



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

**Implicaciones epigenéticas de las regiones
no codificantes de los factores de
transcripción MADS-box en *Arabidopsis
thaliana***

T E S I S

Que para obtener el título de

BIÓLOGO

P R E S E N T A

José María Olvera Herrera

DIRECTORA DE TESIS

Dra. Ma. De La Paz Sánchez Jiménez



Ciudad Universitaria, Cd. Mx., Julio 2022



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

1. Datos del alumno

Apellido paterno
Apellido materno
Nombre(s)
Teléfono
Universidad Nacional Autónoma de México
Facultad de Ciencias
Carrera
Número de cuenta

2. Datos del tutor

Grado
Nombre(s)
Apellido paterno
Apellido materno

3. Datos del sinodal 1

Grado
Nombre(s)
Apellido paterno
Apellido materno

4. Datos del sinodal 2

Grado
Nombre(s)
Apellido paterno
Apellido materno

5. Datos del sinodal 3

Grado
Nombre(s)
Apellido paterno
Apellido materno

6. Datos del sinodal 4

Grado
Nombre(s)
Apellido paterno
Apellido materno

7. Datos del trabajo escrito.

Título
Subtítulo
Número de páginas
Año

1. Datos del alumno

Olvera
Herrera
José María
311 74 62 751
Universidad Nacional Autónoma de México
Facultad de Ciencias
Biología
417008240

2. Datos del tutor

Dra
Ma. de la Paz
Sánchez
Jiménez

3. Datos del sinodal 1

Dra
Diana Belén
Sánchez
Rodríguez

4. Datos del sinodal 2

Dr
Rodrigo
González
Barrios de la Parra

5. Datos del sinodal 3

Dr.
Javier Andrés
Juárez
Díaz

6. Datos del sinodal 4

Dr.
Marco Tulio
Solano
De la Cruz

7. Datos del trabajo escrito

Implicaciones epigenéticas de las regiones no
codificantes de factores de transcripción MADS-box en
Arabidopsis thaliana
66 p
2022

Agradecimientos académicos

Esta tesis se realizó bajo la dirección de la Dra. Ma de la Paz Sánchez Jiménez, en el Laboratorio de Genética Molecular, Epigenética, Desarrollo y Evolución de Plantas del Instituto de Ecología, de la Universidad Nacional Autónoma de México, bajo la coordinación académico-científica de las Dras. Elena Álvarez-Buylla Roces y Adriana Garay Arroyo así como la coordinación administrativa y logística de la Dra. Diana Belén Sánchez Rodríguez.

Además, quiero agradecer al Dr. Marco Tulio Solano de la Cruz por su apoyo en enseñarme a utilizar las herramientas para la búsqueda de motivos que utilicé en este trabajo. También a la Mtra. Laura Rodríguez Casillas y al Dr. Crisanto Gutiérrez Armenta, del Centro De Biología Molecular Severo Ochoa, por su apoyo en la elaboración de Metaplots y Heatmaps. Por último, quisiera agradecer por su tiempo a los miembros del jurado al revisar este trabajo.

El financiamiento para esta investigación provino de proyectos de investigación:

UNAM-DGAPA-PAPIIT: IN203220, IN200920, IN1211721.

CONACyT: CF-2019-102987, CF-2019-102959.

Agradecimientos personales

Primero que nada, gracias a mi mamá, a mi papá y a mi hermana por darme todo el apoyo y amor necesarios para insistir tanto en entrar a la carrera, para terminarla y para escribir esta tesis. Me han inspirado, no me han dejado rendirme, y no creo poder nunca agradecerles lo suficiente. Me hace feliz sentir que puedo verlos a los tres en mis propios ojos.

Hay muchas personas que agradecer para este trabajo tan pequeño y me encuentro sin saber qué escribir en los últimos minutos antes de subir el trabajo a TesiUNAM. Gracias a mis amix del poli, aquellxs del inicio de la carrera, y lxs Mecopteros. Gracias a Ingrid, alguna vez tú dijiste que conocernos fue serendipia y creo que no podría estar más de acuerdo. A Carlos, gracias por siempre estar ahí, gracias por apoyarme, gracias por alejarme del trabajo, y gracias por darme la confianza necesaria para ser quien soy. Paloma, muchas gracias por la diversión, por el chisme, por la distracción, por el buen gusto, y por enseñarme que es posible el balance entre ser bonita y una gran 100tifika. Palma, gracias por las recomendaciones de anime, por las idas a comprar plantas y ver animales, las prácticas de campo improvisadas y las partidas de smash. Pocas veces puede alguien como yo tener la suerte de conocer a un grupo de personas tan especiales como ustedes, gracias.

En el lab, gracias a todxs mis compañerxs, probablemente hubiera durado menos en este lab si no hubiera sido porque me encontré con personas tan cool como ustedes.

Y nada, soy escorpio, así que escribo maldiciones especiales a las personas que hicieron de este trámite un infierno, y a aquellas que desmotivan a sus alumnxs de salir adelante.

