



**INSTITUTO POTOSINO DE INVESTIGACIÓN
CIENTÍFICA Y TECNOLÓGICA**

DIVISIÓN DE BIOLOGÍA MOLECULAR

**Dimerización y localización intracelular de las enzimas
(ZmSPMS1, AtADC1 y AtADC2) involucradas en la biosíntesis
de poliaminas en plantas**

Tesis que presenta:

M. en C. Israel Maruri López

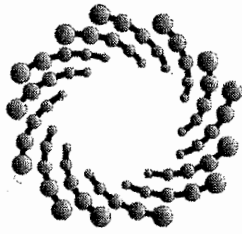
Para obtener el grado de:

Doctor en Ciencias en Biología Molecular

Director de Tesis:

Dr. Juan Francisco Jiménez Bremont

San Luis Potosí, S.L.P., Febrero de 2017



IPICYT

Constancia de aprobación de la tesis

La tesis "***Dimerización y localización intracelular de las enzimas (ZmSPMS1, AtADC1 y AtADC2) involucradas en la biosíntesis de poliaminas en plantas***" presentada para obtener el Grado de Doctor en Ciencias en Biología Molecular fue elaborada por **Israel Maruri López** y aprobada el **tres de febrero del dos mil diecisiete** por los suscritos, designados por el Colegio de Profesores de la División de Biología Molecular del Instituto Potosino de Investigación Científica y Tecnológica, A.C.

Dr. Juan Francisco Jiménez Bremont
Director de la tesis

Dr. Alejandro De Las Peñas Nava
Miembro del Comité Tutorial

Dr. Samuel Lara González
Miembro del Comité Tutorial

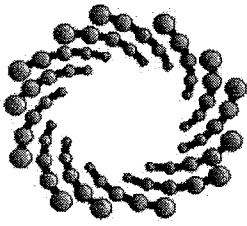
Dra. Margarita Rodríguez y Domínguez Kessler
Miembro del Comité Tutorial



Créditos Institucionales

La presente investigación se realizó en el laboratorio de Biología Molecular de Hongos y Plantas de la División de Biología Molecular del Instituto Potosino de Investigación Científica y Tecnológica, A.C., bajo la dirección del Dr. Juan Francisco Jiménez Bremont.

Durante la realización del trabajo el autor recibió una beca académica del Consejo Nacional de Ciencia y Tecnología (No. de registro 375927) y del Instituto Potosino de Investigación Científica y Tecnológica, A. C



IPICYT

Instituto Potosino de Investigación Científica y Tecnológica, A.C.

Acta de Examen de Grado

El Secretario Académico del Instituto Potosino de Investigación Científica y Tecnológica, A.C., certifica que en el Acta 091 del Libro Primero de Actas de Exámenes de Grado del Programa de Doctorado en Ciencias en Biología Molecular está asentado lo siguiente:

En la ciudad de San Luis Potosí a los 3 días del mes de febrero del año 2017, se reunió a las 16:00 horas en las instalaciones del Instituto Potosino de Investigación Científica y Tecnológica, A.C., el Jurado integrado por:

Dr. Alejandro De Las Peñas Nava	Presidente	IPICYT
Dr. Samuel Lara González	Secretario	IPICYT
Dr. Juan Francisco Jiménez Bremont	Sinodal	IPICYT
Dra. Margarita Rodríguez y Domínguez Kessler	Sinodal externo	UASLP

a fin de efectuar el examen, que para obtener el Grado de:

DOCTOR EN CIENCIAS EN BIOLOGÍA MOLECULAR

sustentó el C.

Israel Maruri López

sobre la Tesis intitulada:

Dimerización y localización intracelular de las enzimas (ZmSPMS1, AtADC1 y AtADC2) involucradas en la biosíntesis de poliaminas en plantas

que se desarrolló bajo la dirección de

Dr. Juan Francisco Jiménez Bremont

El Jurado, después de deliberar, determinó

APROBARLO

Dándose por terminado el acto a las 17:35 horas, procediendo a la firma del Acta los integrantes del Jurado. Dando fe el Secretario Académico del Instituto.

A petición del interesado y para los fines que al mismo convengan, se extiende el presente documento en la ciudad de San Luis Potosí, S.L.P., México, a los 3 días del mes de febrero de 2017.

Mtra. Ivonne Lizette Cuevas Vélez
Jefa del Departamento de Posgrado

Dr. Horacio Flores Zúñiga
Secretario Académico



Dedicatoria

Con todo mi amor, cariño y admiración dedico mi trabajo a mi mamá Martha y mi esposa Itzell. Gracias por creer en mi y apoyarme en todo momento.

Agradecimientos

Mi más sincero agradecimiento al Dr. Juan Francisco Jiménez Bremont, director de esta tesis. Gracias por creer en mi y brindarme la oportunidad de formar parte de su laboratorio.

Al Dr. Alejandro De las Peñas Nava, al Dr. Samuel Lara González y a la Dra. Margarita Rodríguez y Domínguez Kessler por aceptar ser parte de mi comité tutorial, por su tiempo y sus acertadas sugerencias a mi trabajo de tesis.

A la M. en C. Alicia Becerra Flora, por el apoyo técnico brindado para el desarrollo de esta tesis.

Al CONACYT e IPICYT por el financiamiento y recursos para realizar este trabajo.

Contenido

Constancia de aprobación de la tesis	ii
Créditos Institucionales	iii
Acta de examen	iv
Dedicatorias	v
Agradecimientos	vi
Resumen	x
Abstract	xi
I. Introducción	1
1.1. Las Poliaminas y su importancia en las plantas	1
1.2. Ruta de biosíntesis de las poliaminas	3
1.3. Homeostasis de las poliaminas	4
1.4. Catabolismo de poliaminas	6
1.5. Secuencias de degradación proteolítica tipo PEST	7
II. Justificación	8
III. Objetivo general	9
Capítulo 1	10
La secuencia de PEST de la espermina sintasa de maíz (ZmSPMS1) fusionada a la proteína reportera GUS facilita su degradación proteolítica.	
Capítulo 2	13
Caracterización de la ZmSPMS1: Evidencia para su dimerización y ubicación intracelular.	

Capítulo 3	16
Homo y hetedimerización de las enzimas arginina descarboxilasa (AtADC1 y AtADC2) de <i>Arabidopsis thaliana</i> .	
IV. Discusión	19
V. Conclusiones y Perspectivas	24
VI. Referencias	25

1.- Análisis *in silico* de los factores de transcripción WRKY de Arabidopsis:
ocurrencia de secuencias PEST en regiones desordenadas.

Resumen

Las poliaminas (PAs) son metabolitos ubicuos cargados positivamente que desempeñan actividades distintas en procesos celulares fundamentales. La enzima arginina descarboxilasa (ADC) lleva a cabo la producción de la agmatina a partir de arginina, la agmatina es el precursor de la primera PA, la cual es conocida como putrescina. La putrescina se convierte en las PAs superiores, espermidina y espermina. La espermina sintasa (SPMS) cataliza la formación de la PA espermina. En este trabajo de tesis determinamos la localización sub-celular de las enzimas AtADC1 y AtADC2 de *Arabidopsis thaliana* y de la enzima ZmSPMS1 de *Zea mays* a través de fusiones traduccionales con la proteína verde fluorescente (GFP), y la interacción proteína-proteína mediante la complementación bimolecular por fluorescencia (BiFC). Nuestros datos mostraron que las enzimas ADC de *A. thaliana* se localizan en el citoplasma y en el cloroplasto, y son capaces de formar homodímeros. Es interesante mencionar que encontramos la formación de heterodímeros entre AtADC1 y AtADC2. También, demostramos que la enzima ZmSPMS1 de maíz interactúa consigo misma en el citoplasma. Con el fin de obtener una mejor comprensión de la oligomerización de la ZmSPMS1, se generaron construcciones truncas en esta proteína. Encontramos que la delección de la región C-terminal de la ZmSPMS1 abate su dimerización. Además, la versión trunca recombinante fue incapaz de generar espermina en la bacteria *Escherichia coli*. Estos datos sugieren que estas tres enzimas AtADC1, AtADC2 y ZmSPMS1 pueden actuar en complejos para regular su actividad enzimática y su estabilidad. Finalmente, la localización sub-celular específica de estas enzimas implica que la biosíntesis de PAs se lleva a cabo en diferentes compartimentos de la célula y esto sugiere funciones específicas de cada poliamina.

Palabras clave: poliaminas, arginina descarboxilasa, espermina sintasa, oligomerización, localización sub-celular.

Abstract

Polyamines (PAs) are ubiquitous positively charged metabolites that play an important role in fundamental cellular processes. The arginine decarboxylase enzyme (ADC) catalyzes the production of agmatine from arginine, which is the precursor of the first PA known as putrescine. After, putrescine is turned into the higher PAs, spermidine, and spermine; in turn, spermine synthase (SPMS) catalyzes the formation spermine. In the present study, we examined the sub-cellular localization of AtADC1, AtADC2, and ZmSPMS1 enzymes from *Arabidopsis thaliana* and *Zea mays*, respectively, through green fluorescence protein (GFP) translational fusions and its dimerization was evaluated using a Bimolecular Fluorescence Complementation (BiFC) approach. First, our *A. thaliana* ADC data revealed that these enzymes are cytosolic and chloroplastic proteins able to form homodimers with the same localization. Interestingly, we found the formation of AtADC1/AtADC2 heterodimers. Second, we demonstrated that the *Z. mays* ZmSPMS1 enzyme interacts with itself in the cytoplasm. In order to gain a better understanding of ZmSPMS1 protein interaction, two truncated constructs of ZmSPMS1 protein were generated. We found that the deletion of C-terminal ZmSPMS1 region dramatically abated the ZmSPMS1-ZmSPMS1 protein interaction. We found that the C-terminal ZmSPMS1 truncated version fail to generate spermine in *Escherichia coli* cells. These data suggest that these three AtADC1, AtADC2, and ZmSPMS1 enzymes can act as complexes to regulate its enzymatic activity and protein regulation, also the specific localization of these enzymes indicates that biosynthesis takes place in different compartments suggesting specific functions of each polyamine.

Keywords: polyamines, arginine decarboxylase, spermine synthase, oligomerization, and sub-cellular localization.

I. Introducción

1.1. Las Poliaminas y su importancia en las plantas

Las poliaminas (PAs) son policationes nitrogenados alifáticos de bajo peso molecular, y su ruta biosintética está conservada con pequeñas variaciones en todos los organismos vivos (Handa y Mattoo, 2010). La función biológica de las PAs fue inicialmente asociada con su capacidad de unirse a diferentes macromoléculas aniónicas, como DNA, RNA, cromatina y proteínas, esto sugiere un papel estructural (Kusano et al., 2008). Las PAs también han sido implicadas en muchos procesos celulares fundamentales como en: la regulación de la expresión génica, la traducción, la división y la diferenciación celular, así como también, en el funcionamiento de ciertos canales iónicos (Igarashi y Kashiwagi, 2000; Childs et al., 2003; Pottosin et al., 2014).

En plantas, los PAs más abundantes son la diamina putrescina (Put), la triamina espermidina (Spd), la tetramina espermina (Spm) y el isómero estructural de la Spm denominado termoespermina (tSpm) (Knott et al., 2007). Las PAs se pueden encontrar en formas libres o conjugados con ácidos hidroxicinámicos. Estos conjugados han sido implicados en las respuestas de defensa contra patógenos, en la detoxificación de compuestos fenólicos, y también como reserva intracelular de PAs (Grienenberger et al., 2009, Luo et al., 2009). A la fecha, existen varios reportes que evidencian que cada PA participa en funciones biológicas específicas, sin embargo, sus mecanismos de acción aún no ha sido elucidados (Groppa y Benavides, 2008; Takahashi y Kakehi 2009).

La biosíntesis y catabolismo de PAs está asociada con el crecimiento y en las respuestas a estímulos ambientales externos (Flores y Galston 1982; Moschou et al., 2008). En plantas modelo como *Arabidopsis thaliana*, estudios de genética molecular basados en plantas transgénicas y mutantes con alteraciones específicas en genes que participan en la ruta de biosíntesis de las PAs proporcionaron pruebas convincentes que afirman la importancia de estas aminas como moduladores críticos de diversos procesos fisiológicos del crecimiento (Kusano et al., 2008, Alcázar et al., 2010).

En *A. thaliana*, la biosíntesis de Put es necesaria para el crecimiento y desarrollo, participando en la elongación del tallo, el inicio de la floración y la morfología de la semilla (Sánchez-Rangel et al., 2016). La diamina Put es esencial para la supervivencia, ya que mutantes de *A. thaliana* incapaces de producir este polication no logran completar su desarrollo embrionaria (Urano et al., 2005). En respuesta al estrés, la acumulación de Put brinda tolerancia al estrés por sequía y congelamiento en plantas de *A. thaliana* (Cuevas et al., 2008; Alcázar et al., 2010). En especies de tabaco y en otras Solanáceas, la diamina Put es sustrato para la síntesis nicotina y alcaloides tropano relacionados con la defensa al ataque de herbívoros (Biastoff et al., 2009).

Una de las funciones más estudiadas de la triamina espermidina (Spd) es su participación como precursor en la síntesis de la hipusina, que es un aminoácido poco común que se encuentra en eucariotes y en algunas especies de arqueas (Jansson et al. 2000; Park, 2006). Hasta el momento se sabe que la única proteína que contiene hipusina es el factor de iniciación de la traducción eucariótica 5A (eIF5A). La hipusina es el resultado de una modificación post-traducciona poco común, donde un grupo aminobutilo proveniente de la Spd es transferido a la lisina localizada en la posición 50 (Lys50) del eIF5A. Esta modificación es esencial para la función biológica de éste factor, que consiste en el transporte de mRNAs del núcleo hacia el citoplasma (Chattopadhyay et al., 2003). La ausencia de Spd en plantas de *A. thaliana* causa un efecto letal en el desarrollo del embrión, similar a lo observado en mutantes de *Arabidopsis* deficientes de Put (Imai et al., 2004).

In vitro, la tetramina espermina (Spm) funciona como un agente secuestrador de radicales libres y es capaz de proteger al DNA del daño causado por moléculas como el peróxido de hidrógeno (H₂O₂) (Ha et al., 1998). En *A. thaliana*, la Spm está implicada en las respuestas al estrés por sequía y salinidad. La ausencia de este polication provoca a una lenta respuesta del cierre estomático durante el estrés por sequía (Yamaguchi et al., 2007). Es conocido que las PAs, incluyendo la Spm, pueden modular la actividad de ciertos canales iónicos, en especial los de Ca²⁺ y K⁺ (Dobrovinskaya et al., 2000; Pottosin y Shabala, 2014). Se conoce que cambios en la concentración

citoplasmática de Ca^{2+} en las células guarda de los estomas en *A. thaliana* están involucrados en los procesos de cierre y apertura estomática (Rob et al., 2005). Esto sugiere que la Spm podría controlar el flujo de Ca^{2+} , lo cual afecta el movimiento estomático.

La tetramina termoespermina (tSpm) se identificó por primera vez en las bacterias termófilas y se relacionó con la síntesis de proteínas en condiciones ambientales extremas (Oshima, 2007). En *A. thaliana*, la tSpm está relacionada con la proliferación de los vasos del xilema y la elongación del tallo (Hanzawa et al., 1997). La ausencia de este isómero de la Spm, en mutantes de *A. thaliana*, presenta un fenotipo enano, el cual se atribuye principalmente a la excesiva diferenciación del xilema. Esto sugiere que la tSpm regula negativamente el desarrollo del xilema a través de un proceso dependiente de auxinas (Yoshimoto et al., 2012a,b; Tong et al. 2014).

1.2. Ruta de biosíntesis de las Poliaminas

En las plantas, el primer paso para la biosíntesis de PAs es la formación de la Put, y esta se lleva a cabo por dos rutas: una vía directa que utiliza la ornitina como sustrato, y otra indirecta que emplea la arginina. En la vía directa, la Put se sintetiza a partir de la descarboxilación de la ornitina, por medio de la enzima ornitina descarboxilasa (ODC; EC 4.1.1.17). En la vía indirecta, la Put se sintetiza de la descarboxilación de la arginina por la enzima arginina descarboxilasa (ADC; EC 4.1.1.19). El producto intermedio de esta reacción es la agmatina que es subsecuentemente convertida en N-carbamoylputrescina por la agmatina iminohidrolasa (AIH; EC 3.5.3.12) y finalmente en Put por la N-carbamoylputrescina amidohidrolasa (NCPAH; EC 3.5.1.53). La Spd, la Spm y la tSpm se sintetizan por la transferencia a la Put y Spd de un grupo aminopropilo proveniente de la descarboxilación de la S-adenosilmetionina (SAM) se sintetizan por la acción de la S-adenosilmetionina descarboxilasa (SAMDC; EC 4.1.1.50). Este proceso de transferencia de grupos aminopropilos es catalizado por las enzimas aminopropil transferasas, tales como la espermidina sintasa (SPDS; EC 2.5.1.16), la espermina sintasa (SPMS; EC 2.5.1.22) y la termospermina sintasa (TSPMS; EC 2.5.1.79), respectivamente (Fuell et al., 2010) (Figura 1).

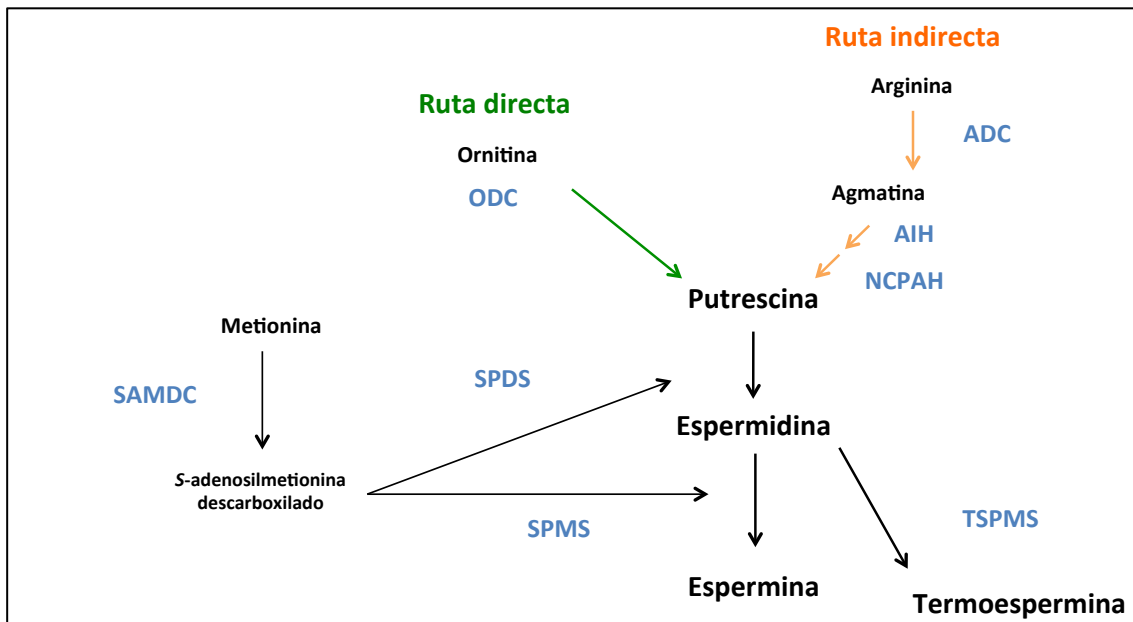


Figura 1. La ruta directa e indirecta de la biosíntesis de las poliaminas. ODC, ornitina descarboxilasa; ADC, arginina descarboxilasa; AIH, agmatina iminohidrolasa; NCPAH, *N*-carbamoil-putrescina amidohidrolasa; SAMDC, *S*-adenosilmetionina descarboxilasa; SPDS, espermidina sintasa; SPMS, espermina sintasa; TSPMS, termoesperimina sintasa.

