

Cristian Becerra Baeza

Académico Departamento de Biotecnología, Facultad de Ciencias Naturales, Matemática y del Medio Ambiente, Universidad Tecnológica Metropolitana.

Carlos M. Vicient Sánchez

Científico titular del Consejo Superior de Investigaciones Científicas (CSIC, España), Centro de Investigación en Agrigenómica (CRAG) (CSIC-IRTA-UAB-UB).

EXPRESIÓN DE PROTEÍNAS DE LA FAMILIA I ANKTM EN *Arabidopsis thaliana*

RESUMEN

Un análisis *in silico* de la Familia I de genes *AtAnkTm* en *Arabidopsis thaliana* L. ha aportado nuevos datos de expresión génica basada en análisis transcripcionales e información sobre proteínas que potencialmente interactúan con las proteínas codificadas por estos genes. Además, se han revisado los conocimientos acerca de los 11 genes integrantes de esta familia al revisar y actualizar el registro de publicaciones asociadas a ellos, que fue ampliado consistentemente.

ABSTRACT

An *in silico* analysis of Family I *AtAnkTm* genes from *Arabidopsis thaliana* L. reported new information of the expression of the genes from transcripts databases and proteins that putatively interact with the proteins encoded by them. In addition, the knowledge of the 11 members of this gene family was updated and significantly expanded by reviewing the publication records associated with them.

INTRODUCCIÓN

El número total de genes estimado en el genoma de la planta modelo *Arabidopsis thaliana* es de unos 26.000 (Berardini et al., 2004), repartidos en unas 7.500 familias, considerando como familia de genes a aquél conjunto de genes que codifican proteínas cuyas secuencias son similares (*The Arabidopsis Genome Initiative*, 2000). De acuerdo a lo anterior, el número de tipos de proteínas codificadas por los genes de *Arabidopsis* se estima en unas 11.600, similar a otros organismos multicelulares. Un 65% de los tipos de proteínas son codificados por familias de genes.

Una forma de predecir la función de un gen es determinar la presencia y organización de dominios proteicos conservados (Sessions et al., 2002). Se entiende por dominio proteico una porción de una proteína con estructura terciaria definida y, en general, asociado a alguna función.

Dominios transmembrana (*Tm*): están compuestos por entre 15 y 30 residuos mayoritariamente hidrofóbicos y se encuentran en proteínas que se localizan en alguna membrana celular. Unos 4.600 genes de *Arabidopsis thaliana* (18%) contienen dos o más dominios transmembrana (Ward, 2001; Schwacke et al., 2003). Las proteínas integrales de membrana pueden tener importantes funciones en transporte de solutos a través de la membrana, percepción y transducción de señales y actividades biosintéticas y metabólicas.

Dominios anquirina: Breeden y Nasmyth (1987) describieron la presencia de unas secuencias de 33 aminoácidos repetidas en tándem (una al lado de la otra) en dos proteínas reguladoras de ciclo celular de *Schizosaccharomyces pombe* y *Saccharomyces cerevisiae* (CDC10 y SW16). Posteriormente, se descubrieron 24 copias

de esta repetición en una proteína humana denominada anquirina, relacionada con el citoesqueleto, que fue la que dio el nombre a la repetición anquirina (ANK). Desde entonces se han identificado numerosas proteínas que contienen dominios ANK con variadas funciones como la regulación del ciclo celular, enzimas mitocondriales, interacciones con el citoesqueleto, traducción de señal o toxinas (Sedgwick y Smerdon, 1999). Existen proteínas con dominios ANK en un amplio espectro de organismos incluyendo Eucariotas (animales, hongos, algas y plantas), Procariotas, Archea y algunos virus (Al-Khodor et al., 2009; Sonnberg et al., 2011). Existe un número creciente de ejemplos en los que las repeticiones ANK funcionan como lugares de unión proteína-proteína, por lo que se considera que esa es su función principal (Sedgwick y Smerdon, 1999).

Los genomas de plantas contienen un número moderado de genes que codifican proteínas con dominios ANK. En el año 2004 Becerra et al. realizaron una amplia descripción de los genes de la planta modelo *Arabidopsis thaliana* L. que codificaban proteínas con repeticiones Anquirina. El estudio permitió identificar un total de 105 genes de los cuales 40 codificaban además dominios transmembrana, por lo que a estos genes se les denominó *AtAnkTm*. La comparación de sus secuencias proteicas permitió agruparlos en seis familias (I a VI). El genoma de arroz contiene 175 genes que codifican proteínas con dominio ANK de los cuales 37 tienen además dominios transmembrana (Huang et al., 2009), en maíz se identificaron 71 en total y 15 con dominio transmembrana (Jiang et al., 2013) y en tomate 130 en total y 25 con dominios transmembrana (Yuan et al., 2013). De todos ellos existen muy pocos de los que se haya estudiado su función. Algunos ejemplos son ANK1 de tabaco (*Ankyrin repeat protein 1*) y su homóloga AKR2 de *Arabidopsis*, involucradas en la defensa frente a patógenos; CAMTAs (*Calmodulin-binding transcription activators*) en