Resumen

Los mecanismos epigenéticos participan en diferentes procesos del desarrollo de las plantas, siendo fundamentales para los procesos de morfogénesis. Los complejos epigenéticos Polycomb (PcG) y Trithorax (TrxG) regulan la expresión genética a través de la remodelación de la cromatina, mientras que PcG reprime la transcripción, TrxG regula la activación transcripcional; ambos complejos participan en procesos de proliferación y diferenciación celular, reprogramación celular, en la respuesta a estrés y en la generación de memoria epigenética. La represión por PcG se da principalmente a través de la deposición de la marca de trimetilación de la lisina 27 de la histona H3 (H3K27me3), la cual es catalizada por el complejo PRC2 y la ubiquitinación de la lisina 121 de la histona H2A (H2AK121ub) catalizada por el complejo PRC1. De esta manera, PcG regula la represión transcripcional de un gran número de genes homeóticos, dentro de los cuales están algunos de los factores de transcripción tipo MADS-box. En *Arabidopsis thaliana* se han descrito más de 100 genes tipo MADS-box y se han dividido en dos tipos, el tipo II agrupa a aquellos cuyas proteínas del dominio MADS tienen una estructura de dominio modular conocida como MIKC. La gran mayoría de los genes MIKC tienen como característica que su estructura incluye un intrón largo (<500 pb), que se predice contiene elementos importantes para su propia regulación transcripcional, como se ha descrito para el gen *FLOWERING LOCUS C (FLC)*. Para que PcG reprima a *FLC* mediante la deposición de la marca de H3K27me3, se requiere un recambio de la variante de histona H3.3 por la histona canónica H3.1, ya que se ha reportado que el enriquecimiento de la variante H3.1 es necesario para mantener un estado transcripcionalmente silenciado. A pesar de que la diferencia entre la histona canónica H3.1 y la histona H3.3 solo consiste en la sustitución de cuatro aminoácidos, es suficiente para que específicamente se monometile la lisina 27 de la H3.1, marca que sirve de sustrato; para que posteriormente, sea trimetilada por PcG. Además, en esta regulación se requiere de la unión de los factores de transcripción VAL1 y VAL2 (VIVIPAROUS1/ABI3-LIKE) a motivos RY encontrados dentro del intrón largo de *FLC*; para que PcG se asocie a este locus. Debido a que un gran número de genes MIKC tiene un intrón largo, es interesante conocer si conservan los mismos elementos y mecanismos similares de regulación epigenética. Por lo tanto, en esta tesis se revisó, a través de análisis *in silico*, el papel de las regiones no codificantes de los factores de transcripción MADS-box tipo MIKC en su regulación epigenética. Específicamente, la influencia de motivos RY en la represión mediada por la maquinaria de PcG, dentro de los intrones largos de dichos factores de transcripción. Se encontró una relación entre niveles altos de marca de H3K27me3, en aquellos genes MADS-box con dos o más motivos RY dentro de su intrón largo, así como niveles mayores de distribución de la histona H3.1 respecto a la variante de histonas H3.3.

ABREVIATURAS Y CONTRACCIONES ÚTILES

CLF	CURLY LEAF
Col-0	Ecotipo Columbia
CME	Cold Memory Element
FLC	FLOWERING LOCUS C
H2AK121ub	Monoubiquitinación de la lisina 121 de la Histona H2A
H3K27me3	Trimetilación de la lisina 27 de la Histona H3
H3K36me3	Trimetilación de la lisina 36 de la Histona H3
H3K4me3	Trimetilación de la lisina 4 de la Histona H3
MSI1	MULTICOPY SUPPRESSOR OF IRA1 VERNALIZATION (VRN2), EMBRYONIC
pb	Pares de bases
PcG	Polycomb
SWN	SWINGER
TrxG	Trithorax
VAL1/VAL2	VIVIPAROUS1/ABI3-LIKE

Índice:

Introducción	8
Arabidopsis como modelo de estudio	9
Regulación epigenética	10
Variantes de histonas.....	12
Características de la histona H3.1.....	13
Características de la variante H3.3	14
Modificaciones postraduccionales de histonas	15
Activación transcripcional por Trithorax	17
Represión transcripcional por Polycomb.....	19
Factores de transcripción MADS-box	21
Función de los genes MADS-box en el desarrollo	23
Papel de los MADS-box en el desarrollo de la raíz.....	24
Importancia del intrón largo en la regulación de un MADS-box <i>FLC</i>	25
Regulación epigenética de <i>FLC</i>	26
Hipótesis:	31
Objetivos:	31
Materiales y métodos	32
Resultados:	36
La expresión de los genes MADS-box tipo MIKC tienen un patrón espacial y temporal diferencial en los diferentes tejidos de Arabidopsis.	36
Los motivos RY están conservados en la mayoría de los intrones largos de los MADS-box expresados en raíz.	38
Una aparente relación se encuentra entre la cantidad de motivos RY dentro del intrón largo, la deposición diferencial de variantes y las marcas de histonas.	40
Discusión:.....	49
Conclusiones.....	56
Anexos	58
Referencias:	59

Introducción

La morfología que podemos apreciar en todas las plantas sigue un patrón de control genético básico, y la manera en que éste se coordina, da lugar a una extensa cantidad de formas y arquitecturas en plantas (McGarry & Ayre, 2012; Poethig, 2003). Además, los eventos de duplicación del genoma a lo largo de su evolución han permitido cambios increíbles en su complejidad biológica (Della Pina et al., 2014; Cezary Smaczniak et al., 2012). De hecho, los eventos de duplicación genética antes de la diversificación funcional han sido fenómenos generalizados con familias de genes que se encargan del desarrollo en organismos (Nam et al., 2003). A diferencia de conceptos anteriores, la amplia diversidad de formas que vemos en las plantas no se debe a la existencia de un gran número de genes que dan lugar a la formación de estructuras particulares, sino a la conservación de genes importantes para ciertos procesos, así como a su regulación (Della Pina et al., 2014). Estos conceptos han ayudado a pasar de una perspectiva en la biología centrada en los genes, a una que se centra en el genoma (Riechmann & Ratcliffe, 2000) y su regulación epigenética.

Estas nuevas perspectivas proveen de condiciones propicias para realizar estudios sobre aquellas proteínas (*i.e.* factores de transcripción) que regulan la expresión génica a través de la interacción específica con elementos en *cis* de los genes (Gonzalez, 2015), y de proteínas (factores epigenéticos) que regulan la expresión de los genes a través de regular la estructura de la cromatina (Lyle, 2014).