En el genoma de *A. thaliana* se han identificado 12 genes responsables para la biosíntesis de PAs. Para la síntesis de Put, existen dos copias de gen *ADC* (*AtADC1* y *AtADC2*), un gen para la *AIH* y otro para la *NCPAH*, cuatro copias para la *SAMDC* (*AtSAMDC1* a *4*), dos copias para la *SPDS* (*AtSPDS1* y *AtSPDS2*), una sola copia para la *SPMS* y para la *TSPMS* (Kusano et al., 2008).

1.3. Homeostasis de las Poliaminas

La modulación de los niveles de poliaminas en respuesta al estrés se ha observado en diferentes especies de plantas (Bouchereau et al., 1999, Walters 2003, Groppa y Benavides 2008; Kusano et al. 2008; Alcázar et al. 2010). A menudo esta regulación se acompaña de cambios en los niveles de expresión de genes implicados en la biosíntesis de PAs (Alcázar et al. 2010). Sin embargo, la síntesis de estas poliaminas también está controlada a otros

niveles, como el post-transcripcional, traduccional y post-traduccional (Carbonell y Blázquez, 2009).

Durante las etapas del desarrollo temprano en *A. thaliana*, la actividad del promotor de gen *AtADC1* es baja; sin embargo, esta se incrementa considerablemente en las raíces y en las hojas de plántulas de dos semanas de edad bajo estrés por frío (Hummel et al., 2004). Por el contrario, la actividad del promotor del gen *AtADC2* está más relacionada con la germinación de las semillas, el desarrollo de raíces y hojas, y también este puede responder a presencia de luz, sacarosa y etileno (Hummel et al., 2004). El análisis de los promotores de los genes *AtADC1* y *AtADC2* mostró diferencias sorprendentes, tales como la presencia de elementos reguladores en *cis*, como elementos de regulación específica del polen, elementos de respuesta a baja temperatura, elemento sensible a la sacarosa, elementos de respuesta a la deshidratación, y la presencia de un elemento transponible en el promotor del gen *AtADC1* (El Amrani et al., 2002). Además, las mutantes sencillas en los genes *adc1* y *adc2* de *A. thaliana* exhibieron sensibilidad al estrés por congelamiento y la adición de Put exógena revirtió este efecto de sensibilidad (Cuevas et al., 2008). Además, la acumulación de Put confiere tolerancia a la sequía en plantas de *A. thaliana* que sobre-expresan el gen *AtADC2* (Alcázar et al., 2010).

Por otro lado, el metabolismo de las PAs superiores (Spd y Spm) se regula a través del control de la traducción de la enzima SAMDC1. Esta regulación se lleva a cabo por medio de dos pequeños marcos de lectura abiertos (uORF) localizados en el extremo 5'-UTR del RNA mensajero de la SAMDC (Ivanov et al., 2009). Estos uORFs controlan la traducción de la SAMDC1 dependiente de los niveles de PAs (Franceschetti et al., 2001; Hanfrey et al., 2005; Ivanov et al., 2010). Además, la *AtSAMDC1* de *A. thaliana* contiene una secuencia de degradación proteolítica tipo PEST [regiones ricas en prolina (P), ácido glutámico (E), serina (S) y treonina (T)], las cuales están presentes en proteínas de vida media corta (Rogers et al., 1986; Zhang et al. 2011).

Además, se conoce que la homeostasis de la Spm está regulada a nivel de su biosíntesis a través de la actividad de la enzima espermina sintasa (SPMS), y por medio de las enzimas poliamino oxidasas que catalizan la desaminación

oxidativa de la Spm (Guerrero-González, et al, 2014; Ahou et al., 2014). La dimerización de las enzimas aminopropil transferasas esta relacionada a su actividad enzimática y la estabilidad de la proteína (Korolev et al., 2002, Wu et al., 2008).

1.4. Catabolismo de Poliaminas

Se ha demostrado que las PAs pueden ser catalizadas a través de dos clases de enzimas, denominadas diaminas oxidasas (DAOs) dependientes de cobre y poliaminas oxidasas (PAOs) dependientes de flavín adenín dinucleótido (FAD) (Cona et al., 2006). Los DAOs catalizan la oxidación de Put produciendo 4-aminobutanal, peróxido de hidrogeno (H_2O_2) y amoníaco (Moschou et al., 2008). Adicionalmente, las enzimas DAO de *A. thaliana* también pueden utilizar Spd como sustrato (Planas-Portell et al., 2013). En *A. thaliana* hay al menos 10 genes que codifican para enzimas tipo DAO, de los cuales cuatro de estos ya se encuentran caracterizados (*AtAO1*, *AtCuAO1*, *AtCuAO2* y *AtCuAO3*). Estos genes presentan un perfil de expresión diferente en respuesta a tratamientos exógenos, tales como la aplicación de ácido abscísico, jasmonato, ácido salicílico, flagelina y por heridas (Møller y McPherson, 1998, Planas-Portell et al., 2013). Además, *AtCuAO1* y *AtAO1* son proteínas extracelulares, mientras que *AtCuAO2* y *AtCuAO3* se localizan en los peroxisomas (Planas-Portell et al., 2013).

En contraste las PAOs tienen una alta afinidad por Spd y Spm (Alcázar et al., 2010). En la plantas, las PAO se dividen en dos grupos principales, el primer grupo lleva acabo el catabolismo terminal de Spd y Spm para producir 1,3-diaminopropano (DAP), H_2O_2 , y N-(3-aminopropil)-4-aminobutanal (Spm), o 4-aminobutanal (Spd), (Cona y et al., 2006; Moschou et al., 2008). El segundo grupo es responsable de las interconversiones de PA, en las que la Spm se convierte de nuevo a Spd, y Spd a Put (Moschou et al., 2008; Mo et al., 2015). En *A. thaliana*, hay cinco genes que codifican para enzimas PAO (*AtPAO1* a *AtPAO5*) (Moschou et al., 2008). La *AtPAO1*, tiene una localización citosólica y oxida preferentemente a la Spm y tSpm pero no Spd (Tavladoraki et al., 2006). Las *AtPAO2*, *AtPAO3* y *AtPAO4* tienen una localización peroxisomal y forman una subfamilia distinta que tiene una alta similitud entre sí (Fincato et al., 2011),

estas enzimas oxidan tanto Spd como Spm (Moschou et al., 2008, Kamada-Nobusada et al., 2008, Fincato et al., 2010). Finalmente, la enzima citosólica AtPAO5 cataliza la interconversión de tSpm en triamina espermidina (Fincato et al., 2010; Kim et al., 2014).

1.5. Secuencias de degradación proteolítica tipo PEST

Las secuencias PEST son uno de los principales motivos en la degradación de proteínas eucariotas mediante la ruta ubiquitina/proteasoma 26S (Rogers et al., 1986). Estas secuencias son definidas como regiones hidrofílicas de 10 o más aminoácidos enriquecidos en prolina (P), ácido glutámico (E), ácido aspártico (D), serina (S) y treonina (T), las cuales están flanqueados con los aminoácidos cargados positivamente como arginina (R), histidina (H) y lisina (K) (Rogers et al., 1986). Estos residuos contribuyen al enriquecimiento de regiones desordenadas en las proteínas que las contienen (Singh et al., 2005). Debido a la flexibilidad conformacional de las secuencias PEST su estructura secundaria no ha sido resuelta; además, con base en su naturaleza hidrofílica, las secuencias PEST posiblemente son regiones expuestas, bucles o extensiones (Rechsteiner y Rogers, 1996; Singh et al., 2005, Gsponer et al., 2008), lo cual les permite ser blanco de fosforilación, ubiquitinación y mediar la interacción proteína-proteína (Sandhu y Dash, 2006).

En plantas, la regulación en las enzimas involucradas en la biosíntesis o el catabolismo de PAs ha sido poca explorada. Sin embargo, Zhang et al. (2011) reportaron que una secuencia PEST presente en la enzima S-adenosilmetionina descarboxilasa 1 de *A. thaliana* (AtSAMDC1) está implicada en su degradación. Específicamente, durante la infección del virus BSCTV, la proteína viral C2 interactúa con la AtSAMDC1. Los autores sugieren que esta interacción detiene la degradación proteolítica de la enzima AtSAMDC1 con la finalidad establecer un ambiente hipometilado que facilite la acumulación viral. Este mecanismo revela la importancia de las secuencias PEST en las enzimas biosintéticas de PAs.

II. Justificación

Las PAs son vitales para diversos procesos celulares, por lo cual su homeostasis esta controlada de manera estricta. El estudio de la degradación, los mecanismos de acción, la dimerización y la localización sub-celular de las enzimas involucradas en la producción de PAs es de gran interés científico ya contribuye a elucidar el papel de cada PA de acuerdo al compartimento celular en el cual se sintetizan.

La degradación de enzimas de biosíntesis de PAs a través de secuencias PEST no ha sido estudiada para enzimas de plantas. Durante mi trabajo de maestría se caracterizó una secuencia PEST de la espermina sintasa ZmSPMS1 de maíz. Posteriormente, durante mi doctorado se estudió la interacción proteína-proteína y localización sub-celular del dímero de la ZmSPMS1. Además, se determinaron las regiones implicadas en dicha interacción. Nuestros datos muestran que el dominio de C-terminal, que contiene la secuencia PEST, está involucrado en la estabilidad de la ZmSPMS1, la capacidad de homodimerización y regula su actividad enzimática. Estos datos contribuyen a explicar donde se lleva a cabo la síntesis Spm y los mecanismos moleculares que regulan la producción de este policatión.

A pesar de la importancia de las arginina descarboxilasas AtADC1 y AtADC2 de *A. thaliana*, las cuales son las únicas enzimas que comienzan la síntesis de PAs en *A. thaliana*, su localización sub-celular y su capacidad de formar complejos enzimáticos de cada proteína es una pregunta relevante que permanece sin respuesta. En este trabajo se exploró la localización sub-celular *in vivo* de ambas proteínas AtADC1 y AtADC2, y también se estudió la localización y formación de complejos enzimáticos.

III. Objetivo general

Determinar la dimerización y localización sub-celular *in vivo* de las enzimas ZmSPMS1, AtADC1 y AtADC2.

Objetivos específicos:

- Determinar si la secuencia PEST de la espermina sintasa 1 de maíz (ZmSPMS1) funciona como una etiqueta de degradación proteolítica.
- Determinar si la proteína ZmSPMS1 de maíz es capaz de formar homodímeros.
- Identificar la localización sub-celular del homodímero de la ZmSPMS1 de maíz.
- Analizar si las regiones N y C-terminal están involucradas en la interacción ZmSPMS1-ZmSPMS1.
- Determinar la localización sub-celular de las proteínas de *A. thaliana* AtADC1 y AtADC2.
- Examinar la capacidad para formar homo y heterodímeros entre las proteínas de *A. thaliana* AtADC1 y AtADC2.

Capítulo 1

La secuencia de PEST de la espermina sintasa de maíz fusionada a la proteína reportera GUS facilita su degradación proteolítica.

Previamente, en nuestro grupo de investigación Rodríguez-Kessler et al. (2006), analizaron la respuesta de genes involucrados en la biosíntesis de PAs en hojas de maíz bajo estrés salino. Este trabajo muestra que el gen de la espermina sintasa 1 (ZmSPMS1) se induce por este tipo de estrés. Además, observaron que el contenido de Spm se incrementó considerablemente. La enzima ZmSPMS1 difiere de sus ortólogos de dicotiledóneas por una extensión de 31 residuos en su extremo C-terminal. Dicha extensión se conserva entre las espermina sintasas de especies monocotiledóneas y codifica para un potencial sitio PEST involucrado en la degradación específica de proteínas vía la ruta ubiquitina/proteosoma 26S (Rodríguez-Kessler y Jiménez-Bremont, 2008; Rogers et al., 1986).

Posteriormente, un análisis de complementación funcional con una mutante de *Saccharomyces cerevisiae* incapaz de producir Spm, reveló que una versión trunca en el extremo C-terminal de la ZmSPMS1 fue capaz de complementar a la mutante de levadura para producir Spm comparada con la versión completa (Rodríguez-Kessler et al., 2010). Este resultado sugiere que la ausencia de la secuencia PEST genera una versión funcional de la enzima ZmSPMS1 en levadura.

Durante mi tesis de maestría caractericé la secuencia PEST de la ZmSPMS1 mediante una fusión entre el gen reportero de la β -glucoronidasa (*GUS*) y 123 pb del marco de lectura abierto del gen *ZmSPMS1*, los cuales codifican los últimos 40 residuos e incluyen la secuencia PEST. Esta fusión traduccional se clonó en el vector de expresión para plantas pMDC32, bajo el control del promotor 2x35S. Posteriormente, la construcción pMDC32-GUS::PEST se evaluó en un sistema de expresión transitoria en monocapas de cebolla y en líneas transgénicas de *A. thaliana*. Nuestros datos revelaron que la secuencia PEST de la enzima ZmSPMS1 condujo a la degradación específica de la proteína reportera GUS en ambos sistemas de expresión.

Para examinar si la fusión GUS::PEST se degrada vía el proteosoma 26S, se analizó su expresión en monocapas de cebolla en presencia del inhibidor del proteosoma, MG132. En este experimento se observó un incremento en la expresión de la señal de GUS en el tratamiento con MG132. Estos resultados sugieren que el proteosoma 26S puede estar involucrado en la degradación de la proteína ZmSPMS1 en monocapas de cebolla.

Durante mi primer año de doctorado analizamos si la disminución en la señal de GUS en las plantas transgénicas de *A. thaliana* que expresan la fusión GUS::PEST se debía a una degradación vía el proteosoma 26S. Para ello se analizó la expresión de GUS en dichas plantas en presencia del inhibidor MG132. Nuestro experimento reveló que la señal GUS se recuperó en presencia de MG132 en comparación a las plantas de *A. thaliana* no incubadas con dicho compuesto. Durante mi segundo año de doctorado se publicó el artículo titulado "A maize spermine synthase 1 PEST sequence fused to the GUS reporter protein facilitates proteolytic degradation" en la revista *Plant Physiology and Biochemistry*. Este trabajo constituye el primer reporte acerca de una secuencia PEST funcional en una espermina sintasa de plantas.



Research article

A maize spermine synthase 1 PEST sequence fused to the GUS reporter protein facilitates proteolytic degradation



Israel Maruri-López^a, Margarita Rodríguez-Kessler^b, Aída Araceli Rodríguez-Hernández^a, Alicia Becerra-Flora^a, Juan Elías Olivares-Grajales^c, Juan Francisco Jiménez-Bremont^{a,*}

^aDivisión de Biología Molecular, Instituto Potosino de Investigación Científica y Tecnológica A. C., 78216 San Luis Potosí, SLP, Mexico

^bFacultad de Ciencias, Universidad Autónoma de San Luis Potosí, Av. Salvador Nava s/n, Zona Universitaria, 78290 San Luis Potosí, SLP, Mexico

^cDepartamento de Biología Molecular de Plantas, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Av. Universidad 2001, Colonia Chamilpa, 62210 Cuernavaca, Morelos, Mexico

ARTICLE INFO

Article history:

Received 19 December 2013

Accepted 18 February 2014

Available online 27 February 2014

Keywords:

Aminopropyl transferases

PEST

Polyamines

26S proteasome

Spermine synthase

ABSTRACT

Polyamines are low molecular weight aliphatic compounds involved in various biochemical, cellular and physiological processes in all organisms. In plants, genes involved in polyamine biosynthesis and catabolism are regulated at transcriptional, translational, and posttranslational level. In this research, we focused on the characterization of a PEST sequence (rich in proline, glutamic acid, serine, and threonine) of the maize spermine synthase 1 (ZmSPMS1). To this aim, 123 bp encoding 40 amino acids of the C-terminal region of the ZmSPMS1 enzyme containing the PEST sequence were fused to the GUS reporter gene. This fusion was evaluated in *Arabidopsis thaliana* transgenic lines and onion monolayers transient expression system. The ZmSPMS1 PEST sequence leads to specific degradation of the GUS reporter protein. It is suggested that the 26S proteasome may be involved in GUS::PEST fusion degradation in both onion and *Arabidopsis*. The PEST sequences appear to be present in plant spermine synthases, mainly in monocots.

© 2014 Elsevier Masson SAS. All rights reserved.

1. Introduction

Spermidine and spermine synthase (SPDS; EC 2.5.1.16 and SPMS; EC 2.5.1.22) belong to a class of aminopropyl transferase enzymes that uses decarboxylated S-adenosylmethionine (dcSAM) as substrate for the transfer of an aminopropyl group to a diamine or triamine acceptor molecule producing the higher polyamines (PAs) spermidine (Spd) and spermine (Spm), respectively (Rodríguez-Kessler et al., 2010; Shao et al., 2012). Because of their positive charge, PAs can bind to anionic molecules, such as DNA, RNA, proteins, and chromatin. PAs act as regulatory molecules in various cellular processes including cell division, synthesis of DNA and proteins, gene expression, membrane stabilization, and the functioning of certain ion channels (Igarashi and Kashiwagi, 2000; Childs et al., 2003; Kusano et al., 2008; Zepeda-Jazo et al., 2011).

Due to these multiple functions, PA homeostasis is crucial and must be strictly maintained by regulation of their biosynthesis,

catabolism, conjugation, and transport (Wallace et al., 2003; Carbonell and Blazquez, 2009; Mattoo et al., 2010). In particular, the biosynthesis of these polycations is controlled at transcriptional, posttranscriptional, translational, and posttranslational level (Malmberg and Cellino, 1994; Zhang et al., 2003; Rodríguez-Kessler et al., 2006; Ivanov et al., 2009; Takahashi and Kakehi, 2010).

In mammals, intracellular levels of PAs are maintained through proteolytic degradation of the first biosynthetic enzyme ornithine decarboxylase (ODC; EC 4.1.1.17), which catalyzes putrescine (Put) synthesis (Pegg, 2006). Degradation of the ODC enzyme occurs when it complexes with a protein known as "antizyme", thereby exposing its C-terminus, which is a region rich in hydrophilic amino acids, such as proline, glutamic acid, serine, and threonine (PEST) (Li and Coffino, 1992, 1993a; 1993b). This PEST sequence is a target for specific protein degradation via the 26S proteasome (Rechsteiner and Rogers, 1996).

Recently, our research group characterized the maize SPMS1 (ZmSPMS1) enzyme, involved in Spm synthesis. *In silico* analysis of ZmSPMS1 identified a potential PEST sequence in the C-terminus. Later, in a functional complementation analysis using a mutant strain of *Saccharomyces cerevisiae* unable to produce Spm (Δ spe4), we found that only the truncated C-terminal version was able to

* Corresponding author.

E-mail addresses: jbremont@ipicyt.edu.mx, jbremont@yahoo.com (J.F. Jiménez-Bremont).

Capítulo 2

Caracterización de la proteína ZmSPMS1: Evidencia de su dimerización y su localización sub-celular.

Dado que se conoce que las proteínas que contienen secuencias PEST son estables como homodímeros (Li y Coffino, 1993), determinamos si la proteína ZmSPMS1 forma homodímeros mediante el sistema de dos híbridos (Y2H) “split-ubiquitin” en *S. cerevisiae*. Además, confirmamos esta interacción mediante ensayos de complementación Bimolecular de Fluorescencia (BiFC) en hojas de *Nicotiana benthamiana*. Nuestros experimentos demostraron que la enzima ZmSPMS1 es capaz de interactuar consigo misma en ambos sistemas. Además, esta interacción ZmSPMS1-ZmSPMS1 se localizó en el citoplasma de las células epidérmicas de las hojas tabaco. Estos datos concuerdan con la localización del homodímero de la SPMS de *A. thaliana*, que se localiza también en el citoplasma de las células de las hojas de tabaco (Belda-Palazón et al., 2012).