Arabidopsis y *Brassica napus*, probablemente implicadas en cascadas de regulación de la transcripción; la familia AKT de Arabidopsis y SKT1, un homólogo en patata, que codifican canales de potasio dependientes de voltaje de la familia Shaker; APKs (*ankyrin protein kinases*), quinasas descritas en *Medicago*; ART2 de Arabidopsis, interviene en la defensa frente a patógenos; CAO, (*chlorophyll a/b binding protein harvesting-organelle specific protein*) descritas en Arabidopsis; AKR (*Ankyrin Repeat gene*) de Arabidopsis, interviene en los procesos de diferenciación celular asociada a luz; NPR1 (*non-expresser of PR genes*) de Arabidopsis, que interviene en el control de la respuesta frente a patógenos SAR (*Systemic Acquired Resistance*); ACBP2 (*cytosolic acyl-CoA-binding*) de Arabidopsis, es una proteína de unión a acil-CoA; EMB506 de Arabidopsis, es necesaria para el correcto desarrollo embrionario.

Dentro de las proteínas con dominios ANK y transmembrana se han estudiado algunos casos (Becerra, 2006) como ACD6 de Arabidopsis, un posible regulador y efector de la señal del ácido salicílico en la respuesta de defensa (Lu et al., 2003); ITN1 de Arabidopsis, involucrado en la respuesta a salinidad (Sakamoto et al., 2012); BDA1 de Arabidopsis, involucrado en reacciones de defensa frente a patógenos (Yang et al., 2012); e IGN1 de *Lotus japonica*, relacionado con la fijación de nitrógeno.

Se realizó un análisis *in silico* con el objetivo de actualizar la información de la Familia I de genes *AtAnkTm* a partir de las bases de datos de acceso público para actualizar la información acerca del patrón de expresión de estos genes. Adicionalmente, se actualizaron los antecedentes bibliográficos disponibles para cada uno de los 11 genes pertenecientes a esta familia.

MATERIALES Y MÉTODOS

Mediante un análisis *in silico*, se realizó una búsqueda y actualización de ESTs (del inglés, *Expressed Sequence Tags*) a través de la base de datos de TAIR (*The Arabidopsis Information Resource*; <http://www.arabidopsis.org/>) para cada uno de los 11 genes *AtAnkTm* pertenecientes a la familia I de la especie vegetal *Arabidopsis thaliana* L. y, de las citas bibliográficas asociadas a la caracterización de cada uno de ellos. Adicionalmente, utilizando la misma base de datos, se identificaron las referencias bibliográficas presentes en el apartado “Detalles de Anotaciones” (del inglés, “*Annotation details*”), sección que corresponde a la descripción de la función (o posible función) de cada uno de los genes, su expresión y localización del producto del gen y, de aquellas disponibles en el apartado “Publicaciones” (del inglés, “*Publication*”), para cada uno de los genes señalados, con el propósito de comparar y actualizar los resultados obtenidos en 2004. La búsqueda de proteínas que interactúan se realizó a través de “*The Arabidopsis Interactions Viewer*”, basado en interacciones confirmadas en los proyectos BIND (*The Biomolecular Interaction Network Database*), basado en micromatrices proteicas y ensayos doble híbrido en levadura (Popescu et al., 2007; Popescu et al., 2009), el *Arabidopsis Interactome* 2011 (Lin et al., 2011), el MIND (*Membrane protein INteractome Database*) (Loqué et al., 2007), y 1190 citas bibliográficas.

RESULTADOS Y DISCUSIÓN

Algoritmos informáticos permiten detectar genes basándose en la conservación de ciertas secuencias, pero dichos genes no pueden ser considerados más que hipotéticos hasta que existan pruebas experimentales de su función. Una de estas pruebas es la existencia de mRNAs (cDNAs). Se han invertido considerables esfuerzos en la secuenciación masiva de moléculas de cDNA, las denominadas ESTs (*Expressed Sequence Tags*). En la base de secuencias dbEST existían 421.027 entradas para esta especie al 20 de enero de 2006, las que se han incrementado a 1.529.700 (01/01/2013) (http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html). A pesar de ello, no se han detectado ESTs correspondientes a muchos de los genes predichos *in silico* a partir de la secuencia genómica. Rudd (2003) calculó que existían ESTs para solo unos 16.115 genes, por lo tanto, el resto de genes continuaban siendo hipotéticos. Se puede suponer que un número de genes todavía escapa a la verificación experimental debido a su baja abundancia de transcritos y/o a la severa restricción espacio-temporal del patrón de expresión, pero en otros casos pueden tratarse de pseudogenes o genes no activos, o, simplemente, de errores de predicción. La familia I de *AtAnkTm* está constituida por 11 genes. En el año 2004 sólo existían 20 EST depositados en las bases de datos para cuatro de estos genes. El 80% de los ESTs correspondía a información disponible del gen *At4g14400*, precisamente el único que había sido descrito previamente en términos de publicaciones (Becerra et al., 2004). En el Cuadro 1 se encuentra el número de Atg (N° de identificación) de cada uno de los genes y los ESTs identificados el año 2004 y los datos actualizados al año 2014 (09/OCT/2014).

