

## EVOLUCIÓN Y CARACTERIZACIÓN MOLECULAR DE FACTORES DE TRANSCRIPCIÓN CYS2-HIS2 ZINC FINGER DE GRAMÍNEAS

Bellino Carolina<sup>A</sup>

<sup>A</sup> Laboratorio de Genética Evolutiva y del Desarrollo, Instituto de Agrobiotecnología del Litoral  
IAL-UNL-CONICET  
Facultad de Bioquímica y Ciencias Biológicas UNL

Área: Ciencias Biológicas

Sub-Área: Biotecnología

Grupo: X

Palabras clave: Angiospermas, ramificación, meristemas axilares

### INTRODUCCIÓN

La familia gramíneas (Poaceae o informalmente llamadas “pastos”) abarca aproximadamente unas 11.000 especies entre las que se encuentran destacados cereales que constituyen una parte muy importante de la base nutricional de la humanidad (GPWGII, 2012). Las especies de gramíneas se distribuyen en 13 subfamilias, tres de ellas evolucionaron tempranamente y el resto se agrupan en dos grandes linajes: el clado BEP y el clado PACMAD (GPWGII, 2012) (Figura 1).

Las gramíneas son morfológicamente únicas entre las Angiospermas, dado que se caracterizan por tener flores altamente modificadas, agrupadas en estructuras llamadas espiguillas (Cheng y col., 1983; Clifford 1987; Ikeda y col., 2004).

Las espiguillas se encuentran distribuidas sobre distintos sistemas de ramificación de inflorescencias. La morfología de la inflorescencia de las gramíneas es conocida por ser extremadamente variable entre las especies, compleja en su desarrollo, además de ser genética y agrónomicamente importante (Malcomber y col., 2006). La forma adulta de las inflorescencias de las gramíneas condiciona la producción de granos, en particular el número, tamaño y calidad. La forma final de una inflorescencia es el resultado de un balance entre la capacidad de los meristemas axilares para formar ramas o terminar en espiguillas (que llevan flores). Actualmente en nuestro laboratorio estamos caracterizando factores de transcripción CYS2-HIS2 (C2H2) ZINC FINGER

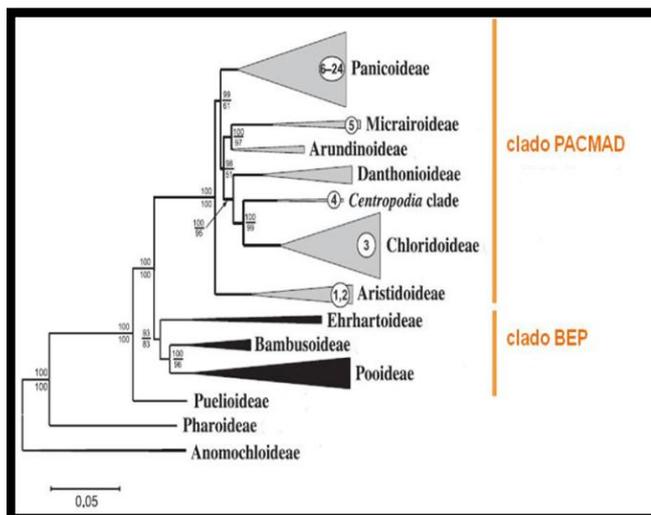


Figura 1: Filogenia de las gramíneas (Adaptado de GPWG II, 2012).

que parecen tener un rol preponderante en esta etapa clave en el desarrollo de las inflorescencias de las gramíneas (Reinheimer y col., en prep.). Estas proteínas contienen un motivo de unión al ADN compuesto por 30 aminoácidos e incluye dos residuos de Cys y dos His conservadas y próximas a los dedos de Zinc (Ciftci-Yilmaza y Mittler, 2008). La mayoría de las proteínas ZINC FINGER de plantas tienen un motivo invariante QALGGH en la hélice de dedos de zinc. Este motivo es crítico para el rol de la proteína (Ciftci-Yilmaza y Mittler, 2008). En particular, las proteínas ZINC FINGER del tipo C2H2 actúan como represores mediante el dominio EAR ubicado hacia el C-terminal (*ethylene-responsive element-binding factor (ERF)-associated amphiphilic repression*) (Ciftci-Yilmaza y Mittler, 2008).

Debido a la importancia de los factores de transcripción ZINC FINGER de tipo C2H2 en el desarrollo de las inflorescencias, nos propusimos explorar el grado de divergencia molecular de estos factores de transcripción en gramíneas.

## OBJETIVOS

Caracterizar el grado de divergencia molecular que existe entre los factores de transcripción C2H2 ZINC FINGER de gramíneas.

