 Objetivos específicos	29
<b>V. Métodos</b>	<b>30</b>
5.1 Análisis de expresión mediante RT-PCR	30
5.2 Análisis de las secuencias promotoras	30
5.3 Expresión heteróloga de las proteínas recombinantes At5g11420 y At4g32460 (Proteína completa y truncada):	31
5.3.1 Amplificación y clonación de los genes	31
5.3.2 Expresión y purificación de proteínas	32
5.3.3 Análisis de las fracciones y western blots	33
5.3.4 Cromatografías de afinidad	33
5.4 Obtención de plantas mutantes por inserción de T-DNA para el locus At4g32460	34

<b>VI. Resultados</b>	<b>35</b>
6.1 Patrones de expresión mediante RT-PCR para los genes pertenecientes a la familia DUF642	35
6.2 Sitios de regulación <i>cis</i> en las secuencias promotoras de los genes pertenecientes a la familia DUF642	38
6.3 Interacciones proteína-proteína	40
6.3.1 Interacción de las proteínas At5g11420 completa y truncada con extractos crudos de flores de <i>A. thaliana</i>	40
6.3.1.1 Interacción de la proteína completa At5g11420	40
6.3.1.2 Interacción de la proteína truncada At5g11420	41
6.3.2 Interacción con las proteínas At5g11420 completa y truncada utilizadas como ligando inmovilizado con extractos crudos de hojas de <i>A. thaliana</i>	42
6.3.2.1 Interacción de la proteína completa At5g11420	42
6.3.2.2 Interacción de la proteína truncada At5g11420	43
6.3.3 Interacción de la proteína At4g32460 completa con extractos crudos de flores de <i>A. thaliana</i>	44
6.3.4 Interacción de la proteína completa At4g32460 utilizada como ligando inmovilizado con proteínas de extractos crudos de hojas de <i>Arabidopsis</i>	45
6.4 Líneas mutantes por inserción de T-DNA	47
<b>VII. Discusión</b>	<b>48</b>
<b>VIII. Conclusiones</b>	<b>56</b>
<b>IX. Perspectivas</b>	<b>57</b>
<b>X. Bibliografía citada</b>	<b>59</b>
<b>Anexo I</b>	<b>65</b>
<b>Anexo II</b>	<b>76</b>
<b>Anexo III</b>	<b>88</b>



## Abreviaturas

<b>ABA</b>	Ácido absícico
<b>ABF</b>	Factor de unión a ABRE (ABRE binding factor)
<b>ABP</b>	Proteína de unión a Auxinas (Auxin Bindig Protein)
<b>ABRE</b>	Elementos de respuesta a ABA (ABA-responsive elements)
<b>AG</b>	AGAMOUS
<b>cDNA</b>	Ácido desoxiribonucleico complementario
<b>ChIP</b>	Inmunoprecipitación de cromatina (Chromatin immunoprecipitation)
<b>Col</b>	Columbia ecotipo
<b>CTR1</b>	Constitutive Triple Response 1
<b>ddH<sub>2</sub>O</b>	Agua bidestilada
<b>DNA</b>	Ácido desoxiribonucleico
<b>DPBF</b>	Factores de unión al promotor Dc3 (Dc3 promotor binding factors)
<b>DUF642</b>	Dominio de Función desconocida 642
<b>ETR1</b>	Receptor de etileno 1 (Ethylene Receptor 1)
<b>EVH1</b>	Dominio homólogo a VASP (vasodilator inducer stimulated phospho protein)
<b>FHA</b>	Dominio asociado a cabeza bífida (forkhead–associates)
<b>FRET</b>	Fluorescencia por transferencia de energía resonante (Fluorecence Resonante Energy Transfer)
<b>FYVE</b>	Dominio rico en cisteinas que se une a dos iones Zn <sup>2+</sup>
<b>GLP</b>	Germin- like proteins
<b>GPA</b>	Proteínas ancladas a glicosil fosfatidil inositol (GPI-anchored proteins)
<b>IPTG</b>	Isopropil tiogalactosido
<b>LC-MS/MS</b>	Cromatografía líquida acoplada a espectrometría de masas en tandem
<b>LRR</b>	Dominio Rico en Leucinas (Leucine rich repeat)
<b>MAPK</b>	Proteína Cinasa Activada por Mitogeno (Mitogen-activated protein kinase)
<b>mRNA</b>	Ácido Ribonucleico mensajero
<b>MS</b>	Murashige/Skoog
<b>pb</b>	Pares de bases
<b>PCR</b>	Reacción en cadena de la polimerasa
<b>PG</b>	Poligalacturonasa
<b>PH</b>	Dominio homólogo a pleckstrina (Pleckstrin domain)
<b>PME</b>	Pectin metil esterasa
<b>PMEI</b>	Inhibidor de Pectin metil esterasa
<b>PTB</b>	Dominio de unión a una tirosina fosforilada (phosphotyrosine-binding domain)
<b>PvPGIP</b>	Proteína inhibidora de Poligalacturonasa de <i>Phaseolus vulgaris</i>
<b>RLK</b>	LRR- receptors like kinase
<b>RLP</b>	Receptor- like protein
<b>RNAt</b>	Ácido Ribonucleico total
<b>RT-PCR</b>	Reacción en Cadena de la Polimerasa de Transcripción Reversa
<b>SIMKK</b>	Salt-stress-inducible MAPK
<b>SOD</b>	Superóxido dismutasa
<b>VSP1</b>	Proteína de almacenamiento vegetativo 1 (Vegetative Storage Protein 1)
<b>WD40</b>	Dominio de 7 repeticiones de hoja β
<b>WT</b>	Tipo Silvestre
<b>WW</b>	Dominio con 2 residuos de triptofano conservados (WW domain)

## Índice de Figuras y Tablas

<b>Figura</b>	<b>Nombre</b>	<b>Pág.</b>
1	Modelo de la ruta de la respuesta a etileno en la regulación de la expresión genética	16
2	Un dominio LRR típico tomado de PvPGIP <sub>2</sub>	17
3	Representación Ribbon de la estructura cristalográfica de PGIP <sub>2</sub> de <i>Phaseolus vulgaris</i>	18
4	Gráfica que muestra el análisis funcional del genoma de <i>Arabidopsis thaliana</i>	20
5	Representación esquemática del dominio DUF642	21
6	Alineación múltiple de las secuencias de aminoácidos correspondientes a 9 de las 10 proteínas registradas como proteínas DUF642 en <i>Arabidopsis thaliana</i>	22
7	Reconstrucción filogenética para los genes de la familia de dominio DUF642	23
8	Reconstrucción filogenética mediante análisis de Máxima Parsimonia basada en la secuencia de aminoácidos de las proteínas pertenecientes a la familia DUF642	27
9	Secuencia de aminoácidos de la proteína At5g11420 completa y truncada	32
10	Secuencia de aminoácidos de la proteína At432460 completa y truncada	32
11	Análisis de expresión de los genes pertenecientes a la familia DUF642 mediante la técnica de RT-PCR en <i>Arabidopsis thaliana</i>	36
12	Gel de poliacrilamida al 12% mostrando las distintas fracciones obtenidas a partir de un ensayo de cromatografía de afinidad con un extracto crudo de flores y la proteína At5g11420 completa utilizada como ligando inmovilizado	41
13	Gel de poliacrilamida al 12% mostrando las fracciones obtenidas a partir de una columna de afinidad con extractos crudos de flores y la proteína truncada At5g11420	41
14	Secuencia de aminoácidos de la proteína FLOR1 mostrando los péptidos identificados por LC-MS/MS	42
15	Secuencia de aminoácidos de la proteína VSP1 mostrando los péptidos identificados por LC-MS/MS	42
16	Gel de poliacrilamida al 12% mostrando las fracciones obtenidas a partir de un ensayo utilizando a la proteína completa At5g11420 y un extracto de hojas	43
17	Gel de poliacrilamida al 12% mostrando las fracciones obtenidas a partir de la cromatografía de afinidad con extractos crudos de hojas y la proteína At5g11420 trunca	43
18	Gel de poliacrilamida al 12 % mostrando las distintas fracciones obtenidas a partir	

	de un extracto de flores y la proteína completa At4g32460 como ligando inmovilizado	44
19	Gel de poliacrilamida al 10% mostrando las distintas fracciones obtenidas a partir de un extracto de flores y la proteína completa At4g32460 como ligando inmovilizado	44
20	Gel de poliacrilamida al 12% mostrando la fracciones proteicas obtenidas a partir de una columna de afinidad con la proteína inmovilizada At4g32460 completa y un extracto crudo de hojas	45
21	Secuencia de aminoácidos de la pectin metil esterasa con el número de registro Q9LUL7 At3g14310 (PME)	46
22	Estructura general de las PME	46

<b>Tabla</b>	<b>Nombre</b>	<b>Pág.</b>
1	Distribución de los genes de <i>A. thaliana</i> órgano-específicos de flor	26
2	Consenso que resume la expresión de los 10 genes pertenecientes a la familia DUF642 de <i>A. thaliana</i> en diferentes tejidos a partir de 3 experimentos independientes	37
3	Análisis de las secuencias promotoras de los genes que codifican para proteínas pertenecientes a la familia DUF642	39

## RESUMEN

Los análisis genéticos hechos en diferentes especies han sido una de las vías de investigación mediante las cuales se han hecho avances considerables hacia el entendimiento de los mecanismos involucrados en el desarrollo de distintos organismos. Con esta finalidad se han caracterizado genes, que son miembros de familias multigénicas agrupados con base en la homología de sus secuencias. Existen, sin embargo, familias completas de proteínas cuya función es totalmente desconocida y cuyos dominios, algunos posiblemente catalíticos, son tan distintos al resto de los dominios conocidos que no pueden ser clasificados dentro de los grupos descritos. En *Arabidopsis thaliana* se han identificado por lo menos 1800 familias de proteínas con dominios desconocidos (Domain Unknown Function, DUF).

Experimentos preliminares nos indican la posibilidad de que algunos de los miembros de una de estas familias, la familia DUF642, estén involucrados en el desarrollo vegetal. En un estudio anterior, se implementó una técnica para determinar posibles proteínas interactoras *in vivo* de AGAMOUS, factor de transcripción de la familia MADS-box. Utilizando una columna de afinidad con un decapeptido derivado del dominio MADS de AGAMOUS, se purificó, a partir de flores de *Arabidopsis thaliana*, una proteína con un dominio rico en leucinas perteneciente a la superfamilia de proteínas LRR's (Leucine Rich Repeat) denominada FLOR1 y una proteína con actividad de fosfatasa ácida VSP1. A partir del mismo enfoque utilizando el decapeptido de AGAMOUS pero con maíz, se purificó una proteína homóloga a una proteína de arroz y a proteínas pertenecientes a la familia de proteínas DUF642 en *Arabidopsis thaliana*. El dominio DUF642 es un dominio muy conservado y se ha descrito en distintas especies que van desde Gimnospermas hasta Angiospermas basales. Las secuencias de aminoácidos de las proteínas DUF642 contienen un alto porcentaje de residuos hidrofóbicos por lo que su función principal podría ser el establecimiento de interacciones proteína-proteína.

El objetivo principal de este trabajo fue determinar la función general de los integrantes de la familia de proteínas DUF642 en *Arabidopsis thaliana*. Para cumplir el objetivo, se realizó el análisis de expresión mediante RT-PCRs en diferentes tejidos para cada uno de los genes DUF642, el análisis de las regiones promotoras *in silico*, la búsqueda de una línea mutante de T-DNA homocigota nula para el gen At4g32460 y el análisis de posibles proteínas interactoras de dos de los miembros de esta familia (At5g11420 y At4g32460) mediante cromatografías de afinidad. Los análisis de expresión muestran que los miembros de esta familia, se expresan a nivel de mRNA en la mayor parte de los tejidos analizados. A través de los ensayos de cromatografía de afinidad se determinó que la proteína completa codificada por el locus At5g11420 que interactuó *in vitro* tanto con la proteína LRR FLOR1 como con VSP1 al utilizar extractos proteicos de flores y con una pectin metil esterasa al utilizar extractos de hoja, mientras que la proteína At5g11420 truncada solo interactuó con FLOR1 al utilizar extractos florales; sugiriendo este hecho que el carboxilo terminal de la proteínas At5g11420 es importante para la interacción con VSP1. La proteína codificada por el locus At4g32460 interactuó *in vitro* con la misma PME y con VSP1 al utilizar extractos de flores. Los resultados arrojados por los análisis de espectrofotometría de masas muestran que los péptidos identificados de la PME, una proteína de pared celular, corresponden únicamente a la región catalítica de esta enzima. Por otro lado, la línea mutante homocigota para el gen At4g32460 resultó no ser una mutante nula por lo que es necesario hacer la búsqueda de una nueva línea mutante por inserción de T-DNA para este y otros genes DUF642. La expresión de los genes pertenecientes a la familia DUF642 puede estar regulada por distintos tipos de señales tanto endógenas (ABA y auxinas) como exógenas (luz y estrés), ya que en sus regiones promotoras se encontraron motivos regulatorios reconocidos principalmente por elementos *in trans* relacionados con este tipo de señales. Estos resultados, junto con los análisis de expresión y los datos obtenidos de la literatura, nos permiten sugerir que la función principal de las proteínas DUF642 es el establecimiento de interacciones específicas con proteínas que participan en importantes procesos del desarrollo tanto reproductivo como vegetativo y con proteínas relacionadas con la respuesta a estrés. Es necesario realizar trabajos mucho más detallados con el fin de conocer la función específica de cada uno de los miembros de la familia DUF642.

## Abstract

The genetic analysis in different plant species has allowed the progress in many aspects related with the developmental mechanisms in plants. It has been possible through these analyses to study the functions of many different multigenic families; however there are some families with domains that don't share any similitude with the proteins that were functionally characterized. In *Arabidopsis thaliana* it has been identified at least 1800 multigenic families of genes coding for proteins with domains of unknown function (DUF). Functional characterization of these families will be crucial for a better understanding of different aspects of plant development.

In previous studies focused to identify putative protein interactors of the transcription factor AGAMOUS (AG), floral extracts of *Arabidopsis thaliana* were passed through an affinity chromatography column with a deca-peptide of AG. This method permitted the isolation of two peptides, FLOR1 a protein with a LRR domain (Leucine Repeat) and VSP1, an acid phosphatase. When the same approach was followed with a maize female floral extract, it was possible to purify a peptide subsequently identified as a protein member of the DUF642 family. DUF642 family is specific for plant and their members have been reported in different species including Gymnosperm and basal Angiosperms. The members of this family share a conserved region that is not homologous to any other catalytic domain or to another domain with known function. The amino acid sequence of this domain contains a high percentage of hydrophobic residues, which suggests its involvement in protein-protein interaction.

The main goal of this work was to study the general function of DUF642 family members in *Arabidopsis thaliana*. We show the expression patterns of members of the DUF642 family, *in silico* analyses of its promoters, screening of T-DNA mutants in the At4g32460 locus and the identification of putative protein interactors of At5g11420 and At4g32460 proteins. In experiments using flower extracts, the complete recombinant proteins At5g11420 and At4g32460 interact *in vitro* with FLOR1 and VSP1, while a truncated At5g11420 only interacts with the former. This suggests that the carboxyl terminus of At5g11420 is important for the VSP1 interaction. Using leaf extracts, the recombinant proteins At4g32460 and At5g11420 interact *in vitro* with the PME, a cell wall enzyme. The peptides identified by Mass Spectrometry are all located in the mature PME region at the carboxyl terminus, the catalytic domain. We obtained a mutant line in the gene At4g32460; however, it was not a null mutant, by this reason it is necessary to search others mutant lines. The *in silico* analysis with the DUF642 promoter regions show that this genes could be regulated not only by internal cues (ABA and auxins) but also by external cues (light and stress). These data suggest a possible participation of DUF642 family members in stress responses, notwithstanding other roles as in as reproductive or vegetative developmental signaling pathways can also be possible. We consider that it is necessary to do more studies with the object of to know the specific function of each DUF642 family members in the plant signaling pathways.

## I. Introducción

### 1.1 La importancia de los dominios de interacción proteína-proteína

Las plantas, a diferencia de los animales no pueden evadir los cambios en las condiciones ambientales que les rodean, y en consecuencia, han desarrollado a través de la evolución, una serie de mecanismos moleculares que les permiten lidiar con estos cambios a pesar de su naturaleza sésil. Los procesos fisiológicos que sustentan las respuestas moleculares son el resultado de la integración de múltiples eventos que convergen en redes de señalización o transducción (Shao *et al.*, 2006). Los sistemas regulatorios y las rutas de señalización en las células eucariontas están controlados por proteínas, con múltiples dominios de interacción, las cuales determinan la respuesta de los receptores a las diferentes señales tanto externas como internas. En muchas ocasiones, estas proteínas se encuentran asociadas formando complejos multiproteicos que pueden ser estables o temporales (Pawson *et al.*, 2002). Las interacciones de los diferentes componentes que conforman estos complejos dependen de regiones discretas en las proteínas denominadas dominios de interacción. Un mismo proceso de señalización intracelular puede requerir la actividad de varios de estos complejos. En ocasiones, las proteínas implicadas en estas rutas de señalización se encuentran asociadas entre si por medio de proteínas que funcionan como adaptadoras o andamios (*scaffold proteins*), formándose así los conocidos “transducisomas”. La mayor parte del conocimiento acerca de las proteínas andamio se debe al hallazgo de la proteína cinasa Ste5, una proteína de levadura involucrada en la ruta de la MAPK Fus3 (Burack & Shaw, 2000).

De acuerdo con ciertas observaciones, se han adjudicado dos funciones principales a las proteínas de andamiaje. La primera está relacionada con el mantenimiento de la especificidad de la señal y la segunda propone que las proteínas de andamiaje catalizan la activación de otros componentes de la ruta. Otra perspectiva argumenta que las proteínas andamio funcionan como co-localizadores de un grupo de moléculas que participan en el mismo proceso de señalización en ciertas áreas de las células, aumentando la eficiencia en la propagación de la señal. Se reconoce que las funciones asignadas a las proteínas andamio no son mutuamente excluyentes ya que una misma proteína de andamiaje catalítica puede también co-localizar componentes de la ruta (Burack & Shaw, 2000).

En animales, la importancia de los dominios de interacción es muy claro y su caracterización ha permitido entender mejor la forma en que se realiza la regulación de la información en las cascadas de señalización, es decir, gracias al hallazgo de muchos de estos dominios de anclaje la visión de la transducción de señales pasó de ser un tanto lineal a entenderse como un proceso que requiere de la formación de complejos multiproteicos, muchos de ellos preformados (Pawson *et al.*, 2002 en Zentella, 2003).

Asimismo, la descripción de dominios tales como el dominio SH2, desde la década de los 80's, llevó a la búsqueda de nuevos dominios de interacción tales como PTB, FHA, WD40, WW, EVH1, PH, FYVE, PX, ENTH y Tubby (Burack & Shaw, 2000).

### **1.1.1 Los dominios de interacción proteína-proteína en plantas y la transducción de señales**

En plantas la descripción de dominios de interacción y de proteínas andamio no es tan abundante como en animales; sin embargo, algunos estudios han permitido el reconocimiento de proteínas andamio tales como la proteína NPH3, la cual contiene al menos dos dominios de interacción proteína-proteína, uno de los cuales promueve su interacción con el receptor primario de respuesta al fototropismo NPH1 (Motchoulski, 1999). También se ha descrito el dominio WPP específico de plantas, el cual está implicado en el direccionamiento hacia la envoltura nuclear (Rose & Meier, 2001). Otro ejemplo son las proteínas con motivos de 8 cisteínas (8CM), las cuales comparten un péptido señal y participan en funciones muy diversas que van desde almacenamiento, protección e inhibición enzimática hasta estructura de la pared celular, pero cuya función general se ha determinado como de andamiaje (Estanyol *et al.*, 2004).

La ruta de transducción de señales relacionada con la respuesta a etileno ilustra la importancia de los dominios de interacción en las plantas. El control de la floración y la interacción entre las vías del etileno, las auxinas y el ácido jasmónico están estrechamente relacionados. La ruta de respuesta al etileno incluye receptores con actividad de cinasa los cuales interactúan con el etileno y pasan la señal hasta que uno o varios factores de transcripción son activados (Zhou *et al.*, 2006). Se han identificado diferentes elementos de señalización del etileno. El primer elemento regulatorio caracterizado fue nombrado CTR1 (Constitutive Triple Response 1). Tanto las plántulas como las plantas adultas mutantes *ctr1* presentan una activación constitutiva de la respuesta al etileno, lo cual indica que CTR1 codifica para una proteína que actúa como regulador negativo de la respuesta al etileno. El C-terminal de dicha proteína es similar al dominio catalítico de la familia de proteínas-cinasas Raf, pero el N-terminal despliega poca similitud con Raf o con cualquier otra familia de proteínas registradas en el banco de datos, excepto por algunos otros genes similares a CTR1 (Huang *et al.*, 2003).

Se sabe que los dominios proteicos fusionados a otros tipos de cinasas, como en el caso de CTR1, actúan como reguladores de la función cinasa ya sea por autorregulación y/o por interacción con otros elementos reguladores. Mediante experimentos de doble híbrido se comprobó que el dominio del N-terminal de CTR1 era necesario para la interacción con el receptor de etileno ETR1. Experimentos posteriores de inmunoprecipitación y arrastre (pull down) demostraron que mutaciones en el dominio N-terminal de CTR1 no afectaban su actividad catalítica de ser/thr cinasa, pero que sí afectaban la interacción con el receptor de etileno ETR1 (Huang, 2003). A

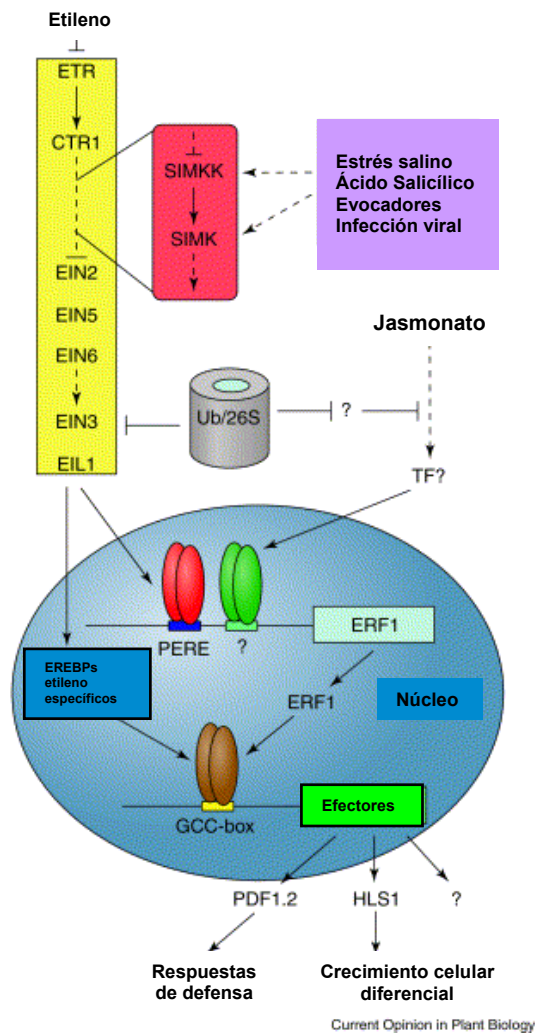
partir de la expresión de CTR1 en *Escherichia coli* se comprobó que esta proteína era capaz de fosforilar mielina básica, un sustrato para numerosas MAPKs (quinasas activadas por mitógeno). Eventualmente CTR1 se identificó como una proteína parecida a las MAPKKK, y a partir de esto se propuso que la ruta de señalización del etileno podría involucrar una cascada de MAPK. A pesar de múltiples intentos por dilucidar esta cascada, la redundancia o letalidad mutacional ha impedido su identificación. Sin embargo, otros estudios han confirmado la participación de una cascada MAPK en la ruta de señalización del etileno; por ejemplo, Chen y colaboradores (2005) mostraron que la expresión de CTR1 en protoplastos de maíz activa una cinasa cinasa activada por mitógeno (MAPKK) aún no identificada (Chen *et al.*, 2005).

Un grupo dedicado a la caracterización de MAPKK, en *Medicago sativa*, involucradas en respuesta a varios tipos de estrés tanto bióticos como abióticos, observaron la activación de ciertas proteínas, en particular la MAPK SIMK (SALT-STRESS-INDUCIBLE MAPK) y *Medicago* MAPK3 (MMK3). Ellos observaron que estas enzimas son activadas en respuesta a estrés salino y a patógenos por la MAPKK SIMKK (SIMK KINASE (SIMKK)). De igual manera, observaron que un precursor del etileno, el ACC (ácido 1-carboxílico 1-aminoaciclopropano), induce la activación de SIMK y de MMK3. Dicha activación es mediada por SIMKK cuya actividad es a su vez estimulada por el ACC. En *Arabidopsis* el etileno induce la actividad de la MAPK6 un ortólogo de *Medicago* SIMK (Ouaked *et al.*, 2003 en Benavente & Alonso, 2006).

La ruta de respuesta a etileno (figura 1) además de ser un ejemplo claro de la importancia de la formación de varios complejos proteicos, también es un ejemplo que nos permite entender cómo diferentes vías de señalización pueden converger o entrecruzarse. En plantas como en otros organismos no existen rutas aisladas que generen un solo tipo de respuesta. Es claro que la convergencia y entrecruzamiento entre diferentes rutas permiten que la planta tenga una o más respuestas en función de uno o más estímulos. Las plantas, comúnmente están sujetas a cambios ambientales muy drásticos, por lo que deben presentar estrategias de adecuación tales como la plasticidad fenotípica, la cual a su vez es un reflejo de la capacidad genotípica que éstas tienen para generar fenotipos variados. Dichas adaptaciones fenotípicas pueden asociarse directamente con una regulación fina durante los procesos de señalización, regulación genética y/o redundancia funcional (Gilroy & Trewavas, 2001).

Durante al menos dos décadas, las aproximaciones genéticas tales como el análisis de mutantes han sido muy exitosas en la determinación de los genes involucrados durante el desarrollo de especies tales como *Arabidopsis*; sin embargo, uno de los retos actuales es encontrar la relación que existe entre los datos genéticos y los mecanismos moleculares que soportan las redes genéticas. El conocer tanto la composición como la estructura de estos complejos podrá en consecuencia ayudar a entender con mayor claridad los procesos celulares relacionados con el desarrollo en plantas.



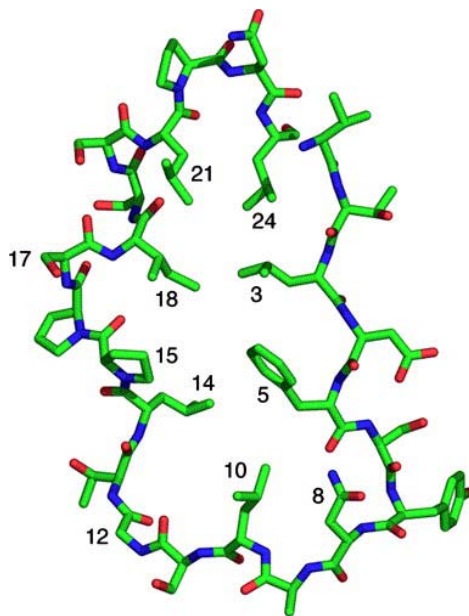


**Figura 1:** Modelo de la ruta de la respuesta a etileno en la regulación de la expresión génica. El etileno es percibido por una familia de receptores asociados al Retículo Endoplásmico (ETR). El etileno unido al receptor inhibe su función. Se ha propuesto que CTR1 inhibe la interacción del receptor con el etileno y que se activa cuando interactúa directamente con este tipo de receptores. La cascada MAPK la cual consiste de SIMKK y SIMK, actúa río abajo de CTR1, sin embargo, los eventos bioquímicos que ocurren como una consecuencia de la activación de la ruta de las MAPK no es clara aún. Debido a que la ruta SIMKK/SIMK se activa en respuesta a muchos estímulos tanto bióticos como abióticos, aún es necesario determinar si su activación es dependiente de las funciones del receptor de etileno y de CTR1. Los componentes reguladores río abajo de la ruta de la respuesta a etileno incluyen varios reguladores positivos: EIN2, EIN5, EIN6 y factores transcripcionales EIN3 y EIL1. El nivel de la proteína EIN3 es controlado positivamente por el etileno vía el proteosoma (Ub/26S). Los componentes primarios de la ruta del etileno (en amarillo) se requieren para todas las respuestas conocidas y no se han relacionado hasta ahora con otra ruta diferente a la del etileno. Algunas ramificaciones de la ruta pueden ocurrir río abajo de EIN3/EIL1. Se sabe que varios factores de transcripción como EREBP (Ethylene Response Element Binding Protein) son blancos inmediatos de EIN3/EIL1 los cuales pueden unirse a elementos de respuesta primaria a etileno (PERE) en los promotores de genes *EREBP*. Un EREBP, llamado ERF1, está también involucrado en la regulación de genes relacionados con el ácido jasmónico, por lo tanto es probable que ERF1 funcione integrando señales tanto del etileno como de ácido jasmónico. Muchas proteínas EREBP regulan la expresión de otros genes a través de la interacción con el elemento en *cis* llamado caja CGG, la cual se puede encontrar en diferentes genes incluyendo *PDF1.2* y *HOOKLESS (HLS1)*. Estos genes codifican para proteínas efectoras que son necesarias para ejecutar una gran variedad de respuestas al etileno; resistencia contra ciertas enfermedades, crecimiento y diferenciación celular (Guo & Ecker, 2004).

## 1.2 Los dominios LRR y el establecimiento de interacciones proteína-proteína

La superfamilia de proteínas con dominios ricos en leucinas (LRR, Leucine-Rich-Repeat, por sus siglas en inglés) es una de las grandes familias multiprotéicas cuyos miembros pueden ser encontrados en diferentes organismos tales como bacterias, protistas, levaduras, animales y en plantas, y cuya función es el establecimiento de interacciones proteína-proteína. Se ha demostrado que los dominios LRR son críticos para la función de estas proteínas. Cada LRR está compuesto de una unidad  $\beta/\alpha$  que forma estructuras elongadas no globulares con forma de herradura. Dicha estructura es la que les permite establecer interacciones durante el reconocimiento de ligandos

específicos y también permite el anclaje con otros componentes en diferentes rutas de señalización (figura 2) (Juge, 2006).



**Figura 2:** Dominio LRR típico de PvPGIP2 de acuerdo con el consenso xxLxLxxNxLt/sGxIPxxLxxLxxL. Posiciones específicas (3, 5, 10, 18, 21 y 24) en el dominio LRR son ocupadas por aminoácidos hidrofóbicos, principalmente leucinas, que apuntan hacia el interior del andamio proteico, estabilizando la estructura a través de interacciones van der Waals (Di Matteo *et al.*, 2006).