Para los genes *At1g03670*, *At4g03470*, *At4g03480* y *At4g14390* no existe información disponible respecto de ESTs que hayan sido depositados

en las bases de datos públicas. Para los genes *At4g03460*, *At4g03490* y *At4g03500*, de los que no existía registro de ESTs (2004), se han depositado secuencias asociadas a la expresión de ellos, uno para cada uno en los dos primeros casos y diez ESTs para el gen *At4g03500*, principalmente asociados a diferentes órganos de la planta y condiciones de estrés. Para *At4g03440*, *At4g03450* y *At4g05040*, ha ocurrido un notable ingreso de registros en las bases de datos de ESTs (Cuadro 1). De los 4 ESTs totales descritos para estos genes en 2004, este número aumentó a 22 ESTs en 2014. La situación es completamente distinta para el gen *At4g14400* ya que en el año 2004 era el único gen *AtAnkTm* que había sido caracterizado y que codifica la proteína ACD6, que es un posible regulador y efector de la señal del ácido salicílico en la respuesta de defensa (Lu et al., 2003). En términos exclusivos de ESTs, se ha incrementado de 16 a 135, lo que está directamente asociado con el estudio y profundización del conocimiento asociado a este gen. No obstante, 79 de los 135 ESTs depositados en *GenBank* corresponden al trabajo desarrollado por Weber et al. (2007), quien utilizando tecnología de pirosecuenciación, permitió la detección de nuevos EST tanto para este gen, como para otros integrantes dentro de esta familia.

De acuerdo a los antecedentes disponibles al revisar cada una de las secuencias depositadas, las citas asociadas a los ESTs descritos para estos genes, en general, están centradas en tan solo 7 publicaciones. De este grupo de publicaciones, destacan dos que han permitido caracterizar en mayor detalle el patrón de expresión de los genes pertenecientes a esta familia, particularmente los trabajos de Seki et al. (2002) y Weber et al. (2007).

Las interacciones entre proteínas son esenciales para la mayoría de las funciones de las células y por ello se han establecido programas para el estudio sistemático de las interacciones de todas

las proteínas de una especie basados en datos experimentales como ensayos doble híbrido en levadura o micromatrices de proteínas (Lin et al., 2011). Estudios de este tipo se han realizado en especies modelo como *Saccharomyces cerevisiae*, *Homo sapiens*, *Drosophila melanogaster*, *Caenorhabditis elegans* y *Arabidopsis thaliana*. El establecimiento de bases de acceso abierto permite obtener información de que proteínas interactúan con la proteína de interés, siempre teniendo en cuenta las limitaciones de los métodos utilizados, es decir, que no todas las

interacciones de una proteína serán observadas. Este análisis es especialmente interesante en el caso de proteínas con dominios ANK ya que la función de este dominio se ha definido precisamente como la de facilitar la interacción con otras proteínas. Una búsqueda con todos los miembros de la familia I de *AtAnkTm* demuestra que existen interacciones para 5 de las 11 proteínas (Cuadro 1) y que los interactores son, con especial abundancia, proteínas reguladoras y transportadores de membrana. Destacar que dos de las proteínas estudiadas interactúan entre sí, las proteínas 2 y 3.

CUADRO 1.- DETALLE DE ESTS, GENOTECAS Y PROTEÍNAS QUE INTERACCIONAN CON PROTEÍNAS CODIFICADAS POR LA FAMILIA I DE GENES *AtAnkTm*