Los estudios en plantas han brindado contribuciones fundamentales tanto en el campo de la genética como en el de la epigenética. El ejemplo más famoso es el de las leyes de la herencia, basadas en los trabajos de Mendel con plantas de chícharos (revisado en Gayon, 2016). Además, fue en *Marchantia polymorpha* en donde se describió por primera vez la distinción entre eucromatina y heterocromatina (Heitz, 1929), así como la impronta parental, que fue observada por primera vez en *Zea mays* (Arteaga-Vazquez & Chandler, 2010). Características distintivas del desarrollo y morfología de plantas las hacen únicas en el estudio de procesos epigenéticos. Debido a que la línea germinal no se define durante la embriogénesis, si no hasta etapas posteriores del desarrollo, las modificaciones epigenéticas pueden ser adquiridas por células

meristemáticas en respuesta a la interacción con el ambiente (Pikaard & Scheid, 2014). Permitiendo que dichas células tengan el potencial de transmitir las nuevas características a las células germinales y a generaciones posteriores (Pikaard & Scheid, 2014). *Arabidopsis thaliana* (*Arabidopsis*), es una de las plantas mejor estudiadas que cuenta con mayores herramientas para los estudios genéticos y epigenéticos, por lo que es la especie estudiada en este proyecto.

Arabidopsis como modelo de estudio

Arabidopsis pertenece a la familia Brassicaceae, y las características que hicieron a esta especie una especie modelo involucran tanto rasgos de desarrollo, como estructurales y genéticos. En su madurez, *Arabidopsis* presenta hojas pequeñas arregladas en roseta, con un tallo que presenta un escapo floral. Llega a medir de 30 a 40 cm de altura, aunque el tamaño depende de las condiciones de crecimiento. Esto permite su fácil cultivo en espacios pequeños. Sus flores miden 3 mm de largo, y consisten en cuatro sépalos alternados a cuatro pétalos. Después del verticilo de pétalos se encuentran seis estambres y, justo en el centro, un ovario con dos carpelos. Sus flores normalmente son auto fecundadas antes de la antesis, lo cual dificulta su fecundación cruzada, una ventaja para su cultivo en laboratorio. Esto facilita la propagación e identificación de ejemplares mutantes homocigotos entre la progenie de una única planta mutagenizada. Sus frutos (silicuas) contienen de 30-60 semillas en su madurez, y pueden llegar a sumar miles en un solo ejemplar (Bowman, 2012). El tiempo de floración depende de un gran número de factores (Andrés & Coupland, 2012; Meyerowitz, 1987), por ejemplo, bajo condiciones de luz continua, a 22 °C, el ecotipo Columbia tarda alrededor de cuatro semanas en florecer. Debido a su ciclo de vida corto, es fácil obtener un gran número de generaciones en poco tiempo (Meyerowitz, 1994).

La raíz es generada a partir de un pequeño número de células troncales (también llamadas iniciales) localizadas en la punta de la raíz. De manera interna a dichas células, y conectándolas entre sí, se encuentra un pequeño número de células centrales con baja actividad mitótica conocidas como células del Centro Quiescente y que dan origen a las células troncales (Scheres et al., 2002).

Se han descrito cuatro tipos de células troncales que dan origen a todos los tipos celulares en la raíz: 1) las células iniciales de la epidermis/cofia lateral que, dan lugar a la epidermis y a la porción externa de la cofia, conocida como la cofia lateral; 2) las células iniciales de columela que se diferencian para formar las células de columela; 3) las células iniciales del córtex/endodermis que dan lugar al tejido fundamental; y, por último, 4) el tejido vascular y el periciclo que cuentan también con sus propias células iniciales. Esta organización celular relativamente simple hace a la raíz un sistema útil para el estudio epigenético del desarrollo en plantas.

El empleo de *A. thaliana* como especie modelo para investigaciones epigenéticas en plantas, ha logrado una fusión de estudios clásicos y moleculares de la genética con el desarrollo de plantas (Meyerowitz, 2001). Esta transición también ha logrado pasar de una biología del desarrollo descriptiva a implementar análisis que buscan las causas que llevan a un desarrollo particular (Meyerowitz, 2001). Una estrategia que ha cambiado y permitido un gran avance para elucidar la manera en que se regula el desarrollo en plantas ha sido la generación de plantas mutantes. Las líneas de plantas mutantes en un solo gen (mutantes simples) han sido parte de la tradición descriptiva de la biología del desarrollo, sin embargo, el uso de líneas mutantes en distintos genes a la vez ha dado paso a la realización de análisis de epistasis (Meyerowitz, 1994). El genoma nuclear diploide de *Arabidopsis* es muy pequeño, y consta de alrededor de 140,000 kb, en cinco cromosomas. Por lo tanto, la realización de librerías de fragmentos cromosomales es relativamente fácil, que facilita la realización de ciertos procedimientos de clonación (Bevan & Walsh, 2005; Meyerowitz, 1987)(Meyerowitz, 1987). Así, *Arabidopsis* se ha convertido en un modelo para la investigación epigenética debido a la disponibilidad de plantas mutantes casi para cada uno de sus genes, a su amplia variación natural, a la información sobre la secuencia del genoma de numerosas accesiones, y a una gran cantidad de datos disponibles de expresión y de marcas de cromatina a lo largo del genoma (Pikaard & Scheid, 2014).

Regulación epigenética

La información genética se encuentra organizada, dentro del núcleo en células eucariontes, en una estructura altamente conservada conocida como cromatina (Fischle et al., 2003). La

cromatina, es capaz de reordenarse al momento de percibir señales tanto internas como externas en procesos como la replicación del DNA, el ciclo celular, eventos de recombinación, la coordinación espacial y temporal de la expresión de genes para el control homeostático, desarrollo bajo condiciones de estrés, y la reparación del DNA (Fischle et al., 2003; Kim et al., 2015). Se distinguen dos organizaciones principales de la cromatina: eucromatina, relajada y transcripcionalmente activa, y heterocromatina, compacta y transcripcionalmente inactiva. La mayoría de los genes se encuentran localizados en regiones de eucromatina, donde pueden ser transcritos, mientras que las regiones de heterocromatina están enriquecidas en regiones que deben de permanecer transcripcionalmente silenciadas (e.g. regiones con elementos transponibles, centrómeros y telómeros) (Doğan & Liu, 2018).