Las proteínas LRR llevan a cabo diversas funciones en los procesos de transducción de señales en plantas. Las proteínas RLK's (LRR-receptor-like kinases) y RLP's (receptor-like proteins) constituyen una gran familia de receptores de membrana plasmática con dominios LRR extracelulares y participan en numerosos procesos de desarrollo y defensa contra patógenos. Otro ejemplo de proteínas LRR involucradas en señalización, son las proteínas F-box/LRR las cuales participan en diferentes procesos como el crecimiento del tubo polínico, el desarrollo de la raíz, la activación de la GTPasa Ran, la regulación transcripcional y la organización de células meristemáticas (Belkhadir *et al.*, 2004; Luo *et al.*, 2005; Osakabe *et al.*, 2005; van der Hoorn *et al.*, 2005).

### 1.2.1 Clasificación de las proteínas LRR

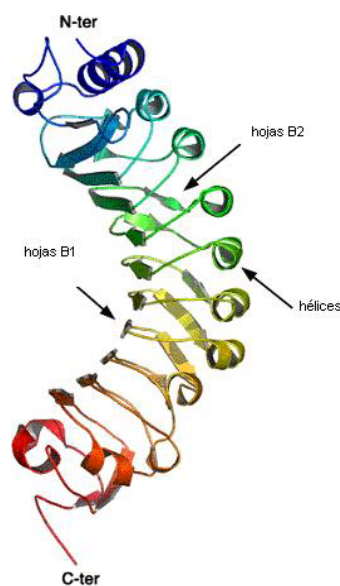
Clasificaciones previas de las proteínas LRR específicas de plantas, las han agrupado en aquellas con motivos parecidos a proteínas involucradas en la defensa contra patógenos con dominios extracelulares y proteínas intracelulares. Un ejemplo de proteínas LRR extracelulares son las PGIP's (Proteínas Inhibidoras de Poligalacturonasas, *Polygalacturonase Inhibitory Proteins*) (figura 3). Al menos 25 genes que codifican para PGIP's se han caracterizado y se han propuesto como genes de resistencia, que funcionan como componentes de la superficie celular formando parte del "sistema inmune de la planta" (Li, 2003). La función principal de las PGIPs es la inactivación de las poligalacturonasas (PG), principalmente de hongos, a través de la interacción proteína-proteína. Existe evidencia directa de la participación de las PGIPs en procesos de defensa en plantas, pues la sobre-expresión de una PGIP tanto en *Nicotina tabacum*, en *Lycopersicon*

*esculentum*, en *Vitia vinifera* y en *Arabidopsis thaliana* limita la colonización por *Botrytis cinerea* y reduce los efectos de la infección (Juge, 2006).

Las proteínas PGIP's no solo regulan procesos relacionados con la resistencia a la invasión por patógenos, sino que también participan en importantes procesos del desarrollo y crecimiento del fruto. Se ha observado una mayor expresión de los transcritos de PGIPs durante la maduración del fruto en especies como la manzana y se ha demostrado, por otro lado, que plantas transgénicas de tabaco y jitomate que sobreexpresan PGIPs de chícharo presentan un mayor crecimiento con respecto a las plantas silvestres (Di Cuixia *et al.*, 2006). Las proteínas PGIPs no sólo son diferentes en términos de la especificidad en el reconocimiento de sus sustratos, sino también en la forma en que se regulan diferencialmente a nivel transcripcional. Se ha reportado una regulación espacio-temporal de las PGIPs durante procesos del desarrollo (tal como maduración floral), y en respuesta a diferentes tipos de estímulos como son infecciones fúngicas, estrés por frío, heridas en el tejido y cambios en la expresión inducidos por moléculas de señalización como son el ácido salicílico y el jasmonato (Li *et al.*, 2003; Ahsan *et al.*, 2005).

Recientemente se ha descrito una nueva clase de proteínas LRR de plantas relacionadas estrechamente con proteínas de hongos y animales que se unen a Ras las cuales han sido denominadas PIRL's (Plant Intracellular Ras group related LRR proteins). Basados en las características estructurales de este nuevo grupo de proteínas se ha propuesto que éstas podrían estar interactuando con otros componentes celulares a través de sus dominios LRR y de esta manera formar parte de los componentes involucrados en las redes de señalización intracelular (Forsthofel *et al.*, 2005). Alternativamente, se ha propuesto la participación de estos dominios en la regulación del transporte núcleo-citoplasma y remodelación de la cromatina (Rose & Maeier, 2001; Askjaer *et al.*, 2002; Kusano *et al.*, 2004).

**Figura 3:** Representación Ribbon de la estructura cristalográfica de PGIP2 de *Phaseolus vulgaris*. Las hojas  $\beta_1$  y  $\beta_2$  se presentan en verde y las hélices  $\alpha$  en azul (azul claro para el N- terminal del  $\alpha$ -hélice y azul oscuro par las  $3_{10}$ -hélices en la porción central LRR de la molécula (Di Matteo *et al.*, 2003; 2006).



## II. Antecedentes

### 2.1 El genoma de *Arabidopsis* y las proteínas DUF

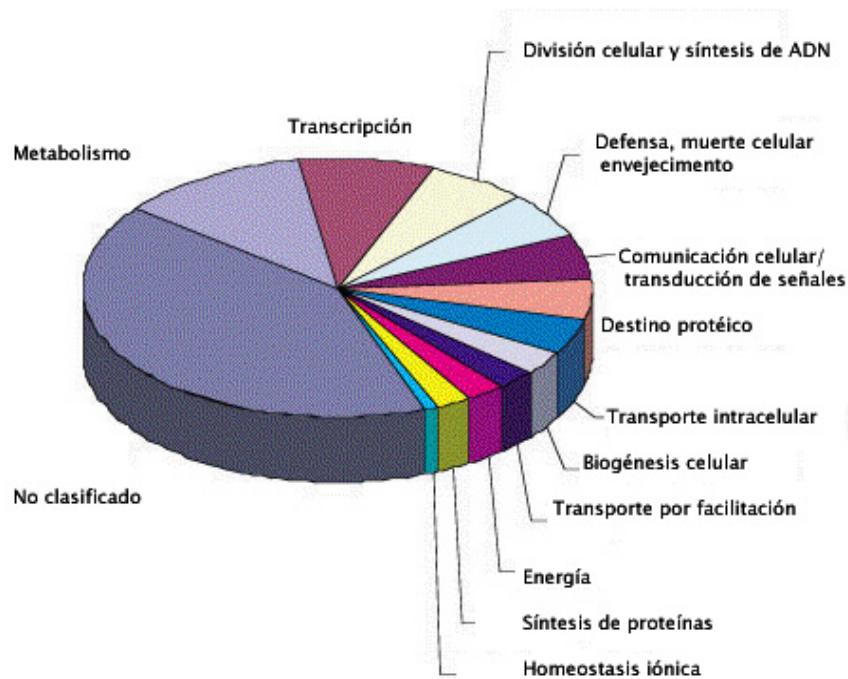
La comunidad de investigación de *A. thaliana* en diciembre del año 2000 (The Arabidopsis Genome Initiative) anunció el término de la secuenciación del genoma de esta especie. En resumen, el genoma de *A. thaliana* está estructurado de la siguiente manera:

- ✓ Contiene cerca de 125 megabases.
- ✓ Codifica para aproximadamente 25,500 genes.
- ✓ Puede ser clasificado bajo especificaciones funcionales genéticas parecidas a las de otros genomas tales como el de *Drosophila melanogaster* y *Caenorhabditis elegans*.
- ✓ El 35% son genes únicos.
- ✓ El 37.5% de los genes pertenecen a familias génicas de 5 ó más miembros.
- ✓ Presenta genes específicos de plantas; aproximadamente 150 familias de proteínas que incluyen 16 familias de factores de transcripción.
- ✓ Contiene varias clases de proteínas que son utilizadas en sistemas animales para procesos no descritos en plantas. Esto implica la utilización de los mismos mecanismos bioquímicos en procesos que podrían ser diferentes.

A pesar de que más de 25,000 genes de *Arabidopsis thaliana* han sido secuenciados, mapeados y sus patrones de expresión han sido descritos para varios de ellos mediante análisis de microarreglos, tanto en diferentes tejidos como bajo diversas condiciones fisiológicas; únicamente se conoce la función para menos del 15% de ellos. Aproximadamente un 30 % de los genes predecidos no han podido ser clasificados, o no se les ha asignado alguna función biológica o bioquímica definitiva (Figura 4) (Ostergaard & Yanofsky, 2004). Además, debido a que en el genoma de las plantas superiores el nivel de complejidad es muy alto y dado que la duplicación genética es un evento común en estos organismos; es evidente que la existencia de familias multigénicas incrementa la probabilidad de redundancia funcional.

Muchas de las grandes familias multigénicas que no han sido caracterizadas se han agrupado por la homología entre sus secuencias. El análisis de las secuencias y el reconocimiento de dominios catalíticos similares u homólogos entre ellas, han permitido hacer propuestas de una función en algunos casos. Sin embargo, existen familias completas agrupadas con base en la homología sus secuencias, cuyos dominios, algunos posiblemente catalíticos, son tan distintos a los previamente descritos en la literatura que no han podido ser clasificados como parte de los grupos ya existentes y, por lo tanto, no ha sido posible asignarles una función. En *Arabidopsis thaliana*, se han identificado por lo menos 1800 familias de proteínas con uno o varios dominios desconocidos

(Domain Unknown Function, DUF) (Lin *et al.*, 1999; Mayer *et al.*, 1999; Salanoubat *et al.*, 2000; Tabata *et al.*, 2000; Theologis *et al.*, 2000).



**Figura 4:** Gráfica que muestra el análisis funcional del genoma de *Arabidopsis*. Proporción de genes predecidos y clasificados en diferentes categorías funcionales (The Arabidopsis Genome Initiative (AGI), 2000).

El análisis de éstas familias de proteínas con dominio de función desconocida (DUF) permitirá entender y expandir el conocimiento en los procesos del desarrollo que se llevan a cabo para el adecuado crecimiento y ciclo de vida de esta especie modelo (Acosta-García, *et al.*, 2004). Actualmente existen diferentes estrategias cuyo objetivo principal es entender las funciones específicas de las proteínas que presentan dominios DUF. Una de las aproximaciones más simple pero que cuya información es muy valiosa, es el análisis de la secuencia primaria con el fin de determinar la naturaleza de los aminoácidos que conforman la proteína o proteínas de estudio. El reconocimiento de motivos específicos permite hacer inferencias sobre su estructura y en ocasiones sobre el papel que desempeñan. La información producida por este tipo de análisis es útil cuando los datos para una determinada familia son escasos o nulos. Este es el caso de la mayor parte de las proteínas DUF.

### 2.1.2 FLOR1, VSP1 y una proteína DUF642 interactúan *in vitro* con un decapeptido del dominio MADS de AGAMOUS

Gamboa de Buen en colaboración con Álvarez Buylla desarrolló una técnica para determinar proteínas candidato a interactuar *in vivo* con AGAMOUS, el factor de transcripción de la familia MADS-box involucrado en el desarrollo de estambres y carpelos en *Arabidopsis*. Utilizando una columna de afinidad con un decapeptido derivado del dominio MADS de AGAMOUS, y a partir de extractos de flores de *Arabidopsis*, se purificaron dos proteínas: FLOR1 y VSP1. Con el mismo método se aisló a partir de flores de maíz una proteína identificada como perteneciente a la familia DUF642. El análisis de la secuencia de aminoácidos de FLOR1 permitió determinar que esta es una proteína con un solo dominio con motivos repetidos ricos en leucinas perteneciente a la superfamilia de LRR's (Leucine Rich Repeat) (Gamboa *et al.*, 2001). *FLOR1* se expresa en el meristemo apical, el meristemo floral y en las estructuras reproductivas, compartiendo patrones de expresión con *AGAMOUS* y otros factores de transcripción de la familia MADS-box. La proteína FLOR1 se localiza tanto en los estambres como en el carpelo en la flor y a nivel celular se encuentra tanto en el núcleo como en el citoplasma. Se confirmó la interacción *in vitro* entre el complejo FLOR1-VSP1 y AGAMOUS mediante una columna de afinidad con la proteína recombinante, análisis de doble híbrido e inmunoprecipitaciones. Los análisis de expresión de *FLOR1* sugieren la interacción *in vivo* entre FLOR1 y AGAMOUS (Acevedo *et al.*, 2004; Páez-Valencia, 2001). FLOR 1 presenta un único dominio el cual está involucrado en el establecimiento de interacciones proteína-proteína, por lo que se sugiere que FLOR1 funciona como una proteína de andamiaje involucrada en diferentes cascadas de señalización intracelular.

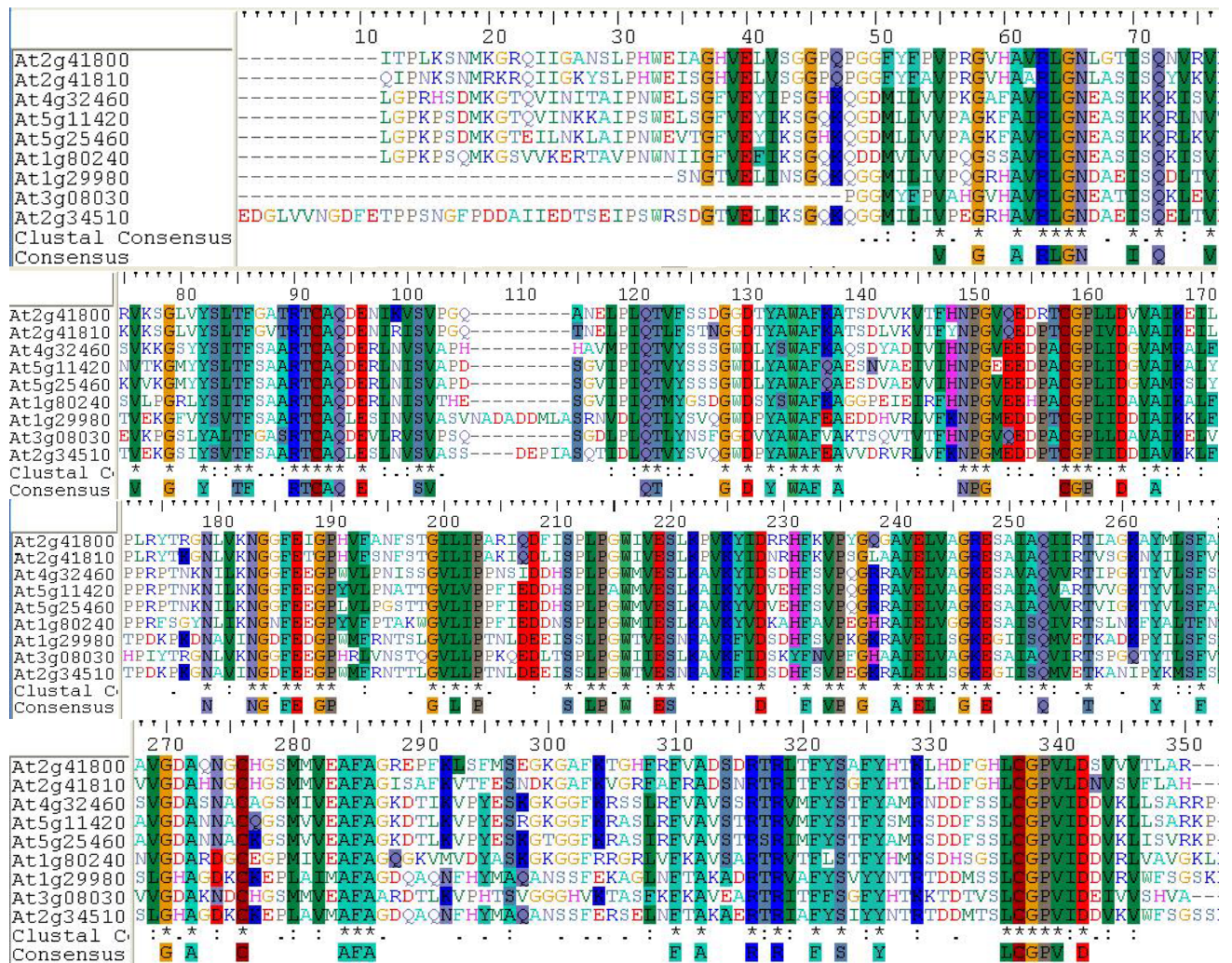
Por otro lado, la proteína DUF642 purificada a partir de extractos crudos de flores femeninas de maíz con un peso molecular de aproximadamente 36 kDa se identificó por medio de un análisis de MALDI-TOF (Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry analysis, por sus siglas en inglés). Se obtuvo un péptido de 19 residuos de aminoácidos: HSQDAEGMLINGNFETAPR. Dicha secuencia fue identificada en el Gen Bank como homóloga a la de la proteína de arroz registrada como OSAAP21410 y a proteínas pertenecientes a la familia DUF642 en *A. thaliana*. Las proteínas DUF642 se caracterizan por la presencia de un único dominio no catalítico y un péptido señal (Figura 5).



**Figura 5:** Representación esquemática del dominio DUF642. Este dominio está altamente conservado en diferentes especies de plantas y está constituido por un péptido señal y un único dominio no catalítico.



Hasta el 2004, se habían descrito en total 21 genes que codifican para proteínas clasificadas como DUF642 pertenecientes a diferentes especies como son: *Oriza sativa* (arroz), *Cicer arietinum* (garbanzo), *Medicago sativa* (alfalfa) y *Ricinus communis* (Zúñiga, 2004). En *A.thaliana* se han reportado 10 genes pertenecientes a la familia DUF642, y a través de la alineación de las secuencias de aminoácidos codificados por estos genes y las proteínas pertenecientes a otras especies, se pudo establecer que estas secuencias estaban muy conservadas. En la figura 6 se muestra la alineación de 9 de las 10 proteínas DUF642 de *A. thaliana*. La proteína codificada por el gen At5g14150 fue la más divergente entre ellas y no se muestra en esta alineación; sin embargo, de acuerdo con la clasificación Pfam (Plant families database) este locus codifica también para una proteína DUF642.



**Figura 6:** Alineación múltiple de las secuencias de aminoácidos correspondientes a 9 de las 10 proteínas registradas como proteínas DUF642 en *Arabidopsis thaliana*. Las barras sombreadas en colores muestran identidad y similitud de aminoácidos (88% de umbral).

La reconstrucción filogenética realizada con los genes de diferentes especies reportados hasta el 2004 como miembros de la familia DUF642 (figura 7), muestra la formación de clados que

agrupan genes tanto de *Arabidopsis* como de *Oriza sativa* (clado B y clado C), así como clados que agrupan genes de *Arabidopsis*, de *Medicago sativa*, de *Ricinus communis* y de *Cicer arietinum* (clado C). Esto supone una mayor similitud de ciertos genes entre diferentes especies que entre genes de la misma especie en algunos casos (Zúñiga, 2004).

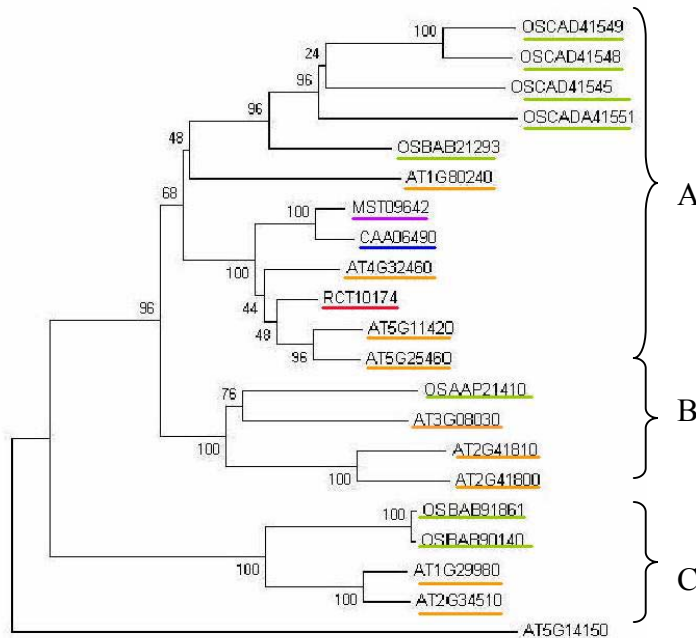


Figura 7: Reconstrucción filogenética mediante análisis de Máxima Parsimonia basada en la secuencia de aminoácidos de las proteínas pertenecientes a la familia DUF642 en diferentes especies de plantas. OS: *Oriza sativa*, MS: *Medicago sativa*, CA: *Cicer arietinum*, RC: *Ricinus comunis*, AT: *Arabidopsis thaliana*. La filogenia muestra 3 clados (A, B y C) agrupando genes pertenecientes a distintas especies (Zúñiga, 2004).

### 2.1.3 Nuevos hallazgos de las proteínas DUF642

A partir del 2002, pero principalmente durante el transcurso de esta investigación, se empezó a determinar la posible función de los elementos de la familia DUF642. El reciente progreso en el campo de la biología molecular (especialmente en microarreglos de DNA), genómica, proteómica y metabolómica ha permitido obtener nuevas pistas sobre las redes de regulación genética y los factores ambientales y señales del desarrollo que regulan la expresión de los genes que conforman estas redes. Existen algunos trabajos de este tipo en los que se incluye a los genes pertenecientes a la familia DUF642. Soeda y colaboradores (2005) han descrito un incremento en los niveles de RNA de los genes At3g08030 y At5g25460 durante la germinación de semillas sometidas a un tratamiento de acondicionamiento o priming. En este mismo trabajo se ubica al gen At4g32460 dentro de un grupo de genes cuya expresión aumenta hasta cuatro veces durante la germinación, lo cual refleja su posible participación durante este proceso, aunque no se descarta su importancia funcional en otros órganos o tejidos ya que Tung y colaboradores (2005) lo describen como una gen cuya proteína puede contribuir a reforzar, modificar o remodelar la estructura de la matriz extracelular durante la interacción polen-pistilo en el proceso de polinización. La expresión de este gen disminuye drásticamente en las células de la papila estigmática de plantas transgénicas cuyas células epidérmicas del estigma y del tracto de



transmisión del estilo fueron específicamente eliminadas a través de la expresión de una toxina celular. Los autores reportan en este mismo artículo a los genes At2g41800 y At5g14150 dentro de este grupo.

Otros resultados han mostrado que el estrés salino, estimula específicamente la expresión del gen At2g41810 en las raíces, según un análisis de expresión mediante microarreglos para 8100 genes de *Arabidopsis* realizado por el grupo de Kreps (2002). Sin embargo, Che y colaboradores (2006), encontraron que la expresión de este mismo gen y su gen hermano At2g41800 es regulada negativamente en la raíz bajo condiciones de cultivo de tejidos. Otro trabajo realizado por el grupo de Osakabe y cols (2005), presenta al gen At5g11420 dentro de un grupo cuya expresión está regulada negativamente en una línea de plantas que sobreexpresan la proteína quimérica RL-BK, construida únicamente a partir de la región N-terminal del receptor con dominios ricos en leucinas RPK1 (por sus siglas es inglés, Receptor-like kinase 1). Esta región está constituida por el dominio transmembranal LRR y el dominio juxtamembranal de este receptor, fusionado a su dominio cinasa (BRI1). Dicho receptor está involucrado en la ruta principal de señalización y percepción temprana de la fitohormona ácido absísico (ABA) en *Arabidopsis*. El ABA regula muchos de los eventos fisiológicos relacionados con la adaptación del tejido vegetativo a estrés abiótico tales como sequía y alta salinidad, así como la maduración de las semillas, dormancia y desarrollo (Osakabe *et al.*, 2005). También se ha descrito que los brasinosteroides (BRs), hormona que promueve la elongación del tallo e inhibe la elongación de la raíz y cuya actividad es muy similar a la de otras hormonas, especialmente a la de las auxinas, regulan negativamente la expresión de At2g41800 y positivamente la expresión de At2g34510 (Goda *et al.*, 2004).

Las proteínas codificadas por At1g29980, At2g34510 y At5g14150 son los únicos miembros de la familia DUF642 de *Arabidopsis* que presentan un péptido señal de unión a membrana [de acuerdo con la página de Internet Plant Membrane Protein database (<http://aramemnon.botanik.uni-koeln.de/>)]. Estos datos *in silico* se confirman con los ensayos proteómicos en los que se aislaron proteínas de membrana de *Arabidopsis* ancladas a glicosil fosfatidil inositol, GPAs (*GPI-anchored proteins*, por sus siglas en inglés). A través de estos ensayos se purificaron estas tres mismas proteínas de la familia DUF642 (Borner, 2003).

Asimismo, de acuerdo con los análisis de expresión por RT-PCRs realizados previamente para el gen At2g41810, se pudo detectar a este gen únicamente en flores. Las plantas homocigotas mutantes por inserción de T-DNA en este gen, no presentaron ninguna variación en el fenotipo con respecto a las silvestres; la redundancia funcional y homología entre las secuencias de las proteínas que conforman esta familia son la posible causa de la obtención de dicho resultado (Zúñiga, 2004).

Por otro lado, el único reporte relacionado con proteínas DUF642 en otras especies es el realizado por Abrahams y colaboradores (1995). En dicho trabajo se describen los patrones de

expresión de tres proteínas que se expresan abundantemente en el tallo de *Medicago sativa*, una de las cuales, la proteína denominada *MsaS1*, corresponde a la única proteína DUF642 de *Medicago sativa* cuya secuencia se puede encontrar en el banco de genes. De acuerdo con los autores, la proteína *MsaS1* codifica para una proteína de 375 amino ácidos cuya secuencia incluye una región hidrofóbica en la región N-terminal, lo que sugiere un procesamiento dentro del sistema de endomembranas. Además, se ha sugerido que esta proteína puede ser exportada de la célula via apoplasto o permanecer dentro de ella dependiendo de la presencia de señales adicionales en la secuencia proteica. *MsaS1* no presenta dominios funcionales característicos de algún grupo específico de enzima, es decir, la identidad y función de esta proteína se reportó como desconocida. El gen *MsaS1* además de expresarse en hipocótilos de 7 días de edad, en botones florales se expresa de manera variable y en las hojas. *MsaS1* presenta una expresión temporal regulada espacio-temporalmente tanto en las hojas como en el tallo; en este último tejido la expresión de *MsaS1* decrece progresivamente de la región de elongación hacia los internodos más viejos. La expresión de *MsaS1* puede en parte estar regulada por sacarosa y ser reprimida por heridas. En conclusión la expresión *MsaS1* está asociada con el proceso de elongación del tallo o la síntesis de la pared primaria (Abrahams *et al.*, 1995).

#### **2.1.4 El Proyecto del Genoma Floral y las proteínas DUF642**

La secuenciación del genoma tanto de *Arabidopsis thaliana* como de *Oriza sativa*, ha promovido grandes avances en la investigación en plantas. Sin embargo, el entendimiento de la evolución de los genes y genomas de las plantas requiere comparaciones que deben partir más allá de los hallazgos hechos tanto en *A.thaliana* como en arroz. Uno de estos grandes proyectos, realizados con la finalidad de cerrar la brecha entre especies modelo como *Arabidopsis* y el resto de las especies vegetales, es el proyecto del genoma floral FGP (*Floral Genome Project*, por sus siglas en inglés). Este proyecto tiene como objetivo principal describir los patrones de conservación y diversificación del transcriptoma floral entre las angiospermas, con la finalidad de elucidar el papel de la duplicación génica y los cambios de los patrones de expresión en el origen y diversificación de las angiospermas. La comparación de los genomas de eudicotiledóneas y monocotiledóneas ha dado muchas pistas acerca de la historia evolutiva de dichos linajes. La realización de análisis filogenéticos con especies adicionales que representan angiospermas basales, eudicotiledoneas y monocotiledóneas basales ha sido valiosa, pues ha permitido generar una visión más amplia de la evolución del genoma en las angiospermas (Albert *et al.*, 2005). El proyecto del genoma floral posee una gran colección de bibliotecas de cDNA de tejidos reproductivos para especies seleccionadas dentro del grupo de las eudicotiledoneas, las angiospermas y las gimnospermas basales. Dichas especies representan los nodos clave en la filogenia de las angiospermas y su grupo hermano, las gimnospermas, por lo que contribuyen a entender la diversidad de estructuras reproductivas y los patrones de desarrollo en este reino. De los 1453 genes de *Arabidopsis* que han sido identificados como genes con expresión órgano-específica en flor (Wellmer *et al.*, 2004), dos

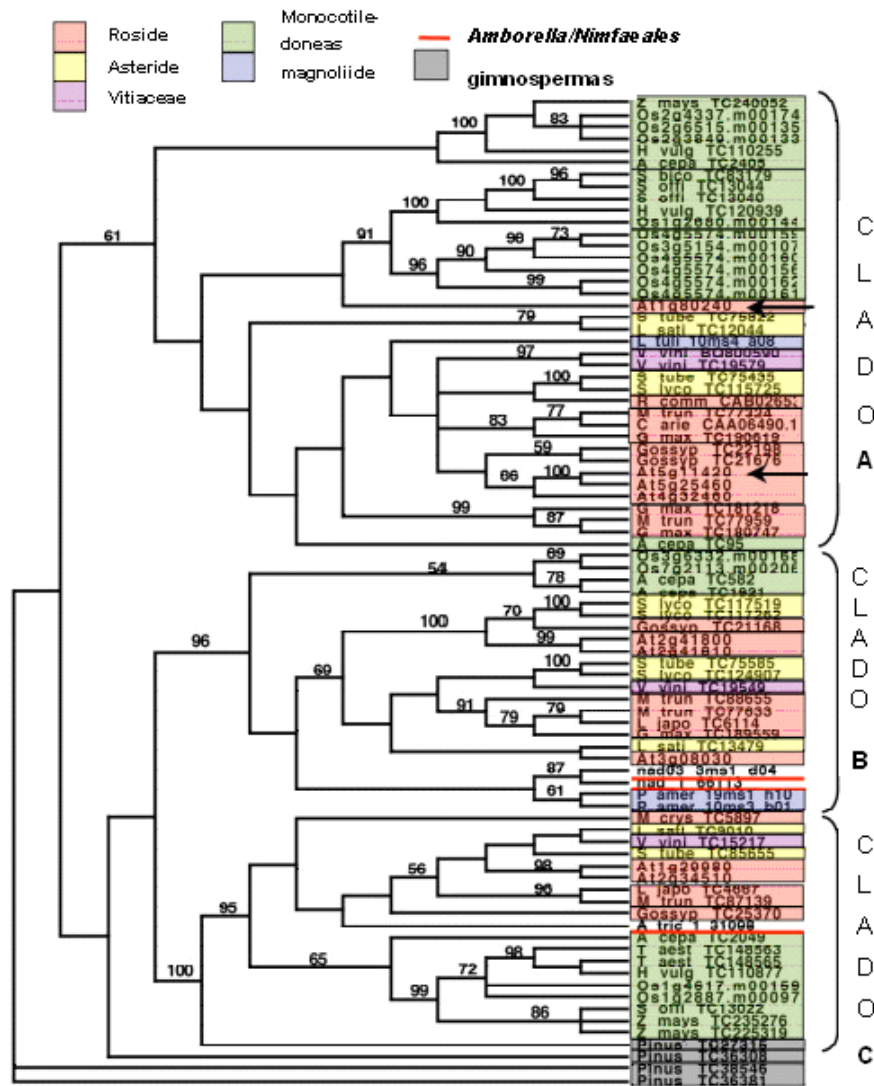
de ellos pertenecen a la familia DUF642 (expresión específica en pétalos) siendo uno de los grupos de genes mejor conservados en los genomas de las especies basales (Tabla 1) tales como: *Asparagus officinalis* (Aof), *Amborella trichopoda* (Atr), *Eschscholzia californica* (Eca), *Liriodendron tulipifera* (Ltu), *Nuphar advena* (Nad) (Albert *et al.*, 2005). También se analizaron los genomas de otras especies [*Acorus americanus* (Aam), *Persea americanus*, (Pam), *Saruma henryi* (She) y *Welwitschia mirabilis* (Wmi)] en las que no se reportaron genes pertenecientes a la familia DUF642.