Nuestro siguiente paso fue determinar si el N- o C-terminal de la ZmSPMS1 participan en la formación del homodímero mediante la construcción de dos versiones truncas en el extremo N- y C-terminal de la ZmSPMS1 de maíz. Es interesante mencionar que se reportó que la SMPMS de *Homo sapiens* es capaz de interactuar consigo misma a través de su extremo N-terminal; sin embargo, esta región no está presente en las enzimas SPMS de hongos ni de plantas (Wu et al., 2008). Para ello, nosotros diseñamos dos versiones truncas en el extremo N- y C-terminal de la ZmSPMS1 de maíz. Nuestros datos muestran que la eliminación de los primeros 74 residuos del dominio N-terminal afectaron la formación del dímero ZmSPMS1 en los sistemas de Y2H y BiFC. Además, la eliminación de la región C-terminal, que incluye la secuencia PEST, resultó en la pérdida de la interacción ZmSPMS1-ZmSPMS1 en ambos sistemas. Estos sugieren que la región C-terminal de la ZmSPMS1, que contiene la secuencia de PEST, es la que principalmente está involucrada en su homodimerización.

Finalmente, dado que la bacteria *E. coli* carece de Spm (Panagiotidis et al., 1987), se analizó la capacidad para sintetizar Spm con la versión completa y las versiones trucas en el extremo N- y C-terminal de la ZmSPMS1 en esta bacteria. Observamos que las cepas que expresan la enzima ZmSPMS1 completa y la versión trunca en el extremo N-terminal son capaces de sintetizar Spm mientras que la construcción trunca en el extremo C-terminal de la ZmSPMS1 no pudo hacerlo. Con estos resultados publicamos el artículo científico titulado “Characterization of maize spermine synthase 1 (ZmSPMS1): Evidence for dimerization and intracellular location” en la revista *Plant Physiology and Biochemistry*.



Research article

Characterization of maize spermine synthase 1 (ZmSPMS1): Evidence for dimerization and intracellular location



Israel Maruri-López^a, Itzell E. Hernández-Sánchez^a, Alejandro Ferrando^b,
Juan Carbonell^b, Juan Francisco Jiménez-Bremont^{a,*}

^a Laboratorio de Estudios Moleculares de Respuesta a Estrés en Plantas, División de Biología Molecular, Instituto Potosino de Investigación Científica y Tecnológica AC, Camino a la Presa de San José 2055, C.P. 78216, AP 3-74 Tangamanga, San Luis Potosí, San Luis Potosí, Mexico

^b Instituto de Biología Molecular y Celular de Plantas, Universidad Politécnica de Valencia-Consejo Superior de Investigaciones Científicas, Avda de los Naranjos s/n, 46022, Valencia, Spain

ARTICLE INFO

Article history:

Received 21 August 2015

Received in revised form

6 October 2015

Accepted 10 October 2015

Available online xxx

Keywords:

Bimolecular fluorescence complementation

Yeast two-hybrid

Spermine synthase

C-terminal

Homodimer

Subcellular localization

Spermine synthesis

ABSTRACT

Polyamines are ubiquitous positively charged metabolites that play an important role in wide fundamental cellular processes; because of their importance, the homeostasis of these amines is tightly regulated. Spermine synthase catalyzes the formation of polyamine spermine, which is necessary for growth and development in higher eukaryotes. Previously, we reported a stress inducible spermine synthase 1 (*ZmSPMS1*) gene from maize. The *ZmSPMS1* enzyme differs from their dicot orthologous by a C-terminal extension, which contains a degradation PEST sequence involved in its turnover. Herein, we demonstrate that *ZmSPMS1* protein interacts with itself in split yeast two-hybrid (Y2H) assays. A Bimolecular Fluorescence Complementation (BiFC) assay revealed that *ZmSPMS1* homodimer has a cytoplasmic localization. In order to gain a better understanding about *ZmSPMS1* interaction, two deletion constructs of *ZmSPMS1* protein were obtained. The Δ N-*ZmSPMS1* version, where the first 74 N-terminal amino acids were eliminated, showed reduced capability of dimer formation, whereas the Δ C-*ZmSPMS1* version, lacking the last 40 C-terminal residues, dramatically abated the *ZmSPMS1*-*ZmSPMS1* protein interaction. Recombinant protein expression in *Escherichia coli* of *ZmSPMS1* derived versions revealed that deletion of its N-terminal domain affected the spermine biosynthesis, whereas C-terminal *ZmSPMS1* truncated version fail to generate this polyamine. These data suggest that N- and C-terminal domains of *ZmSPMS1* play a role in a functional homodimer.

© 2015 Elsevier Masson SAS. All rights reserved.

1. Introduction

Polyamines (PAs) are ubiquitous molecules that have wide-ranging function in a number of physiological processes (Handa and Mattoo, 2010). The complexity of their metabolism and the fact that PAs homeostasis is tightly regulated support the idea that PAs are essential to cell survival (Moinard et al., 2005). In eukaryotes, the most abundant PAs are putrescine (Put), and the higher order PAs, such as spermidine (Spd) and spermine (Spm), all of them with specific biological functions. Moreover, in archaea, bacteria and plants, a structural isomer of spermine termed thermospermine (TSpm) has been found (Knott et al., 2007). Spd, Spm,

and TSpm are synthesized by the action of aminopropyl transferase enzymes, spermidine synthase (SPDS; EC 2.5.1.16), which converts Put into Spd and spermine synthase (SPMS; EC 2.5.1.22) or thermospermine synthase (TSPMS, ACL5; EC 2.5.1.79), that transforms Spd into Spm or TSpm, respectively. The aminopropyl transferases use decarboxylated S-adenosylmethionine as aminopropyl donor. Because of their positive charge, PAs can bind to anionic molecules, such as DNA, RNA, proteins, and chromatin; thus, PAs are implicated in many critical cellular processes including cell division, differentiation, replication, transcription, translation, enzymatic activity alterations, membrane stabilization, and the functioning of certain ion channels (Igarashi and Kashiwagi, 2000; Childs et al., 2003; Kusano et al., 2008; Zepeda-Jazo et al., 2011).

In particular, Spm may function as a free radical scavenger, protecting DNA from free radical attack (Ha et al., 1998). In mammals, Spm synthesis is essential for normal growth and development (Wang et al., 2004). In *Arabidopsis thaliana*, the *AtSPMS* gene is

* Corresponding author.

E-mail addresses: jbremont@ipicyt.edu.mx, jbremont@yahoo.com (J.F. Jiménez-Bremont).

Capítulo 3

Homo y hetedimerización de las enzimas arginina descarboxilasa (AtADC1 y AtADC2) de *Arabidopsis thaliana*.

La enzima arginina descarboxilasa (ADC) lleva a cabo la producción de agmatina a partir de la arginina, y la agmatina es precursor para la síntesis de la primera poliamina (PA), conocida como putrescina. En *A. thaliana*, la producción de PAs sólo se lleva a cabo a partir de arginina. En el genoma de *A. thaliana* existen dos genes (*AtADC1* y *AtADC2*) que codifican para enzimas ADC. Ambos genes ADC son esenciales para el desarrollo correcto del embrión en *A. thaliana* (Urano et al., 2005). El estudio de las enzimas *AtADC1* y *AtADC2* permitirá entender el papel de las PAs durante el desarrollo de las plantas.

La actividad del promotor del gen *AtADC1* es baja durante el desarrollo vegetativo temprano de *A. thaliana*, pero en condiciones de estrés por frío, aumenta considerablemente en las raíces y las hojas. Por el contrario, la actividad del promotor *AtADC2* está más relacionada con la germinación de las semillas, el desarrollo de raíces y hojas, y también responde a los tratamientos de luz, sacarosa y etileno (Hummel et al., 2004).

Reportes anteriores sobre la localización sub-celular de arginina descarboxilasas revelaron que la ADC de la avena (*Avena sativa*) se localiza en el interior del cloroplasto unido a membranas tilacoidales (Borrell et al., 1995); y la ADC de tabaco tiene una localización tanto como en el núcleo y como en el cloroplasto (Bortolotti et al., 2004). Un análisis *in silico* de las secuencias de aminoácidos de las proteínas *AtADC1* y *AtADC2* confirmó la presencia de péptidos señal para la localización en cloroplasto (Illingworth et al., 2003).

En este trabajo, analizamos la localización sub-celular de las proteínas *AtADC1* y *AtADC2* a través de una fusión traduccional con la proteína verde fluorescente GFP. Nuestros datos revelaron que ambas enzimas de *A. thaliana* exhiben una localización sub-celular tanto como en el citoplasma como en el

cloroplasto de células epidérmicas de hojas de *N. benthamiana*. Además, se examinó la capacidad para formar homodímeros de las proteínas AtADC1 y AtADC2 mediante el sistema BiFC. Nuestros análisis reveló que ambas enzimas, AtADC1 y AtADC2, son capaces de interactuar consigo mismas en el citoplasma y en el cloroplasto de células de hojas de *N. benthamiana*. Es interesante mencionar que nosotros encontramos la formación de heterodímeros entre las enzimas AtADC1 y AtADC2, con la misma localización sub-celular que en el caso de los homodímeros. Este estudio revela que ambas proteínas ADC de *A. thaliana* se encuentran localizadas en los mismos compartimentos celulares, y son capaces de formar complejos homo y heterodiméricos. Con estos resultados publicamos el artículo titulado “Hetero- and homodimerization of *Arabidopsis thaliana* Arginine Decarboxylase AtADC1 and AtADC2”.



Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Hetero- and homodimerization of *Arabidopsis thaliana* arginine decarboxylase AtADC1 and AtADC2

Israel Maruri-López, Juan F. Jiménez-Bremont*

Laboratorio de Biotecnología Molecular de Plantas, División de Biología Molecular, Instituto Potosino de Investigación Científica y Tecnológica AC, San Luis Potosí, México

ARTICLE INFO

Article history:

Received 17 January 2017

Accepted 17 January 2017

Available online xxx

Keywords:

BiFC

Heterodimer

Homodimer

Chloroplast localization

Polyamines

ABSTRACT

The arginine decarboxylase enzyme (ADC) carries out the production of agmatine from arginine, which is the precursor of the first polyamine (PA) known as putrescine; subsequently, putrescine is turned into the higher PAs, spermidine and spermine. In *Arabidopsis thaliana* PA production occurs only from arginine and this step is initiated by two ADC paralogues, AtADC1 and AtADC2. PA production is essential for *A. thaliana* life cycle. Here, we analyzed the sub-cellular localization of AtADC1 and AtADC2 enzymes through GFP translational fusions. Our data revealed that the *A. thaliana* arginine decarboxylase enzymes exhibit a dual sub-cellular localization both in the cytosol and chloroplast. Moreover, we examined the protein dimer assembly using a Bimolecular Fluorescence Complementation (BiFC) approach, which showed that AtADC1 and AtADC2 proteins were able to form homodimers in the cytosol and chloroplast. Interestingly, we found the formation of AtADC1/AtADC2 heterodimers with similar sub-cellular localization than homodimers. This study reveals that both ADC proteins are located in the same cell compartments, and they are able to form protein interaction complexes with each other.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

Putrescine (Put), spermidine (Spd), and spermine (Spm) are broad distributed low-molecular weight and positively charged compounds known as polyamines (PAs) [1]. In higher plants, the major pathway for Put production is through arginine. In *Arabidopsis thaliana*, Put is synthesized by arginine decarboxylase (ADC, EC 4.1.1.19), which exists as couple isoforms AtADC1 and AtADC2 [2]. These isoforms convert arginine to agmatine, which is subsequently turned into Put by a N-carbamoylputrescine transitional compound through two reactions, which involves agmatine deiminase/iminohydrolase (AIH, EC 3.5.3.12) and N-carbamoylputrescine amidase/amidohydrolase (NCPAH EC 3.5.1.53) [3]. Put is the precursor of Spd and Spm, through reactions catalyzed by spermidine synthases (AtSPDS1 and 2; EC 2.5.1.16) and spermine synthase (AtSPMS; EC 2.5.1.22), respectively.

In *A. thaliana*, Put synthesis has been related to a wide range of physiological processes throughout its growth and development, such as stem length, flowering rate, and seed morphology [4].

AtADC1 and *AtADC2* genes are essential for seed development. In *A. thaliana*, the double mutant *adc1^{-/-} adc2^{-/-}* is embryo lethal [5]. *AtADC1* and *AtADC2* mRNAs accumulate during development and stress responses [6]. The *A. thaliana* *adc1^{-/-}* and *adc2^{-/-}* single mutants exhibit an increased freezing sensitivity phenotype in contrast to the wild-type plants; which was alleviated by exogenous Put addition [7]. Furthermore, in an *A. thaliana* *AtADC2* overexpressing line, drought tolerance was attributed to Put accumulation [8]. In tobacco species and other members of the Solanaceae family, Put also functions as a substrate for the synthesis of nicotine, tropane, and nortropane alkaloids, which are all thought to deter herbivores attack [9]. It is well known that the *A. thaliana* *AtADC1* and *AtADC2* genes are differentially expressed. In fact, during early vegetative development, the *AtADC1* promoter activity is low but it is considerably increased in roots and leaves under chilling treatment. In contrast, *AtADC2* promoter activity is more related to seed germination, root and leaf development, and also in response to light, sucrose and ethylene treatments [6]. Previous reports about ADC protein sub-cellular localization revealed that the ADC protein from oat (*Avena sativa*) has been localized inside the chloroplast bound to thylakoid membranes [10]; whereas, the tobacco ADC protein showed dual chloroplast and nuclear localization [11]. An *in silico* analysis of *A. thaliana* *AtADC1* and *AtADC2* protein sequences

* Corresponding author.

E-mail addresses: jbremont@ipicyt.edu.mx, jbremont@yahoo.com (J.F. Jiménez-Bremont).

IV. Discusión

En las plantas con flores las PAs más abundantes son la diamina putrescina, la triamina espermidina y las tetraminas espermina y termospermina, cada una de ellas con funciones biológicas específicas (Takahashi et al., 2009). De acuerdo a los papeles fisiológicos y moleculares asignados a las PAs su homeostasis debe de estar controlada de manera estricta, de hecho estos compuestos están sujetos a controles biosintéticos, catabólicos y a través de su conjugación (Fuell et al., 2010; Moschou et al., 2008; Luo et al., 2009). En particular, las enzimas involucradas en la biosíntesis de las PAs están sujetas a diversos niveles de control que comprenden desde el transcripcional hasta el post-traducciona (Carbonell y Blázquez, 2009).

En esta tesis doctoral se caracterizó una región proteolítica tipo PEST de la espermina sintasa 1 de maíz (ZmSPMS1). Para ello, se fusionó el extremo C-terminal que incluye la secuencia PEST de la enzima ZmSPMS1 con el gen reportero GUS, lo cual condujo a la degradación de la proteína reportera GUS en líneas transgénicas de *A. thaliana* y en monocapas de cebolla (Maruri-López et al., 2014). Previamente, Rodríguez-Kessler et al. (2010) reportaron que la expresión una versión corta en el extremo C-terminal de la ZmSPMS1, que solo contenía 8 de los 17 residuos que conforman la secuencia PEST, fue capaz de producir Spm en la cepa mutante *spe4*Δ de *S. cerevisiae*. En conjunto estos resultados sugieren que la secuencia PEST regula la degradación de la enzima ZmSPMS1 de maíz, y que la versión trunca (con una secuencia PEST incompleta) de la ZmSPMS1 dio lugar a una versión estable en levadura capaz de producir Spm.

En mamíferos, uno de los mecanismos más estudiado de degradación a través de una secuencia PEST, es el de la enzima ornitina descarboxilasa (ODC) de ratón, la cual se encarga de iniciar la síntesis de PAs. La ODC es catalíticamente activa en forma de un homodímero y en su extremo C-terminal presenta una secuencia de degradación tipo PEST. Cuando los niveles de PAs son elevados la ODC interactúa con una proteína conocida como antizima, lo cual expone su secuencia PEST conduciendo a su rápida degradación a través del proteosoma 26S en una vía independiente ubiquitinación. La anti-enzima es

traducida completamente por un cambio en el marco de lectura (+1) en la traducción de su mRNA que conduce a la síntesis de una proteína mas larga, este desplazamiento en el marco de lectura también se regula de acuerdo con la concentración celular de PAs (Pegg 2006). La anti-enzima se ha encontrado en *Drosophila melanogaster* y en *S. cerevisiae*; sin embargo, esta no sido identificada en *A. thaliana* (Salzberg et al., 1996; Palanimurugan et al., 2004; Illingworth y Michael, 2011).

La expresión heteróloga de una versión trunca en el extremo C-terminal de la ODC de ratón en plantas de tabaco y la transferencia de esta secuencia PEST a la ODC de Tripanosoma (la cual no contiene secuencias PEST), demuestran que la degradación de la enzima ODC es debido a esta secuencia PEST (DeScenzo y Minocha 1993; Li y Coffino 1993). En plantas, la degradación de las enzimas relacionadas con la biosíntesis o el catabolismo de PAs ha sido poca explorada. Se ha reportado que una secuencia PEST presente en la enzima S-adenosilmetionina descarboxilasa 1 de *A. thaliana* (AtSAMDC1) está implicada en su degradación durante la infección del virus BSCTV establecer un ambiente hipometilado que facilite la acumulación viral. Este mecanismo revela la importancia de las secuencias PEST en las enzimas biosintéticas de PAs (Zhang et al. 2011).

Posteriormente, exploramos ¿cómo esta etiqueta de degradación no es reconocida en la ZmSPMS1? La oligomerización de este tipo de enzimas SPMS ya se ha reportado en varios organismos, tales como mamíferos, plantas y hongos. En *A. thaliana*, la enzima SPMS es capaz de interactuar consigo misma en el citoplasma de células de hojas de *N. benthamiana* (Belda-Palazón et al., 2012). También, la resolución de la estructura cristalográfica de la SPMS de *Homo sapiens* (HsSPMS) reveló que esta enzima puede formar un homodímero a través de su dominio N-terminal (Wu et al., 2008).

Mediante el sistema de Y2H “split-ubiquitin” demostramos que la enzima ZmSPMS1 de maíz interactúa con ella misma, y dicha interacción se confirmó a través de un ensayo de BiFC en hojas de *N. benthamiana* (Maruri-López et al., 2015). Además, con el análisis BiFC determinamos que el homodímero ZmSPMS1 se localiza en el citoplasma de las células epidérmicas de tabaco

(Maruri-López et al., 2015). Además, se evaluó la interacción de dos versiones truncas en el extremo N- y C-terminal de la enzima ZmSPMS1 de maíz y se observó que la delección de los primeros 74 residuos N-terminales afectaron la formación del dímero ZmSPMS1 en células de levadura y tabaco. Sin embargo, la delección de la región C-terminal resultó en una pérdida de la interacción ZmSPMS1-ZmSPMS1. Nosotros reportamos que el dominio C-terminal de la ZmSPMS1 contiene una secuencia de degradación tipo PEST.

La conformación flexible de las secuencias PEST confieren un papel primordial en las interacciones proteína-proteína (Sandhu y Dash, 2006). En mamíferos, un ensayo de Y2H nuclear mostró que la región C-terminal de la quinasa 2 (HIPK2), que contienen una secuencia PEST, es necesaria para la interacción con el supresor de proteína supresora de tumores p73- α (Kim et al., 2002). Nuestros resultados sugieren que el extremo C-terminal de la ZmSPMS1 que contiene la secuencia PEST de la ZmSPMS1 está implicado en su homodimerización.