Gen	Número Atg	Número de ESTs* (2004)	Número de ESTs (2014)	Números de Acceso de EST (GenBank) y descripción de la genoteca (09/10/2014)	Citas Bibliográficas	Proteínas que interactúan con proteína codificada por el gen de interés
1	At1g03670	No	No	-	-	-
2	At4g03440	1	1	AV798477 ; Diferentes estados, deshidratación-frío	-	At3g54860, ATVPS33_VPS33 Superfamilia de proteínas tipo Sec1/munc18 (SM) At4g03450, Proteína de la familia I ANKTM
			1	ES160611; Óvulos	-	
			1	BP625216; Deshidratación-Rehidratación	Seki et al. (2002)	
			1	COo48790; Planta completa	-	
3	At4g03450	2	1	AI997466 ; Raíces (4 - 7 semanas)	-	At1g18660, Familia de proteínas dedos de zinc (C3HC4-tipo anillo de dedo) At5g03730, AtCTR1_CTR1_SIS1 Proteína de la superfamilia de proteínas quinasas
			1	BE662952 ; Ozono	-	
			2	EG451685, EG451687; Diferentes tejidos y tipos de estrés	-	
			1	COo48791; Planta completa	-	
4	At4g03460	No	1	EH956114; Plántulas 8 días	Weber et al. (2007)	At1g18660, Familia de proteínas dedos de zinc (C3HC4-tipo anillo de dedo) At5g03730, AtCTR1_CTR1_SIS1 Proteína de la superfamilia de proteínas quinasas
5	At4g03470	No	No	-	-	-
6	At4g03480	No	No	-	-	-
7	At4g03490	No	1	COo48792; Planta completa	-	-
8	At4g03500	No	2	DR380737, DR251868; Botones florales y raíces	Alexandrov et al. (2006)	At3g24300, AMT1;3, ATAMT1;3 Transportador de amonio1;3
			2	EG457527, EG457528; Diferentes tejidos y tipos de estrés	-	
			3	EL162997, EH966640, ELO61396; Plántulas 8 días	Weber et al. (2007)	
			1	AV783353; Roseta, deshidratación	-	
			1	AV822553; Roseta, deshidratación	-	
			1	BP655801; Silicuas y flores	Seki et al. (2002)	

9	At4g05040	1	3	AV793119 , AV825974, BP793571; frío	-	143 proteínas
			1	DR230259; Flores, hojas y raíces	Alexandro et al. (2006)	
			2	BX835291, BX838828; Botones florales y flores	-	
			4	BP668309 ^(a) , BP868925, BP851846, BP856126; Estrés y hormonas	Seki et al. (2002)(a)	
			1	BP830779; Flores y silicuas	-	
			2	BP807328, BP808686; Oscuridad	-	
10	At4g14390	No	No	-	-	-
11	At4g14400 (ACD6)	16	135	BU634999; CA781914; CA781971; Hojas infectadas con <i>Erysiphe cichoracearum</i>	-	-
				A1100255, H37081, T42092, T46084 ; Diferentes tejidos	Newman et al. (1994)	
				A1998090 ; Roseta (4 - 7 semanas)	-	
				AV440557 ; Parte aérea (2 - 6 semanas)	Asamizu et al. (2000)	
				BE662757 ; Ozono	-	
				EG421697, EG450297, EG485142, EG485143, EG494607, EG513027, EG513028; Diferentes tejidos y tipos de estrés	-	
				EL095310, EL095991, EL192547, EL147004, EL147099, EL178727, EL125548, EL185222, EL147454, EL084301, EL147887, EL115744, EL162090, EL185938, EL159766, EL173042, EL157059, EL185254, EL152639, EL177298, EL185046, EL112274, EL114075, EL188720, EL332228, EL330177, EL295258, EL337334, EL337334, EL292167, EL334235, EL297974, EL332392, EL207327, EL299725, EL287363, EL297690, EL320921, EL233577, EL223232, EL339673, EL327859, EL262231, EH896943, EH899397, EH844824, EH927090, EH903474, EH920133, EH808332, EH828332, EH861293, EH826732, EH810577, EH904979, EH825276, EH938091, EH829071, EH986457, EH944107, EL016954, EH945922, EL016599, EH951002, EL053457, EH949187, EL028657, EH956460, EL043434, EL032781, EH956719, EH957868, EL068731, EH980959, EL056988, EH981220, EH949781, EL051481, EL002177; Plántulas 8 días.	Weber et al. (2007)	
				CB263352; Plántulas	Schmid et al. (2003)	
				AV782706, AV822072 ; Roseta, frío	-	
				AV793013 , BP561505 ^(a) ; BP778726, BP788301, BP788748, BP790043, BP792619, BP792951, BP794196, BP794969, BP795026, BP795145, BP779663, BP780119, BP782081, BP782202, BP782370, BP783345, BP783860, BP786804, BP787324; Roseta, frío	Seki et al. (2002) ^(a)	
				AV794986, AV826452 ; Roseta, deshidratación	-	
				AV807835, AV830548, AV815244 , AV813132, AV812070; Diferentes estados, deshidratación y frío	-	
				AU231325 , BP666894 ^(a) , BP862282, BP868477, BP850613, BP858110, BP859370; Estrés y hormonas	Seki et al. (2002) ^(a)	
CB074323; Hojas infectadas con <i>Peronospora</i> parasítica	Mahalingam et al. (2003)					

*: Becerra C, Jahrman T, Puigdomenech P, Vicent CM. 2004. *Gene* **340**: 111-121.

(a): corresponde al EST específicamente citado.

Obs.:

1) Datos en negrilla (N° de acceso, 5ta. Columna) corresponden a los EST descritos el año 2004.