Tabla 1: Distribución de los genes de *Arabidopsis* órgano-específicos de flor, entre 7 conjuntos de unigenes. En negritas los genes con dominio DUF642 con expresión diferencial en pétalos (Albert *et al.*, 2005).

Organ	Gene ID	Annotation	Tribe	Ath	Osa	Aam	Aof	Atr	Eca	Ltu	Nad	Pam	She	Wmi	Tot
carpel	At5g44635	similar to putative CDC21 protein	549	8	10				1		1	1			3
carpel	At1g71691	hypothetical protein with GDSL-like motif	40	84	79	1			1	1		2	1		6
carpel	At5g53120	spermidine synthase	233	3	3			1		1	3	1			6
carpel	At5g07280	receptor-like protein kinase-like protein	8	194	478	2	1		2					1	7
carpel	At3g51860	Ca <sup>2+</sup> /H <sup>+</sup> -exchanging protein-like	823	6	7		1	1		1	2	2	1		8
carpel	At5g06860	polygalacturonase inhibiting protein (PGIP1)	8	194	478	1	1		1	1	1				8
carpel	At5g02540	putative protein	348	11	14	1		1	1	1	1	1	2	1	9
carpel	At5g59320	nonspecific lipid-transfer protein precursor – like	277	11	18		1	1	1	2	2	2		1	10
petal	At3g62700	glutathione-conjugate transporter, putative	192	16	22	1		1	2	1					5
<b>petal</b>	<b>At5g11420</b>	<b>putative protein (DUF642 Domain)</b>	<b>328</b>	<b>10</b>	<b>16</b>		<b>1</b>	<b>1</b>	<b>2</b>	<b>1</b>	<b>1</b>				<b>6</b>
<b>petal</b>	<b>At1g80240</b>	<b>putative protein (DUF642 Domain)</b>	<b>328</b>	<b>10</b>	<b>16</b>		<b>1</b>	<b>1</b>	<b>2</b>	<b>1</b>	<b>1</b>				<b>6</b>
sepal	At1g69120	floral homeotic gene APETALA1	65	46	51							1	1		3
stamen	At5g14780	formate dehydrogenase (FDH)	4097	1	3		1	1		1	1	3	1		8
stamen	At1g52570	phospholipase D, putative	300	10	17				4	1	1	2	1		9
stamen	At3g09390	metallothionein-like protein	6057	2	0		1		1	2	1	2	2		10
stamen	At3g03080	putative NADP-dependent oxidoreductase	315	14	13	1		2	1	1	1	3			11
stamen	At3g62290	ADP-ribosylation factor-like protein	197	19	17	4	3	2	1	1		2		1	14
stamen	At5g43330	cytosolic malate dehydrogenase	1443	4	4	3	2	1	1	3	1	3			14
stamen	At1g13950	initiation factor 5A-4	1208	3	7		1	3	1	3	2	5			18
stamen	At5g45775	Expressed protein	923	4	7	2		1	7	2	1	3	1	1	18
stamen	At5g14670	ADP-ribosylation factor – like protein	197	19	17	3		2	6	3	3	4		2	24

En la filogenia realizada recientemente a partir de estos grandes proyectos relacionados con el estudio del origen y la diversificación de las angiospermas, se muestran genes con el dominio DUF642 en especies tales como: *Arabidopsis thaliana*, *Gossypium spp.* (algodón), *Glycine max* (soya), *Lotus japonicus*, *Medicago trunculata*, *Vitis vinifera* (uva), *Lactuca sativa* (lechuga), *Solanum tuberosum* (papa), *Lycopersicon esculentum* (jitomate), *Mesembryanthemum crystallinum* (planta hielo), *Allium cepa* (cebolla), *Triticum aestivum* (trigo), *Hordeum vulgare* (cebada), *Sorghum bicolor*, *Saccharum officinarum* (caña de azúcar), *Zea mays* (maíz), *Liriodendron tulipifera*, *Persea americana* (aguacate), *Nuphar advena*, *Amborella trichopoda* (Figura 8). Esta filogenia reveló tres subfamilias, una de ellas con un índice de soporte débil y otras dos con un buen índice de soporte, referidas como clados A, B y C respectivamente. En ésta filogenia se propone al gen de pino como el gen hermano de todos los genes de las angiospermas del clado C indicando que el origen de esta subfamilia precede al ancestro común de angiospermas y gimnospermas. Los genes de Asteride Roside y monocotiledóneas pueden ser identificados en cada una de las tres subfamilias; los genes de las Magnoliide se encontrarían tanto en el clado A como en el B, mientras que las angiospermas más basales (*Amborella* y *Nymphaeales*) estarían representadas tanto en el clado B como en el C (Albert *et al.*, 2005). Los autores sugieren que de

los dos genes que se expresan exclusivamente en pétalos, el At5g11420 y el At1g80240, el primero proviene de una reciente duplicación, sin embargo, su gen hermano no se expresa diferencialmente en pétalos. La determinación del patrón de expresión de los genes DUF642 en *Arabidopsis thaliana*, arroz u otras especies y el registro de esta información en una filogenia podría ser el primer paso para el entendimiento de la función actual y la evolución funcional de esta familia. De acuerdo con Albert y cols, (2005), la familia DUF642 se encuentra entre un grupo de familias poco estudiadas cuya caracterización funcional podría proveer de pistas importantes para entender su historia en relación a la evolución y diversificación de las adaptaciones de las plantas al medio ambiente.



**Figura 8:** La reconstrucción filogenética para los genes de la familia de dominio DUF642 indica que dos genes con expresión diferencial en pétalos no son el producto de un evento de duplicación reciente. Los genes con el dominio DUF642, específico de plantas pero funcionalmente no caracterizado, forman tres clados A, B y C. El clado A además de agrupar los genes con expresión específica de pétalos de *Arabidopsis* incluye genes de Asterídes, Vitis, Magnoliídes y monocotiledóneas, indicando que el clado precede a la divergencia de estos linajes; las flechas indican expresión específica en pétalos. Los genomas de los linajes de angiospermas más basales (*Amborella* y *Nymphaeales*) y gimnospermas pueden también contener genes del clado A no muestreados. Los genes del clado B fueron muestreados de *Nuphar* (*Nymphaeales*) y los genes del clado C fueron identificados de *Amborella* y *Pinus*. Los valores de soporte o bootstrap (>50%) se muestran por arriba de cada rama; At, *Arabidopsis thaliana*; Gossyp, *Gossypium* spp. (algodón); G. max, *Glycine max* (soya); L. japo, *Lotus japonicus*; M. tren, *Medicago truncatula*; V. vini, *Vitis vinifera* (uva); L. sati, *Lactuca sativa* (lechuga); S. tube, *Solanum tuberosum* (papa); S. escu, *Solanum esculentum* (tomate); M. crys, *Mesembryanthemum crystallinum* (planta hielo); A. cepa, *Allium cepa* (cebolla); T. aest, *Triticum aestivum* (trigo); H. vulga, *Hordeum vulgare* (cebada); S. bico, *Sorghum bicolor*, S. offi, *Saccarum officinarum* (caña de azúcar); Z. mays, *Zea mays* (maíz); L. tuli, *Liriodendron tulipifera*; P. amer, *Persea americana* (aguacate); N. adve, *Nuphar advena*; A. tric, *Amborella trichopoda* (Albert et al., 2005).

### **III. Hipótesis**

Las proteínas de la familia DUF642, las cuales son específicas de plantas, presentan un alto grado de similitud tanto intraespecífica como interespecífica y el único dominio que presentan no comparte ninguna característica con los dominios catalíticos descritos, por lo que se considera que esta familia de proteínas podría participar en diferentes procesos de desarrollo vegetal a través de la interacción con otras proteínas.

## IV. Objetivo

### 4.1 Objetivo general

- Determinar la función de las proteínas con dominio DUF642 de *Arabidopsis thaliana*.

### 4.2 Objetivos particulares

1. Determinar la presencia de mRNA de los genes que integran la familia DUF642 en *A. thaliana*: At2g41800, At5g25460, At2g34510, At1g29980, At3g08030, At4g32460, At1g80240, At5g11420 y At5g14150 y analizar sus patrones de expresión en diferentes tejidos.
2. Hacer un análisis de las secuencias promotoras de estos 10 genes con el fin de conocer cuales son las posibles proteínas (factores de transcripción) que estén regulando su expresión.
3. Determinar a partir de columnas de afinidad, la interacción entre proteínas de la familia DUF642 con la proteína AGAMOUS o con otras proteínas MADS-box, o más aún, con proteínas de familias distintas.
4. Realizar la búsqueda de líneas mutantes homocigotas por inserción de T-DNA de los genes de la familia DUF642, preferentemente de aquellas cuyos genes se expresen en flor y/o en inflorescencia. Para esto:
  - 4.1 Establecer la línea mutante homocigota
  - 4.2 Determinar el sitio de inserción del T-DNA
  - 4.3 Analizar mediante RT-PCR si la mutante es nula o no
  - 4.4 Caracterizar la línea mutante homocigota obtenida

### 5.1 Análisis de expresión mediante RT-PCR

Se diseñaron diez oligonucleótidos específicos para la amplificación de cada uno de los genes pertenecientes a la familia DUF642 (Tabla 1, anexo III). Se realizaron extracciones de RNA de diferentes tejidos de plantas silvestres ecotipo Columbia de *A. thaliana* tales como flores, inflorescencias, raíces, hojas caulinas, hojas de roseta, tallos y plántulas. Para ello fue necesario germinar semillas previamente vernalizadas durante 3 días, en placas con medio (MS) bajo condiciones de luz continua y/o fotoperíodo largo (16 h luz y 8 h oscuridad) durante un período de 15 días. La parte aérea de varias plántulas de esta edad fueron colectadas, las raíces también se colectaron, el resto se transplantaron a macetas con tierra. La colecta de los tejidos se realizó una vez que las plantas presentaban el eje principal y varias ramas laterales, aproximadamente 35 días post-germinación. El tejido colectado se congeló inmediatamente en nitrógeno líquido y se guardó a -80 en un ultracongelador para su uso posterior.

Los tejidos colectados de plantas silvestres fueron utilizados para la extracción de RNA total (extracción RNA, anexo III). A partir de RNA se sintetizó cDNA utilizando el oligo dT (Invitrogen) y la enzima SuperScript II retrotranscriptasa (Invitrogen) (síntesis de cDNA, anexo III). Una vez sintetizado el cDNA se realizaron reacciones de PCR para cada uno de los genes pertenecientes a la familia DUF642 de *Arabidopsis*, utilizando para cada uno de ellos el par de oligos específicos (9 genes). Al menos tres experimentos independientes fueron realizados con el fin de determinar con mejor claridad la expresión de los diferentes genes en los diferentes tejidos. En algunas ocasiones fue necesario utilizar otros oligonucleótidos (ver anexo III) para la amplificación del gen como fue el caso del At5g11420 y At2g41800. Los programas de PCR utilizados en las amplificaciones se muestran en la tabla 2 del anexo III. Las bandas amplificadas fueron clonadas y secuenciadas para cada uno de los genes. El análisis de las secuencias obtenidas se realizó utilizando programas de las siguientes páginas: Sequence Manipulation Suit [www.bioinformatics.org/sms/](http://www.bioinformatics.org/sms/), Salk Institute Genome Analysis Laboratory <http://signal.salk.edu/> y National Center for Biotechnology Information [www.ncbi.nlm.nih.gov/BLAST](http://www.ncbi.nlm.nih.gov/BLAST).

### 5.2 Análisis de las secuencias promotoras

En *Arabidopsis* se han registrado diez genes pertenecientes a la familia DUF642. Las secuencias de estos diez genes fueron obtenidas del banco TAIR (The Arabidopsis Information Resource) cuya página de internet es <http://www.arabidopsis.org>. Dicha dirección permite obtener la secuencia genómica de un gen dado con anotaciones tales como: promotores, regiones o secuencias no codificantes (UTRs), así como el inicio y el término de exones e intrones. El análisis de las regiones promotoras anotadas con base en los análisis *in silico* en los bancos de genes se

realizó con el programa SignaScan con la dirección <http://www.gene-regulation.com>. Este programa permite realizar la búsqueda de secuencias específicas encontradas en las regiones promotoras de plantas. Para estos fines se copió la region promotora marcada por el SeqViewer Nucleotide View de la página de TAIR, la cual se encuentra río arriba del sitio del inicio de la transcripción sin tomar en cuenta la región señalada como UTR o región no codificante. Otro programa con la dirección <http://arabidopsis.med.ohio-state.edu/AtcisDB/> (*Arabidopsis Gene Regulation Information Server* por sus siglas en inglés) es también una herramienta útil para la búsqueda de elementos *cis*-regulatorios *in silico* en el genoma de *A. thaliana* ya que permite la búsqueda de motivos reguladores en la región 5' en intrones y regiones 3' en menor grado DE 29,308 genes de *A. thaliana* (Davuluri *et al.*, 2003).

### **5.3 Expresión heteróloga de las proteínas recombinantes At5g11420 y At4g32460 (Proteínas completa y truncada)**

#### **5.3.1 Amplificación y clonación de los genes**

Se realizó la amplificación de los genes At5g11420 y At4g32460 a partir de cDNA de tejido de hojas caulinas y flores de plantas silvestres de *A. thaliana* ecotipo Columbia de 35 días de edad. Los oligonucleótidos diseñados para la expresión de las proteínas completas y truncadas se muestran en la tabla 4 del anexo III. Los oligonucleótidos fueron diseñados con los sitios de restricción para las enzimas NcoI en el oligo sentido y BglII en el oligo antisentido. Dichos oligos se diseñaron con la finalidad de amplificar las secuencias completas e incompletas de los genes At5g11420 y At4g32460 (eliminando aproximadamente 200 pb y 100 pb del extremo 3' respectivamente) ambas sin su péptido señal. Los productos de PCR fueron analizados en un gel de agarosa al 1%, teñido con bromuro de etidio y las bandas correspondientes fueron purificadas y clonadas en el plásmido TOPO PCR II para su secuenciación.

Las secuencias clonadas en el plásmido TOPO PCR II fueron digeridas con las enzimas de restricción NcoI y BglII. Una vez digeridas las bandas fueron purificadas y ligadas en el vector de expresión pQE60. Células competentes de *E. coli* de la cepa DH5 $\alpha$  fueron transformadas con la reacción de ligación (50  $\mu$ l de células + 5  $\mu$ l de reacción de ligación) (Protocolo de transformación, anexo III). Una vez obtenidas las clonas seleccionadas por resistencia, se realizó la búsqueda de plásmidos con el inserto del tamaño adecuado para cada una de las secuencias.

Se obtuvieron los plásmidos con los insertos correspondientes a la secuencia completa y truncada del gen At5g11420 y del gen At4g32460. Dichos plásmidos fueron purificados mediante el kit QIAprep Spin Miniprep de QIAGEN y secuenciados para confirmar que las secuencias codificantes estuvieran en el marco de lectura correcto.



### 5.3.2 Expresión y purificación de proteínas

Una vez purificados los plásmidos con los insertos en el marco de lectura correcto, se utilizaron células competentes *E. coli* M15 para la transformación con el plásmido puro pQE60-At5g11420 (secuencia completa y truncada) y pQE60-At4g32460 (secuencia completa y truncada) (figura 9 y 10). Las células transformantes se crecieron toda la noche a 37° C con agitación en 6 mL de medio LB que contenía ampicilina (100 µg/ml) y kanamicina (25 µg/mL).

mkggslsflfvlliatitsvMGFSDGMLPNGDFELGPKPSDMKGTQVINKKAIPSWELSGFVEYIKSGQKQGD  
MLLVVPAGKFAIRLGNEASIKQRLNVTKGMYYSLTFSARTCAQDERLNISVAPDSGVIPIQTVYSSSG  
WDLYAWAFQAESNVAEIVIHNPGEEDPACGPLIDGVAIKALYPPRPTNKNILKNGGFEEGPYVLPN  
ATTGVLVPPFIEDDHSPLPAWMVESLKAIKYVDVEHFSVPQGRRAVELVAGKESAIQVARTVVGKT  
YVLSFAVGDANNACQGS MVVEAFAGKDTLKVPYESRGKGGFKRASLRFVAVSTRTRVMFYSTFYSM  
RSDDFSSLCGPVIDDVKLLSARKPRSHHHHH\*

**Figura 9:** Secuencia de aminoácidos de la proteína At5g11420. En café se muestra el péptido señal no presente en las proteínas expresadas. En verde oscuro una metionina sintética. En rosa se muestran los aminoácidos que fueron eliminados en la secuencia truncada. En rojo se muestra la colita de HISTIDINAS y los aminoácidos mutados.

mkemgviivllllhsffiyafMGFNDGLLPNGDFELGPRHSDMKGTQVINITAIPNWELSGFVEYIPSGHKQ  
GDMILVVPKGAFAVRLGNEASIKQKISVKKGSYYSITFSARTCAQDERLNVSVAPHHAVMPIQTVY  
SSSGWDLYSWAFKAQSDYADIVIHNPGEEDPACGPLIDGVAMRALFPPRPTNKNILKNGGFEEGP  
WVLPNISSGVLIPPNSIDDHSPLPGWMVESLKAVKYIDSDFSVQGRRAVELVAGKESAVAQVVR  
TIPGKTYVLSFSVGDASNACAGSMIVEAFAGKDTIKVPYESKGGKGGFKRSSLRFVAVSSRTRVMFY  
TFYAMRNDDFSSLCGPVIDDVKLLSARRPHHHHH\*

**Figura 10:** Secuencia de aminoácidos de la proteínas At432460. En café se muestra el péptido señal no presente en las proteínas expresadas. En verde oscuro una metionina sintética. En azul claro los aa modificados. En rosa se muestran los a.a que fueron eliminados en la secuencia truncada. En rojo se muestra la colita de HISTIDINAS y los aa mutados.

Cada uno de los precultivos se utilizó para inocular 200 mL de medio LB (10 g/L de triptona peptona, 5 g/L extracto de levadura 10 g/L NaCl) con antibiótico (ampicilina 100 µg/mL y kanamicina 25 µg/mL). Los cultivos se incubaron a 37° C con agitación vigorosa hasta alcanzar una OD600 que podía ir de 0.5 a 0.7. Se tomó una muestra de cada cultivo y se marcaron como control no inducido. Una vez alcanzada la absorbancia se indujo la expresión de las proteínas con 2 mL de IPTG (concentración final 1 mM).

Los cultivos con IPTG se incubaron durante 4 horas adicionales en agitación a 37°C. Se tomó una muestra de cada cultivo y se marcaron como control inducido. Dichos cultivos fueron centrifugados durante 25 min. a 4000 x g (aproximadamente 6000 rpm). El pellet de cada uno de los cultivos fue resuspendido en 5 mL de buffer de lisis para purificación de la proteína bajo condiciones nativas (50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8; 300 mM NaCl; 10 mM de Imidazol). Para romper las células se agregaron 500 µL de lisozima a una concentración inicial de 10 mg/mL para obtener una concentración final de 1 mg/mL. Esta resuspensión se incubó en hielo durante 30 a 40

minutos. Transcurrido este periodo se observó la viscosidad del cultivo y en caso de no haberse roto las células se sometió el cultivo a cambios de temperatura de -80 a 4°C. El cultivo lisado fue centrifugado a 14,000 x g a 4° C durante 30 minutos. El sobrenadante fue decantado (extracto crudo con proteína soluble) y almacenado. El pellet fue resuspendido en 5 mL de buffer de lisis (extracto crudo, proteína no soluble).

Las proteínas recombinantes fueron purificadas bajo condiciones nativas a través de columnas de cromatografía de afinidad Ni-NTA (matriz de Ni-ácido Nitrilotriacético acoplado a sefarosa) (Hochuli, 1989). Dichas columnas se prepararon con 5 mL de extracto crudo de proteína soluble y 1 mL de agarosa Ni-NTA previamente equilibrada. Las columnas se agitaron durante una hora a 4° C y posteriormente el extracto crudo de proteína soluble se dejó eluir, la fracción fue colectada como control (fracción no unida).

A continuación las columnas fueron lavadas con amortiguador de lavado 2x (50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8, 300 mM de NaCl y 20 mM de imidazol) esta fracción también fue colectada. Finalmente, la proteína recombinante fue eluída con 0.5 mL de amortiguador de elusión 4x (50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8.0, 300 mM NaCl, 250 mM de imidazol).

### **5.3.3 Análisis de las fracciones por electroforesis y western blots**

Se realizaron geles de poliacrilamida al 12 % para separar las fracciones obtenidas, tanto de las muestras control como de las proteínas purificadas (figuras en el anexo II). La mitad de la fracción eluída fue utilizada en una prueba para confirmar la colita de histidinas utilizando un conjugado que reconoce las histidinas mediante Ni-NTA acoplado a fosfatasa alcalina de intestino de ternera. Dicho conjugado puede ser utilizado para la detección de cualquier proteína recombinante con una colita de histidinas (Western Blot anexo III). La prueba dio positiva para todas las proteínas excepto para la proteína At4g32460 truncada. Este hecho se consideró como una consecuencia de su baja expresión.

### **5.3.4 Cromatografías de afinidad**

Plantas silvestres ecotipo Columbia se crecieron bajo condiciones controladas de luz y temperatura (luz continua a 20° C). Las flores e inflorescencias de dichas plantas fueron colectadas y congeladas en nitrógeno líquido y almacenadas a -80 ° C para su posterior utilización. Aproximadamente 7 g de tejido floral fue molido en mortero. El polvo obtenido se disolvió en 14 mL de amortiguador de extracción (50mM Tris-HCl y 10 mM de Mg Cl<sub>2</sub>, pH 7.5) y se centrifugó durante 30 minutos a 6000 rpm.

El sobrenadante de éste extracto crudo se colocó en una columna de Sefacel (5 x 10 cm.) la cual había sido previamente equilibrada con amortiguador de extracción a 4° C. La fracción resultante fue utilizada en la columna de afinidad.

Las proteínas recombinantes At5g11420 completa y truncada (extracto de proteína soluble) fueron inmovilizadas en columnas de agarosa Ni-NTA y equilibrada con 20 mM de Tris-HCl, pH 7.5 y 5 mM de MgCl<sub>2</sub>. La fracción obtenida a partir de la columna de Sefacel, es decir, el extracto de flores, fue vertido en las columnas y mantenidas en agitación durante 1 hora a temperatura ambiente (aproximadamente 25 ° C) en una proporción de 10 mL de extracto crudo / 0.2 mL de resina.

Posteriormente, se dejó pasar el extracto crudo y las columnas de Ni-NTA fueron lavadas con 50 volúmenes del mismo amortiguador de extracción para remover las proteínas no unidas. La(s) proteína(s) unida(s) fue removida con diferentes concentraciones de NaCl 100 mM, 200 mM y 500 mM. Se realizaron de dos a tres replicas por experimento.

Las diferentes fracciones de proteína fueron precipitadas con acetona fría (1:1) y corridas en un gel de poliacrilamida al 12%. Las proteínas obtenidas fueron identificadas mediante espectrofotometría de masas LC-MS/MS en el Centre Proteomique de l'Est du Québec. También se realizó la cromatografía de afinidad utilizando 10 g de hojas de roseta tanto para la proteína completa como para la proteína trunca de At5g11420.

Este mismo procedimiento se realizó utilizando como ligando inmovilizado a la proteína recombinante completa At4g32460 con extracto crudo de flores y con extracto crudo de hojas de roseta.

#### **5.4 Obtención de plantas mutantes por inserción de T-DNA en el locus At4g32460**

En el banco de genes del SALK INSTITUTE se buscaron posibles mutantes por inserción de T-DNA en genes de la familia DUF642. Se encontraron líneas mutantes para los genes At2g41800, At5g11420 y At4g32460, que correspondían a las líneas registradas como SALK\_138024, SALK\_038436 y SALK\_142260 respectivamente. Estas semillas (aproximadamente 30) fueron germinadas bajo condiciones de luz continua a 20° C en placas con medio (MS) y a los 15 días las plántulas fueron trasplantadas en macetas con tierra. Se colectaron y congelaron para el análisis hojas de roseta de plantas maduras tanto mutantes, (de 10 a 17 individuos por línea), como silvestres (2 individuos) y se les extrajo DNA genómico (Extracción de DNA anexo III).

El programa utilizado para la amplificación del DNA fue el siguiente: temperatura inicial 94° C 5'. 35 ciclos de 94° C 30", 60 ° C 45" y 72° C 1' 30". Finalmente, las reacciones se sometieron a una temperatura de 72° C durante 5'. Para identificar las plantas con la inserción se utilizó el oligo Lbb1 con la secuencia 5'GCGTGGACCGCTTGCTGCAACT3' que se ancla en el borde izquierdo del inserto de T-DNA y oligonucleótidos específicos MET32460 y FIN32460 en el caso del gen At4g32460 (las bandas de la amplificación del T-DNA en plantas de la generación T1 se puede observar en el figura 1, del anexo III).Las regiones amplificadas se clonaron en el plásmido Topo PCR II y fueron enviadas secuenciar para localizar el sitio de inserción del T-DNA en el gen.

### 6.1 Patrones de expresión mediante RT-PCR para los genes pertenecientes a la familia DUF642 de *Arabidopsis thaliana*

El análisis mediante RT-PCR es un método específico, sensible y útil para la detección de transcritos raros o para el análisis de muestras de cantidad limitada (Marone, 2001). Los patrones de expresión para cada uno de los genes pertenecientes a la familia DUF642 se pueden observar en la figura 11 y en resumen en la tabla 2. En la figura 11-A se presenta la expresión para el gen At2g41800, el cual tiene un 56.49% de identidad y un 85.14% de similitud con la proteína codificada por el gen At3g08030, y un 79.189% de identidad y un 83.78% de similitud con la proteína codificada por el locus At1g41810. Como se puede observar este gen se expresa en todos los tejidos y es abundante tanto en las inflorescencias como en las raíces, sin embargo, en inflorescencias la secuencia amplificada es aproximadamente 100 pb más grande que la del resto de los tejidos. Las bandas secuenciadas están señaladas con un \*. La secuencia de la banda de inflorescencia presenta el primero de los dos intrones encontrados en la secuencia genómica pero la banda amplificada a partir de cDNA de raíz no presenta ninguno de los intrones. La secuencia del gen At2g41800 con el primer intrón no codifica para ninguna proteína en ninguno de los tres marcos de lectura posibles.

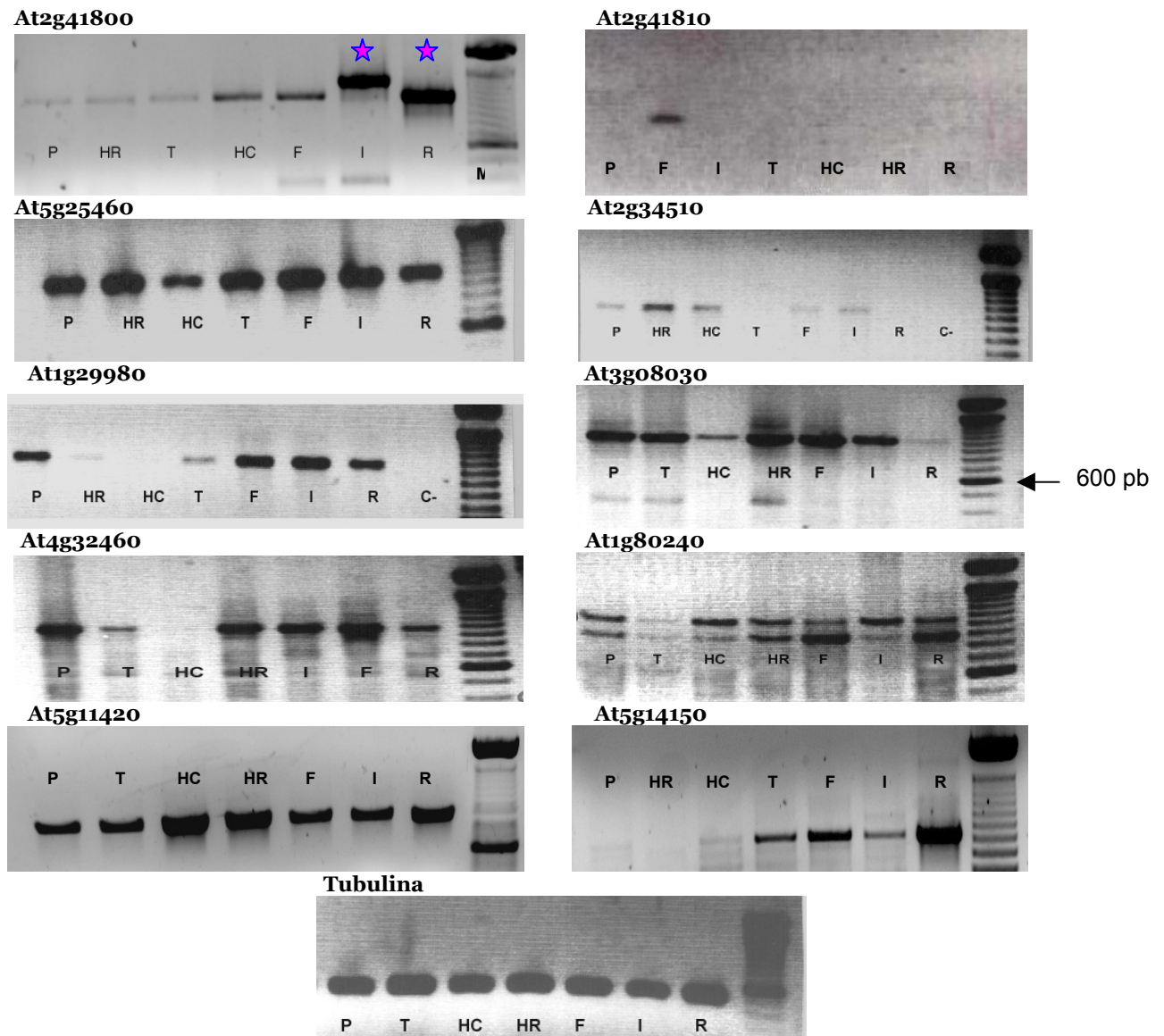
Sólo el gen At2g41810 es específico de un solo tejido (figura 11) cuya expresión ya había sido analizada previamente. Dicho gen se expresa únicamente en flores, pero en otras ocasiones se detectó en cDNA de inflorescencias. No se descarta la posibilidad de que se exprese en las raíces como se ha descrito en los distintos análisis de transcriptoma.