## METODOLOGÍA

Con el objetivo de estudiar la evolución molecular de proteínas del tipo C2H2 ZINC FINGER en plantas, reconstruimos un árbol filogenético con secuencias de la región codificante obtenidas de dos fuentes: (1) búsquedas BLAST en genomas de gramíneas y otras Angiospermas depositados Phytozome v.12 (<https://phytozome.jgi.doe.gov/pz/portal.html>) utilizando la región Cys2-His2 del dominio *zinc finger* y (2) secuencias adicionales generadas en nuestro laboratorio. Para ello, las plantas fueron cultivadas en las cámaras de crecimiento del Instituto de Agrobiotecnología del Litoral. Para las RT-PCR, el RNA fue extraído de inflorescencias usando RNazol@RT (Molecular Research Center) siguiendo las instrucciones de fabricación. Para las reacciones de RT se usó la enzima M-MLV (PBL) siguiendo el protocolo de fabricación. El oligo-dT se optimizó a 100µM. UBIQUITINA se usó como control positivo. Utilizando las secuencias disponibles en la web, se diseñaron oligos degenerados para amplificar la secuencia en otras especies de gramíneas usando el software Primacade (Gadberry y col. 2005). Los productos amplificados fueron subclonados en el vector pGEM-T (Promega) y verificados mediante secuenciación usando el servicio de MACROGEN Inc.

El dominio *zinc finger*, de aproximadamente 30 aminoácidos, de todas las secuencias obtenidas, se alineó usando el software MAFFT (Katoh et al., 2002). A partir de este alineamiento se reconstruyó la evolución molecular de todas las secuencias *zinc finger* obtenidas. Para ello, se empleó el método algorítmico de Cadena de Markov Monte Carlo implementado en MrBayes v.3.1.2 (Huelsenbeck y Ronquist, 2001) y el modelo GTR + G + I inferido en MrModeltest v.2.3 (Nylander, 2004) basado en el criterio Akaike. La convergencia y el tamaño efectivo de la muestra para cada réplica se verificó empleando el software Tracer v.1.5 (Rambaut y Drummond, 2007). Como resultado de este análisis obtuvimos un árbol dividido en dos grandes linajes. Uno de estos linajes está integrado por secuencias de gramíneas, al que decidimos llamar *GrassZFP*, hermanas a un clado formado por secuencias de monocotiledóneas-gramíneas y no-gramíneas y dicotiledóneas entre las que se destacan *SUP*, *RABBIT EARS (RBE)*, *ZINC-FINGER PROTEIN 10 y 11 (ZFP10, ZFP11)* de Arabidopsis. A

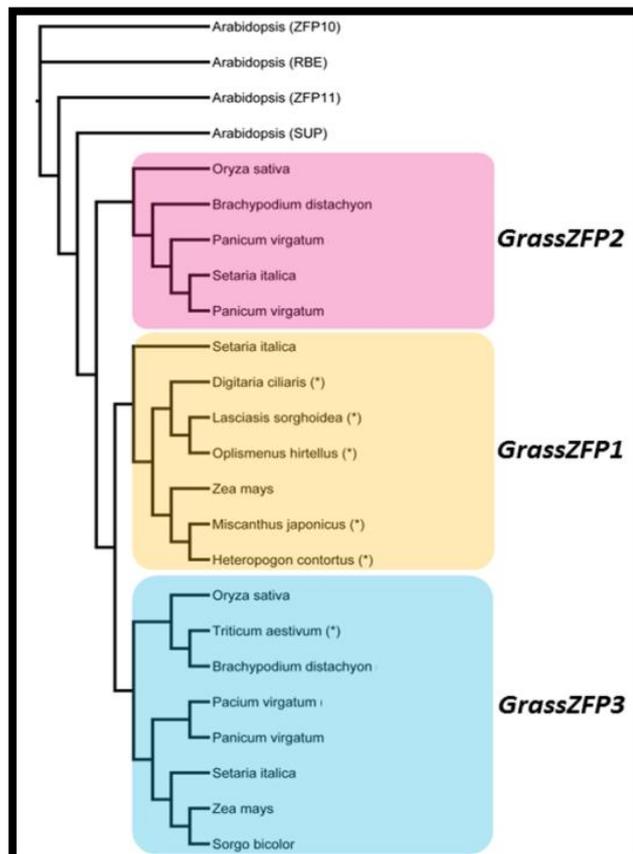
partir de estos resultados, decidimos enfocar nuestros estudios en los genes que codifican para las proteínas GrassZFP. Para ello, se generó un nuevo conjunto de datos con las secuencias que integran exclusivamente el clado *GrassZFP*. Las secuencias completas de esta nueva matriz fueron convertidas a secuencias peptídicas y alineadas manualmente en MEGA v.6.06 (Tamura et al., 2013). Luego, la matriz alineada se convirtió a secuencias de nucleótidos para su posterior análisis. Los árboles se reconstruyeron empleando el método algorítmico de Cadena de Markov Monte Carlo implementado en MrBayes v.3.1.2 (Huelsenbeck y Ronquist, 2001) y el modelo GTR + G + I inferido en MrModeltest v.2.3 (Nylander, 2004). Finalmente se reconstruyó un árbol de consenso de mayoría basado en 45002 árboles. Las secuencias de los genes que codifican para proteínas ZFP10, ZFP11, RBE y SUP se emplearon como secuencias de referencia y *outgroup*.