2) EST **T46084** del gen *At4g14400* registrado en 2004 no se encuentra disponible en la base de datos consultada; se confirmó directamente a través de NCBI, *National Center for Biotechnology Information* (<http://www.ncbi.nlm.nih.gov/>).

En el “Detalle de anotaciones” presentado en el Cuadro N°2, para cada uno de los genes pertenecientes a la Familia I *AtAnkTm*, los resultados de la búsqueda arrojaron la existencia de una serie de referencias bibliográficas para todos los genes, excepto *At1g03670* y *At4g03460*, de los que no se tenía registro previo en su conjunto. En términos generales, la mayor información disponible acerca del estudio de expresión de genes no es un ámbito exclusivo del análisis tan sólo de ESTs, sino que también, puede incluir estudios a gran escala a través de técnicas de secuenciación masiva de última generación, las que permiten una mayor cantidad de datos respecto de la presencia de transcritos de un gen en un momento dado de la vida de la planta y de cuyo nivel de expresión sea no detectable utilizando los ESTs. Además, del estudio de las micromatrices a través de los cuales se puede conocer la expresión de un gran número de genes utilizando diferentes tejidos de la planta y diferentes condiciones de estrés (por ejemplo, deshidratación, frío, etc.), es otro ámbito a destacar dentro de estos trabajos. Los continuos avances de estas técnicas, así como las relacionadas con las bases de datos y la utilización de programas de análisis masivos que permiten integrar datos de diferentes fuentes, han contribuido a dilucidar mayores antecedentes relativos a los genes *AtAnkTm* pertenecientes a esta familia.

Ahora bien, el análisis de las publicaciones relacionadas con esta familia de genes, señala que existe una serie de datos relevantes en todos los genes en estudio, como se puede observar al revisar la cantidad de trabajos científicos en los que han sido incluidos. Destaca entre ellos, lo desarrollado por Du y colaboradores (2007) quien analizó la diversidad genética y evolución de los genes que se encuentran en tándem, dentro de esta familia.

Dentro de este grupo, el gen *At4g14400* es el que presenta la mayor profundización y más amplia descripción (14 publicaciones) como motivo de estudio, cuestión que evidencia el trabajo sostenido a través del tiempo de diversos grupos de investigación involucrados en el estudio de la respuesta de las plantas bajo condiciones de estrés y en la determinación de vías metabólicas específicas.

CUADRO 2.- REFERENCIAS BIBLIOGRÁFICAS ASOCIADAS A LOS GENES *AtANKTM* DE LA FAMILIA I

Gen	Número Atg	Detalles de Anotación (Referencias)	Publicaciones
1	At1g03670	-	Du et al. (2007)
2	At4g03440	Obulareddy et al. (2013)	Du et al. (2007)
3	At4g03450	Heyndrickx et al. (2012) Schmid et al. (2005)	Ahlfors et al. (2009) Ascencio-Ibañez et al. (2008) Du et al. (2007) Whitham et al. (2003)
4	At4g03460	-	Hanada et al. (2011) Du et al. (2007)
5	At4g03470	Sugiyama et al. (2008)	Du et al. (2007)
6	At4g03480	Sugiyama et al. (2008)	Du et al. (2007)
7	At4g03490	-	Du et al. (2007)
8	At4g03500	Obulareddy et al. (2013) Schmid et al. (2005)	Luhua et al. (2013) Hanada et al. (2011) Staal et al. (2008) Du et al. (2007)
9	At4g05040	Obulareddy et al. (2013) Borges et al. (2008) Schmid et al. (2005)	Lu et al. (2003)
10	At4g14390	Borges et al. (2008) Sugiyama et al. (2008) Schmid et al. (2005)	Ascencio-Ibañez et al. (2008) Che et al. (2006) Goda et al. (2004) Lu et al. (2003)
11	At4g14400 (ACD6)	Obulareddy et al. (2013) Chen et al. (2010) Schmid et al. (2005) Lu et al. (2005) Lu et al. (2003) Rate et al. (1999)	Okuma et al. (2014) Todesco et al. (2014) Yang et al. (2013) Meinke (2012) Todesco et al. (2010) Miura y Ohta (2010) Chen et al. (2010) Lu et al. (2009) Lee et al. (2008) Ascencio-Ibañez et al. (2008) Lu et al. (2005) Lu et al. (2003) Vanacker et al. (2001) Rate et al. (1999)