La expresión del gen At5g25460 prácticamente no varía entre un tejido y otro (figura 11) aunque es posible que la variación en los niveles de la expresión de este gen en los distintos tejidos no se haya detectado mediante esta técnica y un análisis de PCR cuantitativo podría ser más adecuado.

Por otro lado, la expresión del gen At2g34510 se detectó en la mayoría de los tejidos con excepción de la raíz, y el nivel de expresión en tallos es muy bajo (figura 11).

El gen At1g29980 se expresó en todos los tejidos analizados con excepción de las hojas caulinas. Su expresión en hojas de rosetas es muy baja.

De acuerdo con la figura el gen At3g08030 se expresa abundantemente en todos los tejidos aunque en la raíz la detección fue mucho menor. El gen At4g32460 también se expresa en todos los tejidos y aunque en la figura no se observa su expresión en las hojas caulinas fue posible visualizar niveles muy bajos de expresión en este tejido en otras repeticiones sin poder plasmarse en la fotografía.



**Figura 11:** Análisis de la expresión por RT-PCR de los genes pertenecientes a la familia DUF642 en *Arabidopsis thaliana* en los siguientes tejidos: P: plántulas, T: tallos, HC: Hojas Caulinas, HR: Hojas de Roseta, F: Flores, I: Inflorescencias, R: Raíces, C-: Control negativo M: marcador de peso molecular 100 pb. El análisis de expresión se realizó en tres experimentos independientes. Diferentes programas de PCR fueron utilizados para la amplificación de las bandas en cada caso. El consenso de los patrones de expresión obtenidos a partir de tres experimentos independientes se muestra en la tabla 2.

En el caso del análisis de expresión del gen At1g80240 se observan dos bandas una de 900 pb que corresponde a la amplificación del cDNA y otra banda de aproximadamente 1200 pb que corresponde al gen con los dos intrones encontrados en la secuencia genómica. Es probable que se haya amplificado la banda superior de aproximadamente 1200 pb a partir de DNA genómico, debido a que el RNA no fue tratado con DNAsas; sin embargo, en reacciones de PCR para otros

genes con el mismo cDNA no se amplificaron bandas de DNA genómico. No se descarta la posibilidad de que este gen también se exprese como mRNA con sus dos intrones al igual que en el caso del gen At2g41800 que se expresa con el primer intrón. El gen At5g11420 se expresa en todos los tejidos y el nivel de expresión casi no cambia entre un tejido y otro. Finalmente el gen At5g14150 se detectó en todos los tejidos pero con niveles mucho mayores en tallos, flores, inflorescencias y raíces.

**Tabla 2:** Resumen de la expresión de los 10 genes pertenecientes a la familia DUF642 de *A. thaliana* en diferentes tejidos a partir de 3 experimentos independientes. +++: Expresión clara en ese tejido; +: baja expresión; nd: no detectado en experimentos independientes con respecto a lo reportado en la literatura.

Tejido Locus	Plántulas	Tallos	Hojas caulinas	Hojas de roseta	Flores	Inflorescencias	Raíces
At2g41800	+++	+	+++	+++	+++	+++	+++
At2g41810					+++	+++	nd
At5g25460	+++	+++	+++	+++	+++	+++	+++
At2g34510	+++	++	+	+++	++	++	+
At1g29980	+++	+	nd	+	+++	+++	+++
At3g08030	+++	+++	+++	+++	+++	+++	+++
At4g32460	+++	++	+	+++	+++	+++	+++
At1g80240	+++	+++	+++	+++	+++	+++	+++
At5g11420	+++	+++	+++	+++	+++	+++	+++
At5g14150	+	++	+	+	+++	+++	+++

## 6.2 Sitios de regulación *cis* en las secuencias promotoras de los genes pertenecientes a la familia DUF642

Las regiones promotoras son regiones localizadas usualmente río arriba del sitio de inicio de la transcripción y se caracterizan por la presencia de motivos *cis* reconocidos específicamente por factores de transcripción (elementos *trans*), los cuales pueden regular los niveles de expresión positiva o negativamente. El control transcripcional es un paso clave para la regulación de los procesos biológicos determinantes para un adecuado desarrollo. Dado que la expresión de los genes que conforman la familia DUF642 de *A. thaliana* presentaron un patrón de expresión diferencial en los tejidos, siendo para algunos de ellos más específica, se realizó un análisis de las secuencias promotoras de los 10 miembros de esta familia, con la finalidad de conocer posibles reguladores de la expresión de plantas. Este tipo de análisis es interesante porque proporciona pistas sobre cuales son las señales que determinan que un gen presente un patrón de expresión determinado, sin embargo, ninguno de estos análisis provee datos concluyentes, y es necesario llevar a cabo experimentos mucho más detallados con la finalidad de comprobar que la información obtenida a partir de dichos programas ocurra *in vivo*. En la tabla 3 se resumen algunos de los motivos encontrados en las regiones promotoras y el número de motivos para cada gen, las características de los sitios reconocidos se presentan de manera más amplia en la tabla I del anexo II.

Como se puede observar en el tabla 3 y de acuerdo con el análisis que se realizó para cada uno de los promotores de los genes pertenecientes a la familia DUF642, existen secuencias en *cis*, reconocidas por factores de transcripción o elementos en *trans*, encontradas únicamente en uno o varios genes ubicados en el mismo clado. Es el caso de los motivos CARG (o caja CARG) los cuales fueron encontrados únicamente en la secuencia promotora del gen At2g41800; o la presencia de un motivo LTRE y un motivo ABF encontrados en el promotor del gen At2g41810 y el gen At2g41800, respectivamente. Otros ejemplos son los motivos Z-box, localizados en los promotores de los genes pertenecientes al mismo clado, el At2g41800 y el At3g08030. Podemos mencionar también la presencia de motivos reconocidos por factores de transcripción MYB tanto en el gen At2g41800 como en el gen At5g14150.

De igual manera es posible aseverar, en términos generales, que el gen At5g14150 es también el más divergente en cuanto a sitios de regulación en *cis*, ya que este gen es el único que no presenta las secuencias de regulación encontradas en la mayoría de los genes pertenecientes a esta familia. Estos motivo de regulación son: secuencias de reconocimiento por el factor de transcripción MADS-box LFY (encontrado en las secuencias promotoras de 8 de los 10 genes DUF642), el motivo GATA, el motivo T-box, el motivo de unión a ARFs, el motivo de unión al factor de transcripción MYB4 y el sitio de reconocimiento por RAV-1.

Otro detalle interesante es el hallazgo de varias secuencias de reconocimiento relacionadas con la regulación por ABA en los genes pertenecientes al mismo grupo: At5g11420, At4g32460 y At5g25460. Hay que recordar que tanto el gen At4g32460 como el At5g25460 han sido descritos como genes cuyos niveles de expresión se incrementan en semillas sometidas a un tratamiento previo de osmocondicionamiento (Tung *et al.*, 2005), y que estos genes también presentan secuencias de reconocimiento por factores de transcripción bZIP relacionados con la expresión específica en el embrión.

**Tabla 3:** Resumen del análisis realizado con las secuencias promotoras de los genes que codifican para proteínas pertenecientes a la familia DUF642 en *A. thaliana*, los números representan la cantidad de veces que se encontró ese motivo en el promotor, los detalles sobre los motivos se describen en el anexo II.

Motivos reconocidos/loci	41800	41810	08030	80240	32460	11420	25460	29980	34510	14150
CArG	2									
Bellringer/re IN AG.	4	2			2		2	1		1
Homeobox										
Sitio consenso reconocido por LFY	6	1		3	3	4	3	1	10	
ATB2/AtbZIP53/AtbZIP44			1			3	2		4	
/GBF5 BS in ProDH										
SORLIP1			1	3			1	1		
SORLIP2	1	2	1		3	1	2	2		
SORLIP3						2				
SORLIP4									1	
SORLIP5							1	1	1	1
lbox	1	4	1		1	2	2			
Z-box	2		1							
T-box	4	5	1	3	5	1	2		2	
BoxII	6			3	2	1				
GATA	11	8	8	2	9	9	8	6	5	
CCA1	1	1	2				1	1	2	
Opaque	1			1		2	2			
RY-repeat			2		4	2				
L1-box	1				1	1	1			
G-box	2					4	2			
GBF1/2/3 BS in ADH1	2					2	2			
ABFs	1									
CBF2 AP2-EREBP	2					2	2			
ABRE	1	1								
ABRE-like	5		1			2				
GCC-box										
LTRE		1								
AtMYC2 BS en RD22	4	5	3	1	5	5	2	2		
AtMYB2 BS in RD22										
DPBF1&2	5	6	3		6	4	5	1		
bZip										
BBF1	2		6	2	3	4		5	2	1
ARF1	1		1	1	2	3	1	1	10	
MYB	2									3
MYB4	3	4	2	2	6	2	5	2		
DRE-like	1			1		2				
RAV1-A	8	9	9		7	9	8	5	9	
ABI3VP1										
W-box	2	4	1		5	5	2	5	6	1
PBF					2	1	1	1	1	



## **6.3 Interacciones proteína-proteína**

### **6.3.1 Interacción de las proteínas At5g11420 completa y truncada utilizadas como ligandos inmovilizados y extractos crudos de flores de *A. thaliana*.**

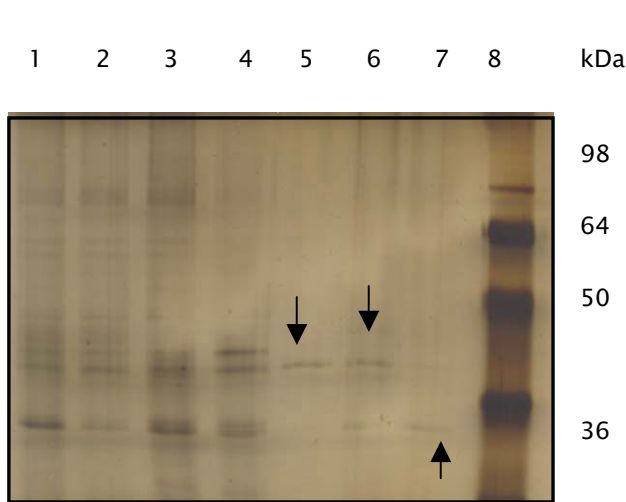
#### **6.3.1.1 Interacción de la proteína completa At5g11420**

Otra de las estrategias para determinar el papel que desempeñan estas proteínas fue la búsqueda de posibles proteínas interactoras mediante ensayos de cromatografía de afinidad. Para este estudio se consideraron dos genes de dos clados diferentes. Los genes At2g41800 y At2g41810 cuya expresión es más específica, y los genes de otro clado el At5g11420 y el At4g32460 que presentan una expresión generalizada. Considerando que también el análisis de las regiones promotoras de estos cuatro genes mostró diferencias en cuanto a los motivos de regulación de la expresión, es posible que la expresión diferencial de estos cuatro genes en los tejidos analizados sea un reflejo de la regulación a nivel transcripcional y esto a su vez este relacionado con las funciones específicas que llevan a cabo estos genes en un tejido u otro. Por otro lado se cuenta con las líneas mutantes de SALK para los cuatro genes. Sin embargo, dado dadas las dificultades en el laboratorio al obtener las proteínas recombinantes de los genes At2g41800 y At2g41810 se realizaron los ensayos de cromatografía de afinidad para los genes At5g11420 y At4g32460. Actualmente, se está trabajando en el laboratorio en la obtención de las proteínas recombinantes At2g41800 y At2g41810.

También, se realizó la purificación de las proteínas completa y truncada de At5g11420 y de la completa de At4g32460. Las fracciones obtenidas a partir de los procesos de purificación fueron analizadas en geles de poliacrilamida al 12% y teñidos con azul de Coomassie (figuras 5-A, 6-A y 7-A del anexo II). También se realizaron western blots, con el fin de determinar si las proteínas puras correspondían a las proteínas deseadas (figuras 5-B, 6-B y 7-B del anexo II). Un primer análisis de interacción se realizó utilizando un extracto crudo de flores de *A. thaliana* con la proteína At5g11420 como ligando inmovilizado en la agarosa Ni-NTA.

En la figura 12 se muestran las diferentes fracciones obtenidas a partir de la columna de afinidad con la proteína completa At5g11420 y un extracto de flores. En los carriles 5 y 6 se pueden observar dos bandas de aproximadamente 38 kDa las cuales interaccionaron específicamente con la proteína recombinante At5g11420 además de una proteína de aproximadamente 34 kDa en el carril 7. Las proteínas señaladas con flechas fueron enviadas para su identificación mediante espectrofotometría de masas LC-MS/MS. Las bandas que se observan en los carriles 5 (100 mM de NaCl) y 6 (200 mM de NaCl) fueron perfectamente identificadas como la proteína LRR FLOR1, mientras que la proteína de aproximadamente 34 kDa presente en la fracción eluída con una

concentración de 500 mM de NaCl (carril 7) fue identificada como la proteína VSP1 (Vegetative Storage Protein 1).

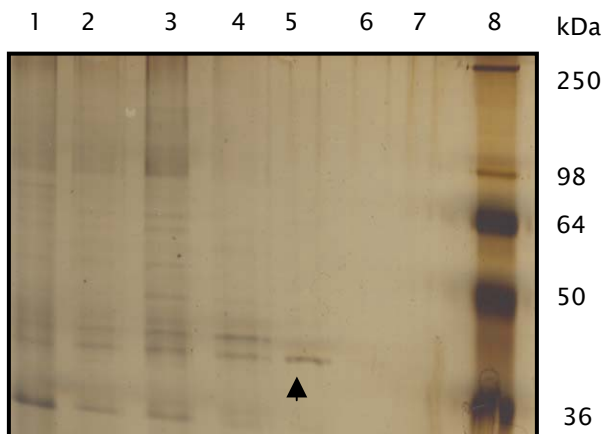


**Figura 12:** Gel de poliacrilamida al 12% (tinción con plata) mostrando las distintas fracciones obtenidas a partir de un ensayo de cromatografía de afinidad con un extracto crudo de flores y la proteína At5g11420 completa como ligando inmovilizado. Se realizaron tres repeticiones independientes de este ensayo

- 1) Extracto crudo de flores
- 2) Extracto de flores después de la columna de DEAE-sefacel
- 3) Fracción no unida
- 4) Fracción de lavado
- 5) Elusión con 100 mM de NaCl
- 6) Elusión con 200 mM de NaCl
- 7) Elusión con 500 mM de NaCl
- 8) Marcador de peso molecular.

### 6.3.1.2 Interacción de la proteína truncada At5g11420

Se realizó un análisis de interacciones utilizando extractos crudos de flores y la proteína trunca At5g11420. En la figura 13 se muestra un gel de poliacrilamida al 12% con las diferentes fracciones obtenidas. Como se observa en el carril 5 únicamente se evidenció una proteína de aproximadamente 40 kDa. La otra proteína de menor peso molecular, obtenida a partir de la utilización de la proteína completa como ligando inmovilizado, no se observó en ninguno de los dos carriles correspondientes a las fracciones eluidas con 200 y 500 mM de NaCl. La banda señalada en el carril 5 se identificó por LC-MS/MS y el análisis mostró que la proteína pegada corresponde a FLOR1. La proteína VSP1 no se purificó en ninguna de las fracciones de elusión, por lo que se sugiere que VSP1 no interactúa *in vitro* con la proteína truncada en el carboxilo terminal At5g11420.



**Figura 13:** Gel de poliacrilamida al 12% (tinción con plata) mostrando las fracciones obtenidas a partir de una columna de afinidad con extractos crudos de flores y la proteína truncada At5g11420. Se realizó una réplica más de este mismo ensayo, obteniéndose los mismos resultados.

- 1) Extracto crudo de flores
- 2) Extracto de flores después de la columna de DEAE-sefacel
- 3) Fracción no unida
- 4) Fracción de lavado
- 5) Elusión con 100 mM de NaCl
- 6) Elusión con 200 mM de NaCl
- 7) Elusión con 500 mM de NaCl
- 8) Marcador de peso molecular

En la figura 14 se muestra la secuencia de aminoácidos de la proteína FLOR1 y en distintos colores los péptidos reconocidos mediante cromatografía líquida y espectrofotometría de masas (LC-MS/MS). La figura 15 corresponde a la secuencia de aminoácidos de la proteína VSP1 y en colores se muestran los péptidos identificados a partir del mismo tipo de análisis. En ambas secuencias proteicas se muestran los péptidos obtenidos a partir de las reacciones de triptinización de tres experimentos independientes.

```

1  MKLFLVHLSIF FSILFITLPS SYSTENDKN ALLQIKKALG NPPLLSSWNP RTDCCTGWTG
                                     KALG NPPLLSSWNP RT
61  VECTNRRVTG LSVTSGEVSG QISYQIGDLV DLRTLDFSYL PHLTGNIPRT ITKLNLTNL
                                               NLNTL
121 YLKHTSLSGP IPDYISELKS LTFDLDFNQ FTGPIPGSLS QMPKLEAIQI NDNKLTGSIP
    YLKH
181 NSFGSFGNV PNLVLSNNKL SGKIPESLSK YDFNAVDLSG NGFEGDAFMF FGRNKTIVRV
                                     KIPESLSK Y
241 DLSRNMENFD LVKVKFARSI VSLDLSQNH IYKIPPALTK LHLEHFNVD NHLCGKIPSG
    DLSRNMENFD LVKVK SI VSLDLSQNH YGK

```

**Figura 14:** Secuencia de aminoácidos de la proteína FLOR1 mostrando los péptidos identificados por LC-/MS/MS en colores.

```

1  MKILSLSLLL LLAATVSHVQ SSASVPGLIE LLESNTIFGN EAELLEKEGL SINYPCRSW
61  HLGVETSNII NFDTPANCK AYVEDYLITS KQYQYDSKTV NKEAYFYAKG LALKNDTVNV
                                     AYVEDYLITS KQYQYDSKTV NKEAYFYAK
121 WIFDLDDTLL SSIPYAKYG YGTENTAPGA YSWLESSES TPGLPETLHL YENLLELGIE
181 PIIISDRWKK LSEVTVENLK AVGVTKWHL ILKPNGSKLT QVVYKSKVRN SLVKKGYNIV
    KK LSEVTVENLK A KLT QVVYKS
241 GNIGDQWADL VEDTPGRVEK LPNPLYVPS
    VFK LPNPLYVPS

```

**Figura 15:** Secuencia de aminoácidos de la proteína VSP1 mostrando en colores los péptidos identificados por LC-MS/MS

### 6.3.2 Interacción con las proteínas At5g11420 completa y truncada utilizadas como ligandos inmovilizados con extractos crudos de hojas de *A. thaliana*.

#### 6.3.2.1 Interacción de la proteína completa At5g11420

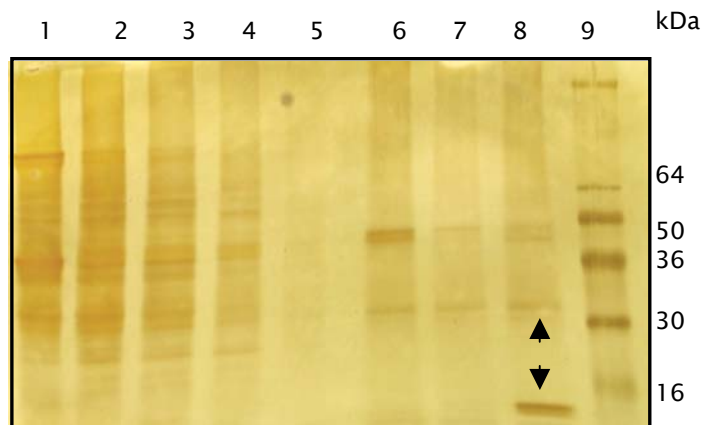
Con el fin de conocer a los posibles interactores de la proteína At5g11420 en otros tejidos se llevaron a cabo ensayos de interacción en columnas de afinidad utilizando extractos crudos de hojas como fuente proteica y como ligando inmovilizado la proteína completa At5g11420.

En la figura 16 se muestra un gel de poliacrilamida al 12 % teñido con plata con las diferentes fracciones obtenidas a partir de la cromatografía de afinidad. En los carriles 6, 7 y 8 se observan las fracciones eluidas a las diferentes concentraciones de NaCl. Los péptidos identificados

a partir de estas dos bandas señaladas correspondieron a una pectin metil esterasa con el código de registro Q9LUL7.

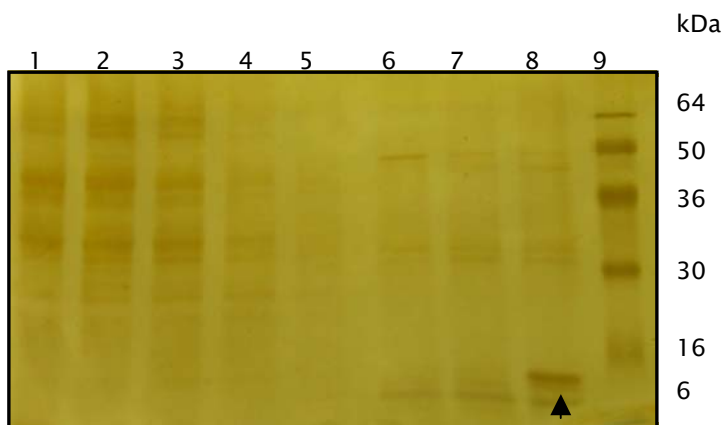
**Figura 16:** Gel de poliacrilamida al 12%, (tinción con plata) mostrando las fracciones obtenidas a partir de un experimento utilizando a la proteína completa At5g11420 y un extracto de hojas. Se realizaron dos réplicas más de este experimento.

- 1) Extracto crudo de hojas
- 2) Extracto crudo de hojas después de la columna de DEAE-sefacel
- 3) Fracción no unida
- 4) Fracción de lavado
- 5) Fracción de lavado
- 6) Elusión con 100 mM de NaCl
- 7) Elusión con 200 mM de NaCl
- 8) Elusión con 500 mM de NaCl
- 9) Marcador de peso molecular.



### 6.3.2.2 Interacción de la proteína truncada At5g11420

Se realizó un experimento utilizando como fuente de proteínas un extracto crudo de hojas y como ligando inmovilizado la proteína At5g11420 truncada. En la figura 17 se presenta el gel con las distintas fracciones obtenidas. La banda señalada con una flecha se identificó por LC-MS/MS. El resultado del análisis demostró que esta proteína correspondía a la pectin metil esterasa con el número de acceso Q9LUL7 (no publicado).

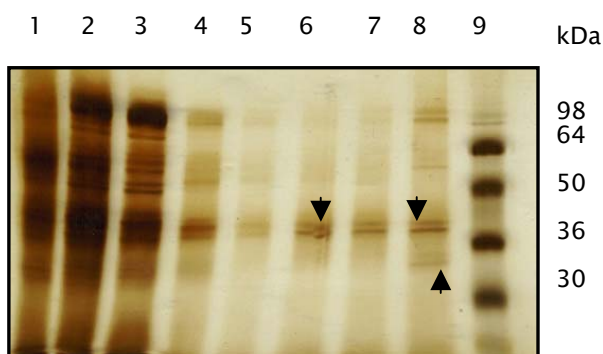


**Figura 17:** Gel de poliacrilamida al 12% (tinción con plata) mostrando las fracciones obtenidas a partir de la cromatografía de afinidad con extractos crudos de hojas y la proteína At5g11420 trunca.

- 1) Extracto crudo de hojas
- 2) Extracto crudo de hojas después de la columna de DEAE-sefacel
- 3) Fracción no unida
- 4) Fracción de lavado
- 5) Fracción de lavado
- 6) Elusión con 100 mM de NaCl
- 7) Elusión con 200 mM de NaCl
- 8) Elusión con 500 mM de NaCl
- 9) Marcador de peso molecular

### 6.3.3 Interacción de la proteína At4g32460 completa y extractos crudos de flores de *A. thaliana*

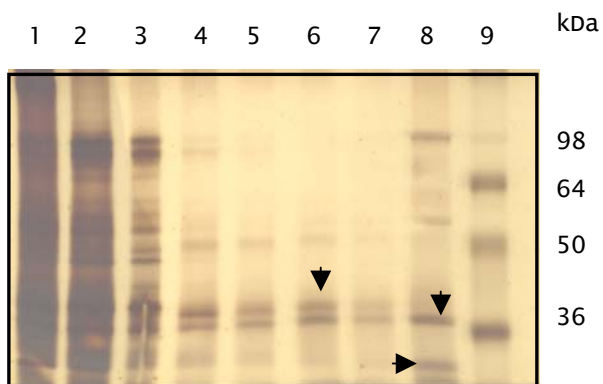
Al igual que con la proteína At5g11420 se realizaron análisis de interacción por cromatografía de afinidad, con el fin de conocer las proteínas que pudieran estar interactuando *in vitro* con la proteína At4g32460 utilizando un extracto crudo de flores como fuente proteica. En la figura 18 se observan las fracciones obtenidas a partir de un primer experimento, las bandas señaladas corresponden a las proteínas que fueron identificadas por LC-MS/MS.



**Figura 18:** Gel de poliacrilamida al 12 % (tinción con plata) mostrando las distintas fracciones obtenidas a partir de un extracto de flores y la proteína completa At4g32460 como ligando inmovilizado.

- 1) Extracto crudo de flores antes de la columna DEAE-Sefacel
- 2) Extracto crudo de flores después de la columna de DEAE-Sefacel
- 3) Fracción no unida
- 4) Fracción de lavado
- 5) Fracción de lavado
- 6) Elusión con 100 mM de NaCl
- 7) Elusión con 200 mM de NaCl
- 8) Elusión con 500 mM de NaCl
- 9) Marcador de peso molecular

En la figura 18 se observa la presencia de algunas proteínas con un peso molecular muy similar (carriles 7 y 8 flechas superiores), por lo que fue necesario correr estas mismas fracciones en un gel al 10% de poliacrilamida con el fin de separar de mejor manera dichas proteínas (Figura 19).



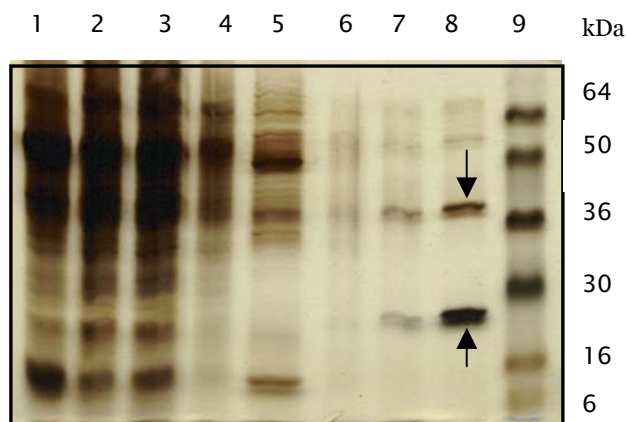
**Figura 19:** Gel de poliacrilamida al 10% mostrando las mismas fracciones del gel anterior. Las flechas indican las proteínas que fueron identificadas mediante LC-MS/MS.

De acuerdo con los resultados de espectrofotometría de masas los péptidos obtenidos a partir de las proteínas purificadas corresponden a péptidos de FLOR1 (proteína de aproximadamente 38 kDa), a péptidos de la pectin metil esterasa Q9LUL7 (proteína de aproximadamente 37 kDa), purificada anteriormente a partir de extractos de hojas y a péptidos de

la proteína VSP1 (proteína de aproximadamente 34 kDa). Los péptidos identificados por LC-MS/MS se encuentran entre los señalados en las figuras 14 y 15 respectivamente.

### 6.3.3.1 Interacción de la proteína completa At4g32460 utilizada como ligando inmovilizado y proteínas de extractos crudos de hojas de *Arabidopsis*

Al igual que con la proteína completa y truncada At5g11420 se realizó una columna de afinidad con la finalidad de conocer las proteínas de hoja que podrían estar interactuando *in vitro* con la proteína At4g32460. Los resultados obtenidos a partir de dicho ensayo se muestran en la figura 20.



**Figura 20:** Gel de poliacrilamida al 12% (tinción de plata). Fracciones proteicas obtenidas a partir de una columna de afinidad con la proteína inmovilizada At4g32460 completa y un extracto crudo de hojas.

- 1) Extracto crudo de hojas antes de la columna DEAE-sefacel
- 2) Extracto crudo después de la columna DEAE-sefacel
- 3) Fracción no unida
- 4) Lavado
- 5) Lavado
- 6) Elución con 100 mM de NaCl
- 7) Elución con 200 mM de NaCl
- 8) Elución con 500 mM de NaCl
- 9) Marcador de peso molecular

En el carril 8 se observan dos proteínas, una de aproximadamente 37 kDa la cual se identificó como una PME y una de aproximadamente 29 kDa. La identificación de esta última proteína por LC-MS/MS no fue concluyente, ya que solo se pudo identificar un péptido. De acuerdo con este análisis el péptido podía corresponder a una proteína de unión a Auxinas (*Auxin Binding Protein*) o a una proteína *Germin-like*. Estos resultados deberán ser corroborados.

En los análisis de interacción realizados con las proteínas recombinantes At5g11420 (completa y truncada) y At4g32460 (completa), con extractos de hojas y flores, se identificaron proteínas de diferentes pesos moleculares. En varios de estos experimentos los péptidos identificados correspondieron a una pectin metil esterasa con el número de acceso Q9LUL7 del NCBI (Cheuk *et al.*, no publicada); dicha PME es codificada por el locus At3g14310. Los péptidos identificados por LC-MS/MS se pueden observar en la figura 21.

```

1  MAPSMKEIFS KDNFKKNNKL VLLSAAVALL FVAAVAGISA GASKANEKRT LSPSSHAVLR
61  SSCSSTRYPE LCISAVVTAG GVELTSQKDV IEASVNLITIT AVEHNYFTVK KLIKKRGLT
121 PREKTALHDC LETIDETLDE LHETVEDLHL YPTKKTLREH AGDLKTLISS AITNQETCLD
181 GFSHDDADKQ VRKALLKGQI HVEHMCSNAL AMIKNMTDTD IANFEQKAKI TSNNRKLKEE
241 NQETTVAVDI AGAGELDSEG WPTWLSAGDR RLLQSGVKA DATVAADGSG TFKTVAAAVA
                                     A DATVAADGSG TFKTVAAAVA
301 AAPENSNKRY VIHIKAGVYR ENVEVAKKKK NIMFMGDGRT RTIITGSRNV VDGSTTFHSA
    AAPENSNKRY VIHIKAGVYR ENVEVAK    NIMFMGDGRT    TIITGSRNV VDGSTTFHSA
361 TVAAVGERFL ARDITFQNTA GPSKHQAVAL RVGSDFSAFY NCDMLAYQDT LYVHSNRQFF
    TVAAVGER    DITFQNTA GPSK
421 VKCLIAGTVD FIFGNAAVVL QDCDIHARRP NSGQKNMVTA QGRTPDNQNT GIVIQKCRIG
                                     TDPNQT GIVIQK IG
481 ATSDLQSVKG SFPTYLGRPW KEYSQTVIMQ SAISDVIRPE GWSEWTGTFA LNTLTyreYS
    ATSDLQSVK
541 NTGAGAGTAN RVKWRGFKVI TAAAEAQKYT AGQFIGGGGW LSSTGFPFSL GL
    NTGAGAGTAN RVK    GFKVI TAAAEAQK

```

**Figura 21:** Secuencia de aminoácidos de la proteína Q9LUL7 correspondiente a la pectin metil esterasa (PME) At3g14310. Los péptidos identificados mediante LC-MS/MS de las proteínas obtenidas por cromatografía de afinidad y electroforesis se presentan en colores.