Además, la versión trunca en el extremo C-terminal de la ZmSPMS1 no es capaz de producir Spm en células de *E. coli* (la cual no produce Spm), en comparación con la versión completa y trunca en el N-terminal. La dimerización de las enzimas SPMS es requisito indispensable para su actividad. La enzima SPMS de humano dimeriza a través de su dominio N-terminal. Si se le elimina dicha región, la enzima es incapaz de producir Spm (Wu et al., 2008). El dominio N-terminal de la SPMS de humano no esta presente en las enzimas de plantas y hongos (Wu et al., 2008); sin embargo, las SPMS de plantas y hongos poseen una extensión en su extremo C-terminal. Nuestros datos sugieren que el domino de C-terminal de la enzima ZmSPMS1 está involucrado en la estabilidad, y la capacidad de homodimerización y por consiguiente esto regula la síntesis de Spm.

La importancia de la Spm para el crecimiento celular ha sido reportada en ratón, humano y en otros organismos (Wang et al., 2004). En particular, en ratón la eliminación del gen *SPMS* conduce a síntomas de hiperactividad, agitación, anomalías en el oído interno, sordera, esterilidad, retardo del crecimiento postnatal y propensión a la muerte súbita (Wang et al., 2004). En

humanos, el síndrome de Snyder Robinson es causado por mutaciones puntuales que afectan a la formación de homodímeros de la enzima HsSPMS, lo cual reduce drásticamente su actividad enzimática. Los pacientes con el síndrome Snyder Robinson muestran retardo mental leve a moderado, osteoporosis y asimetría facial (Cason et al., 2003). En este sentido, una versión quimérica de la HsSPMS de humano, a la cual se le transfirió una secuencia de la SPMS de *Thermotoga marítima* dentro de los sitios conocidos que albergan las mutaciones causantes del síndrome Snyder Robinson, fue más estable y catalíticamente activa que la enzima humana (Zhang et al., 2013).

En las plantas, la vía principal para la producción de la diamina putrescina es a partir de la arginina. En *A. thaliana*, la putrescina se sintetiza por dos enzimas arginina descarboxilasas que se encuentran codificadas por los genes parálogos *AtADC1* y *AtADC2* (Hanfrey et al., 2001). Durante este trabajo de doctorado también se analizó la localización sub-celular de las enzimas *AtADC1* y *AtADC2* de *A. thaliana* a través de fusiones tradicionales con la proteína verde fluorescente GFP.

Nuestros resultados demostraron que ambas enzimas están localizadas en el citoplasma y en el cloroplasto de las células epidérmicas de hojas de *N. benthamiana*. Previamente, Borrell et al. (1995) demostraron que la proteína ADC de avena está localizada en el cloroplasto unido a las membranas de tilacoidales. Estos autores hipotetizan que la asociación con los tilacoides podría ser a través de un dominio putativo α hélice anfipático.

Mediante un ensayo de BiFC, demostramos que las proteínas *AtADC1* y *AtADC2* son capaces de formar homodímeros los cuales se localizan tanto en el citoplasma y como en el cloroplasto. Anteriormente, Hanfrey et al. (2001) propusieron que la enzima *AtADC1* podría ser catalíticamente activa como un homodímero. Los autores observaron que una mutante de la *AtADC1* con cambios en los residuos lisina 136 y cisteína 524 de la *AtADC1* redujo drásticamente su actividad enzimática. Estos autores también propusieron que el homodímero tiene dos sitios activos que actúan a través de la interfaz del dímero. Análisis biofísicos sugieren que la proteína recombinante ADC de

Dianthus caryophyllus se ensambla como un tetrámero, producto de dos dímeros (Ha et al., 2004).

Además, se observó que las enzimas AtADC1 y AtADC2 son capaces de formar un heterodímero en el citoplasma y el cloroplasto. Se ha reportado que las espermidina sintetas AtSPDS1 y AtSPDS2 de *A. thaliana*, enzimas implicadas en la síntesis de espermidina, pueden interactuar consigo mismas y también formar heterodímeros (Panicot et al., 2002; Belda-Palazón et al., 2012). Además, se ha propuesto que el heterodímero AtSPDS1 y AtSPDS2 podría ser enzimáticamente activo, lo que podría proporcionar una ventaja biológica que podría facilitar la síntesis de Spd o exhibir funciones reguladoras sobre la actividad de cada dominio enzimático (Panicot et al., 2002).

Finalmente, la localización sub-celular específica de estas enzimas implica que la biosíntesis de PAs tiene lugar en diferentes compartimentos y sugiere funciones específicas para cada poliamina.

V. Conclusiones

- La enzima ZmSPMS1 se degrada a través de una etiqueta PEST presente en su extremo C-terminal vía el proteosoma 26S.
- La enzima ZmSPMS1 es capaz de formar un homodímero in planta localizado en el citosol.
- La eliminación del extremo C-terminal de la ZmSPMS1 abate su homodimerización.
- El dominio C-terminal de la ZmSPMS1 está involucrado en su estabilidad, en su capacidad de homodimerización, y regula su actividad enzimática.
- Las argininas descarboxilasas AtADC1 y AtADC2 de *A. thaliana* están localizadas tanto en el citosol como en el cloroplasto.
- Ambas proteínas AtADC1 y AtADC2 son capaces de formar complejos homo y heterodiméricos con una localización citosólica y cloroplástica.

Perspectivas

- Identificar los residuos involucrados en la degradación de la enzima ZmSPMS1.
- Determinar la actividad enzimática de cada versión de la ZmSPMS1 en un fondo doble mutante *spms^{-/-}/tspms^{-/-}* de *A. thaliana*.
- Estudiar si la localización de ADC puede ocurrir en otros plástidos como cromoplastos o leucoplastos.
- Analizar la formación de homo y heterodímeros de las enzimas ADC en otros tejidos de *A. thaliana*.
- Determinar que papel que tiene la homo o heterodimerización de las ADC en su actividad enzimática.

VI. Referencias

Ahou, A., Martignago, D., Alabdallah, O., Tavazza, R., Stano, P., Macone, A., Pivato, M., Masi, A., Rambla, J.L., Vera-Sirera, F., Angelini, R., Federico, R., Tavladoraki, P., 2014. A plant spermine oxidase/dehydrogenase regulated by the proteasome and polyamines. *J. Exp. Bot.* 65 (6), 1585-1603. <http://dx.doi.org/10.1093/jxb/eru016>.

Alcázar, R., Altabella, T., Marco, F., Bortolotti, C., Reymond, M., Koncz, C., Carrasco, P., Tiburcio, A. F., 2010. Polyamines: molecules with regulatory functions in plant abiotic stress tolerance. *Planta.* 231, 1237-1249.

Belda-Palazón, B., Ruiz, L., Martí, E., Tárraga, S., Tiburcio, A. F., Culiáñez, F., Farràs, R., Carrasco, P., Ferrando, A., 2012. Aminopropyltransferases involved in polyamine biosynthesis localize preferentially in the nucleus of plant cells. *PLoS One* 7:e46907. doi:10.1371/journal.pone.0046907.

Biastoff, S., Brandt, W., Draeger, B., 2009. Putrescine N-methyltransferase- the start for alkaloids. *Phytochemistry* 70, 1708-1718. doi:10.1016/j.phytochem.2009.06.012.

Borrell, A., Culiáñez-Macia, F. A., Altabella, T., Besford, R. T., Flores, D., Tiburcio, A. F., 1995. Arginine decarboxylase is localized in chloroplasts. *Plant Physiol.* 109, 771-776.

Bouchereau, A., Aziz, A., Larher F., Martin-Tanguy, J., 1999. Polyamines and environmental challenges: recent development. *Plant Sci.* 140, 103-125.

Carbonell, J., Blázquez, M.A., 2009. Regulatory mechanisms of polyamine biosynthesis in plants. *Gene Genom.* 31, 107-118.

Chattopadhyay, M. K., Park, M. H., Tabor, H., 2008. Hypusine modification for growth is the major function of spermidine in *Saccharomyces cerevisiae* polyamine auxotrophs grown in limiting spermidine. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6554-6559. doi:10.1073/pnas.0710970105.

Childs, A.C., Mehta, D.J., Gerner, E.W., 2003. Polyamine-dependent gene expression. *Cell. Mol. Life Sci.* 60, 1394-1406.

Cona, A., Rea, G., Angelini, R., Federico, R., Tavladoraki, P., 2006. Functions of amine oxidases in plant development and defence. *Trends Plant Sci.* 11, 80-88.

Cuevas, J. C., Lopez-Cobollo, R., Alcazar, R., Zarza, X., Koncz, C., Altabella, T., et al., 2008. Putrescine is involved in *Arabidopsis* freezing tolerance and cold acclimation by regulating abscísico acid levels in response to low temperature. *Plant Physiol.* 148, 1094-1105. doi:10.1104/pp.108.122945.

Cuevas, J. C., Lopez-Cobollo, R., Alcazar, R., Zarza, X., Koncz, C., Altabella, T., et al., 2008. Putrescine is involved in *Arabidopsis* freezing tolerance and cold acclimation by regulating abscísico acid levels in response to low temperature. *Plant Physiol.* 148, 1094-1105. doi:10.1104/pp.108.122945.

Dobrovinskaya, O. R., Muniz, J., Pottosin, I. I., 2000. Asymmetric block of the plant vacuolar Ca²⁺-permeable channel by organic cations. *Eur. Biophys. J.* 28, 552-563. doi: 10.1007/s002490050237

El Amrani, A., Marie, L., Aïnouche, A., Nicolas, J., Couée, I., 2002. Genome-wide distribution and potential regulatory functions of AtATE, a novel family of miniature inverted-repeat transposable elements in *Arabidopsis thaliana*. *Molecular Genetics and Genomics.* 267, 459-471.

Fincato, P., Moschou, P. N., Spedaletti, V., Tavazza, R., Angelini, R., Federico, R., et al., 2011. Functional diversity inside the *Arabidopsis* polyamine oxidase gene family. *J. Exp. Bot.* 62, 1155-1168. doi: 10.1093/jxb/erq341.

Flores, H.E., Galston, A.W., 1982. Analysis of polyamines in higher plants by high performance liquid chromatography. *Plant Physiol.* 69, 701-706.

Fuell, C., Elliott, K. A., Hanfrey, C. C., Franceschetti, M., Michael, A. J., 2010. Polyamine biosynthetic diversity in plants and algae. *Plant Physiol Biochem.* 48, 513-520.

Grienenberger, E., Besseau, S., Geoffroy, P., Debayle, D., Heintz, D., Lapierre, C., Pollet, B., Heitz, T., Legrand, M., 2009. A BAHD acyltransferase is expressed in the tapetum of Arabidopsis anthers and is involved in the synthesis of hydroxycinnamoyl spermidines. *Plant J* 58, 246-259.

Groppa, M. D., Benavides, M. P., 2008. Polyamines and abiotic stress: recent advances. *Amino Acids*. 34, 35-45.

Gsponer, J., Futschik, M.E., Teichmann, S.A., Babu, M.M., 2008. Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. *Science*. 322 (5906), 1365-1368.

Guerrero-González, M.L., Rodriguez-Kessler, M., Jimenez-Bremont, J.F., 2014. uORF, a regulatory mechanism of the Arabidopsis polyamine oxidase 2. *Mol. Biol. Rep.* <http://dx.doi.org/10.1007/s11033-014-3098-5>.

Ha, B. H., Cho, K. J., Choi, Y. J., Park, K. Y., Kim, K. H., 2004. Characterization of arginine decarboxylase from *Dianthus caryophyllus*, *Plant Physiol. Biochem.* 42, 307-311.

Ha, H.C., Sirisoma, N.S., Kuppusamy, P., Zweier, J.L., Woster, P.M., Casero Jr., R.A., 1998. The natural polyamine spermine functions directly as a free radical scavenger. *Proc. Natl. Acad. Sci.* 95 (19), 11140-11145.

Handa, A.K., Mattoo, A.K., 2010. Differential and functional interactions emphasize the multiple roles of polyamines in plants. *Plant Physiol. Biochem.* 48, 540-546.

Hanfrey, C., Elliott, K. A., Franceschetti, M., Mayer, M. J., Illingworth, C., Michael, A. J., 2005. A dual upstream open reading frame-based autoregulatory circuit controlling polyamine-responsive translation. *J Biol Chem.* 280, 39229-39237.

Hanzawa, Y., Takahashi, T., Komeda, Y., 1997. ACL5: an Arabidopsis gene required for internodal elongation after flowering. *Plant J.* 12, 863-874. doi: 10.1046/j.1365-313X.1997.12040863.x

Hummel, I., Bourdais, G., Gouesbet, G., Couée, I., Malmberg, R.L., El Amrani, A., 2004. Differential expression of arginine decarboxylase ADC1 and ADC2 in *Arabidopsis thaliana*: characterization of transcriptional regulation during seed germination and seedling development. *New Phyt.* 163, 519-531.

Igarashi, K., Kashiwagi, K., 2000. Polyamines: mysterious modulators of cellular functions. *Biochem Biophys. Res. Commun.* 271, 559-564.

Illingworth, C., Mayer, M.J., Elliott, K., Hanfrey, C., Walton, N. J., Michael, A. J., 2003. The diverse bacterial origins of the *Arabidopsis* polyamine biosynthetic pathway. *FEBS Lett.* 549, 26-30.

Imai, A., Matsuyama, T., Hanzawa, Y., Akiyama, T., Tamaoki, M., Saji, H., Shirano, Y., Kato, T., Hayashi, H., Shibata, D., Tabata, S., Komeda, Y., Takahashi, T., 2004. Spermidine Synthase Genes Are Essential for Survival of *Arabidopsis*. *Plant Physiol.* 135, 1565–1573. doi:10.1104/pp.104.041699.

Ivanov, I. P., Atkins, J. F., Michael, A. J., 2010. A profusion of upstream open reading frame mechanisms in polyamine-responsive translational regulation. *Nucleic Acids Res.* 38, 353-359.

Jansson, B. P. M., Malandrin, L., Johansson, H. E., 2000. Cell cycle arrest in archaea by the hypusination inhibitor N1-guanyl-1,7-diaminoheptane. *J. Bacteriol.* 182,1158-1161.

Kamada-Nobusada, T., Hayashi, M., Fukazawa, M., Sakakibara, H., Nishimura, M., 2008. A putative peroxisomal polyamine oxidase, AtPAO4, is involved in polyamine catabolism in *Arabidopsis thaliana*. *Plant Cell Physiol.* 49, 1272-1282. doi: 10.1093/pcp/pcn114.

Kim, D. W., Watanabe, K., Murayama, C., Izawa, S., Niitsu, M., Michael, A. J., Berberich, T., Kusano, T., 2014. Polyamine oxidase5 regulates *Arabidopsis* growth through thermospermine oxidase activity. *Plant Physiol.* 165, 1575-1590.

Knott, J.M., Roemer, P., Sumper, M., 2007. Putative spermine synthases from *Thalassiosira pseudonana* and *Arabidopsis thaliana* synthesize thermospermine rather than spermine. *FEBS Lett.* 581, 3081-3086.

Korolev, S., Ikeguchi, Y., Skarina, T., Beasley, S., Arrowsmith, C., et al., 2002. The crystal structure of spermidine synthase with a multisubstrate adduct inhibitor. *Nat. Struct. Mol. Biol.* 9, 27-31.

Kusano, T., Berberich, T., Tateda, C., Takahashi, Y., 2008. Polyamines: essential factors for growth and survival. *Planta* 228, 367-381.

Li, X., Coffino, P., 1993. Degradation of ornithine decarboxylase: exposure of the Cterminal target by a polyamine-inducible inhibitory protein. *Mol. Cell. Biol.* 13, 2377-2383.

Luo, J., Fuell, C., Parr, A., Hill, L., Bailey, P., et al., 2009. A Novel Polyamine Acyltransferase Responsible for the Accumulation of Spermidine Conjugates in *Arabidopsis* Seed. *Plant Cell.* 21, 318-333.

Maruri-López, I., Hernández-Sánchez I. E., Ferrando A., Carbonell, J., Jiménez-Bremont, J. F., 2015. Characterization of maize spermine synthase 1 (ZmSPMS1): Evidence for dimerization and intracelular location. 97, 264-271. doi:10.1016/j.plaphy.2015.10.017.

Maruri-López, I., Rodríguez-Kessler, M., Rodríguez-Hernández, A. A., Becerra-Flora, A., Olivares-Grajales, J. E., Jiménez-Bremont, J. F., 2014. A maize spermine synthase 1 PEST sequence fused to the GUS reporter protein facilitates proteolytic degradation. *Plant Physiol. Biochem.* 78, 80-87.

Mo, H., Wang, X., Zhang, Y., Zhang, G., Zhang, J. F., Ma, Z. Y., 2015. Cotton polyamine oxidase is required for spermine and camalexin signalling in the defence response to *Verticillium dahliae*. *Plant J.* 83, 962-975. doi: 10.1111/tpj.12941.

Møller, S. G., McPherson, M. J., 1998. Developmental expression and biochemical analysis of the *Arabidopsis atao1* gene encoding an H₂O₂-generating diamine oxidase. *Plant J.* 13, 781-791.

Moschou, P. N., Paschalidis, K. A., & Roubelakis-Angelakis, K. A., 2008. Plant polyamine catabolism: The state of the art. *Plant Signaling & Behavior*, 3(12), 1061-1066.

Moschou, P. N., Sanmartin, M., Andriopoulou, A. H., Rojo, E., Sanchez-Serrano, J. J., Roubelakis-Angelakis, K. A., 2008. Bridging the gap between plant and mammalian polyamine catabolism: A novel peroxisomal polyamine oxidase responsible for a full back-conversion pathway in *Arabidopsis*. *Plant Physiol.* 147, 1845-1857.

Oshima, T., 2007. Unique polyamines produced by an extreme thermophile, *Thermus thermophilus*. *Amino Acids*. 33, 367-372.

Palanimurugan R, et al., 2004. Polyamines regulate their synthesis by inducing expression and blocking degradation of ODC antizyme. *EMBO J* 23(24): 4857-67.

Panagiotidis, C. A., Blackburn, S., Low, K. B., Canellakis, E. S., 1987. Biosynthesis of polyamines in ornithine decarboxylase, arginine decarboxylase, and agmatine ureohydrolase deletion mutants of *Escherichia coli* strain K-12. *Proc. Natl. Acad. Sci. U. S. A.* 84, 4423e4427.

Park, M., 2006. The post-translational synthesis of a polyamine-derived amino acid, hypusine, in the eukaryotic translation initiation factor 5A (eIF5A). *J. Biochem.* 139, 161-169. doi:10.1093/jb/mvj034.

Planas-Portell, J., Gallart, M., Tiburcio, A. F., et al., 2013. Copper-containing amine oxidases contribute to terminal polyamine oxidation in peroxisomes and apoplast of *Arabidopsis thaliana*. *BMC Plant Biol.* 13, 109.

Pottosin, I., Shabala, S., 2014. Polyamine control of cation transport across plant membranes: implications for ion homeostasis and abiotic stress signaling. *Front. Plant Sci.* 5:154. doi:10.3389/fpls.2014.00154.

Pottosin, I., Velarde-Buendia, A. M., Bose, J., Zepeda-Jazo, I., Shabala, S., Dobrovinskaya, O., 2014. Cross-talk between reactive oxygen species and

polyamines in regulation of ion transport across the plasma membrane: implications for plant adaptive responses. *J Exp Bot.* doi:10.1093/jxb/ert1423.