La unidad fundamental de la cromatina es el nucleosoma. Este consiste en 147 pares de bases (pb) de DNA enrolladas alrededor de un octámero de proteínas llamadas histonas, formado por dos subunidades de H2A, dos de H2B, dos de H3 y dos de H4 (**Fig. 1**). En este nivel de empaquetamiento, la cromatina forma un collar de perlas (Fischle et al., 2003; Luger et al., 1997). A su vez, las histonas de clase 1 (H1) se asocian al DNA entre los nucleosomas simples para establecer el siguiente nivel de organización, llamado solenoide (Doğan & Liu, 2018; Fischle et al., 2003; Luger et al., 1997). La conformación de la cromatina no es simplemente una herramienta de compactación, sino que se ajusta de forma dinámica ante señales necesarias para programar las vías celulares adecuadas (Margueron & Reinberg, 2010), a través de la regulación de la expresión génica. Este reordenamiento de la cromatina y su asociación con la regulación de la expresión génica ha llevado a definir a la epigenética como el estudio de aquellos mecanismos que provocan cambios en la expresión de los genes, que son heredados de forma mitótica y/o meiótica, y que no involucran un cambio en la secuencia de DNA (Sanchez et al., 2015). Entre dichos mecanismos se encuentran la metilación del DNA, la regulación por RNAs no codificantes, y modificaciones postraduccionales de histonas (**Fig. 1**) (Margueron & Reinberg, 2010).

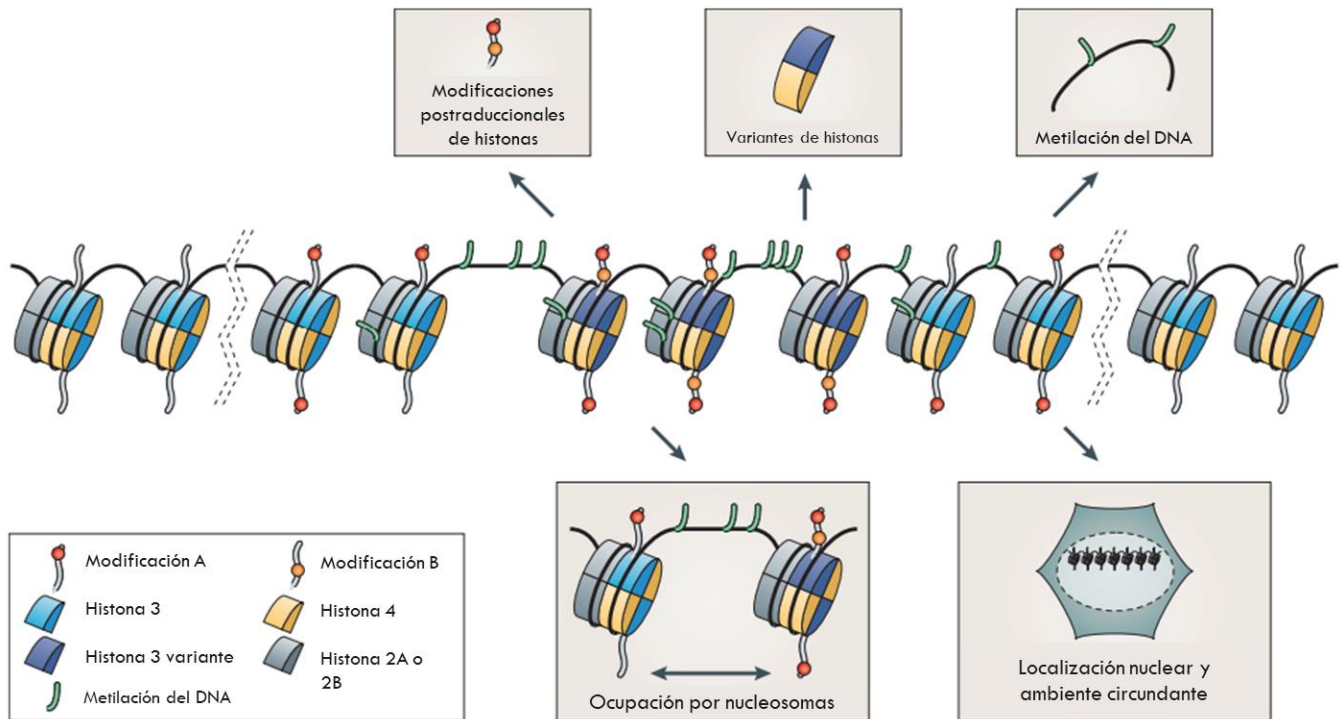


Fig. 1. Esquema tomado de Margueron & Reinberg (2010), sobre la composición de los nucleosomas, así como las modificaciones y condiciones que definen a los distintos estados de la cromatina.

Las modificaciones postraduccionales de las histonas tienen un impacto sobre la expresión génica, no solo por la combinación de modificaciones postraduccionales que pueda tener un nucleosoma, sino por las variantes de histonas que conforman los nucleosomas. Esto es debido a que algunas variantes de histonas determinan el tipo de modificaciones postraduccionales que puedan ocurrir en las mismas (Fischle et al., 2003). Para entender la importancia que tienen las variantes de histonas sobre la regulación epigenética mediada por las modificaciones postraduccionales de las histonas, es necesario primero conocer sus variantes.

Variantes de histonas

Las histonas están altamente conservadas entre los organismos eucariontes, sin embargo, existen variantes no alélicas de las histonas H2A, H3 y H1 con modificaciones como sustituciones de aminoácidos y/o cambios en la cantidad de intrones (Foroozani et al., 2022) (**Fig. 2**). El genoma

de *Arabidopsis* codifica para cuatro tipos principales de variantes de H2A: H2A, H2A.X, H2A.Z y H2A.W. Por otro lado, la familia H3 contiene tres tipos principales de variantes: H3.1, H3.3, y CenH3 (Jiang & Berger, 2017b), siendo CenH3 la variante más divergente del resto, y, como su nombre lo indica, es depositada en centrómeros, siendo esencial para la división nuclear (Ravi et al., 2011). En este trabajo me enfocaré en las variantes H3.1 y H3.3, las cuales únicamente difieren en cuatro aminoácidos en las posiciones 31, 41, 87, y 90 (Cui et al., 2015; Foroozani et al., 2022; Jiang & Berger, 2017b). Estas diferencias crean un sesgo preferencial entre H3.1 y H3.3 sobre sus modificaciones postraduccionales en los dominios amino-terminales (Cui et al., 2015; Jiang & Berger, 2017b).