De acuerdo con el European Bioinformatic Institute, este tipo de pectin metil esterases existe en diferentes especies de plantas. La PME identificada pertenece a una familia de 123 PMEs encontradas en distintas especies de plantas. En la figura 22 se muestra un esquema de los dominios presentes en este tipo de proteínas y específicamente la estructura general de la PME Q9LUL7. Como se puede observar, este tipo de proteínas presentan un dominio que codifica para un inhibidor de la PME y un dominio catalítico con actividad de pectin metil esterasa. Los péptidos reconocidos mediante LC-MS/MS de las PME purificadas a partir de las interacciones tanto con la proteína At5g11420 como con la proteína At4g32460 corresponden a la región N-terminal de la proteína, es decir al dominio catalítico.

A

### 123 proteínas con PMEI, Arquitectura Pectinesterasa

[Q9LPX7\\_ARATH](#) [*Arabidopsis thaliana* (mouse-ear cress)]



B

[Q9LUL7\\_ARATH](#) [*Arabidopsis thaliana* (mouse-ear cress)] pectinesterasa (At3g14310)



**Figura 22:** A. Estructura general de las PME agrupadas como aquellas que presentan un péptido señal para que la proteína pueda ser secretada. Además se señala el dominio inhibidor en verde (PMEI: Pectinmethyl esterase inhibitor, por sus siglas en inglés) y un dominio con la actividad catalítica de PME en rojo. De acuerdo con el Instituto Sanger (<http://www.sanger.ac.uk/>) se han clasificado 123 PME de diferentes especies de plantas en este grupo. B. La PME (Q9LUL7) posee en el N-terminal un dominio PMEI, una región coil-coil formada por dos o tres alfa hélices y el dominio con actividad catalítica de PME en el C-terminal.

#### **6.4 Línea mutantes por inserción de T-DNA**

La genética es una de las principales herramientas de la biología que ha permitido especificar las funciones de genes en diversos organismos. En plantas varios proyectos relacionados con la mutagénesis e inserción de grandes secuencias bacterianas en los cromosomas de especies como *A. thaliana* han tenido como principal objetivo la obtención de grandes bancos de semillas con inserciones en la mayor parte del genoma de este tipo de especies modelo. En muchas ocasiones la dilucidación de la función de un gen no es un trabajo sencillo, sobre todo en especies cuyo genoma ha sufrido un alto grado de duplicación genética y la redundancia funcional es un fenómeno común (Moore *et al.*, 2005). Con el fin de conocer si la mutación en alguno de los genes de la familia DUF642 nos permitía tener alguna información sobre las funciones de estos, se realizó la búsqueda de plantas mutantes por inserción de T-DNA. Como ya se había mencionado antes se encontraron semillas mutantes para varios genes entre ellos el At5g11420, At4g32460, At2g41800 y At2g41810. La obtención de una línea homocigota para la inserción solo fue posible obtenerla a partir de la línea mutante con la inserción de T-DNA en el locus At4g32460. Al determinar si se trataba de una mutante nula mediante ensayos de RT-PCR se confirmó que no era así, los resultados obtenidos a partir de dichos análisis se pueden ver en el anexo II.



## VII. Discusión

La familia de proteínas DUF642 es específica de plantas y presenta una serie de características que sugieren su participación en importantes procesos de transducción de señales durante el desarrollo vegetal y durante la interacción planta-ambiente. Las proteínas que integran esta familia están altamente conservadas en el reino vegetal; presentan un único dominio no catalítico posiblemente involucrado en procesos de interacción proteína-proteína. La identidad entre las secuencias de aminoácidos de los diferentes miembros de esta familia en *A. thaliana* es de 27% (figura 6); las proteínas de esta familia se han descrito tanto en Gimnospermas como en Angiospermas basales (Albert *et al.*, 2005). En este trabajo, los patrones de expresión para los 10 genes descritos en *A. thaliana* mostraron que esta familia se expresa en todos los tejidos estudiados aunque con un patrón de expresión diferencial en cada caso (figura 11). Estos resultados coinciden en su mayoría con los datos obtenidos mediante análisis de microarreglos descritos en el atlas de *Arabidopsis* Gene Investigator <http://www.geneinvestigator.ethz.ch/> (anexo I). Sin embargo, de acuerdo con nuestros resultados, el gen codificado por el locus At2g41810 se expresa únicamente en flores e inflorescencias, mientras que en el atlas de *Arabidopsis* Gene Investigator también se ha descrito su expresión en la raíz (anexo I). Kreps y colaboradores (2002) sugieren que la expresión de este gen en las raíces es estimulada específicamente en respuesta a estrés salino, sin embargo, en otro estudio reciente se muestra que la expresión de este gen en la raíz, es inhibida bajo condiciones de cultivo de tejidos de *A. thaliana* (Che *et al.*, 2006). Dado que el análisis de expresión que realizamos para este gen se llevó a cabo con raíces de plántulas crecidas en placas de petri bajo condiciones estándares, es posible que la luz y otros factores inhibieran su expresión. Por lo tanto, es necesario analizar la expresión de este gen tanto en raíces de plantas crecidas en agrolita (oscuridad) como en raíces de plantas bajo tratamientos diferenciales de NaCl y otros tipos de estrés. El patrón de expresión de este gen es muy específico a diferencia de la del resto de los genes de esta familia. Por esto, el análisis de las regiones promotoras es útil para hacer hipótesis sobre los posibles elementos reguladores de la expresión de este gen y del resto de la familia.

A pesar de que el gen At2g41800 presenta una identidad del 79.19% con el gen At2g41810 este gen no presentó una expresión tejido-específica, ya que se detectó en todos los tejidos pero más abundantemente en flores, inflorescencias y raíces; este resultado coincide con los datos de microarreglos del anexo I. Este último gen había sido propuesto como gen hermano del At2g41800 debido a la homología entre las secuencias de sus aminoácidos y al hecho de que estos dos se encuentren en tandem en el cromosoma 2, sugiriendo esto, un posible evento de duplicación reciente (Zúñiga, 2004). De acuerdo con nuestros resultados, el transcrito del gen At2g41800 amplificado en las inflorescencias (figura 11) es de mayor tamaño ya que el primer intrón del gen está presente. Existen varios codones de término dentro de la secuencia de este intrón, que impiden la traducción de una proteína completa en cualquiera de los tres marcos de lectura, por lo tanto, es posible que la

presencia de sólo el primer intrón refleje una regulación adicional dentro del proceso de edición del mRNA, sin embargo, estos datos necesitan ser confirmados utilizando RNA tratado previamente con DNAsas. La expresión específica del gen At2g41800 en los órganos florales podría ser determinante, ya que se ha descrito que su expresión disminuye en las flores de plantas cuyas células estigmáticas y del tracto de transmisión han sido eliminadas (Tung *et al.*, 2005). Estos datos y nuestros resultados indican que los genes At2g41800 y At2g41810 pueden participar en diversos procesos del desarrollo pero más estrechamente relacionados con el desarrollo de la flor y de la raíz. Es posible que la expresión específica del gen At2g41810 y menos específica del At2g41800 en flores pudiera estar regulada por las cajas de unión al factor de transcripción BELLRINGER, encontradas en sus regiones promotoras ya que como se ha descrito este factor de transcripción regula la expresión de AGAMOUS en los óvulos. Las cajas de unión a LFY podrían ser también importantes para la regulación de la expresión de estos genes en el meristemo.

Los genes del clado A (figura 7) se expresan en todos los tejidos y sus niveles de expresión no varían entre los tejidos, estos genes son: el At1g80240, At5g25460, el At4g32460 y el At5g11420. Los genes At4g32460 y At5g25460 del clado A, y el gen At3g08030 del clado B, podrían participar en procesos de señalización durante la germinación, ya que se ha reportado que los niveles de expresión de sus ortólogos se incrementan en la germinación de semillas de *Brassica oleracea* y la expresión de los genes At3g08030 y At5g25460 también se incrementan durante el priming (Soeda *et al.*, 2005). El incremento en los niveles de expresión de una gran cantidad de genes durante el acondicionamiento o “priming” confirma la hipótesis de que se están llevando a cabo distintos procesos relacionados con la germinación dentro de las semillas sometidas a este tipo de tratamiento, ya que la expresión genética está correlacionada con los eventos fisiológicos que ocurren durante la germinación. Se sabe que el priming promueve la absorción de agua y algunos de los procesos relacionados con la germinación, pero previene la emergencia de la radícula para evitar la pérdida de la tolerancia a la desecación. La maduración y germinación de las semillas también están reguladas por señales hormonales. El ácido absísico (ABA), además de participar en importantes procesos fisiológicos relacionados con el estrés, inhibe la germinación. Dado que uno de los procesos clave dependientes de ABA es la regulación negativa o positiva de la expresión de genes involucrados en respuesta a estrés y en procesos relacionados con la maduración de las semillas, el hecho de que estos tres genes presenten en su región promotora una gran cantidad de secuencias de reconocimiento por factores de transcripción bZIP que interactúan con elementos de respuesta a ABA (ABFs), constituye otra evidencia de que estos genes podrían estar involucrados en la germinación. Estos genes también poseen sitios en *cis* ACACNNG reconocidos por DPBFs por lo que su expresión durante etapas tardías del desarrollo embrionario podría ser importante (Kim, 2002). El trabajo de Soeda y colaboradores (2005) marca la pauta para la realización de ensayos de germinación con semillas de *Arabidopsis*, tanto silvestres como mutantes sometidas a distintos tratamientos de acondicionamiento o priming, con el fin de conocer si alguno de estos tres genes (el At3g08030, el At4g32460 y el At5g25460) están involucrados en procesos relacionados con el acondicionamiento y la germinación. La participación de estos genes en otros procesos del desarrollo

no se descarta, puesto que se expresan abundantemente en todos los tejidos, por ejemplo en el caso del gen At4g32460 se ha reportado que podría también ser importante en procesos de reconocimiento polen-estigma, específicamente durante la remodelación y el mantenimiento de la matriz extracelular (Tung *et al.*, 2006).

Con respecto al gen At5g11420 (perteneciente al clado A), el cual se expresa en todos los tejidos y presenta sitios de regulación por ABA, existe un trabajo realizado por el grupo de Osakabe y cols. (2005), en el que se describe que la expresión de este gen es regulada negativamente en una línea de plantas que sobreexpresa la proteína quimérica RL-BK construida a partir de la región N-terminal del receptor con dominios ricos en leucinas RPK1 (Receptor-like kinase 1). Dicho receptor está involucrado en la ruta principal de señalización y percepción temprana del ABA en *A.thaliana*. Hasta ahora no existe ninguna otra evidencia experimental que asegure la participación de estos genes durante procesos relacionados con ABA por lo que es necesario realizar los ensayos pertinentes para confirmar esta información. Desafortunadamente no se realizó un experimento de germinación con ABA para conocer si los niveles de expresión en esta etapa se modificaban al ser germinadas con esta hormona.

En el caso del gen At1g80240 (clado A) se amplificaron dos bandas en todos los tejidos, la mayor con los dos intrones (figura 11). No se descarta la posibilidad de que esto se relacione con la regulación de su expresión, como pudiera ser el caso del gen At2g41800 que fue amplificado con el primer intrón. At1g80240 es uno de los dos genes que se expresa específicamente en los pétalos (tabla 1) (Wellmer *et al.*, 2004).

Por otro lado, los genes At1g29980, At2g34510 y At5g14150, los cuales se encuentran agrupados en el clado (C) de acuerdo con la filogenia de esta familia presentada en la figura 7, son los únicos que poseen un péptido señal de unión a membrana (<http://aramemnon.botanik.uni-koeln.de/>). Estos análisis están apoyados por ensayos proteómicos en *Arabidopsis* en los que se aislaron proteínas de membrana ancladas a glicosil fosfatidil inositol, GPAs (*GPI-anchored proteins*) incluyendo a estas tres proteínas DUF642. Las GPAs se encuentran en todos los eucariontes, son proteínas-blancas de la superficie celular y están involucradas en la remodelación de la matriz extracelular. Las GPAs pueden existir tanto en su forma soluble como en su forma asociada a la membrana. De acuerdo con Tung y colaboradores (2005), el gen At5g14150 es un gen cuya expresión en células de papila estigmática es muy importante durante el proceso de polinización, por lo que el papel de este gen, como se sugiere en estos dos trabajos, podría estar relacionado con las rutas de señalización involucradas en procesos de remodelación de la membrana plasmática y/o secreción polarizada, tal como ha sido descrita en sistemas animales. Este tipo de eventos son de gran importancia durante el crecimiento del tubo polínico a través del tracto de transmisión del estilo (Borner, 2003).

Otros hallazgos relacionados con el análisis de los promotores, sugieren que la expresión de algunos de los miembros de la familia DUF642 podría estar regulada por distintos tipos de estrés.

El ABA también juega un papel muy importante en la respuesta al estrés oxidativo, regulando genes o activando enzimas involucradas en la remoción de especies reactivas de oxígeno (radical superóxido y H<sub>2</sub>O<sub>2</sub>) que se acumulan bajo varias condiciones de estrés y que pueden causar daño celular oxidativo; aunque también existe un tipo de respuesta a estrés no mediada por ABA. Se ha reportado, por ejemplo, que la regulación de la expresión del gen *ERD1* (*Early Responsive to Dehydration*) está mediada por elementos en *cis* que inducen su expresión bajo condiciones de estrés por deshidratación y senescencia inducida por oscuridad. En el promotor de dicho gen se pueden encontrar secuencias homólogas a elementos de respuesta a ABA (secuencias ABRE-like) del tipo ACGTG, secuencias de reconocimiento por elementos MYB-like (C/TAACNA/G) y MYC-like (CANNTG), entre otros. Algunos genes con elementos de reconocimiento para este tipo de factores de transcripción ABRE-like son los siguientes: At2g41800, At5g11420, At3g08030, At1g80240 y At5g14510. De igual manera se pudieron encontrar secuencias de reconocimiento para factores de transcripción MYC-like, descritos en genes cuya expresión está regulada por fitocromos (Wang, 1997).

Es importante considerar algunos otros motivos de reconocimiento encontrados en la regiones promotoras de los genes DUF642 de *A. thaliana*, ya que en ocasiones la expresión genética está determinada por la acción combinada de diferentes factores de transcripción, dando como resultado un rango diverso de patrones de expresión (Rombauts *et al.*, 2003; De Bodt *et al.*, 2006). Otros elementos regulatorios presentes en los genes de la familia DUF642 son: 1) los elementos G-box (5'-CACGTG-3) los cuales han sido identificados como genes regulados por luz, por ABA e inducidos bajo condiciones de estrés, 2) elementos en *cis* involucrados en regulación órgano específica (Ha, 1988) (en los genes At2g41800 y At3g08030); 3) elementos GATA implicados principalmente en la regulación de la expresión de genes dependientes de luz (en promotores de 9 de los 10 genes pertenecientes a la familia DUF642 de *Arabidopsis*), 4) elementos Box-II también presentes en genes regulados por luz (Menkens & Caschmore, 1994; Le Gourrierc *et al.*, 1999; Li *et al.*, 1999), 5) motivos de reconocimiento para el factor de transcripción MYC, AtMYC2 cuya sobreexpresión en *Arabidopsis* promueve la sensibilidad a ABA con un incremento concomitante en los niveles de expresión de genes regulados por estrés y ABA (presente en los genes: At2g41800, At2g41810, At1g80240, At4g32460, At511420, At525460, At3g08030, At1g29980); 6) elementos de reconocimiento por factores de transcripción MYB, que al igual que los MYC son inducidos por ABA en condiciones de sequía y alta salinidad y su sobre-expresión confiere hipersensibilidad a ABA y 7) motivos descritos en genes específicos de flor (en los promotores de los genes At2g41800, At5g14150) (Sablowski *et al.*, 1994).

Se cree que las diferencias en la regulación genética han influido fuertemente en la gran diversidad de los eucariontes superiores, en particular la divergencia de las regiones promotoras de los genes o de las familias genéticas producto de duplicaciones. La divergencia de las regiones promotoras tiene un profundo efecto en la función y en la regulación genética ya que es una importante fuerza directriz que genera novedades evolutivas, al permitir que los genes diverjan funcionalmente a través de eventuales modificaciones en sus secuencias estructurales, y/o en sus

patrones de expresión (De Bodt *et al*, 2006). Es evidente que los miembros de la familia DUF642 presentan diferencias en sus regiones promotoras; tal es el caso de las regiones promotoras de los genes At2g41800 y At2g41810, las cuales, además de ser distintas entre si, presentan motivos de regulación ausentes en el resto de los promotores de los otros genes. Otro ejemplo es el gen At5g14150 cuya secuencia promotora difiere considerablemente de los otros ya que no cuenta con los motivos de regulación presentes en la mayoría de los genes DUF642. Es posible que las diferencias en los patrones de expresión de los miembros de esta familia, sean el producto de la combinación de distintos elementos que actúan en *trans* en respuesta a diversas señales. La regulación transcripcional de la familia DUF642 debe ser el resultado de la integración de diferentes señales tanto externas (luz, sequía, salinidad, etc.), como internas (hormonas); es probable que el entrecruzamiento entre estas distintas rutas de señalización resulte en la regulación fina de la expresión de los genes miembro.

Por otro lado, las proteínas de la familia DUF642 solo presentan un dominio cuyas características no son similares a ninguna de las proteínas hasta ahora descritas. Es posible que su función esté relacionada con el establecimiento de interacciones específicas proteína-proteína como lo sugieren nuestros resultados. Dado que la expresión del mRNA de los genes At5g11420 y At4g32460 fue abundante en todos los tejidos y al hecho de que se contaba con semillas mutantes para estos dos genes y para el At2g41800 se consideró que la expresión de estas dos proteínas junto con la At2g41800 sería útil para entender la función particular que desempeñarían estas proteínas en distintos órganos (hojas y flores), sin embargo no fue posible obtener células transformantes con la construcción del gen At2g41800 por lo que al final solo se realizaron análisis de cromatografía de afinidad con las proteínas At5g11420 y At4g32460 usando como fuente proteica extractos crudos de hojas y de flores de *A. thaliana*. Las proteínas completas At5g11420 y At4g32460 interaccionaron *in vitro* con las proteínas FLOR1 y VSP1, mientras que la proteína At4g32460 también interaccionó con una pectin metil esterasa al utilizar como fuente de proteínas flores de *Arabidopsis*. La proteína truncada At5g11420 solo interaccionó con FLOR1 lo que sugiere la posibilidad de que la región C-terminal de esta proteína sea necesaria para la interacción con la proteína VSP1. Estos resultados sugieren que el dominio DUF642 permite el establecimiento de interacciones con proteínas específicas y que probablemente sean capaces de interactuar con un rango limitado de ellas. Se ha comprobado en ensayos preliminares la formación de un complejo FLOR1-VSP1-AGAMOUS y la interacción *in vitro* y posiblemente *in vivo* de FLOR1 con AGAMOUS. Aunque no se logró comprobar una interacción de la proteína At5g11420 o de la proteína At4g32460 con AGAMOUS o con algún otro factor de transcripción MADS-box, se estableció una posible relación indirecta con estos factores a través de la interacción con FLOR1. También es posible, de acuerdo con nuestros análisis *in silico*, que ocurran interacciones del tipo DNA-proteína entre factores de transcripción MADS-box y algunos de los genes pertenecientes a la familia DUF642 que presentan motivos de unión a este tipo de proteínas homeóticas.

Por otro lado, la misma pectin metil esterasa (PME) que fue purificada e identificada de extractos florales también fue purificada de extractos de hojas utilizando las columnas de afinidad con las dos proteínas (At5g11420 y At4g32460) como ligandos inmovilizados. Los tres polipéptidos que interaccionaron con las proteínas At5g11420 completa y truncada, al igual que el polipéptido de 37 kDa que interaccionó con la At4g32460 (figura 22), fueron identificados como parte de la secuencia de aminoácidos de la PME At3g14310 (figura 23). Como se puede observar en esta figura todos los péptidos que fueron identificados pertenecen a la región C-terminal de la PME (Q9LUL7), es decir, corresponden a la región catalítica de esta enzima

Desde el punto de vista del desarrollo vegetal, estos resultados son muy interesantes. Las PMEs participan en procesos tales como microesporogénesis, crecimiento del tubo polínico, germinación, desarrollo de la raíz, polaridad de las hojas, elongación del tallo, maduración de los frutos así como en la respuesta a patógenos fúngicos (Michelli, 2003; Bosh *et al.*, 2005; Di Matteo *et al.*, 2005) y son producidas tanto por las plantas como por patógenos microbianos y microorganismos simbioses durante la interacción con la planta (Lievens *et al.*, 2002). Además, se ha descrito que tanto las PME como sus inhibidores se expresan en tejidos florales y en granos de polen; estudios recientes de transcriptoma de polen de *A. thaliana* confirman la expresión de por lo menos 12 diferentes isoformas de PME (Bosh *et al.*, 2005).

Los polisacáridos que conforman la pared celular: la celulosa, la hemicelulosa y las pectinas, son secretados a las paredes celulares en sus formas altamente metilesterificadas y subsecuentemente son modificados por pectinasas tales como las PME. Estas enzimas catalizan la demetilesterificación de homogalacturonanos. Las PME se encuentran como pre-pro-proteínas inactivas en el citoplasma y sólo la parte madura y activa de la PME se encuentra en la pared celular. Se ha sugerido que la región pro puede inhibir a la parte madura durante su secreción al apoplasto para evitar la dimetilesterificación de las pectinas antes de su inserción a la pared celular, pero aún no se ha determinado si las PME son secretadas con la región pro o sin ella (Micheli, 2001). Dado que tanto la proteína At4g11420 como la proteína At5g32460 interactuaron con péptidos pertenecientes a la región C-terminal del dominio catalítico de la PME Q9LUL7 y no con la región N-terminal correspondiente al dominio inhibitorio, es posible que la actividad de este tipo de enzimas pueda estar regulada de alguna forma a través de la interacción con proteínas DUF642, es decir, que las proteínas DUF642 At5g11420 y At4g32460 estén participando en procesos de regulación enzimática y/o de chaperonas.

También es importante mencionar que los inhibidores de las PME (PMEI) actúan de una manera similar a los mecanismos propuestos para los inhibidores de las poligalacturonasas PGIP. En este sentido, considerando que FLOR1 es una proteína LRR cuyo motivo LxxLxxLxxLxLxxNxLxGxIP la incluye dentro de las proteínas inhibitoras de poligalacturonasas (PGIP) (Di Matteo, 2003), se podría sugerir un mecanismo de acción similar para las proteínas DUF642 al interactuar tanto con la PME en diferentes tejidos, como con FLOR1 en el tejido reproductivo. Se ha propuesto un modo de inhibición competitiva similar entre las PGIP -proteínas

inhibidoras de poligalacturonasas fúngicas- y las PMEs, lo cual representaría una estrategia general de las plantas para controlar las enzimas pécticas (Juge, 2006). De esta manera, la interacción de las proteínas DUF642 con este tipo de enzimas o proteínas podría ser un mecanismo de regulación común relacionado, ya sea con inhibición, o con eventos de direccionamiento o señalización y/o transporte. La estructura terciaria de las PMEs y las PGIPs es muy parecida y es posible que la interacción de estas dos enzimas con proteínas DUF642 ocurra de manera general en los diferentes tejidos. Además, si consideramos la participación de éste tipo de enzimas, es decir de las PMEs y las PGIPs, tanto en procesos del desarrollo órgano-específicos como en procesos relacionados con la respuesta a distintos tipos de estrés, podríamos sugerir que las proteínas DUF642 participan durante los mismos tipos de procesos (Li *et al.*, 2003, Ahsan *et al.*, 2005, Di Cuixia *et al.*, 2006). Este tipo de conclusiones relacionadas con estrés, se han visto sustentadas por los resultados obtenidos a partir experimentos realizados en el laboratorio, con dos líneas mutantes *flor1* (una nula y otra con niveles de expresión medios) y dos líneas de sobre-expresión. Los resultados preliminares muestran que las plantas que sobreexpresan FLOR1 germinan y se desarrollan mucho mejor con respecto a las silvestres bajo condiciones de estrés hídrico.

Asimismo, la proteína de 27 kDa que interaccionó con la proteína At4g32460 en hojas, se identificó como una posible proteína de unión a auxinas (*auxin binding protein* ABP) o una *germin-like protein* (GLP) (figura 22). Tanto las proteínas GLPs como las ABPs pertenecen a una superfamilia de proteína denominadas “cupinas”, nombradas así, con base a su estructura en forma de barril  $\beta$ . La identificación de esta superfamilia está basada originalmente en una proteína termoestable de trigo. Esta proteína comparte una secuencia peptídica con una proteína de *Physarum polycephalum* la cual ha sido relacionada con procesos de estrés. Esta similitud se extendió al grupo de las GLPs y de manera mucho menos directa a las proteínas de almacenamiento denominadas globulinas (Dunwell *et al.*, 2004). A pesar de que existe gran cantidad de información sobre las GLPs, su función bioquímica como oxalato oxidasa fue descubierta recientemente. Otros miembros tanto de plantas como de hongos presentan actividad de superóxido dismutasa (SOD) (Nakata *et al.*, 2002; Trornbung *et al.*, 2003). Por esto, se puede relacionar la expresión de las GLPs con los mecanismos de defensa en plantas. Las ABPs a su vez están relacionadas con una gran variedad de respuestas durante el crecimiento de las plantas. La similitud del sitio de unión a auxinas con el sitio activo de otras cupinas sugiere que probablemente esta y otras ABPs tienen una función enzimática aún no descrita (Dunwell *et al.*, 2004).

Finalmente, con base en los resultados obtenidos en este trabajo se puede sugerir que el papel de las proteínas DUF642 puede abarcar distintas rutas de señalización. Sin embargo, es probable que su función principal sea mediar interacciones con proteínas involucradas tanto en respuesta a estrés (FLOR1, VSP1, PME y ABP/GLP) como con proteínas presentes durante el desarrollo de órganos (FLOR1 y la PME e indirectamente con AGAMOUS). Las proteínas con dominios DUF642, podrían ser un nuevo tipo de proteínas específicas de plantas, capaces de interactuar con distintas proteínas involucradas en procesos de defensa y/o desarrollo así como

con proteínas estructurales. Las proteínas DUF642 podrían ser exportadas de la célula o permanecer dentro de ella dependiendo de las señales que posean en su secuencia proteica, tal como lo sugieren Abrahams y colaboradores (1995) en el caso de la proteína DUF642 de *Medicago sativa*; o permanecer o ser exportadas dependiendo de las interacciones que establezcan con otras proteínas. Los dominios de interacción juegan un papel crítico en la activación selectiva de distintas rutas de señalización a través de su capacidad de reclutar proteínas y regular la subsiguiente formación de complejos de señalización en un compartimento subcelular dado. Tales dominios de interacción pueden no solo controlar la especificidad de la señal, si no también, la manera en que las células responden a las diferentes señales externas o internas. Se ha descrito, por ejemplo, que las PGIPS poseen una amplia superficie de interacción dentro de la fase cóncava del dominio LRR ocupado químicamente por diversos residuos tanto hidrofílicos como hidrofóbicos. Además, la superficie “docking” de este tipo de proteínas parece ser extremadamente versátil y capaz de establecer tanto interacciones polares como apolares con superficies proteicas estructuralmente similares, pero al mismo tiempo tener funciones diversas. El caso de las proteínas DUF642 puede ser similar ya que establecen interacciones con proteínas estructuralmente similares pero cuyas funciones pueden ser ~~son~~ muy diversas.



## VIII. Conclusiones

- ✓ Los genes miembros de la familia DUF642 en *Arabidopsis* se expresan diferencialmente en los órganos de esta especie y sólo uno de ellos, el codificado por el locus At2g41810, se expresa en el tejido reproductivo, aunque no se descarta su expresión en la raíz como se ha reportado en análisis de microarreglos con chips de Afimetrix.
- ✓ Los genes pertenecientes a la familia DUF642 presentan en sus secuencias promotoras elementos *cis* regulatorios reportados en otros genes cuya expresión es regulada por luz, estrés, ABA, y factores de transcripción MADS-box; por lo que la expresión de estos puede también estar regulada por señales tanto endógenas como exógenas.
- ✓ La proteína recombinante At5g11420 completa interactuó *in vitro* con las proteínas FLOR1 y VSP1 al utilizar extractos crudos de flores como fuente proteica. La proteína At5g11420 truncada solo interactuó *in vitro* con la proteína FLOR1, por lo que sugiere que el C-terminal de la proteína At5g11420 media la interacción entre ésta y VSP1.
- ✓ La región catalítica codificada por el dominio N-terminal de la enzima PME (locus At3g14310) interactuó *in vitro* con la proteína At5g11420 al utilizar extractos crudos de hojas como fuente de proteínas. Esta misma enzima interactuó con la proteína At4g32460 al utilizarse extractos crudos de flores y hojas.
- ✓ La proteína At4g32460 completa interactuó *in vitro* con una proteína de unión a auxinas (Auxin Binding Protein) o una *germin like protein* al utilizar un extracto de hojas.
- ✓ Las proteínas At5g11420 y At4g32460 podrían participar en las rutas de señalización en respuesta a distintos tipos de estrés tanto biótico como abiótico, durante procesos relacionados con la matriz extracelular y la pared celular, así como también en importantes cascadas de señalización durante el desarrollo, pues las proteínas con las que interactúan *in vitro* median este tipo de procesos.
- ✓ El estudio de la familia DUF642, específica de plantas, a través de distintas aproximaciones proveerá de datos interesantes relacionados con la evolución y la diversificación de las plantas ya que estas proteínas se han conservado durante la evolución en los distintos grupos, desde las Gimnospermas basales hasta las Angiospermas.
- ✓ La línea mutante para el locus At4g32460 es una mutante no nula a pesar de presentar la inserción de T-DNA en el segundo exón; es necesario estandarizar las condiciones para la amplificación del borde izquierdo del T-DNA si es que este está presente en esta parte del genoma, o bien, hacer la búsqueda de otras líneas mutantes para los genes DUF642 en otros bancos de semillas, ya que no se logró amplificar el borde izquierdo de T-DNA en la línea At5g11420 y At2g41800.

## IX. Perspectivas

Para entender la función de esta familia específica de plantas es necesario hacer análisis mucho más detallados de la expresión de cada uno de los genes en los diferentes tejidos. La implementación de los análisis de la expresión mediante hibridaciones *in situ* puede ser útil para generar información mucho más amplia acerca de la expresión de estos genes. Esta información, como lo sugieren Albert y colaboradores (2005), puede ser integrada en una filogenia, proveyendo de importantes pistas acerca de la diversificación de esta familia. Además sería interesante comprobar que los genes At5g11420 y At1g80240 se expresan únicamente en los pétalos de la flor, como se describe en el análisis genómico realizado por Wellmer y colaboradores con los órganos florales de *Arabidopsis*, en comparación con el gen At5g25460, el cual a pesar de ser gen hermano de uno de ellos no fue reportado por Wellmer *et al*, como un gen con expresión órgano-específica de flor.