Rechsteiner, M., Rogers, S.W., 1996. PEST sequences and regulation by proteolysis. *Trends Biochem Sci.* 21 (7), 267-271.

Rob, M., Roelfsema, G., Hedrich, R., 2005. In the light of stomatal opening: new insights into 'the Watergate' . *New Phytol.* 167, 665-691.

Rodríguez-Kessler, M., Alpuche-Solís, A.G., Ruiz, O.A., Jiménez-Bremot, J.F., 2006. Effect of salt stress on the regulation of maize (*Zea mays* L.) genes involved in polyamine biosynthesis. *Plant Growth Reg.* 48, 175-185.

Rodríguez-Kessler, M., Delgado-Sánchez, P., Rodríguez-Kessler, G.T., Moriguchi, T., Jiménez-Bremot, J.F., 2010. Genomic organization of plant aminopropyl transferases. *Plant Physiol. Biochem* 48 (7), 574-590.

Rodríguez-Kessler, M., Jiménez-Bremont, J.F., 2008. *Zmspds2* maize gene: coding a spermine synthase? *Plant Signal Behav.* 3 (8), 551-553.

Rogers, S.W., Wells, R., Rechsteiner, M., 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* 234, 364-368.

Salzberg, A., Golden, K., Bodmer, R., Bellen, H. J. 1996. Gutfeeling, a *Drosophila* Gene Encoding an Antizyme-like Protein, Is Required for Late Differentiation of Neurons and Muscles. *Genetics.* 144 (1): 183-196.

Sánchez-Rangel, D., Chávez-Martínez, A., Rodríguez-Hernández, A. A., Maruri-López, I., Urano, K., Shinozaki, K., Jiménez-Bremont, J. F., 2016. Simultaneous Silencing of Two Arginine Decarboxylase Genes Alters Development in *Arabidopsis*. *Front. Plant Sci.* <http://dx.doi.org/10.3389/fpls.2016.00300>.

Sandhu, K.S., Dash, D., 2006. Conformational flexibility may explain multiple cellular roles of PEST motifs. *Proteins.* 63 (4), 727-732.

Singh, G.P., Ganapathi, M., Sandhu, K.S., Dash, D., 2005. Intrinsic unstructuredness and abundance of PEST motifs in eukaryotic proteomes. *Proteins*. 62, 309-315.

Takahashi, T., Kakehi, J. I., 2009. Polyamines: ubiquitous polycations with unique roles in growth and stress responses. *Ann Bot*. 105, 1-6.

Tavladoraki, P., Rossi, M. N., Saccuti, G., Perez-Amador, M. A., Polticelli, F., Angelini, R., et al., 2006. Heterologous expression and biochemical characterization of a polyamine oxidase from *Arabidopsis* involved in polyamine back conversion. *Plant Physiol*. 141, 1519-1532. doi: 10.1104/pp.106.080911

Tong, W., Yoshimoto, K., Kakehi, J. I., Motose, H., Niitsu, M., Takahashi, T., 2014. Thermospermine modulates expression of auxin-related genes in *Arabidopsis*. *Front Plant Sci*. 5, 94.

Urano, K., Hobo, T., Shinozaki, K., 2005. *Arabidopsis* ADC genes involved in polyamine biosynthesis are essential for seed development. *FEBS Lett*. 579, 1557-1564. doi: 10.1016/j.febslet.2005.01.048.

Walters, D. R., 2003. Polyamines and plant disease. *Phytochemistry*. 64, 97-107.

Wu, H., Min, J., Zeng, H., McCloskey, D.E., Ikeguchi, Y., et al., 2008. Crystal structure of human spermine synthase. *J. Biol. Chem*. 283, 16135-16146.

Yamaguchi, K., Takahashi, Y., Berberich, T., Imai, A., Takahashi, T., Michael, A.J., Kusano, T., 2007. A protective role for the polyamine spermine against drought stress in *Arabidopsis*. *Biochem. Biophys. Res. Commun*. 352, 486-490.

Yoshimoto, K., Noutoshi, Y., Hayashi, K., Shirasu, K., Takahashi, T., Motose, H., 2012a. A chemical biology approach reveals an opposite action between thermospermine and auxin in xylem development in *Arabidopsis thaliana*. *Plant Cell Physiol* 53, 635-645.

Yoshimoto, K., Noutoshi, Y., Hayashi, K., Shirasu, K., Takahashi, T., Motose, H., 2012b. Thermosperminesuppresses auxin-inducible xylem differentiation in *Arabidopsis thaliana* . *Plant Signal Behav.* 7, 937-939.

Zepeda-Jazo, I., Velarde-Buendia, A.M., Enriquez-Figueroa, R., Bose, J., Shabala, S., Muneiz-Murguia, J., Pottosin, I.I., 2011. Polyamines interact with hydroxyl radicals in activating Ca(2+) and K(2+) transport across the root epidermal plasma membranes. *Plant Physiol.* 157 (4), 2167-2180.

Zhang, Z., Chen, H., Huang, X., Xia, R., Zhao, Q., Lai, J., Teng, K., Li, Y., Liang, L., Du, Q., Zhou, X., Guo, H., Xie, Q., 2011. BSCTV C2 attenuates the degradation of SAMDC1 to suppress DNA methylation-mediated gene silencing in *Arabidopsis*. *Plant Cell.* 23, 273-288.

Zhang, Z., Zheng, Y., Petukh, M., Pegg, A., Ikeguchi, Y., Alexov, E., 2013. Enhancing human spermine synthase activity by engineered mutations. *PLOS Comput. Biol.* 9 (2), e1002924.

VII. Anexo

Análisis *in silico* de los factores de transcripción WRKY de *Arabidopsis*: presencia de secuencias PEST en regiones desordenadas.

En las plantas, las proteínas WRKY constituyen una de las familias más grandes de factores de transcripción (Eulgem et al., 2000), los cuales están involucrados con varios procesos biológicos específicos, tales como la resistencia a enfermedades, la respuesta al estrés abiótico, la privación de nutrientes, la senescencia, el desarrollo de semillas y tricomas (Luo et al., 2002; Pandey y Somssich, 2009; Bakshi y Oelmüller, 2014).

Los factores de transcripción WRKY contienen un dominio de unión a DNA altamente conservado de 60 aminoácidos que comprende la secuencia WRKYGQK en el extremo N-terminal y una firma tipo dedo de zinc en el extremo C-terminal, el cual se llamado WRKY (Eulgem et al., 2000). En *A. thaliana*, la familia WRKY se clasifica en tres principales grupos (Ia y b, IIa y b y III), los cuales difieren en el número de dominios WRKY y en el patrón del motivo dedo de zinc. Así, las proteínas WRKY con dos dominios WRKY pertenecen al grupo I, mientras que la mayoría de las proteínas con un dominio WRKY pertenecen al grupo II. Ambos grupos (I y II) tienen el mismo patrón de sitios potenciales de unión al zinc. Además, las proteínas WRKY que tienen sólo un dominio WRKY pero diferentes patrones de motivos de dedo de zinc se clasifican como grupo III (Wang et al., 2011).

Aunque los factores de transcripción WRKY han sido objeto de numerosos estudios, la presencia de secuencias proteolíticas en estas proteínas ha sido un campo menos explorado. Durante mi doctorado examinamos la presencia de secuencias tipo PEST en 66 miembros de la familia WRKY de *A. thaliana*. Nuestro análisis reveló que el 67% de los miembros de la familia WRKY contienen al menos una potencial región PEST, las cuales se pueden encontrar de uno hasta seis veces distribuidas a lo largo de toda su cadena polipeptídica.

Específicamente, se detectó que la mayoría de las proteínas que contienen PEST eran miembros de los grupos WRKY I (82%) y II (62%); además,

observamos que el grupo III poseen sólo el 46% de sus miembros con secuencias de PEST. En particular, nuestro análisis muestra un alto número de presencia de secuencias PEST dentro de las regiones desordenadas de los factores de transcripción WRKY, principalmente en los segmentos que son adyacentes al dominio de unión DNA. Además, llevamos a cabo un análisis del contenido de aminoácidos de las proteínas WRKY de *A. thaliana*, el cual reveló que estos factores de transcripción contienen una baja cantidad de residuos promotores de orden como triptófano, isoleucina, leucina, fenilalanina, valina y tirosina, y son enriquecidos en aminoácidos promotores de desorden, tales como serina, prolina, glutamina y arginina.

Finalmente, se realizó un análisis comparativo sobre la presencia de secuencias PEST entre la familia WRKY de *A. thaliana* y los ortólogos de una especie altamente relacionada, *Arabidopsis lyrata*. Nuestro análisis *in silico* reveló que las secuencias PEST están conservadas en número y localización entre las proteínas WRKY de ambas especies de Brassicaceae. Nuestro estudio abre la posibilidad de nuevos análisis de secuencias PEST como una regulación post-traducciona l de estos importantes factores de transcripción de plantas. Con estos resultados redactamos el manuscrito titulado "Arabidopsis WRKY transcription factors: occurrence of PEST sequences in disordered regions".

Title:

**Arabidopsis WRKY transcription factors: occurrence of
PEST sequences in disordered regions**

Running title:

PEST sequences in Arabidopsis WRKY proteins

Israel Maruri-López, Itzell E. Hernández-Sánchez, and Juan Francisco Jiménez-Bremont

Laboratorio de Biología Molecular de Hongos y Plantas, División de Biología Molecular, Instituto Potosino de Investigación Científica y Tecnológica AC, Camino a la Presa de San José 2055, C.P. 78216, AP 3-74 Tangamanga, San Luis Potosí, San Luis Potosí, México.

*Corresponding author. E-mail addresses: jbremont@ipicyt.edu.mx, jbremont@yahoo.com (J.F. Jiménez-Bremont).

ABSTRACT

The *Arabidopsis thaliana* WRKY transcription factors (TFs) are involved in plant development and its response to biotic and abiotic stresses. Furthermore, these TFs are regulated through their synthesis or degradation rates, post-translational modifications, and subcellular localization. Herein, we examined the PEST sequence occurrence in 66 members of the *A. thaliana* WRKY family. The PEST sequences are one of the major motifs used in protein degradation; these sequences are enriched in disorder-promoting residues, such as proline, glutamic acid, serine, and threonine. Our analysis revealed that 67% of the members of the WRKY family contain potential PEST sequences, which can be found from one up to six times distributed throughout the polypeptide chain. The majority of the PEST-containing proteins were detected to be members of the WRKY groups I (a and b) and II (a and b) that exhibit 82 and 62%, respectively; in addition, we noticed that group III possess only 46% of members with PEST sequences. We found that the PEST sequences in WRKY proteins are located in the predicted disordered regions, which flank the WRKY DNA-binding domains. Analysis of amino acid composition from *A. thaliana* WRKY proteins has revealed that these TFs are diminished in order-promoting residues and enriched in disorder-promoting amino acids. Finally, a comparative analysis of PEST sequence occurrence in *A. thaliana* WRKY family with the highly related *Arabidopsis lyrata* species was conducted. Our *in silico* analysis revealed that PEST sequences have been conserved in number and localization between WRKY TFs from both *Brassicaceae* species. We reported an important incidence of PEST sequences within disordered regions in both WRKY TFs of *Arabidopsis* species, particularly flanking the WRKY DNA-binding domain.

Keywords: WRKY family, PEST sequences, disorder predicted regions, transcription factors, *in silico* analysis.

INTRODUCTION

Transcriptional activation is the first step of gene expression and is modulated by the coordination of various elements, such as transcription factors (TFs), co-activators, and the RNA polymerase II (Cooper, 2000). In particular, TF activity may be modulated by several processes, such as synthesis or degradation, post-translational modifications, and its subcellular location. Although each TF is eventually degraded, TF stability is a strictly regulated process (Desterro et al., 2000). In this sense, ubiquitin-dependent proteolysis may control the life span of TFs and other regulatory proteins; in fact, several sequence motifs have been identified in proteins as degradation signals. The PEST sequences are one of the major motifs for protein degradation (Rogers et al., 1986). They are enriched in disorder-promoting residues, such as proline (P), glutamic acid (E), serine (S) and threonine (T), and contribute to the enrichment of disorder regions (Singh et al., 2006). It has been reported that eukaryotes TFs possess disordered segments; in human TFs, disordered regions flank the DNA-binding domains (Liu et al., 2006; Guo et al., 2012). For this reason, the analysis of the disordered segments present in TFs has important implications for understanding their stability.

In plants, WRKY proteins are highly expanded, constituting one of the largest families of TFs. Several biological processes are regulated by WRKYs, having important functions in disease resistance, abiotic stress response, nutrient deprivation, senescence, seed and trichome development, embryogenesis, and hormone-controlled processes (Johnson et al., 2002; Luo et al., 2005; Pandey and Somssich, 2009; Bakshi and Oelmüller, 2014). Their members contain the highly conserved 60 amino acid long WRKY DNA-binding domain, which comprise the sequence WRKYGQK at the N-terminus and a metal chelating zinc finger signature at the C-terminus. In order to regulate gene expression, the WRKY domain recognizes and binds to the W-box in the DNA promoter sequence. In *Arabidopsis thaliana*, the WRKY family is classified into three major groups (Ia and b, IIa and b, and III), which differ in the number of WRKY domains and in the pattern of the zinc-finger motif. The WRKY proteins with two WRKY domains belong to group I, whereas most proteins with one WRKY

domain belong to group II. Both groups (I and II) have the same pattern of potential zinc binding sites (C-X₄₋₅-C-X₂₂₋₂₃-H-X₁-H). Additionally, WRKY proteins that have only one WRKY domain but different patterns of zinc finger motifs are categorized as group III. While, the groups I and II have a C₂-H₂ pattern, group III WRKY domains contain a C₂-H-C motif (C-X₇-C-X₂₃-H-X₁-C) (Eulgem et al., 2000; Wang et al., 2011).

Although TFs are subject of numerous studies, the study of proteolytic sequences in these proteins has been less explored. In this work, we examine the PEST sequence occurrence in WRKY transcription factors from the highly related *A. thaliana* and *A. lyrata* species. The WRKY orthologous TFs from these *Brassicaceae* species were also analyzed in order to identify a) PEST sequence occurrence and localization, b) presence of disordered regions, and c) protein amino acid composition. In this sense, our analyses show a high degree of PEST sequence occurrence within disordered regions in Arabidopsis WRKY TFs, mainly in segments that are adjacent to WRKY domain.

METHODS

***Arabidopsis* WRKY protein sequences**

First, the *Arabidopsis thaliana* WRKY protein sequences were obtained from the Arabidopsis Gene Regulatory Information Server (AGRIS; Yilmaz et al., 2011). The *A. thaliana* WRKY classification was made according to Wang et al. (2011). The *A. lyrata* orthologous WRKY TFs were found by BLASTP search program in the *A. lyrata* version 1.0 genome sequence database (Goodstein et al., 2002; <http://www.phytozome.net>).

Identification of PEST sequences in WRKY proteins

The *A. thaliana* and *A. lyrata* WRKY protein sequences were analyzed with the pestfind algorithm (emboss.bioinformatics.nl/cgi-bin/emboss/epestfind). The algorithm identifies stretches of 10 or more amino acids, enriched in proline (P), glutamic acid (E), aspartic acid (D), serine (S), and threonine (T), that are flanked with the positively charged amino acids arginine (R), histidine (H) and lysine (K). This parameter is combined with hydrophathy index to obtain a PEST score as expressed by the following equation: $PEST\ score = 0.55 * DEPST - 0.5 * hydrophobicity\ index$ (Rogers et al., 1986). We only selected those PEST sequences with positive scores. To describe the PEST localization, we defined as N- and C-terminal domains the 30% of initial and final amino acids into WRKY proteins, respectively.

Prediction of KEN- and D-boxes in *A. thaliana* WRKY proteins

The 66 *A. thaliana* WRKY protein sequences were submitted in GPS-ARM (Group-based Prediction System) software for prediction of both KEN- and D-boxes using a high threshold (<http://arm.biocuckoo.org>; Liu et al., 2012).

Prediction of disordered regions

The disordered regions were predicted by using the Predictor of Natural Disordered Regions PONDR VL3-BA (<http://www.pondr.com/pondr-tut2.html>) and the PreDisorder methods (<http://sysbio.mnet.missouri.edu/predisorder.html>).

The PONDR VL3-BA is a neural network that was trained on regions of 152 long characterized disorder regions and a set of ordered proteins contained 290 chains from Protein Data Bank (PBDS25), which do not have disorder-promoting residues (Obradovic et al., 2003). In the PreDisorder method, the protein sequence is aligned against several template profiles by a PSI-BLAST search program generating an input profile of the sequence. This profile along with the predicted secondary structure and solvent accessibility is fed into a neural network that makes the disorder predictions (Deng et al., 2009). In fact, the proteins with 50% of overall percent of disordered residues and having a score of 0.5 were classified as disordered proteins.

Amino acid composition analysis

To determine the amino acid composition from both *A. thaliana* and *A. lyrata* WRKY proteins, we used the Composition Profiler software (Vacic et al., 2007). This tool compares two samples of amino acid sequences and generates a graphical output composed of twenty data points (one for each amino acid), where bar heights indicate the residue enrichment or depletion against a baseline. Herein, we compared the *A. thaliana* and *A. lyrata* WRKY protein sequences against the PBDS25 control subset from Protein Data Bank (Berman et al., 2000). The graphical output was generated by 10,000 bootstrap iterations using a significance (α) value of 0.05.

Statistical analysis

Statistics comparisons between non PEST-containing versus PEST-containing WRKY TFs, and ordered versus disordered proteins were done using the unpaired non-parametric *t*-test ($n = 66$) (GraphPad Prism software; GraphPad Software Inc., San Diego, USA). Differences of $P < 0.05$ are considered to be significant.

RESULTS

PEST sequences occurrence in *Arabidopsis thaliana* WRKY transcription factors

In order to explore the existence of PEST sequences in the *A. thaliana* WRKY family, we examined 66 WRKY protein sequences using the pestfind algorithm (Rogers et al., 1986). According to the analysis, a potential PEST sequence was elected as a region with a positive score in agreement with the pestfind algorithm. Our analyses revealed that 67% of *A. thaliana* WRKY TFs contained at least one potential PEST sequences (Table 1). The WRKY I group contained the most matches, with 85 and 79% of positive scores for groups Ia and Ib, respectively (Table 1). In the case of other the WRKY groups, we detected a lower percentage of PEST-containing proteins; for example, group IIa with 64%, IIb with 60%, and group III with 46% (Table 1). The PEST sequence enrichment was found statistical significant using an unpaired *t* test ($p < 0.05$).

When we analyzed the number of PEST sequences per protein, it was found that the WRKY Ia group exhibits a major number of potential PEST sequences within each transcriptional factor. The members that contained more PEST sequences were AtWRKY20 and AtWRKY33, with six and four PEST sequences, respectively. The highest PEST score (29.6) was found in AtWRKY58 from group Ia; this protein showed a unique PEST sequence. On the other hand, the PEST sequence localization in *A. thaliana* WRKY proteins was analyzed; we found that 24 WRKY TFs contained only one PEST sequence; of these, 11 members harbored the PEST region in their N-terminal domains and 5 members in its C-terminal domains (Supplementary Table S1). The remaining WRKY proteins, which contain one or more PEST sequences, were preferentially located in the regions adjacent to the WRKY DNA-binding domain (Figure 1).

Prediction of disordered regions in *A. thaliana* WRKY transcription factors

In order to analyze whether WRKY proteins are enriched in disordered regions, the 66 WRKY proteins sequences were analyzed using the Predictor of Natural

Disordered Regions (PONDR VL3-BA; Obradovic et al., 2003) and PreDisorder (Deng et al., 2009) algorithms. According to Singh et al. (2006), a protein is considered to be disordered when the overall percentage of disordered residues is higher than 50% and the prediction score above 0.5 (Supplementary Table S2).