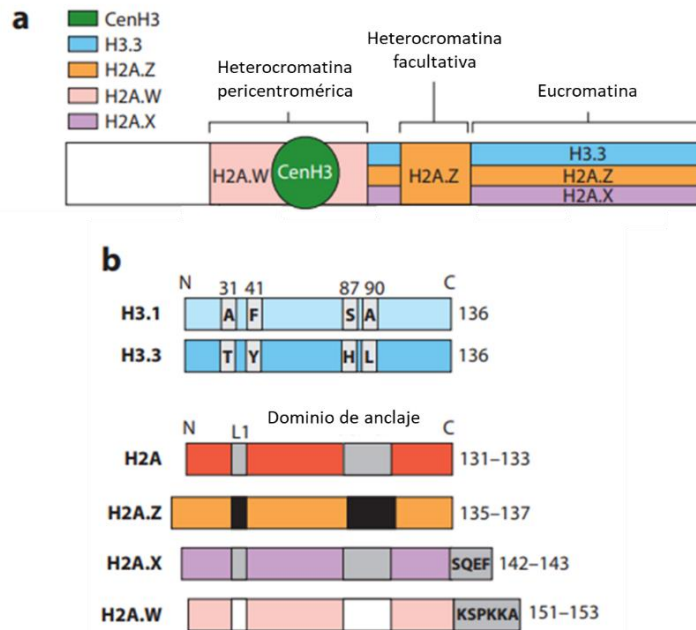


Fig. 2. Esquema modificado de Foroozani et al. (2022), sobre las variantes de histonas y su distribución en la cromatina de plantas. A) Diagrama de un cromosoma, mostrando la distribución de los principales estados de la cromatina y las variantes de histonas asociado con cada uno. B) comparación entre las variantes de histonas y sus contrapartes canónicas. Las regiones en las secuencias donde existen diferencias entre variantes y tipos canónicos se resaltan en forma de cajas.

Características de la histona H3.1

La histona H3.1 se caracteriza por presentar las siguientes sustituciones de aminoácidos respecto a la variante H3.3: Ala³¹, Phe⁴¹, Ser⁸⁷ y Ala⁹⁰ (**Fig. 2b**) (Jiang & Berger, 2017b). Su

expresión es dependiente de la replicación, tiene un pico máximo antes de la fase S del ciclo celular y, posteriormente, es reprimida al completar la replicación del DNA (Cui et al., 2015). La deposición de esta variante es mediada por el complejo Chromatin Assembly Factor 1 (CAF1), que trabaja en conjunto con la maquinaria de replicación, siendo depositada mayormente durante la fase S del ciclo celular (Jiang & Berger, 2017b; Ramirez-Parra & Gutierrez, 2007). Se ha demostrado que la deposición de H3.1 en el proceso de replicación del DNA facilita el restablecimiento de diferentes niveles en las marcas de represión de la histona H3 (*i.e.* H3K27me3: trimetilación de la lisina 27 en la histona H3) tras la deposición de esta variante. La alanina en la posición 31 (Ala³¹) de la variante H3.1 permite que su lisina 27 sea monometilada, gracias a la actividad de las metiltransferasas ARABIDOPSIS TRITHORAX-RELATED PROTEIN5 (ATXR5) y ATXR6; esto provoca que la marca de H3K27me3 disminuya drásticamente durante este periodo pero que sus niveles sean restaurados posteriormente durante la fase G2 del ciclo celular (Jiang & Berger, 2017a). Se ha observado que tanto la eucromatina como la heterocromatina están marcadas por variantes distintas de histonas (Jiang & Berger, 2017b). La variante H3.1 ha sido asociada a regiones transcripcionalmente reprimidas tanto para eucromatina como para heterocromatina (regiones pericentroméricas). Su distribución en ambos tipos de cromatina, sugiere que cumple una función de soporte estático a la cromatina (Jiang & Berger, 2017b; Stroud et al., 2012; Wollmann et al., 2012).

Características de la variante H3.3

La variante de histonas H3.3 se caracteriza por tener las siguientes sustituciones de aminoácidos respecto a la variante H3.1 : Thr³¹, Tyr⁴¹, His⁸⁷ y Leu⁹⁰ (**Fig. 2b**) (Jiang & Berger, 2017b). A diferencia de la variante H3.1, la deposición de la variante H3.3 es independiente de la replicación. Este tipo de variante se mantiene expresada constitutivamente a lo largo del ciclo celular (Cui et al., 2015). Su deposición es mediada por el complejo *Histone regulator A* (HIRA) (Nie et al., 2014), y su patrón de localización en genes se correlaciona con la localización de la RNAPII (RNA Polimerasa II), en los sitios de inicio de la transcripción pero también presenta picos hacia el sitio de término de la transcripción (TTS) de los genes (Jiang & Berger, 2017b). Se piensa que, como histona de reemplazo, H3.3 es ensamblada en el nucleosoma después de la elongación