Por otro lado, son diversas las aproximaciones útiles para conocer cuales son los elementos en *trans* que regulan la expresión de los genes, un ejemplo de ellas son los ensayos de movilidad electroforética (EMSAs) y los ensayos de inmunoprecipitación de cromatina (ChIP). Este tipo de ensayos serían útiles para determinar experimentalmente cuales son las señales que regulan la expresión de los genes de la familia DUF642, pues como lo sugieren nuestros análisis *in silico* el patrón de expresión de los genes de la familia DUF642 es el producto de la integración de diversas señales. Otro tipo de análisis *in silico* que también podría implementarse, es el *phylogenetic footprinting* el cual sería útil para conocer y comparar los sitios de regulación presentes en los promotores de estos genes.

Sobre el establecimiento de las funciones particulares de cada uno de los miembros de la familia DUF642 queda aún mucho trabajo por realizar, sería importante por ejemplo, continuar con la búsqueda de las líneas mutantes homocigotas nulas para estos genes en *Arabidopsis* ya que aún queda la posibilidad de encontrar fenotipos en las mutantes los cuales puedan ser útiles para entender la función de estos genes. La obtención de dobles mutantes en el caso de los genes At4g11420 y At1g80240 los cuales presentan expresión órgano-específica en la flor, por ejemplo, podría evidenciar el papel que desarrollan estas proteínas durante el desarrollo floral. Además, la interacción entre la proteína At4g11420 y la proteína FLOR1 apoya la dirección de la investigación en la obtención de mutantes dobles entre FLOR1 y las DUF642.

También, es necesario confirmar las interacciones descritas *in vitro* mediante otras técnicas. Para ello es necesario hacer uso de estrategias como la de doble-híbrido o de inmunoprecipitación, por lo que la obtención de anticuerpos policlonales también ya ha sido

considerada. Es también un requisito comprobar que las interacciones descritas *in vitro*, ocurren *in vivo*. En este caso sería útil realizar construcciones con las proteínas DUF642 y proteínas fluorescentes con el fin de determinar no solo la localización subcelular de estas proteínas sino también con la finalidad de saber si las proteínas DUF642 colocalizan, por ejemplo con FLOR1 o con la PME. También sería importante determinar si las proteínas DUF642 actúan intra o extracelularmente.

## X. Bibliografía Citada

- Abe H., Yamaguchi-Shinozaki K., Urao T., Iwasaki T., Hosokawa D., Shinozaki K. (1997). Role of *Arabidopsis* MYC and MYB homologs in drought-and abscisic Acid-regulated gene expression. *The Plant Cell* 9:1859-1868.
- Abe M., Takahashi T., Komeda Y. (2001). Identification of a cis-regulatory element for L1 layer-specific gene expression, which is targeted by an L1-specific homeodomain protein. *The Plant Journal* 26(5):487-494.
- Abrahams S., Hayes C., Watson J. (1995). Expression patterns of three genes in the stem of lucerne (*Medicago sativa*). *Plant Molecular Biology* 27:513-528.
- Acevedo G., Gamboa A., Paéz-Valencia J., Izaguirre-Sierra M., Álvarez-Buylla R. (2004). FLOR1, A Putative Interaction Partner Of The Floral Homeotic Protein AGAMOUS, Is A Plant-Specific Intracellular LRR. *Plant Science* 167 (2):225-231.
- Acosta-García G. Aufran D., Vielle-Calzada J.Ph. Recombinant Gene Expression. Review and Protocols, 2a edición. Human Press Totowa USA 2004. 397-414.
- Ahsan N., Yoon H-S., Jo J. (2005). Molecular cloning of a BcPGIP cDNA from *Brassica campestris* and its expression to several stresses. *Plant Science* 169:1081-1089.
- Albert VA., Soltis DE., Carlson JE., Farmerie WG., Wall PK., Ilut DC., Solow TM., Mueller LA., Landherr LL., Hu Y., Buzgo M., Kim S., Yoo M-J., Fröhlich MW., Perl-Treves R., Schlarbaum S., Bliss BJ., Zhang X., Tanksley SD., Oppenheimer DG., Soltis PS., Ma H., de Pamphilis CW., Leebens-Mack JH. (2005). Floral gene from basal angiosperms for comparative genomics research. *Bio Medical Central Plant Biology* 5(5):1-15.
- Andrianopolus A. & Timberlake WE. (1994). The *Aspergillus nidulans abaA* gene encodes a transcriptional activator that acts as a Genetic Switch to control Development. *Molecular and Cellular Biology* 14(4):2503-2515.
- Argüello-Astorga G & Herrera- Estrella L. (1998). Evolution of Light-regulated plant Promoters. *Annual Review Plant Physiology and Plant Molecular Biology* 49:525-555.
- Askjaer P., Galy V., Hannak E., Mattaj W. (2002). Ran GTPase cycle and Importins  $\alpha$  and  $\beta$  are essential for spindle formation and nuclear envelope assembly in living *Caenorhabditis elegans* embryos. *Molecular Biology of the Cell* 13:4355-4370.
- Bao X., Franks RG., Levin JZ., Liu Z. (2004). Repression of *AGAMOUS* by *BELLRINGER* in floral and inflorescence Meristems. *The Plant Cell* 16:1478-1489.
- Baumann K., De Paolis A., Constantino P., Gualberti G. (1999). The DNA Binding site of the Dof Protein NtBBF1 is essential for Tissue-specific and Auxin-regulated Expression of the *rolB* oncogene in Plants. *The Plant Cell* 11:323-333.
- Belkhadir Y., Subramaniam R., Dangl J. (2004). Plant disease resistance protein signaling: NBS-LRR proteins and their partners. *Current Opinion in Plant Biology* 7:391-399.
- Benavente L. & Alonso MJ. (2006). Molecular mechanisms of ethylene signaling in *Arabidopsis*. *Molecular Biosystems* 2:165-173.
- Borner GHH., Lilley KS., Stevens TJ., Dupree P. (2003). Identification of Glycosylphosphatidylinocitol-Anchored Proteins in *Arabidopsis*. A Proteomic and Genomic Analysis. *Plant Physiology*. 132:568-577.
- Bosh M., Cheung A., Hepler P. (2005). Pectin Methyltransferase, a regulator of pollen tube growth. *Plant Physiology* 138: 1334-1346.
- Burack W. and Shaw A. (2000). Signal transduction: hanging on a scaffold. *Current Opinion in Cell Biology* 12:211-216.

- Chan CS., Guo L., Shih MG. (2001). Promoter analysis of nuclear gene encoding the chloroplast glyceraldehyde-3-phosphate dehydrogenase B subunit of *Arabidopsis thaliana*. *Plant Molecular Biology* 46:131-141.
- Chen Y-F., Etheridge N., Schaller E. (2005). Ethylene Signal Transduction. *Annals of Botany* 95:901-915.
- Choi H., Hong J., Ha J., Kang J., Kim SY. (2000). ABFs, a Family of ABA-responsive Element Binding Factors. *The Journal of Biological Chemistry* 275(3) 1723-1730.
- De Bodt S., Theissen G., Van de Peer. (2006). Promoter analysis of MAD-box Genes in Eudicots through Phylogenetic Footprinting. *Molecular Biology and Evolution* 23(6): 1293-1303.
- De Vetten NC & Ferl RJ. (1995). Characterization of a maize G-box binding factor that is induced by hypoxia. *The Plant Journal* 7(4):589-601.
- Di Cuixia, Zhang M., Xu S., Cheng T., An L. (2006). Role of Polygalacturonase-Inhibiting protein in Plant Defense. *Critical Reviews in Microbiology* 32:91-100.
- Di Matteo A., Bonivento D., Tsernoglou D., Federici L., Cervone F. (2006). Polygalacturonase-inhibiting protein (PGIP) in plant defense: a structural view. *Phytochemistry* 67:528-533.
- Di Matteo A., Federici L., Mattei B., Salvi G., Johnson K., Savino C., De Lorenzo G., Tsernoglou D. (2003). The crystal structure of polygalacturonase-inhibiting protein (PGIP), a leucine-rich repeat protein involved in plant defense. *Proceeding of the National Academy of Science* 100 (17):10124-10128.
- Di Matteo A., Giovane A., Raiola A., Camardella L., Bonivento D., De Lorenzo G., Cervone F., Bellincampi D., Tsernoglou D. (2005). Structural Basis for the interaction between Pectin Methylsterases and a Specific Inhibitor Protein. *The Plant Cell* 17: 849-858.
- Dickinson CD., Evans RP., Nielsen NC. (1988) RY repeats are conserved in the 5'-flanking regions of legume seed-protein. *Nucleic Acid Research* 16(1): 371.
- Dunwell JM., Purvis A., Khuri S. (2004). Cupins: the most functionally diverse protein superfamily? *Phytochemistry* 65:7-17.
- Estanyol MJ., Gomis-Rüth F., Puigdomenech P. (2004). The eighth-cysteine motif, a versatile structure in plant proteins. *Plant Physiology and Biochemistry* 42:355-365.
- Forsthoefel N., Cutler K., Port M., Yamamoto T., Vernon D. (2005). PIRLS: A novel class of plant Intracellular Leucine-Reach Repeat proteins. *Plant Cell Physiology* 46 (6):913-922.
- Gamboa A. Paéz-Valencia J., Acevedo Gf., Vázquez-Moreno L., Alvarez-Buylla. (2001). Floral Transcription Factor AGAMOUS Interacts *In Vitro* with A Leucine-Rich Repeat and an Acid Phosphatase Complex. *Biochemical and Biophysical Research Communications* 288(4): 1018-1026.
- Gilroy S. & Trewavas A. (2001). Signal processing and transduction in plant cells: the end of the beginning? *Nature Review* 2:307-314.
- Giuliano G., Hoffman NE., Kenton K., Scolnik PA., Cashmore AR. (1988). A light-entrained circadian clock controls transcription of several plant genes. *The EMBO Journal* 7(12): 3635-3642.
- Goda H., Sawa S., Asemi T., Fujioka S., Shimada Y., Yoshida S. (2004). Comprehensive Comparison of Auxin-Regulates and Brassinoesteroid-Regulated genes in *Arabidopsis*. *Plant Physiology* 134:1555-1573.
- Guiltinan MJ., Marcotte WRJ. Quatrano RS. (1990). A plant leucine zipper protein that recognizes an abscisic acid response element. *Science* 250(4978):267-271.
- Guo H. & Ecker J. (2004). The ethylene signaling pathway: new insights. *Current Opinion in Plant Biology* 7:40-49.
- Ha A-B. & An Gynheung. (1988). Identification of upstream regulatory elements involved in the developmental expression of the *Arabidopsis thaliana cab1* gene. *Proceeding of the National Academy of Science* 85:8017-8021.

- Hepworth SR., Valverde F., Ravenscroft D., Mouradov A., Coupland G. (2002). Antagonistic regulation of flowering-time gene *SOC1* by CONSTANS and FLC via separate promoter motifs. *European Molecular Biology Organization Journal* 21(16): 4327-4337.
- Huang Y., Li Hui., Hutchison CE., Laskey J., Kieber J. (2003). Biochemical and functional analysis of CTR1, a protein kinase that negatively regulates ethylene signaling in *Arabidopsis*. *The Plant Journal* 33:221-233.
- Hudson ME & Quail PH. (2003). Identification of promoter Motifs Involved in the network of Phytochrome A-regulated gene Expression by combined Analysis of Genomic Sequence and Microarray Data. *Plant Physiology* 133:1605-1616.
- José-Estanyol M., Gomis-Ruth X., Puigdomènech P. (2004). The eight-cysteine motif, a versatile structure in plants proteins. *Plant Physiology and Biochemistry* 42: 355-365.
- Juge N. (2006). Plant protein inhibitors of cell wall degrading enzyme. *TRENDS in Plant Science* 11(7):359-367.
- Kagaya Y., Ohmiya K., Hattori T. (1999). RAV, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely founds in higher plants. *Nucleic Acids Research* 27(2):470-478.
- Kim S., Kang JY., Choi DI., Park HY., Kim SY. (2004). ABF2, ABRE-binding bZIP factor, is an essential component of glucose signaling and its overexpression affects multiple stress tolerance. *The Plant Journal* 40: 75-87.
- Kim YS., Ma J., Perret P., Li Z., Thomas TL. (2002). Arabidopsis ABI5 Subfamily Members Have Distinct DNA-Binding and Transcriptional Activities. *Plant Physiology* 130: 688-697.
- Kreps JA., Wu Y., Chang H-S., Zhu T., Wang X., Harper JF. (2002). Transcriptome Changes for Arabidopsis in Response to Salt, Osmotic, and Cold Stress. *Plant Physiology* 130:2129-2141.
- Kusano A., Yoshioka T., Nishijima H., Nishitani H., Nishimoto T. (2004). *Schizosaccharomyces pombe* Ran GAP Homolog, SNrANA1, is required for Centromeric Silencing and Chromosome Segregation. *Molecular Biology of the Cell* 15:4960-4970.
- Le Gourrierc J., Li Y-F., Zhou D-X. (1999). Transcriptional activation by *Arabidopsis* GT-1 may be through interaction with TFIIA-TBP-TATA complex. *The Plant Journal* 18(6): 663-668.
- Li R., Rimmer R., Yu M., Sharpe AG., Séguin-Swartz G., Lydiat D., Hagedus DD. (2003). Two *Brassica napus* polygalacturonase inhibitory protein genes are expressed at different levels in response to biotic and abiotic stresses. *Planta* 217:299-308.
- Lin X., Kaul S., Rounsley S., Shea TP., Benito ML., Town CD., Fujii CY., Mason T., Bowman CL., Barnstead M. (1999). Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* 402, 761-768.
- Luo M., Dennis ES., Berger F., Peacock WJ., Chaudhury A. (2005). *Minissed3 (MINI3)*, a *WRKY* family gene, and *HAIKU2 (IKU2)*, a leucine-rich repeat (LRR) kinase gene, are regulators of seed size in *Arabidopsis*. *Proceeding of the National Academy of Science* 29 (48):17531-17536.
- Marone M., Mozzeti S., De Ritis D., Pierelli L., Scambia G. (2001). Semiquantitative RT-PCR analysis to assess the expression levels of multiple transcripts from the same sample. *Biological Procedures Online* 3(1):19-25.
- Mayer K., Schuller C., Wambutt R., Murphy G., Volckaert G., Pohl T., Dusterhoft A., Stiekema W., Entian KD., Terry N. (1999). Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*. *Nature* 402, 761-768.
- Mele G., Ori N., Sato Y., Hake S. (2003). The *knotted1*-like homeobox gene *BREVIPEDICELLUS* regulates cell differentiation by modulating metabolic pathway. *Genes & Development Research Communication* 17:2088-2093.
- Menkens AE. & Caschmore A R. (1994). Isolation and Characterization of fourth *Arabidopsis thaliana* G-box-binding factor, which has similarities to Fos oncoprotein. *Proceeding of the National Academy of Science* 91:2522-2526.
- Micheli F. (2001). Pectin methylesterases: cell wall enzymes with important roles in plant physiology. *Trends in Plant Science* 6(9): 414-419.

- Moore R., Grand SR., Purugganan MD. (2004). Molecular population Genetics of Redundant Floral-Regulatory Genes in *Arabidopsis thaliana*. *Molecular Biology and Evolution* 22(1):91-103.
- Motchoulski A. & Liscum E. (1999). *Arabidopsis* NPH3: A NPH1 Photoreceptor-Interacting Protein Essential for Phototropism. *Science* 286: 961-964.
- Nordin K., Palva ET. (1993). Differential expression of two related, low-temperature-induced genes in *Arabidopsis thaliana* (L.) Heynh. *Plant Molecular Biology* 21(49):641-653.
- Nurrish S. & Treisman R. (1995). DNA Binding Specificity Determinants in MADS-box Transcription Factors. *Molecular and Cell Biology* 15 (8): 4076-4085.
- Onodera Y., Suzuki A., Wu C-Y., Washida H., Takaiwa F. (2001). A rice functional Transcriptional Activator, RISBZ1, responsible for Endosperm-specific Expression of storage Protein Genes through GCN4 Motif. *The Journal of Biological Chemistry* 276(17):14139-14152.
- Osakabe Y., Maruyama K., Seki M., Satou M., Shinozaki K., Yamaguchi-Shinozaki K. (2005) Leucine- Rich Repeat Receptor- Like kinase1 is a key membrane-bound regulator of Abscisic acid Early Signaling in Arabidopsis. *The Plant Cell* 17:1105-1119.
- Ostergaard L & Yonofsky F. (2004). Establishing gene function by mutagenesis in *Arabidopsis thaliana*. *The Plant Journal* 39, 682-696.
- Páez-Valencia J. (2001). Expresión de FLOR1, una proteína con motivos repetidos ricos en leucinas, durante el desarrollo floral de *Arabidopsis thaliana* (producción de anticuerpos policlonales e inmunolocalización). Tesis de Licenciatura. UNAM.
- Pawson T., Raina M., Nash P. (2002). Interaction domains: from simple binding to complex cellular behavior. *FEBS Letters* 513: 2-10.
- Pla M., Vilardell J., Guitinan MJ., Marcotte WR., Niogret MF., Quatrano RS., Pages M. (1993). The cis-regulatory element CCACGTGG is involved in ABA and water stress responses of the maize gene rab28. *Plant Molecular Biology* 21(2)259-266.
- Reyes JC., Muro-Pastor I., Florencio F. (2004). The GATA family of Transcription Factors in Arabidopsis and Rice. *Plant Physiology* 134:1718-1732.
- Rombautus S., Floquin., Lescot M., Marchal K., Rouzé P, Van de Peer Y. (2003). Computational Approaches to Identify Promoters and cis-regulatory elements in Plant Genomes. *Plant Physiology* 132:1162-1176.
- Rose A. & Meier. (2001). A domain unique to plant Ran GAP is responsible for its targeting to the plant nuclear rim. *Proc. Natl. Acad. Sci.* 98 (26): 15377-15382.
- Sablowski WM., Moyano E., Culianez-Macia A., Schuch M., Martin C., Bevan M. (1994). A flower-specific Myb protein activates transcription phenylpropanoid biosynthetic genes. *The EMBO Journal* 13(1): 128-137.
- Saddic LA., Huvermann B., Bezhani S., Su Y., Winter CM., Kwon CS., Collum RP., Wagner D. (2006). The LEAFY target LMI1 is a meristem identity regulator and acts together with LEAFY to regulate expression of *CAULIFLOWER*. *Development* 133:1673-1682.
- Salanoubat M., Lemcke K., Rieger M., Ansorge W., Unseld M., Fartmann B., Valle G., Blocker H., Pérez-Alonso M., Obermaier B. (2000). Sequence and analysis of chromosome 3 of the plant *Arabidopsis thaliana*. *Nature* 408: 820-822.
- Satoh R., Fujita Y., Nakashima K., Shinozaki K., Yamaguchi-Shinozaki K. (2004). A novel subgroup of bZIP proteins functions as Transcriptional Activators in Hypoosmolarity-Responsive Expression of *ProDH* gene in *Arabidopsis*. *Plant Cell Physiology* 45 (3):309-317.
- Shao H-B., Guo Q-J., Chu L-Y, Zhao X-N., Su Z-L., Hu Y-C., Cheng J-F. (2007). Understanding molecular mechanism of higher plant plasticity under abiotic stress. *Review. Colloids and Surfaces B: Biointerfaces* 54 (1):37-45.
- Shinozaki K. & Yamaguchi-Shinozaki K. (2000). Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Current Opinion in Plant Biology* 3(3).217-223.

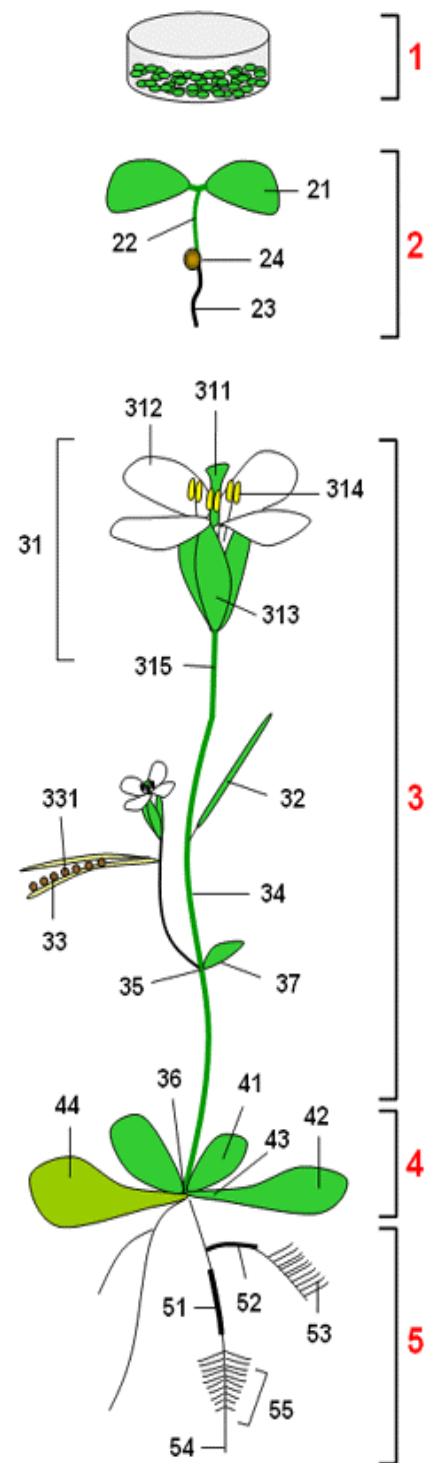
- Simpson SD., Nakashima K., Narusaka Y., Seki M., Shinozaki K., Yamaguchi-Shinozaki K. (2003). Two different novel *cis*-acting elements of *erd1*, a *clpA* homologous *Arabidopsis* gene function in induction by dehydration stress and dark-induces senescence. *The Plant Journal* 33:259-270.
- Soeda Y., Konings M., Vorst O., van Houwelingen A., Stoopen G., Maliepaard C., Kodde J., Bino R., Groot S., van der Geest A. (2005). Gene expression programs during *Brassica oleracea* seed maturation, osmopriming, and germination are indicator of progression of the germination process and the stress tolerance level. *Plant Physiology* 137:354-368.
- Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E., Miyajima N., Sasamoto S., Kimura T., Hosouchi T. (2000). Sequence and analysis of chromosome 5 of the plant *Arabidopsis thaliana*. *Nature* 408: 823-826.
- Tamminen I., Mäkelä P., Heino., Palva TE. (2001). Ectopic expression of *ABI3* gene enhances freezing tolerance in response to abscisic acid and low temperature in *Arabidopsis thaliana*. *The Plant Journal* 25(1):1-8.
- Teakle GR., Manfield IW., Graham JF., Gilmartin PM. (2002). *Arabidopsis thaliana* GATA factors: organization, expression and DNA-binding characteristics. *Plant Molecular Biology* 50:43-57.
- The Arabidopsis genome Initiative. (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796-813.
- Theologis A., Ecker JR., Curtis JP., Federspiel NA., Kaul S., Venter C. (2000). Sequence and analysis of chromosome 1 of the plant *Arabidopsis thaliana*. *Nature* 408: 816-820.
- Treisman R., Marais., Wynner J. (1992). Spatial flexibility in ternary complexes between SRF and its accessory proteins. *EMBO Journal* 11(12): 4631-4640.
- Tung C-W., Dwyer K., Nasrallah ME., Nasrallah JB. (2005). Genome-wide Identification of genes Expressed in Arabidopsis Pistils Specifically along the Path of pollen Tube Growth. *Plant Physiology* 138:977-989.
- Ulmasov T., Hagen G., Guilfoyle JT. (1999). Dimerization and DNA binding of auxin response factors. *The Plant Journal* 19(3): 309-319.
- Urao T., Noji Masa-aki, Yamaguchi-Shinozaki K., Shinozaki K. (1996). A transcriptional activation domain of ATMYB2, a drought-inducible *Arabidopsis* Myb-related protein. *The Plant Journal* 10(6): 1145-1148.
- Van der Hoorn RAL., Wulff BBH., Rivas S., Durrant MC., van der Ploeg A., de Wit JGM., Jones JDG. (2005). Structure-Function Analysis of Cf-9, a Receptor-like protein with extracitoplasmic Leucine Rich Repeat. *The Plant Cell* 17:1000-1015.
- Wang Z., Ueda T., Messing J. (1998). Characterization of the maize prolamin box-binding factor-1 (PBF-1) and its role in the developmental regulation of the zein multigene family. *Gene* 223(1-2):321-32.
- Wang ZY., Kenigsbuch D., Sun L., Harel E., Ong Ms., Tobin EM. (1997). A Myb-Related transcription factors is involved in the phytochrome Regulation of an Arabidopsis *LHCB* gene. *The Plant Cell* 9:491-507.
- Washio K. (2003). Functional Dissections between GAMYB and Dof Transcription Factors Suggest a Role for Protein-Protein associations in the Gibberellin-Mediated Expression of the *RAmy1A* gene in rice *Aleurone*. *Plant Physiology* 133:850-863.
- Wellmer F., Riechemann JL., Alves-Ferreira M., Meyerowitz EM. 2004. Genome-wide analysis of spatial gene expression in *Arabidopsis* flowers. *The Plant Cell* 16 (5):1314-1326.
- Yanagisawa S. (2004). Dof Domain Proteins: Plant-Specific Transcription Factors Associated with Diverse Phenomena Unique to Plants. *Plant Cell Physiology* 45 (4):386-391.
- Yu D., Chen C., Chen Z. (2001). Evidence for an important role of WRKY DNA binding proteins in the regulation of *NPR1* gene expression. *The Plant Cell* 13 1527-1539.
- Zentella A & Alcántara RH. (2003). Importancia de los dominios de interacción proteica en la formación de complejos en los sistemas de transducción. *REB.* 22(3):117-129.



- Zhou H-L., Cao W-H., Cao Y-R., Liu J., Hao Y-J., Zhang J-S., Chen S-Y. (2006). Roles of ethylene receptor NTHK1 domains in plant growth stress response and protein phosphorylation. *FEBS Letters* 580: 1239-1250.
- Zúñiga E. (2004). Estudio de las posibles funciones de dos proteínas que interactúan *in vitro* con un factor de transcripción de la familia MADS: establecimiento de las líneas mutantes homocigas. Tesis de licenciatura. UNAM.

## **ANEXO I**

Anatomy	# of Chips	Mean	Std_Error	260553_at AT2G41800 Linear				
				0	6000	12000	18000	24000
1 cell suspension	<a href="#">42</a>	268	105					
2 seedling	<a href="#">386</a>	715	35	■				
21 cotyledons	<a href="#">3</a>	74	37					
22 hypocotyl	<a href="#">3</a>	115	14					
23 radicle	<a href="#">3</a>	3510	66					
3 inflorescence	<a href="#">139</a>	309	49					
31 flower	<a href="#">58</a>	429	92					
311 carpel	<a href="#">6</a>	2025	441					
312 petal	<a href="#">6</a>	18	6					
313 sepal	<a href="#">6</a>	75	26					
314 stamen	<a href="#">8</a>	22	4					
3141 pollen	<a href="#">2</a>	37	0					
315 pedicel	<a href="#">3</a>	38	12					
32 silique	<a href="#">11</a>	71	13					
33 seed	<a href="#">32</a>	398	120					
34 stem	<a href="#">7</a>	28	7					
35 node	<a href="#">3</a>	12	3					
36 shoot apex	<a href="#">17</a>	42	9					
37 cauline leaf	<a href="#">3</a>	84	15					
4 rosette	<a href="#">604</a>	180	26	■				
41 juvenile leaf	<a href="#">86</a>	165	57					
42 adult leaf	<a href="#">189</a>	326	76					
43 petiole	<a href="#">12</a>	50	11					
44 senescent leaf	<a href="#">3</a>	67	18					
45 hypocotyl	<a href="#">9</a>	102	32					
451 xylem	<a href="#">3</a>	120	94					
452 cork	<a href="#">3</a>	40	8					
5 roots	<a href="#">200</a>	2702	259	■				
52 lateral root	<a href="#">4</a>	830	285					
53 root tip	<a href="#">4</a>	20729	3176					
54 elongation zone	<a href="#">4</a>							
55 root hair zone	<a href="#">4</a>	2045	596					



Patrones de expresión de los genes pertenecientes a la familia DUF642 obtenidos a partir del Atlas de Arabidopsis en la dirección de internet <http://www.geninvesti3009.ethz.ch/> (Cuadros de la pagina 67-75 genes DUF642, pag. 76 expresión de la PME Q9LUL7).