In silico analysis using the PONDR VL3-BA method revealed that all transcription factors of Ia group were disordered proteins (Table 2). In the other subgroups, such as WRKY Ib, IIa and IIb a high percentage of disordered proteins (79, 91, and 87%, respectively) were also detected; in contrast, the subgroup III only 23% of the proteins were disordered, hence few proteins had disordered residues. The disordered protein members were found statistically significant using an unpaired *t* test analysis ($p < 0.05$). In Figure 2, we evidence that PEST sequences were located within predicted disordered regions. In the same way, we observed that the predicted disordered WRKY proteins are enriched in PEST motifs in contrast to ordered ones (Table 2).

Amino acid composition of WRKY transcription factors from *A. thaliana*

Subsequently, we performed a comparative analysis between the 66 WRKY proteins sequences and a control subset from Protein Data Bank (PDBS25; Berman et al., 2000) to compare their amino acid compositions. Both datasets were analyzed using the web-based tool Composition Profiler (Vacic et al., 2007). In Figure 3, the diminution and enrichment of order-promoting, disorder-promoting, or disorder-order neutral residues in the *A. thaliana* TFs dataset are shown; the comparative analysis reveals that *A. thaliana* WRKY family displays a low content of order-promoting amino acids, such as tryptophan, isoleucine, leucine, phenylalanine, valine, and tyrosine and that they are enriched in serine, proline, glutamine, and arginine disorder-promoting residues. Also, disorder-order neutral histidine and threonine were over-represented. Finally, no significant differences were observed in the content of order-promoting cysteine, disorder-promoting glutamic acid and lysine, and disorder-order neutral aspartic acid between *A. thaliana* TFs and the PDBS25 datasets. This result shows that the *A. thaliana* WRKY family is enriched in disorder-promoting residues.

Comparative analysis of PEST sequences occurrence in WRKY family from *A. thaliana* and *A. lyrata*

In order to determine whether PEST sequences are conserved in orthologous WRKY proteins, we focused on finding the existence of this proteolytic target in the highly related species *A. lyrata*. Consecutively, to obtain the *A. lyrata* WRKY protein sequences we used the BLASTP search program in the *A. lyrata* version 1.0 genome sequence database (Goodstein et al., 2002; <http://www.phytozome.net>). The presence of WRKY conserved domain in the obtained *A. lyrata* WRKY sequences were analyzed by the Pfam program (Finn et al., 2014; <http://pfam.sanger.ac.uk/>). We found 64 *A. lyrata* WRKY orthologous sequences to the *A. thaliana* WRKY TFs. In the case of AtWRKY63 and AtWRKY66, no orthologous sequences were identified in *A. lyrata* genome database (Supplementary Table S3).

Our comparative analysis revealed that most PEST sequences were conserved in their number and localization between both *Arabidopsis* WRKY orthologues (compare Figures 1 and 4). Interestingly, our comparative analyses revealed that 73% of the *A. lyrata* WRKY TFs contained potential PEST sequences. Particularly, the *A. lyrata* WRKY Ia, IIa, and III groups contained a higher percentage of members with PEST sequences (92, 73, and 91%, respectively) than in *A. thaliana* (82, 62, and 46%, respectively), respectively (compare Tables 1 and 3). The PEST sequence enrichment in *A. lyrata* WRKY proteins was found to be statistically significant using an unpaired *t* test analysis ($p < 0.05$). As in *A. thaliana*, the orthologous TFs AIWRKY20 and AIWRKY33 from *A. lyrata* were the most PEST-containing proteins with six and four PEST sequences (Figure 4). Similarly, the highest PEST score (21.7) was found in the unique PEST sequence present in AIWRKY58 (Figure 4). Additionally, we found that 26 WRKY TFs contain only one PEST sequence, of which 11 and 10 members have the PEST region located in their N-terminal and C-terminal domains (Supplemental Table S4). The remaining *A. lyrata* WRKY proteins, containing one or more PEST sequences, were preferentially located in the regions adjacent to the WRKY DNA-binding domain (Figure 4).

The PONDR VL3-BA algorithm revealed an important increase in the percentage of predicted disordered proteins in *A. lyrata* WRKY Ib, IIb, and III groups with 93, 93, and 36%, respectively, in comparison to *A. thaliana* orthologues that showed 79, 87, and 23%, respectively. The WRKY Ia and IIa groups did not display differences in their percentage of predicted disordered members (100 and 91%, correspondingly) (Tables 4 and 2). In addition, the percentage of disordered proteins that contain PEST sequences increased in the *A. lyrata* WRKY Ia, IIa, and III groups, having 92, 80, 100%, respectively, compared to 85, 70, and 67% from *A. thaliana* orthologous groups. In contrast, this value was diminished in the *A. lyrata* WRKY Ib and IIb subclasses with 62 and 57% compared to 91 and 67% from *A. thaliana* WRKY Ib and IIb, respectively. Finally, as we observed in *A. thaliana*, the amino acid composition of the *A. lyrata* WRKY group is diminished in order-promoting residues and enriched in disorder-promoting amino acids (Supplementary Figure S1).

Discussion

The ability of cells to express specific genetic information is a crucial step for adaptation to changing conditions throughout their life span. Transcription factors (TFs) regulate the first step of gene expression. In this sense, the temporal and spatial regulation of TFs is essential for transcriptional control. TFs are regulated from transcription to post-translation. Post-translational modifications, such as phosphorylation, acetylation, methylation, glycosylation or ubiquitination, will impact the nuclear transport, DNA-binding, co-factor interactions, and protein degradation of TFs.

In the *Arabidopsis thaliana* transcriptome, around 6.5% of its genes encode TFs (Sharma et al., 2013; Riechmann et al., 2000), which is higher than many? other eukaryotic organisms, such as *Drosophila melanogaster* (4.7%) or *Caenorhabditis elegans* (3.6%) (Riechmann et al., 2000), and *Homo sapiens* (6.0%) (Venter et al., 2001). The *A. thaliana* TFs are grouped in 50 families, from which 45% are specific to plants (Feller et al., 2011). It has been suggested that in sessile organisms, transcriptional regulation might play a central role in order to face a wide range of environmental stresses (Mitsuda and Ohme-Takagi, 2009).

The WRKYs are one of the largest families of TFs exclusives of plants, with 72 members in *A. thaliana* grouped into three major subclasses (I, II, and III) (Wang et al., 2011). In recent years, there has been increasing interest in the study of the WRKY family, which will give fresh insights about the role of WRKY TFs in biotic and abiotic stress responses, development processes, and hormone signaling in plants (Johnson et al., 2002; Murray et al., 2007; Chen et al., 2010; Bakshi and Oelmüller, 2014). However, little is known about WRKY protein turnover signaling.

We therefore analyzed the occurrence of PEST regions in 66 members from *A. thaliana* WRKY TFs. The PEST sequences are rapid protein degradation tags, which are enriched in proline, glutamic acid, serine and threonine amino acids, and generally flanked by lysine residues. They are predicted to be solvent exposed regions that make them susceptible to post-translational modifications,

such as phosphorylation and ubiquitination (Rechsteiner and Rogers, 1996; Singh et al., 2006). Our analysis revealed that 67% of *A. thaliana* WRKY proteins contain potential PEST sequences, which were found from one up to six times within a single TF protein sequence. In this regard, the WRKY I and II groups were the most PEST-containing TFs with 82 and 62%, respectively, compared to group III with 46%. On the other hand, we analyzed the presence of other destruction motifs, such as KEN- and D-box in the *A. thaliana* WRKY TFs. These sequences are involved in the proteolysis of mitotic cell-cycle regulatory proteins, which ensure the correct progression of the cell cycle (Peters, 2006). We found that putative KEN- and D-box proteolytic sequences were less frequent in this WRKY family with 3 and 18%, respectively; in contrast to 67% PEST sequences (Supplementary Table S5). In particular, the AtWRKY11 (IIb group) and AtWRKY41 (III group) contain a KEN-box; in fact, these WRKYs don't contain potential PEST regions. As regards to the D-box, 12 WRKY proteins contain this tag, of which 5 WRKYs did not contain potential PEST regions. Our data reveal a low occurrence of KEN- and D-box in contrast to PEST sequences.

As well, *A. thaliana* WRKY I and II groups contain a high number of predicted disordered proteins in contrast to WRKY III subclasses (~90% versus 23%). These data revealed that the PEST sequences occurrence in WRKY TFs is highly correlated with disordered regions, mainly flanking the WRKY DNA-binding domain. It has been reported that TFs have disordered regions in their sequences; frequently, these regions are located flanking the DNA-binding domains (Liu et al., 2006; Guo et al., 2012). Duan et al. (2007) and Yamasaki et al. (2005) resolved the secondary structure of the C-terminal domains, comprising the WRKY DNA-binding domain and its flanking regions, from *A. thaliana* WRKY1 and WRKY4 proteins through X-ray or NMR spectroscopy. The WRKY domain structures consist of a five- and four-stranded beta-sheet, respectively; however, the secondary structure of the adjacent regions of WRKY domains were not resolved. This is presumably due to these regions being disordered, which is in accordance with our *in silico* data.

It has been proposed that in humans TFs these disordered regions are essential for transcriptional regulation (Minezaki et al., 2006); TF conformational flexibility might contribute to their ability to recognize target sequences in DNA, to bind to other factors, and to form complexes positioned on the DNA, which may then be involved in transcriptional regulation, or present activation domains to downstream transcriptional regulatory machinery (Guo et al., 2012). In addition, Prakash et al. (2004) reported that these disordered regions are required for efficient proteasome-mediated degradation.

A. thaliana diverged from *A. lyrata* 10 million years ago (Hu et al., 2011), the genomes of these two highly related *Brassicaceae* species have been completely sequenced, providing us an exceptional opportunity to understand how the variation process works over this time within the WRKY TFs family. In this sense, we conducted a comparative analysis between *Arabidopsis* WRKY proteins. Our results revealed that these two *Brassicaceae* species show differences in the number of PEST presence and disordered members. In the *A. lyrata* WRKY protein family, 73% of them were PEST-containing members. Notably, the *A. lyrata* WRKY III group exhibited the highest percentage of PEST-containing sequences (91%) and disordered proteins (36%) when compared to its *A. thaliana* orthologous group (46% and 23%, respectively). For example, the AIWRKY41, 46, 64, and 70 proteins contain PEST sequences compared to their *A. thaliana* orthologous where PEST sequences are missing. In this regard, analyzing the corresponding PEST regions of *A. lyrata* versus *A. thaliana* orthologous, we found that *A. thaliana* PEST regions had more hydrophobic residues, such as alanine and valine (Supplementary Table S6). Although some WRKYs of both species display some differences in their PEST regions, in general, the conservation of PEST sequences in WRKY transcription factor groups of *A. thaliana* and *A. lyrata* suggests that these sequences have an important participation in a mechanism for its post-translational regulation.

Even though there are few reports on PEST sequences in *A. thaliana* TFs, their functionality has been demonstrated. For example, the *A. thaliana* dehydration-responsive DREB2A TF possess a PEST sequence in its central region. The deletion of this motif leads to a constitutively active form of DREB2A, which

indicates that this PEST sequence plays a fundamental role in the stability of this protein. Moreover, the constitutive overexpression of DREB2A improved drought tolerance in transgenic *Arabidopsis* plants (Sakuma et al., 2006). In the same way, the Abscisic Acid-Insensitive 4 (ABI4) TF of *A. thaliana* contains a PEST sequence localized to its N-terminal domain. The authors demonstrated that the degradation of ABI4 protein is mediated by the 26S proteasome through its PEST motif in response to sugars levels (Gregorio et al., 2014). Also, the stability of the transcriptional repressor VNI2 involved in the control of vessel differentiation in *A. thaliana* is regulated by a PEST motif-dependent proteolysis. Moreover, the stability of VNI2 is increased by treating *A. thaliana* seedlings with the proteasome inhibitor MG132. Also, the C-terminally truncated VNI2 version, which lacks part of the PEST sequence, is more stable than full-length protein (Yamaguchi et al., 2010).

The protein interaction of AtWRKY6 with an *A. thaliana* RING-type finger E3 ligase has been observed (Chen et al., 2009). The authors demonstrated a decrease of AtWRKY6 protein in *A. thaliana* seedling growth under low phosphate treatment, which was inhibited by addition of a 26S proteasome inhibitor MG132. Similarly, Miao and Zentgraf (2010) reported that AtWRKY53 protein interacts with the E3 ubiquitin ligase UPL5, which led to its proteolytic degradation in *A. thaliana* plants. Also, over-expression of AtWRKY53 in the *A. thaliana upl5* mutant background causes an accelerated senescence phenotype; these results suggest that UPL5 regulates leaf senescence in *Arabidopsis* through degradation of AtWRKY53. In this study, we found that the *A. thaliana* AtWRKY6 and AtWRKY53 TFs contain a PEST sequence located in its central and N-terminal domain. In this context, the PEST sequences of AtWRKY6 and AtWRKY53 may play a role in protein degradation mediated by the 26S proteasome.

Taken together, the proteasome-mediated rapid turnover of TFs has been demonstrated, and also it has been reported that some PEST-containing TFs are degraded through the ubiquitin-26S proteasome. We found a high occurrence of PEST sequences in WRKY members from two *Brassicaceae* species revealing a significant conservation of PEST regions among WRKY

orthologous. Our study opens the possibility for further analyses of PEST sequence as post-translational regulators for these important plant transcription factors.

5. References

Bakshi, M. and Oelmüller, R., 2014. WRKY transcription factors: Jack of many trades in plants. *Plant Signal Behav.* 9:27700. doi:10.4161/psb.27700.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., Bourne, P. E., 2000. The protein data bank. *Nucleic Acids Res.* 28:235.

Chen, H., Lai, Z., Shi, J., Xiao, Y., Chen, Z., Xu, X., 2010. Roles of arabidopsis WRKY18, WRKY40 and WRKY60 transcription factors in plant responses to abscisic acid and abiotic stress. *BMC Plant Biol.* 10:281; PMID:21167067; <http://dx.doi.org/10.1186/1471-2229-10-281>.

Chen, Y.F., Li, L.Q., Xu, Q., Kong, Y.H., Wang, H., and Wu, W.H., 2009. The WRKY6 transcription factor modulates PHOSPHATE1 expression in response to low Pi stress in Arabidopsis. *Plant Cell.* 21:3554-66.

Cooper, G.M., 2000. *The Cell: A Molecular Approach*. 2nd edition. Sunderland (MA): Sinauer Associates. Eukaryotic RNA Polymerases and General Transcription Factors. <http://www.ncbi.nlm.nih.gov/books/NBK9935>.

Deng, X., Eickholt, J., Cheng, J., 2009. PreDisorder: *Ab Initio* Sequence-based Prediction of Protein Disordered Regions. *BMC Bioinformatics.* 10:436; doi: 10.1186/1471-2105-10-436.

Desterro, J.M., Rodriguez, M.S., Hay, R.T., 2000. Regulation of transcription factors by protein degradation. *Cell Mol Life Sci.* 57, 1207-19.

Duan, M.R., Nan, J., Liang, Y.H., Mao, P., Lu, L., Li, L., Wei, C., Lai, L., Li, Y., Su, X.D., 2007. DNA binding mechanism revealed by high resolution crystal structure of Arabidopsis thaliana WRKY1 protein. *Nucleic Acids Res.* 35(4): 1145-54.

Eulgem, T., Rushton, P.J., Robatzek, S., Somssich, I.E., 2000. The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 5:199-206.

Feller, A., Machemer, K., Braun, E.L., Grotewold, E. 2011. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* 66(1):94-116; doi: 10.1111/j.1365-313X.2010.04459.x.

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L. L., Tate, J., Punta, M., 2014. Pfam: the protein families database. *Nucleic Acids Res.* 42 (Database issue):D222-D230.

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., Rokhsar, D. S., 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40 (Database issue):D1178–D1186; <http://doi.org/10.1093/nar/gkr944>.

Gregorio, J., Hernández-Bernal, A. F., Cordoba, E., León, P., 2014. Characterization of evolutionarily conserved motifs involved in activity and regulation of the ABA-INSENSITIVE (ABI) 4 transcription factor. *Mol Plant* 7(2):422-36; doi: 10.1093/mp/sst132.

Guo, X., Bulyk, M. L., Hartemink, A. J., 2012. Intrinsic disorder within and flanking the DNA-binding domains of human transcription factors. *Pacific Symposium on Biocomputing.* 104-115.

Hu, T. T., Pattyn, P., Bakker, E. G., Cao, J., Cheng, J.F., Clark, R. M., et al., 2011. The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. *Nature Genetics.* 43(5):476-481; <http://doi.org/10.1038/ng.807>.

Johnson, C. S., Kolevski, B., Smyth, D. R., 2002. TRANSPARENT TESTA GLABRA2, a trichome and seed coat development gene of *Arabidopsis*, encodes a WRKY transcription factor. *Plant Cell.* 14:1359-75.

Liu, J., Perumal, N. B., Oldfield, C. J., Su, E. W., Uversky, V. N., Dunker, A. K., 2006. Intrinsic disorder in transcription factors. *Biochemistry.* 45(22):6873-88; <http://doi.org/10.1021/bi0602718>.

Liu, Z., Yuan, F., Ren, J., et al., 2012. GPS-ARM: Computational Analysis of the APC/C Recognition Motif by Predicting D-Boxes and KEN-Boxes. Haslam NJ, ed. PLoS ONE. 7(3) e34370. doi:10.1371/journal.pone.0034370.

Lu, Q. S., Paz, J. D., Pathmanathan, A., Chiu, R. S., Tsai, A. Y., Gazzarrini, S., 2010. The C-terminal domain of FUSCA3 negatively regulates mRNA and protein levels, and mediates sensitivity to the hormones abscisic acid and gibberellic acid in Arabidopsis. *Plant J.* 64(1):100-113; doi: 10.1111/j.1365-313X.2010.04307.x.

Luo, M., Dennis, E. S., Berger, F., Peacock, W. J., Chaudhury, A., 2005. MINISEED3 (MINI3), a WRKY family gene, and HAIKU2 (IKU2), a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in Arabidopsis. *Proc Natl Acad Sci USA.* 102:17531-36.

Maruri-López, I., Hernández-Sánchez, I. E., Ferrando, A., Carbonell, J., Jiménez-Bremont, J. F., 2015. Characterization of maize spermine synthase 1 (ZmSPMS1): Evidence for dimerization and intracellular location. *Plant Physiol Biochem.* 97:264-271; doi: 10.1016/j.plaphy.2015.10.017.

Miao, Y. and Zentgraf, U., 2010. A HECT E3 ubiquitin ligase negatively regulates Arabidopsis leaf senescence through degradation of the transcription factor WRKY53. *Plant J.* 63:179-88; <http://dx.doi.org/10.1111/j.1365-313X.2010.04233.x>.

Minezaki, Y., Homma, K., Kinjo, A. R., & Nishikawa, K., 2006. Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation. *J Mol Biol Mol.* 359(4), 1137-1149.

Mitsuda, N. and Ohme-Takagi, M. 2009. Functional Analysis of Transcription Factors in Arabidopsis. *Plant Cell Physiol.* 50(7):1232-1248; <http://doi.org/10.1093/pcp/pcp075>.

Murray, S.L., Ingle, R.A., Petersen, L.N., Denby, K.J., 2007. Basal resistance against *Pseudomonas syringae* in Arabidopsis involves WRKY53 and a protein

with homology to a nematode resistance protein. *Mol Plant Microbe Interact.* 20:1431-8; PMID:17977154; <http://dx.doi.org/10.1094/MPMI-20-11-1431>.

Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P., Brown, C. J., Dunker, A. K., 2003. Predicting Intrinsic Disorder from Amino Acid Sequence. *Proteins: Struct Funct Gen.* 53:566-572.

Pandey, S. P. and Somssich, I. E., 2009. The role of WRKY transcription factors in plant immunity. *Plant Physiol.* 150:1648-55.

Peters, J.M., 2006. The anaphase promoting complex/cyclosome: a machine designed to destroy. *Nat. Rev. Mol. Cell Biol.* 7:644-656.

Prakash, S., Tian, L., Ratliff, K. S., Lehotzky, R. E., Matouschek, A., 2004. An unstructured initiation site is required for efficient proteasome-mediated degradation. *Nat Struct Mol Biol* 11(9):830-837.

Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., et al., 2000. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science.* 290:2105-2110.

Rogers, S.W., Wells, R., Rechsteiner, M., 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science.* 234:364-368.

Sakuma, Y., Maruyama, K., Qin, F., Osakabe, Y., Shinozaki, K., & Yamaguchi-Shinozaki, K., 2006. Dual function of an Arabidopsis transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression. *Proc Natl Acad Sci USA.* 103(49):18822-18827; <http://doi.org/10.1073/pnas.0605639103>.

Sandhu, K.S. and Dash, D., 2006. Conformational flexibility may explain multiple cellular roles of PEST motifs. *Proteins.* 63(4):727-732.

Sharma, N., Bhalla, P.L., Singh, M.B., 2013. Transcriptome-wide profiling and expression analysis of transcription factor families in a liverwort, *Marchantia polymorpha*. *BMC Genomics.* 14:915.

Singh, G.P., Ganapathi, M., Sandhu, K.S., Dash, D., 2006. Intrinsic unstructuredness and abundance of PEST motifs in eukaryotic proteomes. *Proteins*. 62:309-315.

Song, Y. and Gao, J., 2014. Genome-wide analysis of WRKY gene family in *Arabidopsis lyrata* and comparison with *Arabidopsis thaliana* and *Populus trichocarpa*. *Chin Sci Bull*. 59(8):754-765; doi:10.1007/s11434-013-0057-9.

Vacic, V., Uversky, V. N., Dunker, A. K., Lonardi, S., 2007. Composition Profiler: A tool for discovery and visualization of amino acid composition differences. *BMC Bioinformatics*. 8:211; <http://doi.org/10.1186/1471-2105-8-211>.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., et al., 2001. The sequence of the human genome. *Science*. 291:1304-1351.

Wang, Q., Wang, M., Zhang, X., Hao, B., Kaushik, S. K., Pan, Y., 2011. WRKY gene family evolution in *Arabidopsis thaliana*. *Genetica*. 139(8):973-983. doi: 10.1007/s10709-011-9599-4.

Yamaguchi, M., Ohtani, M., Mitsuda, N., Kubo, M., Ohme-Takagi, M., Fukuda, H., Demura, T., 2010. VND-INTERACTING2, a NAC Domain Transcription Factor, Negatively Regulates Xylem Vessel Formation in *Arabidopsis*. *The Plant Cell* 22:1249-1263

Yamaguchi, M., Ohtani, M., Mitsuda, N., Kubo, M., Ohme-Takagi, M., Fukuda, H., Demura, T. 2010. VND-INTERACTING2, a NAC domain transcription factor, negatively regulates xylem vessel formation in *Arabidopsis*. *Plant Cell*. 22(4):1249-63; doi: 10.1105/tpc.108.064048.

Yamasaki, K., Kigawa, T., Inoue, M., Tateno, M., Yamasaki, T., Yabuki, T., Aoki, M., Seki, E., Matsuda, T., Tomo, Y., Hayami, N., Terada, T., Shirouzu, M., Tanaka, A., Seki, M., Shinozaki, K., Yokoyama, S., 2005. Solution structure of an *Arabidopsis* WRKY DNA binding domain. *Plant Cell*. 17(3):944-56.

Yilmaz, A., Mejia-Guerra, M. K., Kurz, K., Liang, X., Welch, L., Grotewold, E., 2011. AGRIS: the *Arabidopsis* Gene Regulatory Information Server, an update.

Nucleic Acids Res. 39 (Database issue):D1118-D1122;
<http://doi.org/10.1093/nar/gkq1120>.

Legend of figures

Fig. 1. Schematic representation of PEST sequence in the *Arabidopsis thaliana* WRKY family. The *A. thaliana* WRKY family is classified into three major groups, Ia and b, IIa and b, and III. The members of each WRKY group that contain at least one positive PEST sequence are shown. The predicted PEST sequences are represented as black boxes with the PEST score indicated below; the WRKY DNA-binding domains are shown as blue boxes. The scale bar corresponds to 150 residues.

Fig. 2. Prediction of ordered and disordered regions in *Arabidopsis thaliana* WRKY transcription factors. Two WRKY proteins that contain PEST sequences of each group are presented. The graphs indicate the ordered and disordered regions in WRKY proteins using the PreDisorder method; a schematic representation of the protein is shown below the plot. The predicted PEST sequences are represented as black boxes with PEST score indicated below while the WRKY DNA-binding domain are shown as blue boxes. The scale bar corresponds to 150 residues.

Fig. 3. Comparative analysis of amino acid composition between *Arabidopsis thaliana* WRKY transcription factors and a control protein subset from Protein Data Bank (PDBS25) using Composition Profiler software. The bars of the disorder-promoting residues (A, R, S, Q, E, G, K, P) are colored red, order-promoting residues (N, C, I, L, F, W, Y, V) are colored blue, and disorder-order neutral residues (D, H, M, T) are colored gray.

Fig. 4. Schematic representation of PEST sequence occurrence in the *Arabidopsis lyrata* WRKY family. The *A. lyrata* WRKY family is classified into three major groups, Ia and b, IIa and b, and III. The members of each WRKY group that contain at least one positive PEST sequence are shown. The predicted PEST sequences are shown as black boxes with PEST score indicated below, the WRKY DNA-binding domains are shown as blue boxes. The scale bar corresponds to 150 residues.

Table 1. The presence of PEST sequences in the *Arabidopsis thaliana* WRKY transcription factor subgroups.

WRKY group	Number of members	Number of PEST-containing proteins	Percentage of PEST-containing proteins (%)
Ia	13	11	85
Ib	14	11	79
IIa	11	7	64
IIb	15	9	60
III	13	6	46
Total	66	44	67

Table 2. Predicted disordered *Arabidopsis thaliana* WRKY proteins that contain PEST sequences.

WRKY group	Number of members	Number and Percentage of predicted disordered proteins (%)	Number and Percentage of predicted disordered proteins with PEST (%)
Ia	13	13 (100)	11 (85)
Ib	14	11 (79)	10 (91)
IIa	11	10 (91)	7 (70)
IIb	15	13 (87)	8 (62)
III	13	3 (23)	2 (67)
Total	66	50 (76)	38 (76)

Table 3. The presence of PEST sequences in the *Arabidopsis lyrata* WRKY transcription factor subgroups.

WRKY group	Number of members	Number of PEST-containing proteins	Percentage of PEST-containing proteins (%)
Ia	13	12	92
Ib	14	8	57
IIa	11	8	73
IIb	15	9	60
III	11	10	91
Total	64	47	73

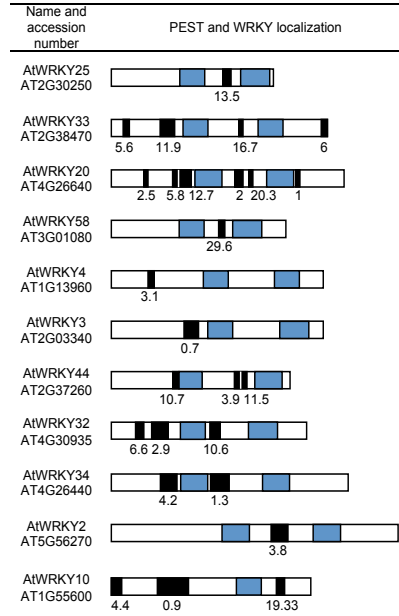
Table 4. Predicted disordered *Arabidopsis lyrata* WRKY proteins that contain PEST sequences.

WRKY group	Number of members	Number and Percentage of predicted disordered proteins (%)	Number and Percentage of predicted disordered proteins with PEST (%)
Ia	13	13 (100)	12 (92)
Ib	14	13 (93)	8 (62)
IIa	11	10 (91)	8 (80)
IIb	15	14 (93)	8 (57)
III	11	4 (36)	4 (100)
Total	64	54 (84)	40 (74)

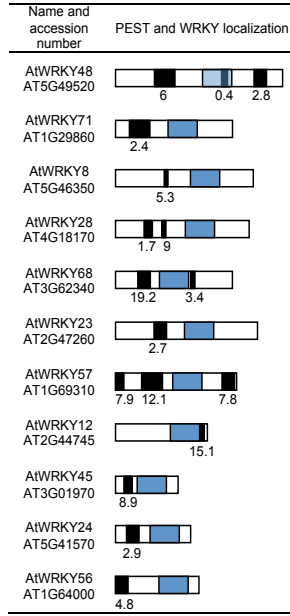
Fig. 1

■ PEST sequence ■ WRKY DNA-binding domain — 150 amino acids

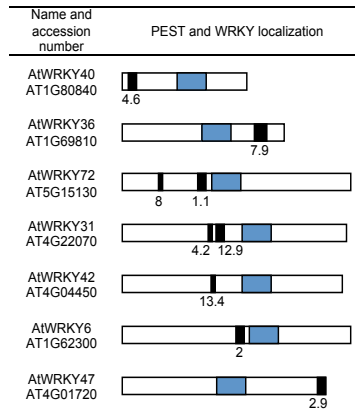
Ia



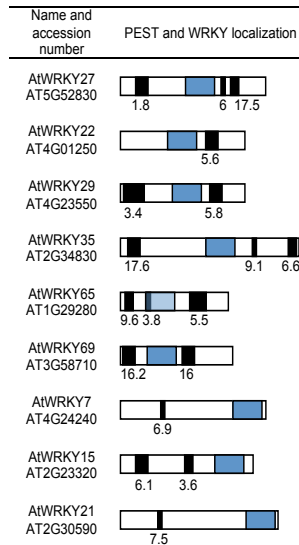
Ib



IIa



IIb



III

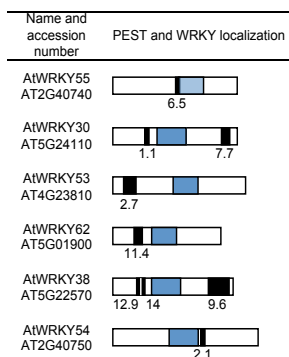


Fig. 2

■ PEST sequence ■ WRKY DNA-binding domain 150 amino acids

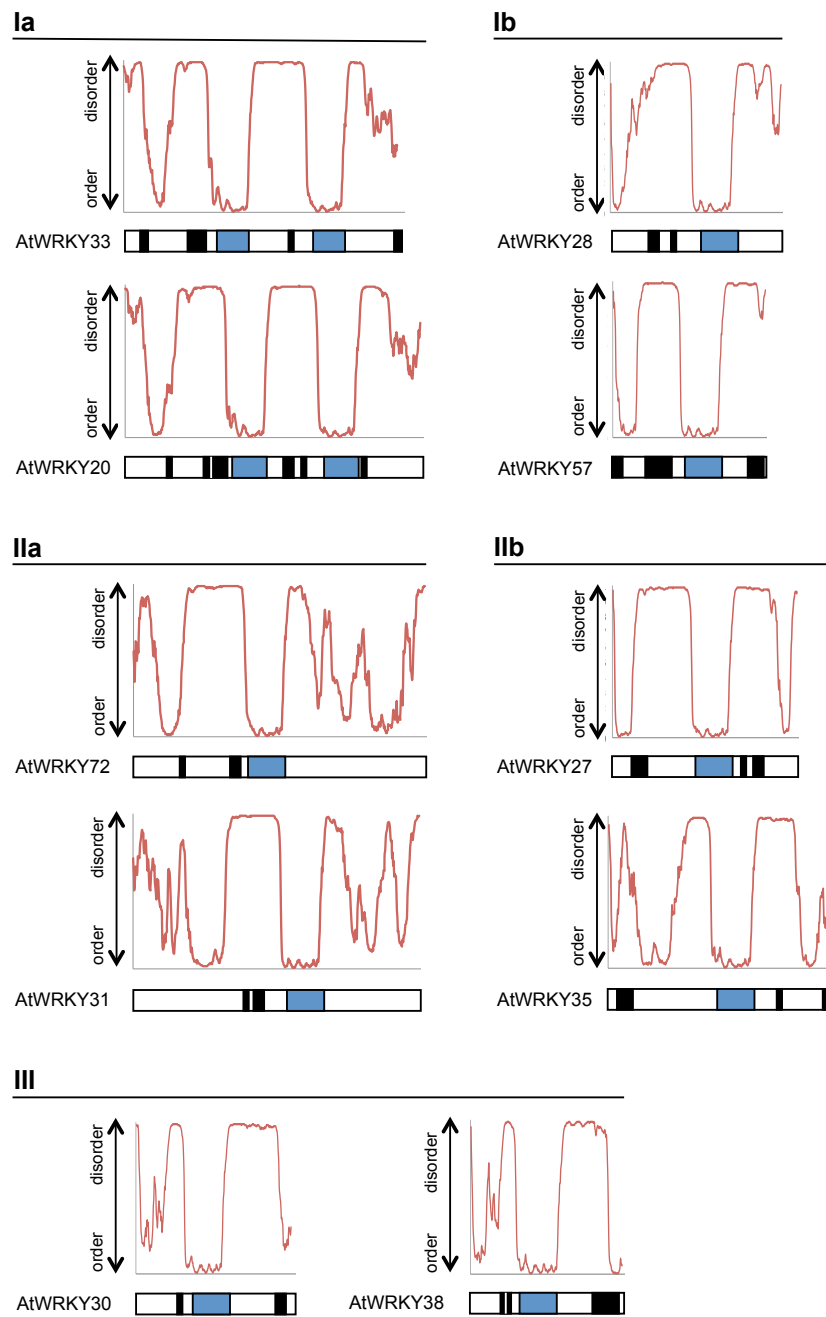


Fig. 3

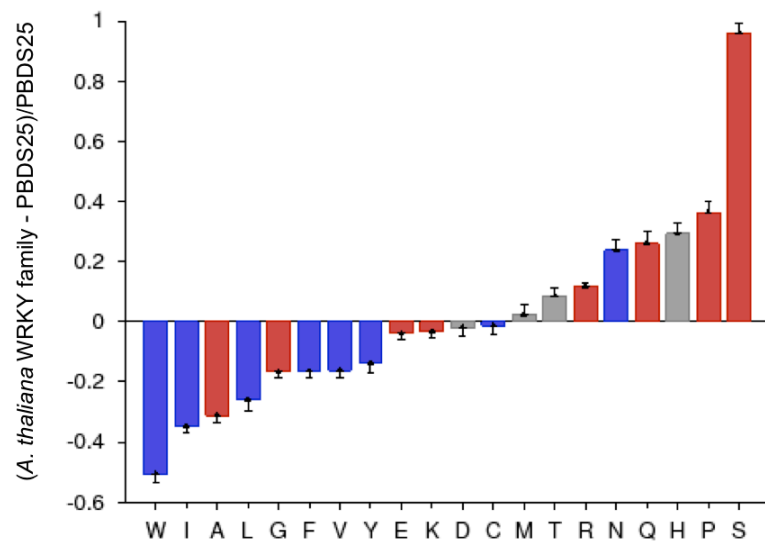
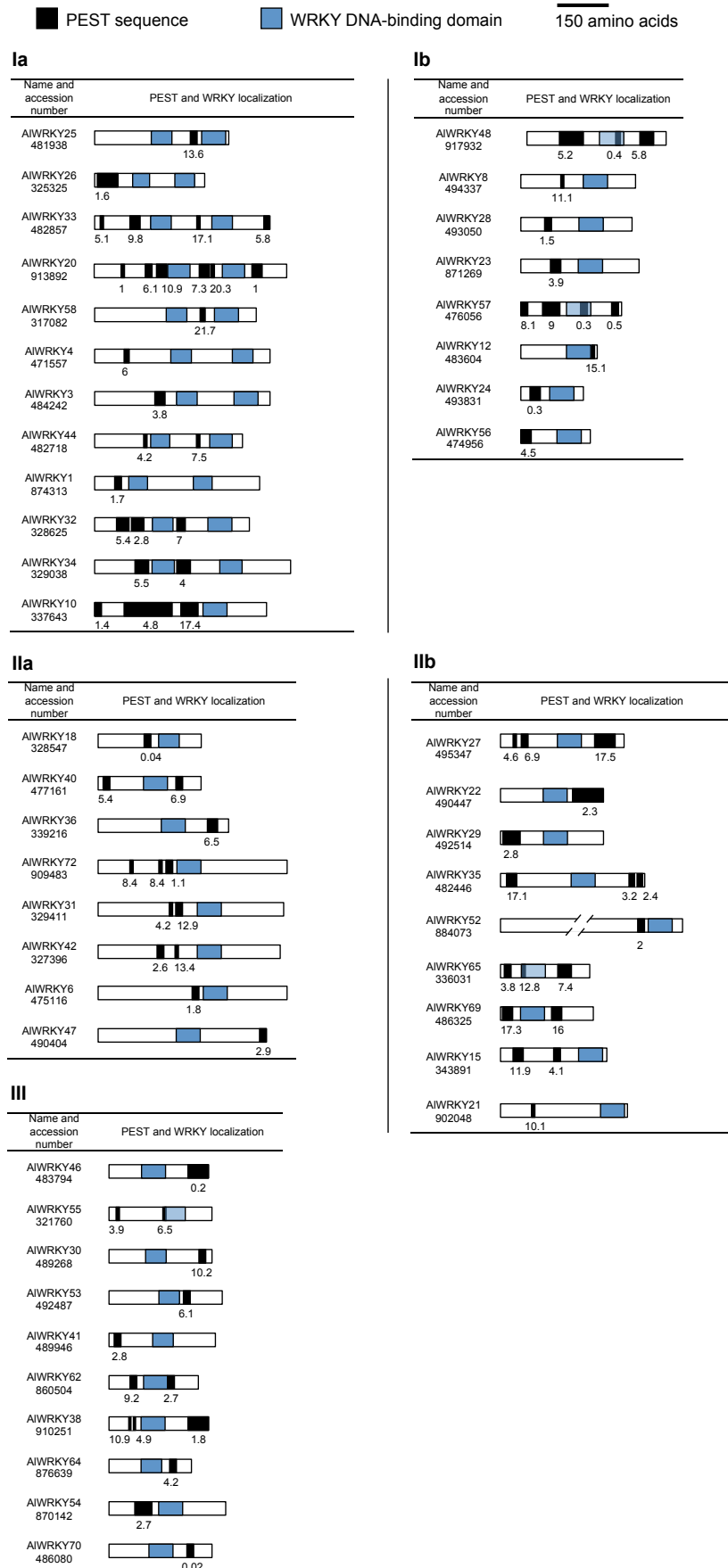
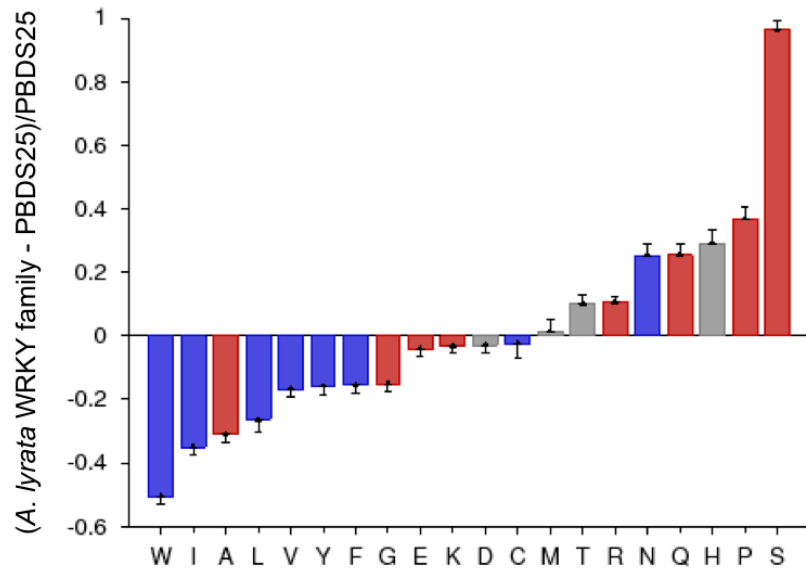


Fig. 4



Supplementary Figure S1



Supplementary Figure S1. Comparative analysis of amino acid composition between *A. lyrata* WRKY transcription factors and a control protein subset from Protein Data Bank (PBDS25) using Composition using Profiler software. Disorder-promoting residues (A, R, S, Q, E, G, K, P) are colored red, order-promoting residues (N, C, I, L, F, W, Y, V) are colored blue, and disorder-order neutral residues (D, H, M, T) are colored gray.