por la RNAPII, que activamente desorganiza nucleosomas durante la transcripción. H3.3 se encuentra más relacionada a regiones transcripcionalmente activas (eucromatina), y es principalmente enriquecida en el extremo 3' de los genes (Stroud *et al.*, 2012). De esta manera, la represión transcripcional está asociada a la reducción en los niveles de H3.3 cerca del TTS de los genes, mientras que la activación transcripcional se relaciona con el aumento de H3.3 en el TTS (Jiang & Berger, 2017b; Probst *et al.*, 2020). Esto fue observado en el proceso de desarrollo del meristemo apical aéreo (SAM, por sus siglas en inglés) y primordios de hojas (Wollmann *et al.*, 2012). Por otro lado, el cambio en la proporción de H3.1/H3.3 a lo largo del eje longitudinal de la raíz de *Arabidopsis* ha sido asociado a distintas tasas de proliferación celular (Otero *et al.*, 2016). En células del meristemo apical de la raíz (RAM, por sus siglas en inglés), durante la fase G2 del ciclo celular, antes de entrar a la etapa de diferenciación, ocurre un intercambio importante de la variante H3.1 por la H3.3 de manera general en la cromatina (Otero *et al.*, 2016). Dicha dinámica de reemplazo entre las variantes de histonas sugiere que presentan una función importante en nuevos programas de expresión durante el desarrollo y diferenciación, probablemente facilitando cambios globales en la estructura de la cromatina y modificaciones de histonas (Otero *et al.*, 2016; Stroud *et al.*, 2012; Wollmann *et al.*, 2012).

La treonina en la posición 31 (Thr³¹) de H3.3 inhibe la actividad de ATRX5 y ATRX6, mientras que la Ala³¹ de H3.1 no. El evadir estas marcas de represión podría contribuir a la protección de regiones transcripcionalmente activas, ocupadas por H3.3, contra el silenciamiento (Jiang & Berger, 2017b).

Modificaciones postraduccionales de histonas

Las histonas centrales (H2A, H2B, H3, H4) presentan dominios amino-terminales flexibles que sobresalen hacia la parte externa del nucleosoma, así como dominios globulares carboxilo-terminales que forman el andamio de los nucleosomas (Lyle, 2014). Se ha demostrado, que los dominios amino-terminales de las histonas H3 y H4 juegan un papel importante en la estabilización de un estado compacto de la cromatina mediando interacciones de atracción internucleosomal. Esto indica que, las modificaciones químicas en los residuos de histonas

pueden alterar la estructura de la cromatina para regular la transcripción de genes (Arya & Schlick, 2009).

Las modificaciones postraduccionales de las histonas en sus dominios amino-terminales se producen en residuos de lisina (K) que se pueden acetilar, metilar y ubiquitinar; en residuos de serina (S) y treonina (T) que pueden ser fosforilados; y en residuos de arginina (R) que se pueden metilar (**Fig. 3**) (Fischle et al., 2003; Kim et al., 2015). A pesar de que las modificaciones postraduccionales de histonas se han correlacionado con funciones definidas, es en realidad la combinación de dichas marcas postraduccionales ya sea en una histona, un nucleosoma, o dominios de nucleosomas, las cuales establecen patrones globales de modificaciones a la cromatina que pueden especificar funciones únicas (Fischle et al., 2003; Kim et al., 2015; Margueron & Reinberg, 2010).

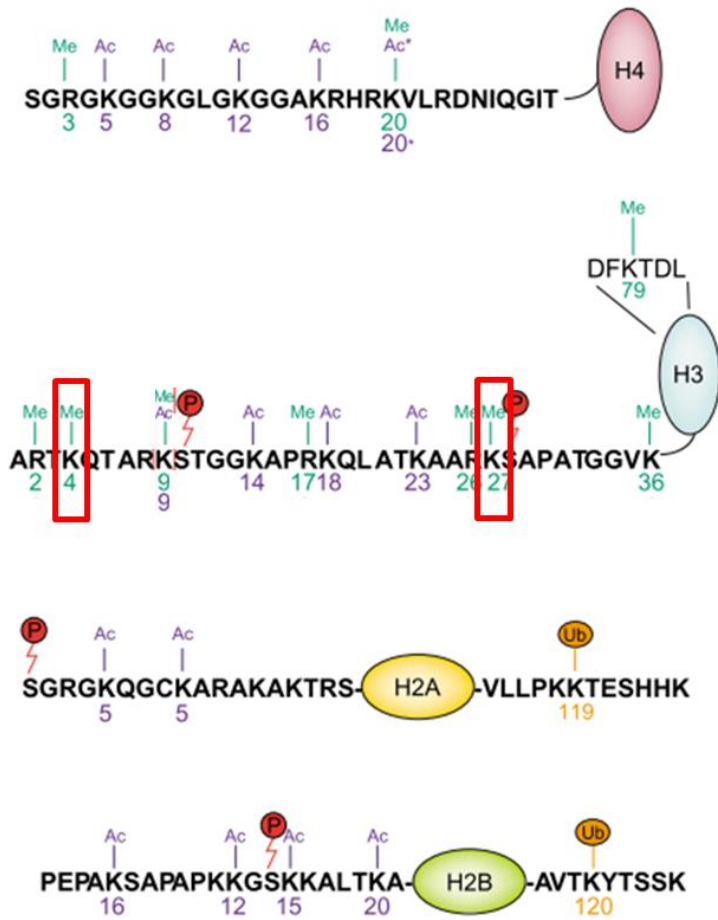


Fig. 3. Esquema modificado de Loidl (2004), que muestra las modificaciones postraduccionales que las histonas pueden sufrir en sus dominios amino-terminales. Las modificaciones que se muestran son acetilación (Ac, morado), metilación (Me, verde), fosforilación (P, rojo) y ubiquitinación (Ub, naranja). En recuadros rojos se resaltan las marcas de H3K4me3 y H3K27me3, depositados por TrxG y PcG respectivamente.

Activación transcripcional por Trithorax

Las proteínas TrxG son aquellas que presentan funciones antagonistas a las de las proteínas PcG, siendo capaces de excluir la represión por PcG, por lo que es asociada a la activación transcripcional (Fletcher, 2017). Se han identificado un gran número de proteínas con la función de TrxG en Arabidopsis. Algunas de las proteínas más estudiadas son: ARABIDOPSIS TRITHORAX1 (ATX1) y SET DOMAIN GROUP2 (SDG2), ambas metiltransferasas; así como la proteína de unión a DNA con dominio SAND, ULTRAPETALA1 (ULT1) (Köhler & Hennig, 2010; Ornelas-Ayala et al., 2021; Sanchez et al., 2015). El mecanismo de función de las proteínas TrxG no ha sido bien descrito hasta ahora, sin embargo, se ha descrito su importancia para diversos procesos en el desarrollo de Arabidopsis.