Anatomy	# of Chips	Mean	Std_Error	260495_at AT2G41810 Linear				
				0	725	1450	2175	2900
1 cell suspension	<a href="#">42</a>	109	20					
2 seedling	<a href="#">386</a>	165	11					
21 cotyledons	<a href="#">3</a>	159	45					
22 hypocotyl	<a href="#">3</a>	144	5					
23 radicle	<a href="#">3</a>	924	65					
3 inflorescence	<a href="#">139</a>	129	7					
31 flower	<a href="#">58</a>	107	11					
311 carpel	<a href="#">6</a>	99	30					
312 petal	<a href="#">6</a>	91	19					
313 sepal	<a href="#">6</a>	91	8					
314 stamen	<a href="#">8</a>	168	59					
3141 pollen	<a href="#">2</a>	412	98					
315 pedicel	<a href="#">3</a>	157	19					
32 silique	<a href="#">11</a>	102	19					
33 seed	<a href="#">32</a>	163	15					
34 stem	<a href="#">7</a>	102	25					
35 node	<a href="#">3</a>	78	21					
36 shoot apex	<a href="#">17</a>	176	24					
37 cauline leaf	<a href="#">3</a>	97	13					
4 rosette	<a href="#">605</a>	137	5					
41 juvenile leaf	<a href="#">86</a>	144	22					
42 adult leaf	<a href="#">189</a>	123	6					
43 petiole	<a href="#">12</a>	258	26					
44 senescent leaf	<a href="#">3</a>	111	30					
45 hypocotyl	<a href="#">9</a>	78	23					
451 xylem	<a href="#">3</a>	70	19					
452 cork	<a href="#">3</a>	17	7					
5 roots	<a href="#">200</a>	465	61					
52 lateral root	<a href="#">4</a>	243	27					
53 root tip	<a href="#">4</a>	1765	354					
54 elongation zone	<a href="#">7</a>	2144	688					

55 root hair zone	<a href="#">4</a>	1179	107	
-------------------	-------------------	------	-----	--

Anatomy	# of Chips	Mean	Std_Error	246919_at AT5G25460 Linear				
				0	6750	13500	20250	27000
1 cell suspension	<a href="#">42</a>	9731	1165					
2 seedling	<a href="#">386</a>	12587	270					
21 cotyledons	<a href="#">3</a>	4447	62					
22 hypocotyl	<a href="#">3</a>	7555	207					
23 radicle	<a href="#">3</a>	1710	18					
3 inflorescence	<a href="#">139</a>	8148	644					
31 flower	<a href="#">58</a>	13083	1019					
311 carpel	<a href="#">6</a>	10715	765					
312 petal	<a href="#">6</a>	24266	2092					
313 sepal	<a href="#">6</a>	6111	2443					
314 stamen	<a href="#">8</a>	12212	4292					
3141 pollen	<a href="#">2</a>	36	15					
315 pedicel	<a href="#">3</a>	19665	349					
32 silique	<a href="#">11</a>	3614	1586					
33 seed	<a href="#">32</a>	2735	764					
34 stem	<a href="#">7</a>	10429	2385					
35 node	<a href="#">3</a>	5575	35					
36 shoot apex	<a href="#">17</a>	3269	353					
37 cauline leaf	<a href="#">3</a>	3012	101					
4 rosette	<a href="#">605</a>	6515	173					
41 juvenile leaf	<a href="#">86</a>	5264	368					
42 adult leaf	<a href="#">189</a>	4970	282					
43 petiole	<a href="#">12</a>	17790	1244					
44 senescent leaf	<a href="#">3</a>	20	6					
45 hypocotyl	<a href="#">9</a>	28	5					
451 xylem	<a href="#">3</a>	16	3					
452 cork	<a href="#">3</a>	41	9					
5 roots	<a href="#">200</a>	9269	318					
52 lateral root	<a href="#">4</a>	3775	807					
53 root tip	<a href="#">4</a>	409	116					

54 elongation zone	<a href="#">7</a>	7356	1527	
55 root hair zone	<a href="#">4</a>	4999	1001	

Anatomy	# of Chips	Mean	Std_Error	266956_at AT2G34510 Linear				
				0	925	1850	2775	3700
1 cell suspension	<a href="#">42</a>	514	142					
2 seedling	<a href="#">386</a>	1204	45					
21 cotyledons	<a href="#">3</a>	3604	37					
22 hypocotyl	<a href="#">3</a>	780	30					
23 radicle	<a href="#">3</a>	455	18					
3 inflorescence	<a href="#">139</a>	915	52					
31 flower	<a href="#">58</a>	949	78					
311 carpel	<a href="#">6</a>	1674	158					
312 petal	<a href="#">6</a>	281	46					
313 sepal	<a href="#">6</a>	524	47					
314 stamen	<a href="#">8</a>	179	25					
3141 pollen	<a href="#">2</a>	142	75					
315 pedicel	<a href="#">3</a>	1641	60					
32 silique	<a href="#">11</a>	1132	112					
33 seed	<a href="#">32</a>	316	34					
34 stem	<a href="#">7</a>	1477	360					
35 node	<a href="#">3</a>	775	54					
36 shoot apex	<a href="#">17</a>	1391	78					
37 cauline leaf	<a href="#">3</a>	1817	70					
4 rosette	<a href="#">605</a>	2127	70					
41 juvenile leaf	<a href="#">86</a>	2049	121					
42 adult leaf	<a href="#">189</a>	2292	172					
43 petiole	<a href="#">12</a>	1448	281					
44 senescent leaf	<a href="#">3</a>	415	23					
45 hypocotyl	<a href="#">9</a>	689	32					
451 xylem	<a href="#">3</a>	633	50					
452 cork	<a href="#">3</a>	699	22					
5 roots	<a href="#">200</a>	173	7					
52 lateral root	<a href="#">4</a>	453	189					






53 root tip	<a href="#">4</a>	118	48	
54 elongation zone	<a href="#">7</a>	106	21	
55 root hair zone	<a href="#">4</a>	259	56	

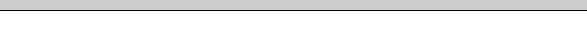

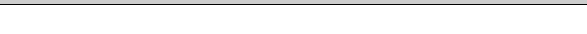










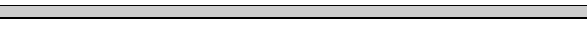












Anatomy	# of Chips	Mean	Std_Error	260028_at AT1G29980 Linear				
				0	1525	3050	4575	6100
1 cell suspension	<a href="#">42</a>	3858	496					
2 seedling	<a href="#">386</a>	1576	28					
21 cotyledons	<a href="#">3</a>	673	24					
22 hypocotyl	<a href="#">3</a>	3354	63					
23 radicle	<a href="#">3</a>	3499	41					
3 inflorescence	<a href="#">139</a>	2909	200					
31 flower	<a href="#">58</a>	2007	181					
311 carpel	<a href="#">6</a>	3583	502					
312 petal	<a href="#">6</a>	1096	341					
313 sepal	<a href="#">6</a>	474	43					
314 stamen	<a href="#">8</a>	320	46					
3141 pollen	<a href="#">2</a>	202	162					
315 pedicel	<a href="#">3</a>	1334	84					
32 silique	<a href="#">11</a>	1485	163					
33 seed	<a href="#">32</a>	4155	579					
34 stem	<a href="#">7</a>	1499	383					
35 node	<a href="#">3</a>	2852	53					
36 shoot apex	<a href="#">17</a>	5897	177					
37 cauline leaf	<a href="#">3</a>	382	12					
4 rosette	<a href="#">605</a>	1312	30					
41 juvenile leaf	<a href="#">86</a>	1989	121					
42 adult leaf	<a href="#">189</a>	1287	51					
43 petiole	<a href="#">12</a>	1443	165					
44 senescent leaf	<a href="#">3</a>	161	31					
45 hypocotyl	<a href="#">9</a>	907	177					
451 xylem	<a href="#">3</a>	298	46					
452 cork	<a href="#">3</a>	902	4					
5 roots	<a href="#">200</a>	2162	60					

52 lateral root	<a href="#">4</a>	1175	178	
53 root tip	<a href="#">4</a>	4239	359	
54 elongation zone	<a href="#">7</a>	1428	188	
55 root hair zone	<a href="#">4</a>	820	142	

Anatomy	# of Chips	Mean	Std_Error	258641_at AT3G08030 Linear				
				0	3750	7500	11250	15000
1 cell suspension	<a href="#">42</a>	12608	1422					
2 seedling	<a href="#">386</a>	7073	188					
21 cotyledons	<a href="#">3</a>	3867	69					
22 hypocotyl	<a href="#">3</a>	12193	265					
23 radicle	<a href="#">3</a>	7304	200					
3 inflorescence	<a href="#">139</a>	5788	330					
31 flower	<a href="#">58</a>	5398	424					
311 carpel	<a href="#">6</a>	8691	192					
312 petal	<a href="#">6</a>	2686	1019					
313 sepal	<a href="#">6</a>	2466	963					
314 stamen	<a href="#">8</a>	758	210					
3141 pollen	<a href="#">2</a>	496	171					
315 pedicel	<a href="#">3</a>	6922	62					
32 silique	<a href="#">11</a>	3997	523					
33 seed	<a href="#">32</a>	4355	847					
34 stem	<a href="#">7</a>	8285	2432					
35 node	<a href="#">3</a>	6304	67					
36 shoot apex	<a href="#">17</a>	9677	115					
37 cauline leaf	<a href="#">3</a>	1387	28					
4 rosette	<a href="#">605</a>	4866	101					
41 juvenile leaf	<a href="#">86</a>	6868	336					
42 adult leaf	<a href="#">189</a>	3755	169					
43 petiole	<a href="#">12</a>	7671	735					
44 senescent leaf	<a href="#">3</a>	275	10					
45 hypocotyl	<a href="#">9</a>	1309	237					
451 xylem	<a href="#">3</a>	572	46					
452 cork	<a href="#">3</a>	1165	46					



5 roots	<a href="#">200</a>	5054	79	
52 lateral root	<a href="#">4</a>	7000	1189	
53 root tip	<a href="#">4</a>	4146	550	
54 elongation zone	<a href="#">7</a>	6483	797	
55 root hair zone	<a href="#">4</a>	5067	935	

Anatomy	# of Chips	Mean	Std_Error	253437_at AT4G32460 Linear				
				0	7000	14000	21000	28000
1 cell suspension	<a href="#">42</a>	1274	131					
2 seedling	<a href="#">386</a>	5785	263					
21 cotyledons	<a href="#">3</a>	693	7					
22 hypocotyl	<a href="#">3</a>	1007	34					
23 radicle	<a href="#">3</a>	2044	47					
3 inflorescence	<a href="#">139</a>	5231	517					
31 flower	<a href="#">58</a>	8393	995					
311 carpel	<a href="#">6</a>	4782	252					
312 petal	<a href="#">6</a>	25939	1220					
313 sepal	<a href="#">6</a>	1151	183					
314 stamen	<a href="#">8</a>	10749	3283					
3141 pollen	<a href="#">2</a>	559	100					
315 pedicel	<a href="#">3</a>	5095	315					
32 silique	<a href="#">11</a>	1370	327					
33 seed	<a href="#">32</a>	3184	768					
34 stem	<a href="#">7</a>	2318	440					
35 node	<a href="#">3</a>	1305	48					
36 shoot apex	<a href="#">17</a>	2823	121					
37 cauline leaf	<a href="#">3</a>	637	63					
4 rosette	<a href="#">605</a>	1445	52					
41 juvenile leaf	<a href="#">86</a>	2126	137					
42 adult leaf	<a href="#">189</a>	957	51					
43 petiole	<a href="#">12</a>	6369	1450					
44 senescent leaf	<a href="#">3</a>	509	37					
45 hypocotyl	<a href="#">9</a>	1107	50					
451 xylem	<a href="#">3</a>	1095	103					

452 cork	<a href="#">3</a>	1215	87	
5 roots	<a href="#">200</a>	2578	153	
52 lateral root	<a href="#">4</a>	3300	1160	
53 root tip	<a href="#">4</a>	1248	103	
54 elongation zone	<a href="#">7</a>	10448	2519	
55 root hair zone	<a href="#">4</a>	7326	1547	




Anatomy	# of Chips	Mean	Std_Error	262045_at AT1G80240 Linear				
				0	1525	3050	4575	6100
1 cell suspension	<a href="#">42</a>	1517	213					
2 seedling	<a href="#">386</a>	1171	63					
21 cotyledons	<a href="#">3</a>	13	2					
22 hypocotyl	<a href="#">3</a>	7	1					
23 radicle	<a href="#">3</a>	1617	39					
3 inflorescence	<a href="#">139</a>	77	18					
31 flower	<a href="#">58</a>	119	37					
311 carpel	<a href="#">6</a>	8	1					
312 petal	<a href="#">6</a>	637	283					
313 sepal	<a href="#">6</a>	16	4					
314 stamen	<a href="#">8</a>	156	42					
3141 pollen	<a href="#">2</a>	20	6					
315 pedicel	<a href="#">3</a>	36	20					
32 silique	<a href="#">11</a>	9	1					
33 seed	<a href="#">32</a>	81	37					
34 stem	<a href="#">7</a>	66	42					
35 node	<a href="#">3</a>	10	1					
36 shoot apex	<a href="#">17</a>	9	1					
37 cauline leaf	<a href="#">3</a>	10	3					
4 rosette	<a href="#">605</a>	41	12					
41 juvenile leaf	<a href="#">86</a>	149	79					
42 adult leaf	<a href="#">189</a>	18	1					
43 petiole	<a href="#">12</a>	25	2					
44 senescent leaf	<a href="#">3</a>	6	1					
45 hypocotyl	<a href="#">9</a>	11	1					

451 xylem	<a href="#">3</a>	10	2	
452 cork	<a href="#">3</a>	10	2	
5 roots	<a href="#">200</a>	3989	142	
52 lateral root	<a href="#">4</a>	3691	1948	
53 root tip	<a href="#">4</a>	2070	370	
54 elongation zone	<a href="#">7</a>	3884	853	
55 root hair zone	<a href="#">4</a>	4817	1270	

Anatomy	# of Chips	Mean	Std_Error	250366_at AT5G11420 Linear				
				0	4250	8500	12750	17000
1 cell suspension	<a href="#">42</a>	580	253					
2 seedling	<a href="#">386</a>	6410	201					
21 cotyledons	<a href="#">3</a>	5714	105					
22 hypocotyl	<a href="#">3</a>	8164	251					
23 radicle	<a href="#">3</a>	1228	33					
3 inflorescence	<a href="#">139</a>	3531	359					
31 flower	<a href="#">58</a>	4664	667					
311 carpel	<a href="#">6</a>	1608	463					
312 petal	<a href="#">6</a>	16318	430					
313 sepal	<a href="#">6</a>	647	167					
314 stamen	<a href="#">8</a>	5777	1973					
3141 pollen	<a href="#">2</a>	89	9					
315 pedicel	<a href="#">3</a>	7983	167					
32 silique	<a href="#">11</a>	738	310					
33 seed	<a href="#">32</a>	2631	814					
34 stem	<a href="#">7</a>	3058	605					
35 node	<a href="#">3</a>	2584	65					
36 shoot apex	<a href="#">17</a>	2383	365					
37 cauline leaf	<a href="#">3</a>	2910	27					
4 rosette	<a href="#">605</a>	5713	132					
41 juvenile leaf	<a href="#">86</a>	4395	235					
42 adult leaf	<a href="#">189</a>	4596	201					

43 petiole	<a href="#">12</a>	15123	1232	
44 senescent leaf	<a href="#">3</a>	294	16	
45 hypocotyl	<a href="#">9</a>	17	3	
451 xylem	<a href="#">3</a>	18	5	
452 cork	<a href="#">3</a>	25	3	
5 roots	<a href="#">200</a>	535	24	■
52 lateral root	<a href="#">4</a>	949	244	
53 root tip	<a href="#">4</a>	23	11	
54 elongation zone	<a href="#">7</a>	1273	186	
55 root hair zone	<a href="#">4</a>	897	266	
Anatomy	# of Chips	Mean	Std_Error	250224_at AT5G14150 Linear 0 525 1050 1575 2100
1 cell suspension	<a href="#">42</a>	227	26	
2 seedling	<a href="#">386</a>	323	9	■
21 cotyledons	<a href="#">3</a>	34	11	
22 hypocotyl	<a href="#">3</a>	142	6	
23 radicle	<a href="#">3</a>	853	44	
3 inflorescence	<a href="#">139</a>	238	16	
31 flower	<a href="#">58</a>	322	28	
311 carpel	<a href="#">6</a>	619	32	
312 petal	<a href="#">6</a>	310	64	
313 sepal	<a href="#">6</a>	75	18	
314 stamen	<a href="#">8</a>	330	82	
3141 pollen	<a href="#">2</a>	36	1	
315 pedicel	<a href="#">3</a>	68	29	
32 silique	<a href="#">11</a>	243	69	
33 seed	<a href="#">32</a>	217	30	
34 stem	<a href="#">7</a>	92	21	
35 node	<a href="#">3</a>	72	22	
36 shoot apex	<a href="#">17</a>	122	10	
37 cauline leaf	<a href="#">3</a>	90	12	
4 rosette	<a href="#">605</a>	73	3	■
41 juvenile leaf	<a href="#">86</a>	145	11	
42 adult leaf	<a href="#">189</a>	66	4	

43 petiole	<a href="#">12</a>	143	23	
44 senescent leaf	<a href="#">3</a>	72	9	
45 hypocotyl	<a href="#">9</a>	183	28	
451 xylem	<a href="#">3</a>	191	80	
452 cork	<a href="#">3</a>	145	34	
5 roots	<a href="#">200</a>	797	26	
52 lateral root	<a href="#">4</a>	1290	516	
53 root tip	<a href="#">4</a>	694	83	
54 elongation zone	<a href="#">7</a>	1911	169	
55 root hair zone	<a href="#">4</a>	1109	193	

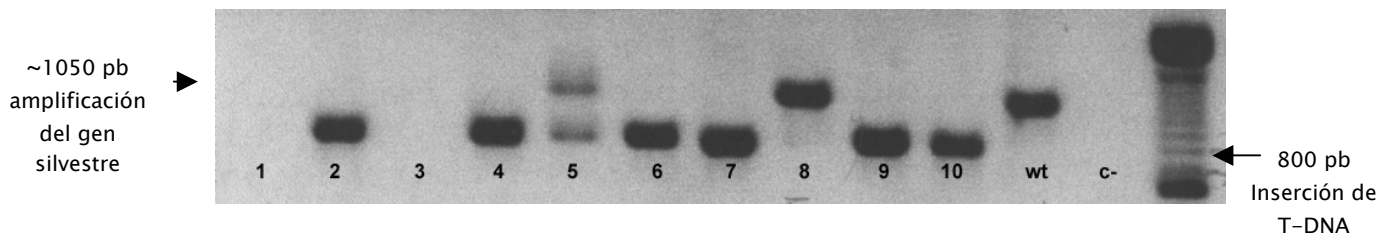
Anatomy	# of Chips	Mean	Std_Error	258369_at AT3G14310 Linear				
				0	6000	12000	18000	24000
1 cell suspension	<a href="#">42</a>	5034	661					
2 seedling	<a href="#">386</a>	10300	244					
21 cotyledons	<a href="#">3</a>	13626	575					
22 hypocotyl	<a href="#">3</a>	23061	586					
23 radicle	<a href="#">3</a>	14626	273					
3 inflorescence	<a href="#">139</a>	4885	311					
31 flower	<a href="#">58</a>	5037	286					
311 carpel	<a href="#">6</a>	4198	106					
312 petal	<a href="#">6</a>	6978	885					
313 sepal	<a href="#">6</a>	3872	1257					
314 stamen	<a href="#">8</a>	3871	986					
3141 pollen	<a href="#">2</a>	256	69					
315 pedicel	<a href="#">3</a>	7545	238					
32 silique	<a href="#">11</a>	1251	247					
33 seed	<a href="#">32</a>	3135	842					
34 stem	<a href="#">7</a>	9824	1934					
35 node	<a href="#">3</a>	8253	121					
36 shoot apex	<a href="#">17</a>	7397	530					
37 cauline leaf	<a href="#">3</a>	9385	112					
4 rosette	<a href="#">605</a>	9968	233					
41 juvenile leaf	<a href="#">86</a>	8735	331					
42 adult leaf	<a href="#">189</a>	11826	436					
43 petiole	<a href="#">12</a>	20546	971					
44 senescent leaf	<a href="#">3</a>	2554	151					
45 hypocotyl	<a href="#">9</a>	1229	270					
451 xylem	<a href="#">3</a>	232	25					
452 cork	<a href="#">3</a>	1399	99					
5 roots	<a href="#">200</a>	6900	238					
52 lateral root	<a href="#">4</a>	9057	2181					
53 root tip	<a href="#">4</a>	2567	175					
54 elongation zone	<a href="#">7</a>	9271	1214					
55 root hair zone	<a href="#">4</a>	14548	648					

## Anexo II

### Análisis de la generación T1 de las plantas mutantes por inserción de T-DNA en el locus *At4g32460*.

Una de las estrategias que ha sido de gran utilidad para la caracterización de muchos genes en *A. thaliana*, ha sido la genética reversa. Esta se basa en mutar el gen que se desea estudiar, esperando que la mutación pueda dar pistas sobre sus funciones a partir del fenotipo.

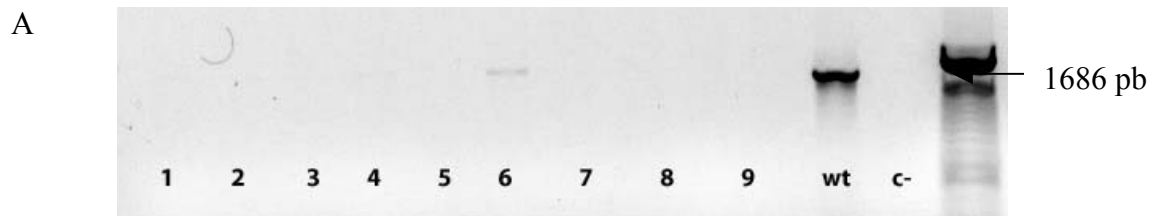
Para ello se realizó el análisis de plantas mutantes por inserción de T-DNA en diferentes genes pertenecientes a la familia DUF642 con el fin de obtener líneas mutantes homocigotas. En la figura I se muestra el análisis que se realizó para la primera generación de plantas (T1) en los genes *At4g32460*.

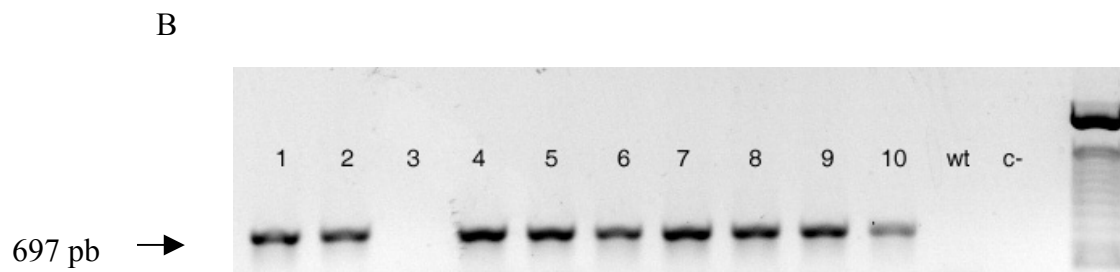


**Figura I:** Búsqueda de plantas mutantes por inserción de T-DNA en el locus *At4g32460* (generación T1). Las reacciones de PCR se amplificaron a partir de DNA genómico de hojas de roseta de 10 plantas mutantes y una planta silvestre. Los oligonucleótidos utilizados para las amplificaciones fueron el LBb1, MET32460 y FIN32460.

### Obtención de una línea homocigota mutante por inserción de T-DNA en el locus *At4g32460*

En la figura II se muestra la segunda generación de plantas mutantes por inserción de T-DNA en el locus *At4g32460*. La segregación de los caracteres permitiría reconocer una línea homocigota para la mutación en este gen.

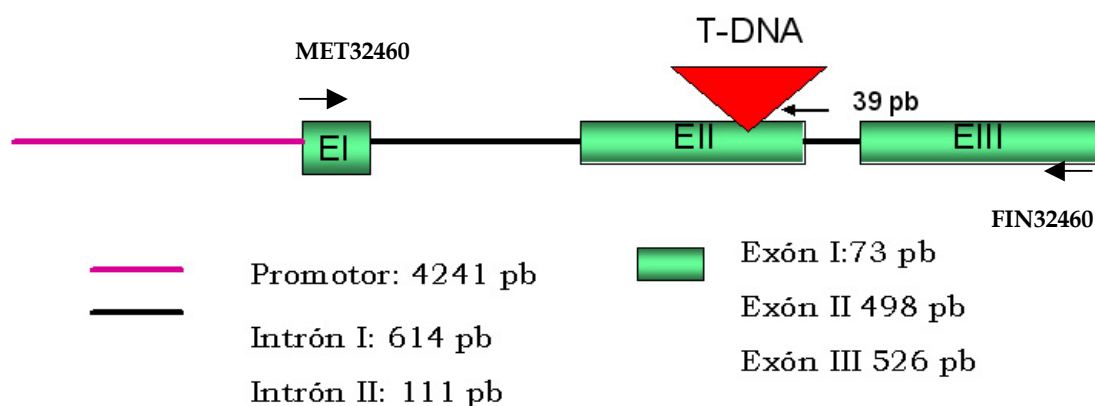




**Figura II:** Análisis de la generación T2 de plantas mutantes por inserción de T-DNA en el locus At4g32460. Se colectaron hojas de roseta de 10 plantas mutantes y una silvestre. **A** Reacciones de PCR con los oligos FAT32460 (primer sentido) y RAT32460 (primer antisentido) que permiten la amplificación del gen completo. **B** Reacciones de PCR con los oligos LB1 (borde izquierdo del T-DNA) y RAT32460.

### Localización del sitio de inserción del T-DNA en una línea mutante homocigota para el gen At4g32460.

La búsqueda de la inserción de T-DNA en la primera generación se realizó para cada una de las líneas, sin embargo, fue imposible la amplificación del borde izquierdo del T-DNA en las líneas mutantes At5g11420 y At2g41800, por lo tanto, únicamente se realizó el análisis de la primera generación (T1) de la línea mutante por inserción de T-DNA del SALK Institute para el locus At4g32460. Mediante el análisis de la secuencia de las bandas amplificadas, a partir de DNA genómico de hojas de roseta y los oligos correspondientes, se determinó que la inserción de T-DNA se localiza 39 pares de bases antes del término del segundo exón del gen (figura III). Una vez obtenida una línea homocigota para la inserción de T-DNA se determinaron los niveles de expresión en las plantas mutantes en este locus.

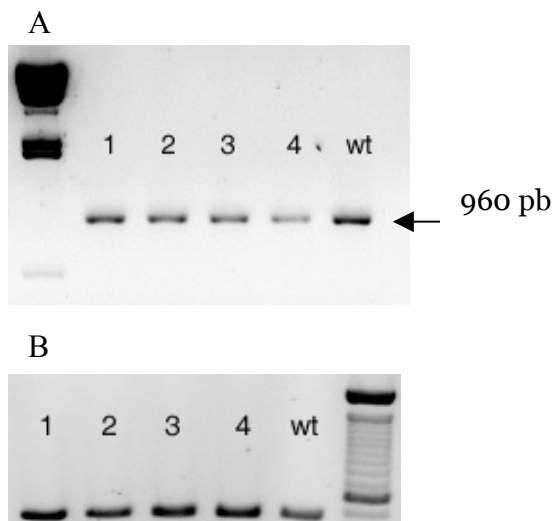


**Figura III:** Esquema que representa el sitio de inserción del T-DNA en el gen At4g32460. Este gen presenta un promotor de 4241 pb, tres exones y dos intrones. La inserción de T-DNA se localiza 39 bases antes del término del segundo exón. Los oligos utilizados para la amplificación del transcrito fueron MET32460 y FIN32460



## Expresión del gen At4g32460 en plantas mutantes homocigotas mediante análisis de RT-PCR

En el caso de la línea mutante en el gen At4g32460 fue posible la obtención de una línea homocigota para este locus. El análisis de expresión para cuatro plantas mutantes de la generación T2 (Figura IV-A) mostró que no hay una diferencia considerable en los niveles de expresión del gen At4g32460 con respecto a la planta silvestre; sin embargo, es necesario realizar un análisis cuantitativo para determinar si la cantidad de mensajero está afectada por dicha inserción. Como se observa en la figura IV-A, el tamaño del fragmento amplificado es el mismo en las plantas mutantes en comparación con el transcrito silvestre. El análisis de la secuencia amplificada a partir de las plantas mutantes demostró que no existe ninguna diferencia entre la secuencia codificante amplificada a partir de plantas mutantes con respecto a la de las plantas silvestres. Este tipo de resultado es raro, debido a que la inserción se localizó en el segundo exón de acuerdo con los análisis de secuenciación que se realizaron. Es posible que a pesar de la inserción de T-DNA el gen se edite correctamente. Otra posibilidad es la existencia de una copia del gen en alguna otra parte del genoma, ya que no se hallaron modificaciones en el transcrito. Es necesario realizar la búsqueda de una nueva línea mutante para este gen e implementar nuevas estrategias para el análisis y la obtención de las líneas homocigotas para los genes At5g11420 y At2g41800, a partir de los lotes de semillas ya existentes en el laboratorio, y en su defecto obtener semillas mutantes de otros bancos.



**Figura IV:** **A** Análisis de expresión de 4 plantas mutantes homocigotas para la inserción de T-DNA en el gen At4g32460 mediante RT-PCR a partir de cDNA de hojas de roseta (1-4) y una planta silvestre (wt). La amplificación de la secuencia codificante se realizó con los oligos MET32460 y FIN32460. **B** control de tubulina.

Tabla I: Características de los motivos encontrados en los promotores de los genes DUF642 de *Arabidopsis thaliana*.