BIBLIOGRAFÍA

- [1] Ahlfors, R., Broché, M., Kollist, H., Kangasjärvi, J. 2009. Nitric oxide modulates ozone-induced cell death, hormone biosynthesis and gene expression in *Arabidopsis thaliana*. The Plant Journal 58: 1–12.
- [2] Alexandrov, N.N., Troukhan, M.E., Brover, V.V., Tatarinova, T., Lu, Y.-P., Flavell, R.B., Feldmann, K.A. 2006. Features of Arabidopsis genes and genome discovered using full-length cDNAs. Plant Mol. Biol. 60 (1): 69–85.
- [3] Al-Khodor, S., Price, C., Kalia, A., Kwaik, Y.A. 2009. Functional diversity of ankyrin repeats in microbial proteins. Trends in Microbiology 18 (3): 132–139.
- [4] Asamizu, E., Nakamura, Y., Sato, S., Tabata, S. 2000. A large scale analysis of cDNA in *Arabidopsis thaliana*: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries. DNA Res. 7 (3): 175–180.
- [5] Ascencio-Ibañez, J., Sozzani, R., Lee, T.-J., Chu, T.-M., Wolfinger, R., Cella, R., Hanley-Bowdoin, L. 2008. Global analysis of Arabidopsis gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. Plant Physiology 148: 436–454.
- [6] Becerra C, Jahrmann T, Puigdomenech P, Vicient CM. 2004. Ankyrin repeat-containing proteins in Arabidopsis: characterization of a novel and abundant group of genes coding ankyrin-transmembrane proteins. Gene 340: 111–121.
- [7] Becerra C, 2006. Tesis doctoral “Genes implicados en el desarrollo de la semilla de *Arabidopsis thaliana* (L.): caracterización de los genes AtAnkTm”, Universidad Autónoma de Barcelona, España.
- <http://www.tdx.cat/handle/10803/127/browse?value=Becerra+Baeza%2C+Cristian+Marcelo+del+Carmen&type=author>
- [8] Berardini, TZ, Mundodi S, Reiser L, Huala E, Garcia-Hernandez M, Zhang P, Mueller LA, Yoon J, Doyle A, Lander G, Moseyko N, Yoo D, Xu I, Zoeckler B, Montoya M, Miller N, Weems D, Rhee SY. 2004. Functional annotation of the Arabidopsis genome using controlled vocabularies. Plant Physiol 135: 745–755.
- [9] Borges, F., Gomes, G., Gardner, R., Moreno, N., McCormick, S., Feijo, J., Becker, J. 2008. Comparative Transcriptomics of *Arabidopsis thaliana* Sperm Cells. Plant Physiology 148: 1168–1181.
- [10] Breeden L, Nasmyth K. 1987. Similarity between cell-cycle genes of budding yeast and fission yeast and the *Notch* gene of *Drosophila*. Nature 329: 651–654.
- [11] Che, P., Lall, S., Nettleton, D., Howell, S. 2006. Gene Expression Programs during Shoot, Root and Callus Development in Arabidopsis Tissue Culture. Plant Physiology 141: 620–637.
- [12] Chen, H., Zhang, Z., Teng, K., Lai, J., Zhang, Y., Huang, Y., Li, Y., Liang, L., Wang, Y., Chu, C., Guo, H., Xie, Q. 2010. Up-regulation of LSB1/GDU3 impacts geminivirus infection by activating the salicylic acid pathway. The Plant Journal 62: 12–23.
- [13] Du J., Wang X., Zhang M., Tian D. and Yang Y.-H. 2007. Unique nucleotide polymorphism of ankyrin gene cluster in Arabidopsis. J. Genet. 86: 27–35.
- [14] Goda, H., Sawa, S., Asami, T., Fujioka, S., Shimada, Y., Yoshida, S. 2004. Comprehensive Comparison of Auxin-Regulated and Brassinosteroid-Regulated Genes in Arabidopsis. Plant Physiology 134: 1555–1573.