Las proteínas ATX1 y SDG2 tienen actividad metiltransferasa, y establecen la trimetilación de la histona H3 en la lisina 4 (H3K4me3), una de las modificaciones de histonas asociadas a regiones transcripcionalmente activas (Alvarez-Venegas et al., 2003). De las metiltransferasas, SDG2 ha sido reportada como la principal responsable de la deposición de la marca de H3K4me3 en *Arabidopsis* (Guo et al., 2010). Por otro lado, los dominios SAND son capaces de unirse al DNA, y se ven involucrados en la regulación transcripcional asociada a la cromatina. De esta manera, es posible que ULT1 reclute a ATX1 a sitios específicos para establecer la marca de H3K4me3 y eventualmente llevar a cabo la activación transcripcional (Xu et al., 2018). Esto ha sido reportado en la parte aérea de *Arabidopsis*, donde ATX1 regula la transcripción de factores de transcripción MADS-box, tales como *APETALA1* (*AP1*), *AGAMOUS* (*AG*), *PISTILATA* (*PI*) y *APETALA3* (*AP3*), durante la transición a la floración y la determinación de meristemas florales (Alvarez-Venegas et al., 2003; Carles & Fletcher, 2009).

Con respecto a ULT1, está reportada su función independiente de ATX1 en la organización del nicho de células troncales (SCN, por sus siglas en inglés) del RAM y en la regulación de genes involucrados en la señalización de auxinas en la raíz (Ornelas-Ayala et al., 2020). Debido a que SDG2 también se ha descrito como importante para el mantenimiento del SCN en la raíz principal, así como en el establecimiento del SCN en raíces laterales (Yao et al., 2013), se cree que podría encontrarse formando un complejo con ULT1 para llevar a cabo sus funciones (Guo et al., 2010; Ornelas-Ayala et al., 2020). Además, ha sido comprobada la interacción de ULT1 con EMF1, una proteína componente del subcomplejo PRC1 de PcG (Xu et al., 2018). Se ha propuesto que la interacción de ULT1 con ATX1 promueve la formación de un complejo con el subcomplejo PRC2 (Polycomb Repressive Complex2) de PcG a través de EMF1, para mantener las marcas de H3K27me3 de represión (Xu et al., 2018). Esto podría indicar que el resultado transcripcional final en genes involucrados en el desarrollo depende del balance entre la acción de las proteínas PcG y TrxG junto con activadores transcripcionales (Köhler & Hennig, 2010; Ornelas-Ayala et al., 2021).

Represión transcripcional por Polycomb

La represión de la transcripción por PcG se da a través de la trimetilación de la lisina 27 de la histona H3 (H3K27me3), la cual es catalizada por el subcomplejo PRC2; y de la ubiquitinación de la lisina 121 de la histona H2A (H2AK121Ub), que es catalizada por el subcomplejo PRC1 (Polycomb Repressive Complex1) (Kralemann et al., 2020; Sanchez et al., 2015). El complejo PRC1 está conformado por las enzimas (ubiquitintransferasas) AtRING1a/b y AtBMIa/b/c, y por las proteínas LIKE HETEROCHROMATIN PROTEIN 1 (LHP1) y EMBRYONIC FLOWER1 (EMF1) que reconocen la marca de H3K27me3 (Molitor & Shen, 2013). Sin embargo, debido a la interacción que se ha probado de LHP1 y EMF1 con proteínas de PRC2, recientemente, algunos autores sugieren que estas dos proteínas forman parte de PRC2 (Baile et al., 2022).

El subcomplejo PRC2, está conformado por al menos una de las tres histonas metiltransferasas CURLY LEAF (CLF), SWN (SWINGER) y MEDEA (MEA) descritas en plantas; y por las proteínas accesorias FERTILIZATION INDEPENDENT ENDOSPERM (FIE) y MULTICOPY SUPPRESSOR OF IRA1 (MSI1) que conforman el núcleo del complejo, las cuales pueden interactuar con VERNALIZATION (VRN2), EMBRYONIC FLOWER2 (EMF2) o FERTILIZATION INDEPENDENT SEED (FIS2) para formar al menos tres complejos tipo PRC2: EMBRYONIC FLOWER (EMF), VERNALIZATION (VRN), y FERTILIZATION INDEPENDENT SEED (FIS) (Hennig et al., 2005; Köhler et al., 2003; Köhler & Hennig, 2010). Estos complejos participan en distintas etapas del desarrollo de Arabidopsis, aunque comparten genes blanco. Esto sugiere que las proteínas de PcG en plantas reprimen programas genéticos requeridos en estadios particulares del desarrollo (Köhler & Hennig, 2010). El complejo FIS, por ejemplo, es importante para el silenciamiento génico que ocurre durante el desarrollo del gametofito femenino y la maduración de la semilla (Köhler, Hennig, Bouveret, et al., 2003), mientras que el complejo EMF regula algunos de los mismos genes blanco del complejo FIS, durante el desarrollo del esporofito (Hennig et al., 2005). Por último, el complejo VRN participa durante un proceso denominado vernalización, donde actúa para silenciar la transcripción de genes represores de la transición a la floración (Baulcombe & Dean, 2014; Derkacheva et al., 2013).