A su vez, para comenzar a explorar el rol de las proteínas GrassZFP, llevamos adelante RT-PCR semi-cuantitativas exploratorias, siguiendo el procedimiento explicado anteriormente, para determinar la preferencia de expresión a lo largo de toda la planta (raíz, tallo, hoja, inflorescencias en diversos estados de desarrollo). Particularmente, los estudios de expresión se llevaron a cabo en la gramínea *Setaria viridis*.

## RESULTADOS

La topología del árbol obtenida sugiere que proteínas GrassZFP son producto de dos duplicaciones sucedidas cuando se originaron las gramíneas. Como resultado de estas duplicaciones se generaron tres linajes de secuencias exclusivas de gramíneas, denominados arbitrariamente GrassZFP1-3. Varios datos surgen de analizar este árbol filogenético: (1) *GrassZFP1* parece ser exclusivo del clado PACMAD, (2) *GrassZFP* son genes exclusivos de gramíneas, (3) las duplicaciones y pérdida de algunas copias de *GrassZFP* sugieren una compleja evolución de estos genes durante la diversificación de las gramíneas (Figura 2).

Al analizar en forma comparativa la secuencias peptídicas de los genes que codifican para los factores de transcripción GRASSZFP, notamos que los homólogos de diversas especies de gramíneas del PACMAD difieren notablemente entre ellas, en especial en el número de motivos EAR. A su vez, identificamos motivos adicionales a



**Figura 2:** Topología del árbol consenso de mayoría de secuencias. Abreviatura: (\*), secuencias generadas en nuestro laboratorio.

los EAR que podrían ser de importancia funcional relevante.

Por último, como resultado de los patrones de expresión, se observa que, en particular, GrassZFP3 de *Setaria viridis* parece expresarse preferentemente en tejidos de inflorescencias que presentan ramas en expansión y aún no se han iniciado los órganos florales. La expresión, a su vez, decae en etapas tardías (durante la iniciación y diferenciación de los órganos florales).

## CONCLUSIÓN

Los resultados muestran que los GrassZFP son exclusivos de gramíneas y sugieren una compleja evolución de estas moléculas durante la diversificación de esta familia. Entender esta diversificación aportará datos valiosos que guiarán futuras investigaciones sobre los mecanismos que le confiere el carácter determinado a los meristemas axilares de gramíneas y su rol en la evolución del desarrollo de los sistemas de ramificación que llevan granos de especies cerealeras de interés.

## BIBLIOGRAFÍA BÁSICA

- Cheng, P.C., Greyson, R.I., Walden, D.B.**, 1983. Organ initiation and the development of unisexual flowers in the tassel and ear of Zea Mays. *Am. J. Bot.* 70: 450–462.
- Ciftci-Yilmaz, S. and Mittler, R.** 2008. The zinc finger network of plants. *Cell. Mol. Life Sci.* 65: 1150 – 1160
- Clifford, H.T.**, 1987. Spikelet and floral morphology. In *Grass Systematics and Evolution*, T.R., Soderstrom, K.W., Hilu, C.S., Campbell, and M.E., Barkworth, eds (Washington, DC: Smithsonian Institute), pp. 21–30.
- Gadberry M. D., Malcomber S. T., Doust A. N., Kellogg E. A.**, 2005. Primaclade - a flexible tool to find primers across multiple species. *Bioinformatics* 21: 1263-1264
- Grass Phylogeny Working Group II.**, 2012. New grass phylogeny resolves deep evolutionary relationships and discovers C4 origins. *New Phytologist* 193: 304-312.
- Huelsenbeck J.P., Ronquist F.**, 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754-755.
- Ikeda, K., Sunohara, H., and Nagato, Y.**, 2004. Developmental course of inflorescence and spikelet in rice. *Breeding Sci.* 54: 147–156.
- Katoh, K., Misawa, K., Kuma, K., Miyata, T.**, 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, Vol. 30 N1. 14: 3059-3066.
- Malcomber, S.T., Preston, J.C., Reinheimer, R., Kossuth, J., and Kellogg E.A.**, 2006. Developmental gene evolution and the origin of grass inflorescence diversity. *Adv. Bot. Res.* 44: 426–481.
- Nylander J.A.A.**, 2004. MrModelTest 2.0. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
- Rambaut A, Drummond AJ.** 2007. Tracer v1.4. Available from <http://beast.bio.ed.ac.uk/Tracer>
- Tamura K., Stecher G., Peterson D., Filipski A., Kumar S.**, 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* 30: 2725-2729.