Supplementary Table S1. The *A. thaliana* WRKY TFs containing one N- or C-terminal PEST sequences.

WRKY group	PEST localization	
	N-terminal	C-terminal
Ia	AtWRKY4	AtWRKY25*
Ib	AtWRKY71 AtWRKY23* AtWRKY45* AtWRKY24* AtWRKY56*	AtWRKY12*
IIa	AtWRKY40	AtWRKY36 AtWRKY47
IIb	AtWRKY7 AtWRKY21	AtWRKY22*
III	AtWRKY53 AtWRKY62*	
Total	11	5

*PEST region adjacent to WRKY DNA-binding domain.

Supplementary Table S2. Prediction of ordered and disordered in *A. thaliana* WRKY proteins

access number/name	VL3-BA method		PreDisorder method
	Overall percent disordered	Average Prediction Score	
la			
AT2G30250/AtWRKY25	65.39	0.59	0.52
AT5G07100/AtWRKY26	72.49	0.60	0.42
AT2G38470/AtWRKY33	63.01	0.56	0.58
AT4G26640/AtWRKY20	77.02	0.65	0.59
AT3G01080/AtWRKY58	66.67	0.62	0.47
AT1G13960/AtWRKY4	91.44	0.67	0.55
AT2G03340/AtWRKY3	87.72	0.66	0.51
AT2G37260/AtWRKY44	68.76	0.60	0.53
AT2G04880/AtWRKY1	66.53	0.61	0.68
AT4G30935/AtWRKY32	85.62	0.74	0.62
AT4G26440/AtWRKY34	53.7	0.52	0.58
AT5G56270/AtWRKY2	84.28	0.66	0.53
AT1G55600/AtWRKY10	78.14	0.66	0.48
lb			
AT5G49520/AtWRKY48	79.45	0.60	0.64
AT1G29860/AtWRKY71	50.71	0.51	0.60
AT5G46350/AtWRKY8	56.75	0.57	0.51
AT4G18170/AtWRKY28	61.64	0.53	0.60
AT3G62340/AtWRKY68	58.12	0.54	0.60
AT2G47260/AtWRKY23	56.38	0.58	0.66
AT1G69310/AtWRKY57	69.69	0.61	0.59
AT4G39410/AtWRKY13	77.3	0.61	0.62
AT2G44745/AtWRKY12	76.15	0.60	0.53
AT5G13080/AtWRKY75	46.9	0.56	0.51
AT3G01970/AtWRKY45	44.22	0.53	0.47
AT5G41570/AtWRKY24	57.54	0.59	0.57
AT1G64000/AtWRKY56	50.26	0.54	0.50
AT2G46130/AtWRKY43	37.61	0.52	0.34
lla			
AT2G25000/AtWRKY60	54.98	0.48	0.41
AT4G31800/AtWRKY18	78.06	0.61	0.47
AT1G80840/AtWRKY40	78.48	0.61	0.48
AT1G69810/AtWRKY36	94.57	0.67	0.51
AT1G18860/AtWRKY61	63.33	0.57	0.40
AT5G15130/AtWRKY72	64.42	0.58	0.52
AT4G22070/AtWRKY31	80.67	0.63	0.51
AT4G04450/AtWRKY42	75.76	0.62	0.46
AT1G62300/AtWRKY6	79.02	0.62	0.48
AT4G01720/AtWRKY47	74.64	0.62	0.54
AT1G68150/AtWRKY9	70.86	0.65	0.63
llb		10	5
AT5G52830/AtWRKY27	83.62	0.76	0.61
AT4G01250/AtWRKY22	49.66	0.51	0.61
AT4G23550/AtWRKY29	90.79	0.71	0.58
AT1G30650/AtWRKY14	60.47	0.56	0.44
AT2G34830/AtWRKY35	69.09	0.59	0.47
AT5G45260/AtWRKY52	11.88	0.26	0.05
AT1G29280/AtWRKY65	100	0.76	0.42
AT3G58710/AtWRKY69	85.29	0.69	0.48
AT4G24240/AtWRKY7	64.87	0.61	0.47
AT2G23320/AtWRKY15	74.76	0.62	0.55
AT2G24570/AtWRKY17	81.93	0.62	0.43
AT4G31550/AtWRKY11	58.46	0.55	0.40
AT2G30590/AtWRKY21	88.42	0.66	0.37
AT5G28650/AtWRKY74	71.52	0.60	0.46
AT3G04670/AtWRKY39	60.91	0.56	0.47
lll		13	4
AT2G46400/AtWRKY46	29.49	0.40	0.48
AT2G40740/AtWRKY55	41.1	0.46	0.52
AT5G24110/AtWRKY30	44.22	0.49	0.57
AT4G23810/AtWRKY53	50.62	0.51	0.54
AT4G11070/AtWRKY41	35.14	0.42	0.52
AT5G01900/AtWRKY62	46.01	0.49	0.47
AT5G22570/AtWRKY38	76.47	0.62	0.48
AT1G80590/AtWRKY66	21.28	0.35	0.37
AT1G66600/AtWRKY63	19.92	0.30	0.45
AT1G66550/AtWRKY67	20.08	0.41	0.48
AT1G66560/AtWRKY64	34.54	0.40	0.41
AT2G40750/AtWRKY54	34.68	0.48	0.49
AT3G56400/AtWRKY70	51.02	0.53	0.46

Supplementary Table S2. Prediction of ordered and disordered in *A. lyrata* WRKY proteins

access number/name	VL3-BA method		PreDisorder method	
	Overall percent disordered	Average Prediction Score	Average Prediction Score	
481938/AIWRKY25	68.53	0.58	0.50	
325325/AIWRKY26	71.11	0.60	0.40	
482857/AIWRKY33	59.21	0.56	0.59	
913892/AIWRKY20	77.18	0.66	0.61	
317082/AIWRKY58	70.73	0.64	0.51	
471557/AIWRKY4	84.31	0.65	0.53	
484242/AIWRKY3	85.1	0.65	0.49	
482718/AIWRKY44	67.67	0.60	0.58	
874313/AIWRKY1	67.44	0.62	0.63	
328625/AIWRKY32	79.16	0.84	0.63	
329038/AIWRKY34	52.82	0.55	0.51	
918771/AIWRKY2	82.65	0.65	0.54	
337643/AIWRKY10	88.89	0.68	0.45	
lb				
917932/AIWRKY48	76.77	0.59	0.62	
313633/AIWRKY71	60.85	0.55	0.54	
494337/AIWRKY8	53.99	0.57	0.63	
493050/AIWRKY28	63.01	0.55	0.62	
486738/AIWRKY68	76.49	0.62	0.60	
871269/AIWRKY23	63.5	0.56	0.56	
476056/AIWRKY57	82.46	0.61	0.56	
490719/AIWRKY13	82.51	0.61	0.53	
483604/AIWRKY12	77.52	0.60	0.50	
488124/AIWRKY75	47.26	0.57	0.48	
896042/AIWRKY45	52.47	0.52	0.51	
493831/AIWRKY24	50.84	0.55	0.56	
474956/AIWRKY56	57.44	0.57	0.49	
322266/AIWRKY43	79.65	0.59	0.26	
IIa				
481352/AIWRKY60	44.28	0.45	0.40	
328547/AIWRKY18	62.79	0.53	0.47	
477161/AIWRKY40	81.46	0.62	0.43	
339216/AIWRKY36	90.31	0.65	0.55	
335144/AIWRKY61	64.11	0.57	0.45	
909483/AIWRKY72	63.78	0.58	0.43	
329411/AIWRKY31	79.55	0.61	0.48	
327396/AIWRKY42	70.13	0.61	0.51	
475116/AIWRKY6	79.39	0.61	0.50	
490404/AIWRKY47	73.82	0.61	0.46	
876849/AIWRKY9	80.83	0.63	0.54	
IIb				
495347/AIWRKY27	85.75	0.76	0.57	
490447/AIWRKY22	55.33	0.52	0.62	
492514/AIWRKY29	86.14	0.67	0.60	
890619/AIWRKY14	66.12	0.59	0.40	
482446/AIWRKY35	66.83	0.56	0.39	
884073/AIWRKY52	10.95	0.27	0.05	
336031/AIWRKY65	100	0.75	0.44	
486325/AIWRKY69	85.98	0.70	0.50	
914166/AIWRKY7	62.75	0.54	0.46	
343891/AIWRKY15	72.93	0.62	0.53	
481313/AIWRKY17	72.73	0.61	0.48	
913314/AIWRKY11	62.09	0.59	0.44	
902048/AIWRKY21	85.95	0.64	0.48	
489657/AIWRKY74	75.76	0.59	0.43	
477731/AIWRKY39	54.82	0.55	0.48	
III				
483794/AIWRKY46	41.98	0.46	0.43	
321760/AIWRKY55	42.81	0.50	0.48	
489268/AIWRKY30	53.29	0.50	0.48	
492487/AIWRKY53	43.25	0.50	0.46	
489946/AIWRKY41	37.7	0.44	0.51	
860504/AIWRKY62	51.15	0.51	0.43	
910251/AIWRKY38	60.96	0.56	0.52	
876639/AIWRKY64	48.13	0.44	0.41	
338870/AIWRKY67	2.99	0.34	0.51	
870142/AIWRKY54	47.48	0.52	0.52	
486080/AIWRKY70	51.34	0.54	0.45	

Supplementary Table S3. Orthologous WRKY proteins of *A. thaliana* and *A. lyrata*.

WRKY Group	Protein name	Accession number		Protein identity (%)
		<i>A. thaliana</i>	<i>A. lyrata</i>	
Ia	WRKY25	AT2G30250	481938	98.7
	WRKY26	AT5G07100	325325	89.6
	WRKY33	AT2G38470	482857	88.4
	WRKY20	AT4G26640	913892	96.9
	WRKY58	AT3G01080	317082	95.7
	WRKY4	AT1G13960	471557	95.9
	WRKY3	AT2G03340	484242	90.3
	WRKY44	AT2G37260	482718	86.9
	WRKY1	AT2G04880	874313	86.0
	WRKY32	AT4G30935	328625	89.3
	WRKY34	AT4G26440	329038	92.4
	WRKY2	AT5G56270	918771	89.5
WRKY10	AT1G55600	337643	82.7	
Ib	WRKY48	AT5G49520	917932	78.4
	WRKY71	AT1G29860	313633	81.9
	WRKY8	AT5G46350	494337	88.7
	WRKY28	AT4G18170	493050	90.3
	WRKY68	AT3G62340	486738	79.4
	WRKY23	AT2G47260	871269	85.8
	WRKY57	AT1G69310	476056	92.3
	WRKY13	AT4G39410	490719	83.9
	WRKY12	AT2G44745	483604	99.1
	WRKY75	AT5G13080	488124	87.6
	WRKY45	AT3G01970	896042	86.41
	WRKY24	AT5G41570	493831	76.0
	WRKY56	AT1G64000	474956	80.5
WRKY43	AT2G46130	322266	98.2	
IIa	WRKY60	AT2G25000	481352	93.4
	WRKY18	AT4G31800	328547	81.0
	WRKY40	AT1G80840	477161	97.7
	WRKY36	AT1G69810	339216	94.8
	WRKY61	AT1G18860	335144	84.0
	WRKY72	AT5G15130	909483	92.7
	WRKY31	AT4G22070	329411	86.1
	WRKY42	AT4G04450	327396	87.3
	WRKY6	AT1G62300	475116	98.0
	WRKY47	AT4G01720	490404	91.2
WRKY9	AT1G68150	876849	89.8	
IIb	WRKY27	AT5G52830	495347	88.5
	WRKY22	AT4G01250	490447	85.6
	WRKY29	AT4G23550	492514	82.6
	WRKY14	AT1G30650	890619	85.6
	WRKY35	AT2G34830	482446	84.8
	WRKY52	AT5G45260	884073	86.5
	WRKY65	AT1G29280	336031	89.2
	WRKY69	AT3G58710	486325	98.9
	WRKY7	AT4G24240	914166	92.6
	WRKY15	AT2G23320	343891	90.2
	WRKY17	AT2G24570	481313	96.9
	WRKY11	AT4G31550	913314	96.6
	WRKY21	AT2G30590	902048	93.7
	WRKY74	AT5G28650	489657	90.0
	WRKY39	AT3G04670	477731	98.2

III	WRKY46	AT2G46400	483794	88.5
	WRKY55	AT2G40740	321760	82.5
	WRKY30	AT5G24110	489268	96.4
	WRKY53	AT4G23810	492487	89.8
	WRKY41	AT4G11070	489946	90.4
	WRKY62	AT5G01900	860504	92.8
	WRKY38	AT5G22570	910251	84.1
	WRKY66	AT1G80590	ND	
	WRKY63	AT1G66600	ND	
	WRKY67	AT1G66550	338870	80.7
	WRKY64	AT1G66560	876639	71.1
	WRKY54	AT2G40750	870142	89.6
	WRKY70	AT3G56400	486080	83.7

ND: not detected

Supplementary Table S4. The *A. lyrata* WRKY TFs containing one N- or C-terminal PEST sequences.

WRKY group	PEST localization	
	N-terminal	C-terminal
la	AIWRKY26 AIWRKY4 AIWRKY1*	AIWRKY25*
lb	AIWRKY28 AIWRKY23 AIWRKY24* AIWRKY56	AIWRKY12*
IIa		AIWRKY36 AIWRKY47
IIb	AIWRKY29 AIWRKY21	AIWRKY22
III	AIWRKY41 AIWRKY54*	AIWRKY46* AIWRKY30 AIWRKY53* AIWRKY64* AIWRKY70*
Total	11	10

*PEST region adjacent to WRKY DNA-binding domain.

Supplementary Table S5. Predicted KEN- and D-boxes in WRKY proteins from *A. thaliana*

WRKY group	Accession number	Position	Peptide	Score		Type
Ia	AT1G55600/AtWRKY10	24 - 27	PPSSPSRPTILAMLEQTD	5.7	5.6	D-box
Ib	-----	-----	-----	-----	-----	-----
IIa	AT4G31800/AtWRKY18	50 - 53	DESASELREELNRVNSEN	5.9	5.6	D-box
	AT1G62300/AtWRKY6	175 - 178	TMDNQKLRELLTQVSNSY	8.3	5.6	D-box
	AT4G04450/AtWRKY42	435 - 438	PLMQFSQRSGLVELNQSV	7.5	5.6	D-box
	AT4G22070/AtWRKY31	129 - 132	KIENQRLRDMLSQATTNF	6.5	5.6	D-box
IIb	AT4G24240/AtWRKY7	154 - 157	CPTPIHRRPPLSHNNNNN	8.5	5.6	D-box
	AT4G31550/AtWRKY11	143 - 145	SAELEFSKENFSVSLNS	18.8	9.1	KEN-box
	AT1G30650/AtWRKY14	280 - 283	NHPWPIQRNALAGSTRSS	5.9	5.6	D-box
	AT4G01250/AtWRKY22	191 - 194	NHPAPTHRNSLAGSTRQK	6.2	5.6	D-box
	AT5G52830/AtWRKY27	228 - 231	THPRPTHRNSLAGSTRNK	6.8	5.6	D-box
	AT2G34830/AtWRKY35	278 - 281	NHPWPTQRNALAGSTRSS	6.9	5.6	D-box
	AT4G11070/AtWRKY41	45 - 47	SYLTTEIKENLLHNIVS	11.8	9.1	KEN-box
III	AT5G01900/AtWRKY62	81 - 84	SPVLESSRKPLHKRGRKT	6.1	5.6	D-box
	AT3G56400/AtWRKY70	90 - 93	FEDSGDSRKRLGPVKGKR	5.9	5.6	D-box

Supplementary Table S6. Comparison of PEST regions between *A. thaliana* and *A. lyrata* WRKY orthologs.

WRKY Name	WRKY Group	PEST sequences and localization	PEST score
AtWRKY1	Ia	64-HTVASQSEVDVASPVSEK-81	-0.7
AIWRKY1		63-HTVASLSEVDVASPTSEK-80	1.7
AtWRKY8	Ib	116-RVSASPSSSEADH-128	5.3
AIWRKY8		116-RVSPSPSSSEADH-128	11.1
AtWRKY40	IIa	234-RSSLTPVTTVDMIESK-250	-5.2
AIWRKY40		234-RSSLTEPVTTVDLTESK-250	6.9
AtWRKY15	IIb	41-RDSFNSDQPSSSSASASAAADLESAR-68	6.1
AIWRKY15		41-RDSFNSDQPSSSSASAATTAEDLESAR-68	11.9
AtWRKY46	III	237-KSNLFLGNFVEDLSPATSGSAITSEVLSAPAAVENSETADSYFSSLDNIIDFGQDWLWS-295	-2.8
AIWRKY46		235-KSNIFSENFMEDLSPATSGSAITSEVISAPASIENSETADSYFSSLDNIIDFGQDWLWS-293	0.2
AtWRKY55		35-KPPESLLTSLDEIVK-49	-3.3
AIWRKY55		25-KPPESLSTSLDEIAK-39	3.9
AtWRKY30		271-HSASNSTVPLESPFESYDPNH-293	7.7
AIWRKY30		272-HSTSNSTVPLESPFESYDPNH-294	10.2
AtWRKY53		215-HTCSQAITRTPPLASPEK-232	*ND
AIWRKY53		215-HTCSQVITPPPPASPEK-232	6.1
AtWRKY41		23-KQLQGSSPSLSASSYLTTTEIK-45	-1.4
AIWRKY41		23-KQLQASPPSPSTSLTMEMK-43	2.8
AtWRKY64		178-KFDQVVSESVMPQLATIDEQAIMEDEAIDH-208	-2.0
AIWRKY64		184-KSVMPQLATIEEQEITMEDEATDH-207	2.9
AtWRKY54		75-KTISVLDSFDPVAVSVPIAVEGSWNASCDDSATPVSCNGGDSGESK-121	-0.1
AIWRKY54		75-KTISVLDSFDPVTVSVPIVSGSWNASCDDSATPASCNGGDSGDSK-121	2.7
AtWRKY70		223-HGSSTENDLSLVWPEMVFEEDYH-245	-0.2
AIWRKY70		227-HGSSTENELSLVWPEMVFEEDYH-249	0.02