<b>Factor que reconoce</b>	<b>Secuencia reconocida</b>	<b>Promotores (loci)</b>	<b>Características</b>	<b>Referencia</b>
<u>CArg promoter motif</u>	ccatatatgg	A2g41800	Regulación por factores de transcripción de tipo MADS-box.	Hepworth SR, <i>et al.</i> , 2002, EMBO J. 21: 4327-4337.
<u>Bellringer/re IN AG. Homeobox</u>	aaattaa y actaattt	At2g41800 At2g41810 At4g32460 At5g25460 At1g29980 At5g14150	La proteína homeodominio BELLRINGER regula la expresión espacio-temporal de AGAMOUS en óvulos	Bao X <i>et al.</i> , 2004, Plant Cell 16: 1478-1489.
Elk-1	NNMMGGAWNN	At5g11420 At4g32460 At5g25460 At1g29980 At1g80240	Formación de complejos ternarios entre el factor de transcripción SRF (Serim Response Factor) y su factor accesorio Elk-1	Treisman R. <i>et al.</i> , 1992, EMBO J. 11:4631-4640
<u>LFY consensus binding site motif</u>	ccagtg	A2g41800 At2g41810 At1g80240 At4g32460 At5g11420 At5g25460 At1g29980 At2g34510	Regulación transcripcional de LMI1, un regulador de la identidad del meristemo, por LEAFY un factor de transcripción MADS-box	Saddic L. A. <i>et al.</i> , 2006, Development 133:1673-1682.
<u>ATB2/AtbZ1 P53/AtbZIP4 4/GBF5 BS in ProDH</u>	actcat	At5g11420 At5g25460 At3g08030 At2g34510	La secuencias ACTCAT actúa como un elemento de regulación <i>cis</i> involucrado en la expresión del gen <i>ProDH</i> ( <i>Proline dehydrogenase</i> ) bajo condiciones de hipoosmolaridad y respuesta a prolina	Satoh, <i>et al.</i> , 2004, Plant Cell Physiol. 45(3):300-317.
<u>SORLIP1</u>	agccac TT[AT]TACTAGT	At1g80240 At5g25460 At3g08030 At1g29980	Motivos Involucrados en la regulación de la expresión genética mediada por fitocromo A	Hudson M <i>et al.</i> , 2003, Plant Physiology 133: 1605-1616.
<u>SORLIP2</u>	gggcc	At2g41800 At2g41810 At4g32460 At5g11420 At5g25460 At3g08030 At1g29980 At5g25460		
<u>SORLIP3</u>	tgtatatat	At5g11420		
<u>SORLIP4</u>	gtatgatgg	At2g34510		

<u>SORLIP5</u>	gagtgag	At5g25460 At1g29980 At2g34510 At5g14150		
<u>lbox promoter motif</u>	gataag	At2g41800 At2g41810 At4g32460 At5g11420 At5g25460 At3g08030	Secuencias de reconocimiento en genes regulados por luz	Giuliano G, <i>et al.</i> , 1988, Proc. Natl. Acad. Sci. USA. 85:7089-7093.
<u>Z-box promoter motif</u>	atacgtgt	At2g41800 At3g08030	Elementos en <i>cis</i> involucrados en la regulación de la expresión órgano específica y regulación inducida por luz.	Ha SB, <i>et al.</i> , 1988, Proc. Natl. Acad. Sci. USA 85: 8017-8021.
<u>T-box promoter motif</u>	actttg	At2g41800 At2g41810 At1g80240 At4g32460 At5g11420 At5g25460 At3g08030 At2g34510	Regiones repetidas relacionadas con la activación de la expresión mediada por luz	Chan CS <i>et al.</i> , 2001, Plant Mol. Biol. 46: 131-141.
<u>BoxII promoter motif</u>	ggtaa	At2g41800 At1g80240 At4g32460 At5g11420	Elementos de regulación en respuesta a luz	Le Gourrierc J. <i>et al.</i> , 1999, Plant J. Jun; 18(6):663-8.
<u>GATA promoter motif [LRE]</u>	tgataa (A/T)GATA(G/A)	At2g41800 At2g41810 At1g80240 At4g32460 At5g11420 At5g25460 At3g08030 At1g29980 At2g34510	Sitios de unión GATA implicados en la regulación de la transcripción en respuesta a luz.	Teakle GR 2002, Plant Mol Biol. Sep; 50(1):43-57.
<u>CCA1 binding site motif</u>	aaaaatct	At2g41800 At2g41810 At5g25460 At3g08030 At1g29980 At2g34510	Regulación de la expresión genética mediada por CCA1, una proteína clave en la ruta de transducción de señales mediada por fitocromo	Wang Z-Y <i>et al.</i> , 1997, Plant Cell 9:491-507.
Opaque	NTGAGTCANN	At2g41800 At5g11420 At5g25460 At1g80240	Elementos <i>cis</i> involucrados en la expresión endospermo-específica	Onodera <i>et al.</i> , 2001 J. Biol. Chem. 276: 14139-14152.
<u>RY-repeat promoter motif</u>	catgcatg	At4g32460 At5g11420 At3g08030	Regiones reguladoras encontradas en los genes que codifican para proteínas de semilla.	Dickinson CD <i>et al.</i> , 1988, Nucleic Acids Res. 16:371.
<u>L1-box</u>	taaagtga	At2g41800 At4g32460	Elementos <i>cis</i> regulatorios	Abe M <i>et al.</i> , 2001, Plant J. 26: 487-

<u>promoter motif</u>	TAAATG(C/T)A	At5g11420 At5g25460	encontrados en genes que se expresan en la capa L1 del meristemo y los primordios de órganos	494.
<u>G-box promoter motif</u>	cacgtg	At5g11420 At5g25460 At2g41800	Los elementos G-box han sido identificados como blancos en genes regulados por luz, genes regulados por ABA y genes inducidos bajo condiciones de estrés	Menkens AE <i>et al.</i> , 1994, Proc. Natl. Acad. Sci. USA 91: 2522-2526.
<u>GBF1/2/3 BS in ADH1</u>	ccacgtgg	At2g41800 At5g11420 At5g25460	Secuencias G-box reconocidas por el factor de transcripción <i>GBF1</i> cuya expresión en es inducida por hipoxia	de Vetten N <i>et al.</i> , 1995, Plant J. 7:589-601.
<u>ABFs binding site motif</u>	cacgtggc	At2g41800	Elementos de regulación por ABA	Guiltinan MJ, 1990, Science 250:267-270.
<u>CBF2 binding site motif</u> <u>AP2-EREBP</u>	ccacgtgg	At2g41800 At5g11420 At5g25460	Elemento regulatorio en <i>cis</i> involucrado en la respuesta a estrés hídrico y ABA	Pla M <i>et al.</i> , 1993, Plant Mol Biol 21:259-266.
<u>ABRE binding site motif</u>	cacgtggc tacgtggc	At2g41800 At2g41810	Elementos regulatorios encontrados en genes cuya expresión es regulada por ABA	Choi H <i>et al.</i> , 2000, J Biol Chem. 275: 1723-1730.
<u>ABRE-like binding site motif</u> <u>GCC-box promoter motif</u>	tacgtgta tacgtgtc  gacgtggc  gccgcc	At2g41800 At5g11420 At3g08030   At1g80240 At5g11420 At3g08030 At2g34510	Regulación de la expresión genética inducida por sequía mediada por ABA	Shinozaki K, <i>et al.</i> , 2000, Curr. Opin. Plant Biol. 3:217-223.
<u>LTRE promoter motif</u>	accgaca	At2g41810	Elementos involucrados en la inducción de la expresión genética mediada por bajas temperaturas	Nordin K <i>et al.</i> , 1993, Plant Mol Biol 21:641-653.
ABA	cattcy  y=pirimidina	At3g08030 At5g11420 At4g32460 At5g25460 At1g29980	Regulación de la expresión mediada por ABA en <i>Aspegillus nidulans</i>	Andrianopoulos A., Timberlake W. E, <i>et al.</i> , 1994, Mol. Cell. Biol. 14:2503-2515.

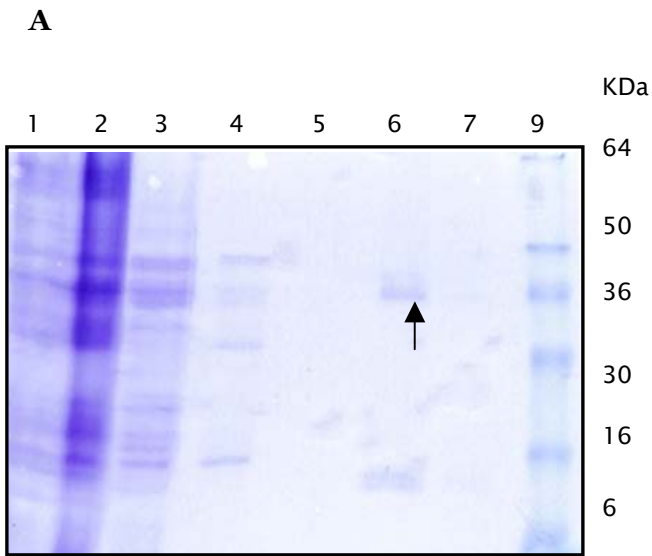
ABA	cattct  agaatg	At5g14150 At1g80240  At3g08030 At5g11420 At4g32460 At5g25460 Atg129980 At5g14150 At1g80240  At5g11420 At4g32460 At5g25460 At1g29980 At5g14150 At1g80240		
<u>AtMYC2 BS en RD22</u>	cacatg	At2g41800 At2g41810 At1g80240 At4g32460 At5g11420 At5g25460 At3g08030 At1g29980	Regulación de la expresión genética por factores de transcripción MYC y MYB en respuesta a ABA y sequía	Abe H. <i>et al.</i> , 1997, Plant Cell 9:1859-1868.
<u>AtMYB2 BS in RD22</u>	ctaacca	At2g41810 At1g80240 At4g32460 At3g08030		
<u>DPBF1&amp;2 binding site motif</u> bZip	acacatg	At2g41800 At2g41810 At4g32460 At5g11420 At5g25460 At3g08030 At1g29980	Motivos reconocidos por DPBF1 y 2 factores de transcripción inducidos por ABA	Kim SY, <i>et al.</i> , 1997, Plant J 11: 1237-1251.
BBF1	acttta	At3g08030 At2g41800 At5g11420 At4g32460 At1g29980 At2g34510 At5g14150 At1g80240	Elemento blanco necesario para la expresión tejido-específico y regulación por auxinas	Baumann K. <i>et al.</i> , 1999, Plant Cell 11:323-334.
<u>ARF1 binding site motif</u>	tgtctc	At2g41800 At1g80240 At4g32460 At5g11420 At5g25460 At3g08030 At1g29980 At2g34510	Elementos de respuesta a auxinas	Ulmasov T, <i>et al.</i> , 1999, Plant J 19:309-319.
<u>MYB binding site promoter</u>	caccaaac (A/C)ACC(A/T)A(A/C)C	At2g41800 At5g14150	Regulación de la expresión por un factor de transcripción MYB específico de flor	Sablowski RWM, <i>et al.</i> , 1994, EMBO J 13:128-137.

<u>MYB4</u> <u>binding site</u> <u>motif</u>	aacaaac	At2g41800 At2g41810 At1g80240 At4g32460 At5g11420 At5g25460 At3g08030 At1g29980	Elemento regulatorio reconocido por el factor de transcripción MYB4 en respuesta a estrés ambiental	Chen W <i>et al.</i> , 2002, <i>Plant Cell</i> . Mar; 14(3):559-574.
<u>DRE-like</u> <u>promoter</u> <u>motif</u>	tgccgacaa  (A/G/T)(A/G)CCGACN(A/T)	At2g41800 At1g80240 At5g11420	Elemento regulatorio reconocido por el factor de transcripción MYB4 en respuesta a estrés ambiental	Chen W <i>et al.</i> , 2002, <i>Plant Cell</i> 14:559-574.
<u>RAV1-A</u> <u>binding site</u> <u>motif</u> <u>ABI3VP1</u>	caaca	At2g41800 At2g41810 At4g32460 At5g11420 At5g25460 At3g08030 At1g29980 At2g34510	Motivos regulatorios reconocidos por el factor de transcripción RAV1 involucrado en la regulación bajo condiciones de estrés y patógenos	Kagaya Y, <i>et al.</i> , 1999, <i>Nucleic Acids Res</i> 27: 470-478.
<u>W-box</u> <u>promoter</u> <u>motif</u>	ttgact	At2g41800 At2g41810 At4g32460 At5g11420 At5g25460 At3g08030 At1g29980 At2g34510 At5g14150	Motivos regulatorios encontrados en genes cuya expresión es regulada por WRINK1, factor de transcripción que se induce específicamente por ácido salicílico	Yu D. <i>et al.</i> , 2001, <i>Plant Cell</i> 13: 1527-1540.
PBF	tgtaaag	At5g11420 At4g32460 At5g25460 At1g29980 At2g34510	Motivo de reconocimiento por el factor de transcripción PBF-1, que regula la expresión de genes zeina	Wang, <i>et al.</i> , 1998, <i>Gene</i> , 223:333-345.

## ***Expresión de proteínas***

### **Expresión y purificación de la proteína At5g11420 completa**

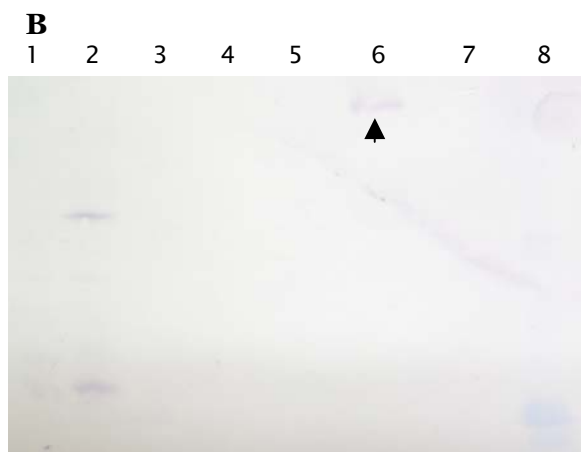
La expresión y purificación de la proteína recombinante At5g11420 completa se llevó a cabo utilizando la columna de cromatografía con agarosa Ni-NTA. En la figura V-A se observa el gel de poliacrilamida al 12% con las diferentes fracciones proteicas obtenidas desde la incubación del cultivo no inducido hasta la elusión de la proteína deseada. En el carril 6 podemos observar que la proteína se eluyó junto con otra proteína de menor peso molecular, el tamaño de esta proteína hizo suponer que probablemente se trataba de lisozima, la cual era abundante en el lisado por ser utilizada durante el proceso de lisis de las células bacterianas. La proteína recombinante correspondería a la banda de aproximadamente 36 kDa que se observa en este mismo carril.



**Figura V-A:** Purificación de la proteína recombinante At4g11420 completa. Gel de poliacrilamida al 12 % con azul de Coomasie.

- 1) Extracto crudo antes de inducir
- 2) Extracto crudo después de inducir
- 3) Extracto bacteriano purificado
- 4) Fracción no unida
- 5) Lavado 10 mM de imidazol
- 6) Elusión 175 mM de imidazol
- 7) Elusión 250 mM de imidazol
- 8) Marcador de peso molecular (kDa)

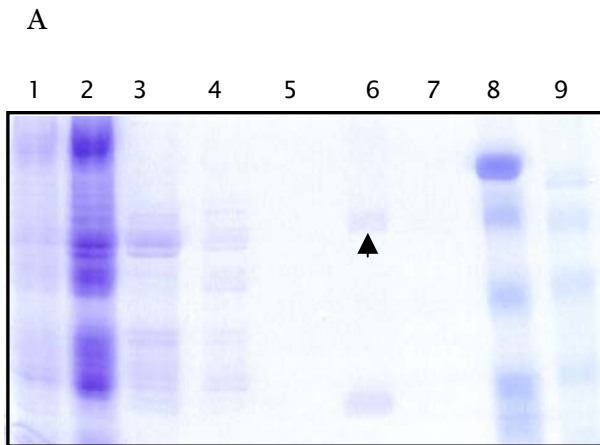
Con el fin de asegurarnos que la banda observada correspondía a la proteína sobreexpresada en *E. coli* se realizó un western-blot; en dicho ensayo se utilizó un conjugado con actividad de fosfatasa alcalina que reconoce colas de histidina. Este ensayo nos permitió confirmar que la banda de aproximadamente 36 kDa correspondía a la proteína recombinante. En la figura V-B se puede observar una leve detección. Se debe recordar que la fracción cargada para la transferencia y la detección era la mitad de la muestra pues primero era necesario percatarse en otro gel si la proteína se había purificado.



**Figura V-B:** Western-blot, detección de la proteína recombinante At5g11420 completa. Las fracciones fueron cargadas en el mismo orden del gel anterior.

## Expresión y purificación de la proteína At5g11420 truncada

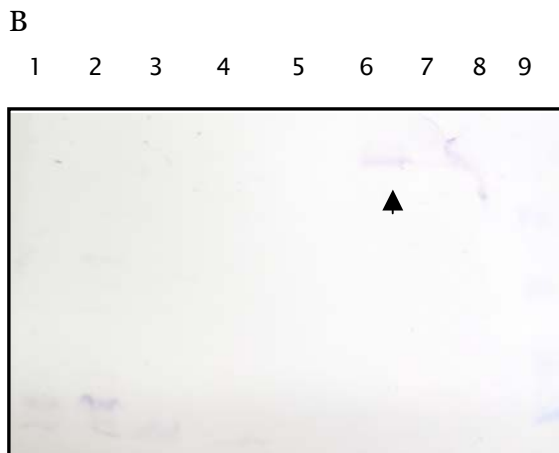
La expresión de la proteína truncada At5g11420 se realizó bajo las mismas condiciones que la proteína completa. Dicha proteína se expresó sin 69 aminoácidos del C-terminal. En el carril 6 de la figura VI-A podemos ver una banda de aproximadamente 30 kDa correspondiente a la proteína At5g11420 truncada y una proteína de aproximadamente 16 kDa.



**Figura VI-A:** Purificación de la proteína At5g11420 truncada. Gel de Poliacrilamida al 12%. Tinción con azul de coomassie

- 1) Extracto crudo antes de inducir
- 2) Extracto crudo después de inducir
- 3) Extracto bacteriano purificado
- 4) Fracción no unida
- 5) Lavado 10 mM de imidazol
- 6) Elución 175 mM de imidazol
- 7) Elución 250 mM de imidazol
- 8 y 9) Marcadores de peso molecular

En la figura VI-B se puede observar la detección de la banda correspondiente a la proteína At5f11420 truncada. La detección se realizó al igual que con la proteína completa con el conjugado que reconoce secuencias de histidina repetidas. El reconocimiento de estas secuencias cortas de histidina puede darse en otras proteínas tal como se ve en los dos primeros carriles con proteínas de bajo peso molecular de los cultivos celulares antes y después de inducir.



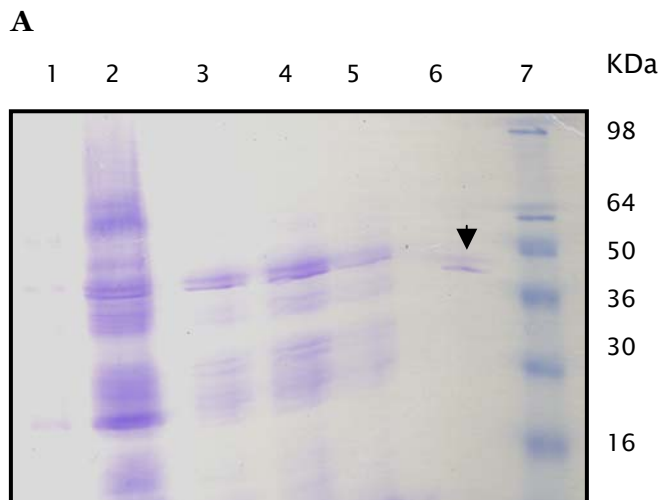
**Figura VI-B:** Western-Blot, detección de la proteína At5g11420 truncada. Las fracciones fueron cargadas en el mismo orden del gel anterior pero sin el carril 9 (marcador).

## Purificación de la proteína At4g32460 completa

Al igual que la proteína At5g11420 completa y truncada, se expresó y purificó la proteína DUF642 completa codificada por el gen At4g32460. La expresión y purificación se realizó bajo las



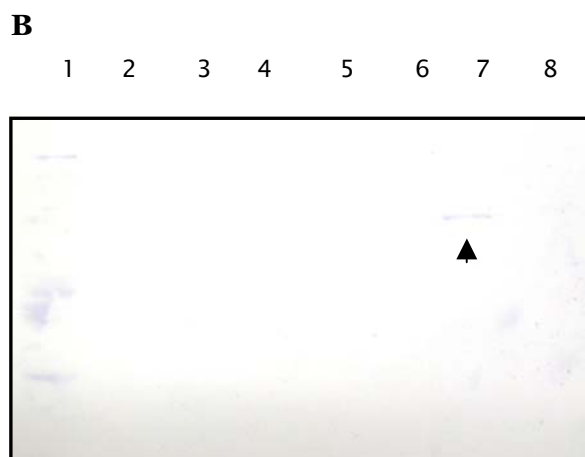
mismas condiciones. En la figura VII-A se presentan el gel de poliacrilamida con las diferentes fracciones obtenidas. Como se observa en el carril 6, correspondiente a la fracción eluída, se purificó una proteína de aproximadamente 39 kDa que corresponde al tamaño de la proteína esperada.



**Figura VII-A:** Purificación de la proteína At4g32460 completa. Gel de poliacrilamida al 12%

- 1) Extracto crudo antes de inducir
- 2) Extracto crudo después de inducir
- 3) Extracto bacteriano purificado
- 4) Fracción no unida
- 5) Lavado con 10 mM de imidazol
- 6) Elusión con 250 mM de imidazol
- 7) Marcadores de peso molecular

En la figura VII-B carril 7 se observa la detección por Western- blot de la proteína completa At4g32460.



**Figura VII-B:** Western-blot, detección de la proteína At4g32460 completa.

- 1) Extracto crudo antes de inducir
- 2) Extracto crudo después de inducir
- 3) Extracto bacteriano purificado
- 4) Fracción no unida
- 5) Lavado 10 mM de imidazol
- 6) Lavado 20 mM de imidazol
- 7) Elusión 250 mM de imidazol
- 8) Marcador de peso molecular (kDa)

### Anexo III

Tabla II: Oligos diseñados para la amplificación de los diez genes miembros de la familia DUF642 de *Arabidopsis thaliana* a partir de cDNA.

Locus	oligonucleótido sentido	oligonucleótido antisentido	Tamaño del producto (pb)
At2g41810	5'ATGGGCCAAAAAACAC3'	5'ATGTCTCTCGTTCTCTCTC3'	1013
At2g41800	5'TCCTCCTCCTATCTCTCTGC3'	5'AAACGGTTCTCTTCTCTGC3'	905
At3g08030	5'GGTCCCAAAGCCATTATTC3'	5'ACAATCTCGTCAATGACAGG3'	1073
At5g25460	5'CTTCCTTCTTTTCATCGCC3'	5'ACGAGAAATCATCGCTCC3'	1027
At5g11420	5'AATCGCCACCATCACTTC3' 5' <b>CCATGG</b> GCTTCAGTGACGGGATG3'	5'CATAACACTTGTGCGGGTC3' 5' <b>AGATCT</b> GAGTGTCTTTTCCCGC3'	963
At4g32460	5'GTGATAGTGCTTCTTCTCCTTCAC3'	5'AGCGACGAATCTCAATGAC3'	959
At1g80240	5'AAAAG CAG CACTCCTCTTAG3'	5'ATCATTGGTCCCTCACAAC3'	879
At1g29980	5'CCGAGCAACAATAGATGC3'	5'ACTGTAGAACGCAACTCTGG3'	1051
At2g34510	5'TTGGTCTCTCCATTGTGGC3'	5'CCTTAACGTCATCAATCACAGG3'	1076
At5g14150	5'TTGCCTCTTTCAGATTTT3'	5'CTTCTCACCAGAGCCAGTCC3'	907

Tabla III. Condiciones de los programas utilizados para la amplificación de los 10 genes que constituyen la familia DUF642.

Locus	Programa de PCR
At2g41810 At3g08030 At1g29980 At2g34510 At5g14150 At2g41800 At5g11420	T inicial 94°C 5' T de desnaturalización 94 °C 30'' T de hibridación o "anealing" 60 °C 30'' T de extensión 72 °C 1' 30'' T final 72 °C 5' T de enfriamiento 4°C
At4g32460 At1g80240	T. de hibridación 61° C 30''
At5g25460	T. de hibridación 61.5°C 30''

Tabla IV: Oligonucleótidos utilizados para la amplificación del inserto de T-DNA en el locus At4g32460

Amplificación	Oligonucleótido sentido	Oligonucleótido antisentido	Tamaño de la banda esperada (pb)
Inserción de T-DNA en el gen At4g32460	LB1 5' GCGTGGACCGCTTGCTGCAACT3'	FIN32460 5'CCATGGGCTTCAATGATGGAGTACTACC3'	800
At4g32460	FAT32460 5'GTGATAGTGCTTCTTCTCCTTCAC3'	FIN32460 5'CCATGGGCTTCAATGATGGAGTACTACC3'	1700

Tabla V. Oligos diseñados para la amplificación y expresión de las proteínas At5g11420 y At4g32460 pertenecientes a la familia DUF642. En rosa se muestra la secuencia añadida para el sitio de restricción de NcoI en los primers sentido y en azul para el sitio de restricción en los primers antisentido.

Locus	Oligo sentido	Oligo Antisentido	Tamaño del producto (pb)
<i>At5g11420</i>	MET11420 5' <b>CCATGGG</b> GCTTCAGTGACGGGAT G3'	11420FIN2 5' <b>AGATCT</b> AGTGTCTTTTCCCGC A3'	858
		11420FIN3 5' AGATCTCGGCTTACGAGCACT GAG 3	981
<i>At4g32460</i>	MET32460 5' <b>CCATGGG</b> CTTCAATGATGGACTA CTACC3'	32460FIN2 5' <b>AGATCT</b> GCGTAAAACGTACTG TAGA 3'	966
		32460FIN3 5' AGATCTCGGCCTCCGAGCACT GAGAAG 3	1038

#### Extracción de RNA total con trizol (Gibco BRL)

- Moler el menos 100 mg de tejido en N<sub>2</sub> líquido, homogenizar con 1 mL de Trizol y mezclar
- Incubar 5 min. a temperatura ambiente
- Añadir 200 µL de cloroformo y mezclar por inversión 15 seg. e incubar a temperatura ambiente por 3 min.
- Centrifugar a 4 ° C por 15 min. a 10 K
- Pasar la fase acuosa a un tubo nuevo y añadir 500 µL de isopropanol, mezclar por inversión e incubar 10 min. a T° A.
- Centrifugar a 4 ° C por 15 min. a 10 K
- Decantar y volver a centrifugar
- Secar 5 min. y resuspender en 20 µL de H<sub>2</sub>O como máximo
- Almacenar a -70 ° C

#### Protocolo para la síntesis de CDNA (Invitrogen)

- En un tubo eppendorf previamente esterilizado agregar:
  - 1µL de oligo dT
  - 1 ng a 5 µg de RNA
  - 1 µL de dNTPs
- Aforar con agua estéril a 12 µl
- Calentar la mezcla a 65 ° C durante 5 minutos y meter al hielo inmediatamente después de la incubación.
- Agregar 4 µL de buffer "First strand",
  - 2 µL de 0.1 M de DTT y
  - 0.7 µL de RNAase
- Mezclar todo suavemente e incubar durante 2 min. a 42 ° C
- Agregar 1 µl de retrotranscriptasa y mezclar mediante pipeteo
- Incubar a 42 ° C durante 50 minutos
- Inactivar la reacción calentándola a 70 ° C durante 15 min.

## Protocolo de transformación en *E. coli*

Se necesita previamente:

- Hielo
- Tubos con medio LB (6 mL)
- Placas con medio LB + antibiótico
- Heat Bloc a 42° C
- Agitador a 37°C a 200-250 rpm

➤ Sacar las células competentes (DH5α o M15) de -70 °C y ponerlas en hielo

En presencia del mechero:

- Añadir 50 µl de células competentes a cada tubo con 5 µl de ligación. Mezclar suavemente por pipeteo.
- Incubar en hielo durante 20 minutos.
- Incubar las células a 42 ° C de 40 a 50 segundos y regresar al hielo inmediatamente.
- Añadir 950 µl de medio LB líquido a cada tubo.
- Incubar 1 hora 30 minutos a 37 ° C en agitación.
- Preparar placas de medio LB con antibiótico, 20 µl de X-GAL 50 mg/mL y 100 µl de IPTG (0.1 M) distribuyendo estos dos últimos con asa de vidrio.
- Centrifugar 10 min. a 6000 rpm
- Retirar 800 µl de sobrenadante. Resuspender el pellet en los 100 µl restantes y plaquear 50 µl por caja de petri con un asa de vidrio.
- Incubar toda la noche a 37° C y después poner en el refrigerador.

## Extracción de DNA genómico

### \*Buffer de extracción de DNA

- 20 mL de Tris- HCl 1 M, pH 7.5
  - 5 mL de NaCl 5 M
  - 5 mL de EDTA 0.5 M, pH 8
  - 5 mL de SDS al 10 %
  - 65 mL de ddH<sub>2</sub>O
- Cortar de 5 a 4 hojas de roseta de una planta y congelarla inmediatamente en nitrógeno líquido.
  - Triturar el tejido con palitos hasta obtener un polvo muy fino, sin descongelar en ningún momento el tejido.
  - Agregar de 500 a 700 µL de buffer de extracción de DNA\* agitar manualmente y dejar durante 10 min. a temperatura ambiente.
  - Agregar de 500 a 700 µL de fenol-cloroformo-isoamil alcohol y agitar invirtiendo el tubo
  - Centrifugar 10 minutos a máxima velocidad.
  - Transferir el sobrenadante a un tubo nuevo y agregar 1/10 de volumen de acetato de sodio 3M.
  - Agregar 350 µL de isopropanol, mezclar por inversión y dejar 5 minutos a temperatura ambiente.
  - Centrifugar 10 minutos a máxima velocidad
  - Lavar el pellet con 150 µL de etanol al 70 %
  - Centrifugar 5 minutos y quitar el etanol.
  - Disolver el pellet en 20 µL de agua estéril

## Protocolo para WESTEN BLOT

### \*Amortiguador de transferencia

25 mM Tris base  
150 mM de glicina  
20% de metanol  
debe estar a 8.3 de pH sin ajustarlo

### Amortiguador pH 9.5

100 mM Tris HCl 9.5 pH  
100 mM NaCl  
5 mM MgCl<sub>2</sub>

### Amortiguador TBS

10 mM Tris-HCl pH 7.5  
150 mM de NaCl

### Amortiguador TBS-tween

20 mM Tris-HCl  
500 mM NaCl  
0.05% (v/v) Tween 20

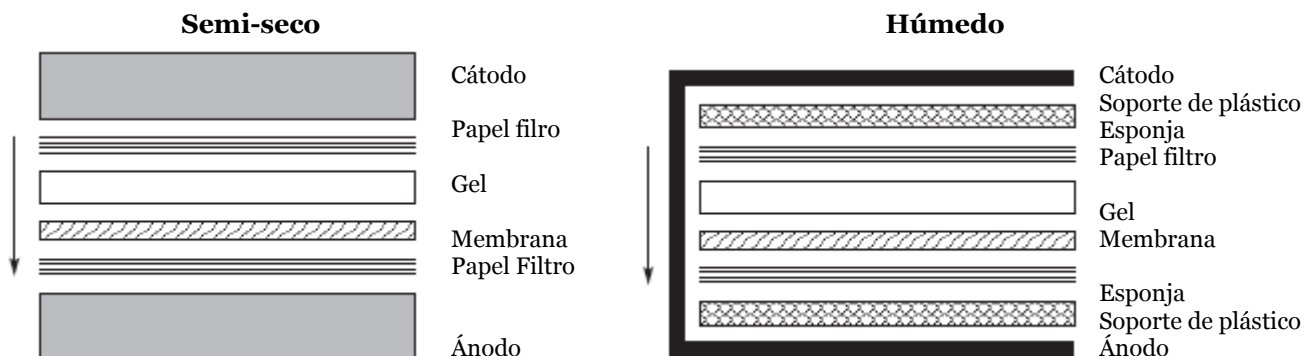
### Solución Reveladora

10 mL de Buffer 9.5  
10 µL de NBT  
10 µL de BCIP

NBT 33 mg/mL de DMS al 70%  
BCIP 17 mg/mL de DMS al 100%

### Trasferencia a la membrana con cámara húmeda

- Cortar 8 piezas de papel filtro u un pedazo de membrana de la misma medida del gel.
- (Evitar la contaminación, siempre manejar el papel filtro y la membrana con guantes)**
- Incubar la membrana durante 10 minutos en buffer de transferencia \*.
  - Mojar el papel filtro con el buffer de transferencia y poner 4 hojas de papel filtro sobre la esponja, seguido por el gel, la membrana, cuatro hojas de papel filtro y finalmente otra esponja (ver figura).
- (Evitar que se formen burbujas pues esto causa problemas de transferencia de proteínas. Las burbujas pueden removerse pasando cuidadosamente un tubo o una pipeta Pasteur sobre cada capa del sándwich)**



- Consultar las instrucciones de manufactura de la cámara de transferencia para determinar el voltaje y el tiempo de transferencia. (El tiempo de transferencia depende del tamaño de las proteínas, el porcentaje de acrilamida y el grosor del gel. El amperaje requerido se determina a partir de la superficie del gel: 0.8 mA/cm<sup>2</sup> (1 hora de transferencia).

### **Detección con conjugado Ni-NTA**

- Lavar el western blot dos veces durante 10 minutos cada vez con buffer TBS\*
- Incubar la membrana durante 1 hora en TBS con 3% de BSA a temperatura ambiente
- Lavar 3 veces durante 10 minutos con buffer TBS-tween\*
- Incubar la membrana durante 1 hora a temperatura ambiente en buffer TBS-tween conteniendo una dilución 1/1000 del conjugado Ni-NTA.

(La membrana puede guardarse en una bolsa de plástico)

### **No incubar en presencia de BSA, leche u otro agente de bloqueo, agentes quelantes o grupos donadores de electrones.**

- Lavar la membrana tres veces durante 10 minutos en buffer TBS- tween a temperatura ambiente.
- Incubar durante 20 minutos con buffer TBS 9.5 pH
- Teñir con la solución de revelado \* durante aproximadamente 5 a 15 minutos.
- Detener la reacción sumergiendo la membrana en agua.
- Secar la membrana y fotografiarla tan pronto como sea posible
- Los colores pueden desaparecer con el tiempo.