- [15] Hanada, K., Sawada, Y., Kuromori, T., Klausnitzer, R., Saito, K., Toyoda, T., Shinozaki, K., Li, W. H., Hirai, M. Y. 2011. Functional compensation of primary and secondary metabolites by duplicate genes in *Arabidopsis thaliana*. *Molecular biology and evolution* 28(1):377–382.
- [16] Heyndrickx KS, Vandepoele, K. 2012. Systematic identification of functional plant modules through the integration of complementary data sources. *Plant Physiol.* 159: 884–901.
- [17] Huang, J., Zhao, X., Yu, H., Ouyang, Y., Wang, L., Zhang, Q. 2009. The ankyrin repeat gene family in rice: genome-wide identification, classification and expression profiling. *Plant Mol Biol* 71: 207–226.
- [18] Jiang, H., Wu, Q., Jin, J., Sheng, L., Yan, H., Cheng, B., Zhu, S. 2013. Genome-wide identification and expression profiling of ankyrin-repeat gene family in maize. *Dev Genes Evol* 223: 303–318.
- [19] Lee, M.W., Jelenska, J., Greenberg, J.T. 2008. Arabidopsis Proteins Important for Modulating Defense Responses to *Pseudomonas syringae* that Secrete HopW1-1. *The Plant Journal* 54: 452–465.
- [20] Lin M, Shen X, Chen X. 2011. PAIR: the predicted Arabidopsis interactome resource. *Nucleic Acids Res.* 39(Database issue):D1134–40.
- [21] Loqué, D., Lalonde, S., Looger, L., von Wirén, N., Frommer, W.F. 2007. Trans-activation and cooperativity in an ammonium transporter complex. *Nature* 446:195–198.
- [22] Lu, H., Rate, D. N., Song, J. T., Greenberg, J. T. 2003. ACD6, a Novel Ankyrin Protein, Is a Regulator and an Effector of Salicylic Acid Signaling in the Arabidopsis Defense Response. *The Plant Cell* 15: 2408–2420.
- [23] Lu, H., Liu, Y., Greenberg, J.T. 2005. Structure-function analysis of the plasma membrane-localized Arabidopsis defense component ACD6. *The Plant Journal* 44: 798–809.
- [24] Lu, H., Salimian, S., Gamelin, E., Wang, G., Fedorowski, J., Lacourse, W., Greenberg, J.T. 2009. Genetic analysis of *acd6-1* reveals complex defense networks and leads to identification of novel defense genes in Arabidopsis. *The Plant Journal* 58: 401–412.
- [25] Luhua, S., Hegie, A., Suzuki, N., Shulaev, E., Luo, X., Cenariu, D., Ma, V., Kao, S., Lim, J., Gunay, M. B., Oosumi, T., Lee, S. C., Harper, J., Cushman, J., Gollery, M., Girke, T., Bailey-Serres, J., Stevenson, R. A., Zhu, J. K., Mittler, R. 2013. Linking genes of unknown function with abiotic stress responses by high throughput phenotype screening. *Physiologia Plantarum* 148: 322–333.
- [26] Mahalingam, R., Gomez-Buitrago, A.M., Eckardt, N., Shah, N., Guevara-Garcia, A., Day, P.M., Raina, R., Fedoroff, N.V. 2003. Characterizing the stress/defense transcriptome of Arabidopsis. *Genome Biology* 4 (3): R20.
- [27] Meinke, D. W. 2012. A survey of dominant mutations in *Arabidopsis thaliana*. *Trends in Plant Science* 18(2): 84–91.
- [28] Miura, K., Ohta, M. 2010. SIZ1, a small ubiquitin-related modifier ligase, controls cold signaling through regulation of salicylic acid accumulation. *Journal of Plant Physiology* 167(7): 555–560.
- [29] Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E., Somerville, C. 1994. Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones. *Plant Physiol.* 106: 1241–1255.

- [30] **Obulareddy, N., Panchal, S., Melotto, M.** 2013. Guard Cell Purification and RNA Isolation Suitable for High Throughput Transcriptional Analysis of Cell-Type Responses to Biotic Stresses. *Mol Plant Microbe Interact.* 26(8): 844–849.
- [31] **Okuma, E., Nozawa, R., Murata, Y., Miura, K.** 2014. Accumulation of endogenous salicylic acid confers drought tolerance to *Arabidopsis*. *Plant Signaling & Behavior* 9: e28085.
- [32] **Popescu SC1, Popescu GV, Bachan S, Zhang Z, Seay M, Gerstein M, Snyder M, Dinesh-Kumar SP.** 2007. Differential binding of calmodulin-related proteins to their targets revealed through high-density *Arabidopsis* protein microarrays. *Proc Nat Acad Sci USA.* 104(11): 4730–5.
- [33] **Popescu SC1, Popescu GV, Bachan S, Zhang Z, Gerstein M, Snyder M, Dinesh-Kumar SP.** 2009. MAPK target networks in *Arabidopsis thaliana* revealed using functional protein microarrays. *Genes Dev.* 23(1): 80–92.
- [34] **Rate, D. N., Cuenca, J. V., Bowman, G. R., Guttman, D. S., Greenberg, J. T.** 1999. The gain-of-function *Arabidopsis acd6* mutant reveals novel regulation and function of the salicylic acid signaling pathway in controlling cell death, defenses, and cell growth. *The Plant Cell* 11 (9): 1695–1708.
- [35] **Rudd S.** 2003. Expressed sequence tags: alternative or complement to whole genome sequences? *Trends Plant Sci* 8: 321–329.
- [36] **Sakamoto, H., Sakata, K., Kusumi, K., Kojima, M., Sakakibara, H., Iba, K.** 2012. Interaction between a plasma membrane-localized ankyrin-repeat protein ITN1 and a nuclear protein RTV1. *Biochemical and Biophysical Research Communications* 423: 392–397.
- [37] **Schmid, K.J., Soerensen, T.R., Stracke, R., Torjek, O., Altmann, T., Mitchell-Olds, T., Weisshaar, B.** 2003. Large-scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in *Arabidopsis thaliana*. *Genome Res.* 13 (6): 1250–1257.
- [38] **Schmid, M., Davison, T., Henz, S., Pape, U., Demar, M., Scholkopf, B., Weigel, D., Lohmann, J., Vingron, M.** 2005. A gene expression map of *Arabidopsis thaliana* development. *Nature genetics* 37 (5): 501–506.
- [39] **Schwacke R, Schneider A, van der Graaff E, Fischer K, Catoni E, Desimone M, Frommer WB, Flugge UI, Kunze R.** 2003. ARAMEMNON, a novel database for *Arabidopsis* integral membrane proteins. *Plant Physiol* 131: 16–26.
- [40] **Sedgwick, S.G., Smerdon, S.J.,** 1999. The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends Biochem. Sci.* 24, 311–316.
- [41] **Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T., Nakajima, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M., Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Shinozaki, K.** 2002. Functional annotation of a full-length *Arabidopsis* cDNA collection. *Science* 296 (5565): 141–145.
- [42] **Sessions A, Burke E, Presting G, aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C, Clarke J, Cotton D, Bullis D, Snell J, Miguel T, Hutchinson D, Kimmerly B, Mitzel T, Katagiri F, Glazebrook J, Law M, Goff S.** 2002. A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell* 14: 2985–2994.
- [43] **Sonnberg, S., Fleming, S.B., Mercer, A.A.** 2011. Phylogenetic analysis of the large family of poxvirus ankyrin-repeat proteins reveals orthologue groups within and across chordo-

poxvirus genera. *Journal of General Virology* 92: 2596–2607.

[44] Staal, J., Kaliff, M., Dewaele, E., Persson, M., Dixelius, C. 2008. RLM3, a TIR domain encoding gene involved in broad-range immunity of *Arabidopsis* to necrotrophic fungal pathogens. *The Plant Journal* 55, 188–200.

[45] Sugiyama N, Nakagami H, Mochida K, Daudi A, Tomita M, Shirasu K, Ishihama Y. 2008. Large-scale phosphorylation mapping reveals the extent of tyrosine phosphorylation in *Arabidopsis*. *Molecular Systems Biology* 4:193.

[46] The Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815.

[47] Todesco, M., Balasubramanian, S., Hu, T. T., Traw, M. B., Horton, M., Epple, P., Kuhns, C., Sureshkumar, S., Schwartz, C., Lanz, C., Laitinen, R. A., Huang, Y., Chory, J., Lipka, V., Borevitz, J. O., Dangl, J. L., Bergelson, J., Nordborg, M., Weigel, D. 2010. Natural allelic variation underlying a major fitness trade-off in *Arabidopsis thaliana*. *Nature* 465(7298): 632–636.

[48] Todesco, M., Kim, S. T., Chae, E., Bomblies, K., Zaidem, M., Smith, L. M., Weigel, D., Laitinen, R. A. 2014. Activation of the *Arabidopsis thaliana* Immune System by Combinations of Common ACD6 Alleles. *PLoS Genetics* 10(7): e1004459.

[49] Vanacker, H., Lu, H., Rate, D. N., Greenberg, J. T. 2001. A role for salicylic acid and NPR1 in regulating cell growth in *Arabidopsis*. *The Plant Journal* 28(2): 209–216.

[50] Ward J. 2001. Identification of novel families of membrane proteins from the model

plant *Arabidopsis thaliana*. *Bioinformatics* 17 (6): 560–563.

[51] Weber, A.P.M., Weber, K.L., Carr, K.M., Wilkerson, C., Ohlrogge, J.B. 2007. Sampling the *Arabidopsis* transcriptome with massively parallel pyrosequencing. *Plant Physiol.* 144 (1): 32–42.

[52] Whitham, S. A., Quan, S., Chang, H. S., Cooper, B., Estes, B., Zhu, T., Wang, X., Hou, Y. M. 2003. Diverse RNA viruses elicit the expression of common sets of genes in susceptible *Arabidopsis thaliana* plants. *The Plant Journal* 33, 271–283.

[53] Yang, L. P., Fang, Y. Y., An, C. P., Dong, L., Zhang, Z. H., Chen, H., Xie, Q., Guo, H. S. 2013. C2-mediated decrease in DNA methylation, accumulation of siRNAs and increase in expression for genes involved in defense pathways in plants infected with beet severe curly top virus. *The Plant Journal* 73(6): 910–917.

[54] Yang, Y., Zhang, Y., Ding, P., Johnson, K, Li, X., Zhang, Y. 2012. The Ankyrin-Repeat Transmembrane Protein BDA1 Functions Downstream of the Receptor-Like Protein SNC2 to Regulate Plant Immunity. *Plant Physiology* 159: 1857–1865.

[55] Yuan, X. Zhang, S. Qing, X., Sun, M., Liu, S., Su, H., Shu, H., Li, X. 2013. Superfamily of ankyrin repeat proteins in tomato. *Gene* 523: 126–136.