El establecimiento de la marca de H3K27me3 en el locus de un gen activo, en respuesta a una señal río arriba, es llevado a cabo a través de tres fases distintas (Menon et al., 2021). La

primera fase consta de la actividad de PRC2, mediada por proteínas accesorias, dirigida a una región específica dentro del locus, conocida como región de nucleación, donde se establece la marca H3K27me3. En la segunda fase, el pico de nucleación de H3K27me3 puede extender esta marca a lo largo de todo el locus, llevándolo a un estado de silenciamiento heredable de forma mitótica. Durante la tercera fase, las proteínas accesorias se retiran de la región de nucleación, dejando un estado epigenético perpetuado en el locus, donde H3K27me3 es mantenida por la retroalimentación de proteínas lectoras del complejo PRC1 y escritoras de PRC2 (Menon et al., 2021; Molitor & Shen, 2013). Sin embargo, la actividad de silenciamiento transcripcional de PRC1, a través de la marca de H2AK121Ub, también se ha encontrado independientemente de PRC2, así como se han encontrado interacciones cruzadas entre miembros de los complejos (e.g. AtRING1A-CLF, LHP1-EMF2, EMF1-MSI1, MSI1-VRN2) (Wang & Shen, 2018). Esto ha llevado a cuestionar el orden de reclutamiento y la separación de acciones entre los complejos para el silenciamiento de PcG.

Un gran número de factores de transcripción importantes para el desarrollo son blancos de PcG (Lafos et al., 2011; Truskina et al., 2021; Zhang et al., 2007). Algunos de estos factores de transcripción se encuentran dentro de las familias MADS-box, WOX, HOMEBOX, YABBY, entre otros (Köhler, Hennig, Spillane, et al., 2003; Sanchez et al., 2015; Schönrock et al., 2006).

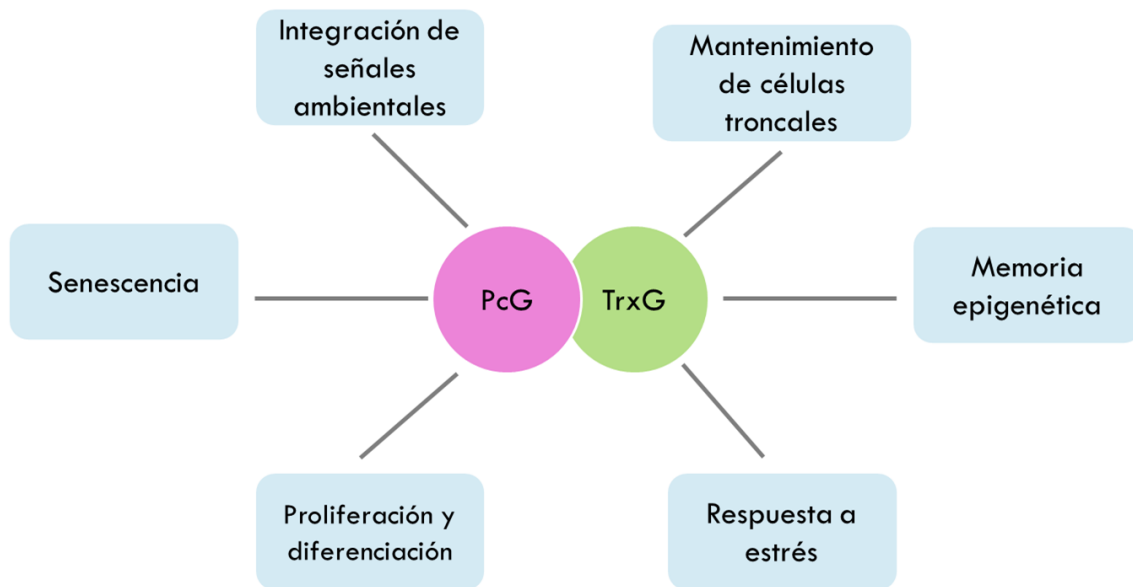


Fig. 4. Resumen de los procesos del desarrollo en los que participan los grupos de reguladores epigenéticos Polycomb (PcG) y Trithorax (TrxG).

Factores de transcripción MADS-box

La superfamilia de factores de transcripción MADS-box se encuentra conservada en todos los reinos de organismos vivos, y participan en varios procesos de desarrollo de animales y plantas (Cezary Smaczniak et al., 2012). Estos factores de transcripción participan en todos los procesos de desarrollo de las plantas, donde un gran número son esenciales para la formación de órganos, el control del tiempo de floración y el mantenimiento de la identidad meristemática (**Fig. 5**) (Ng & Yanofsky, 2001).

En plantas, la superfamilia de estos factores de transcripción ha diversificado a partir de múltiples eventos de duplicaciones y triplicaciones del genoma entero, así como eventos específicos de duplicación de genes (Zhao et al., 2017). La familia de genes MADS-box está formada por más de 100 genes en plantas superiores, y está dividida en dos tipos, basado en la estructura de sus dominios proteicos (Gramzow & Theissen, 2010; C. Smaczniak et al., 2012).

Se consideran de tipo I a aquellos que cuentan únicamente con una secuencia de DNA de alrededor de 180 pb que codifica para el dominio MADS. El tipo II agrupa a aquellos cuyas proteínas del dominio MADS tienen una estructura de dominio modular conocida como MIKC. El dominio modular MIKC es conformado por un dominio MADS de unión a DNA localizado en el extremo N-terminal, seguido de las regiones I (Intervening) y K (keratin-like), las cuales son importantes para la formación de complejos, y, finalmente, un dominio C-terminal altamente variable involucrado en la formación de complejos con proteínas y regulación transcripcional (Cezary Smaczniak et al., 2012). A su vez, los genes pertenecientes al tipo MIKC se clasifican en dos grupos: MIKC^C y MIKC*. Esta última clasificación se debe a la alteración de la estructura del dominio K en la proteína. En *Arabidopsis thaliana* han sido descritos 45 genes MADS-box tipo MIKC, de los cuales únicamente 5 pertenecen a al tipo MIKC*. A excepción de *AP3* y *AGL15*, los genes MIKC^C tienen como característica que su estructura incluye un intrón largo (<500 pb). Estos intrones largos parecen contener elementos importantes para su regulación transcripcional. Su relevancia ya ha sido descrita para genes como *FLOWERING LOCUS C (FLC)* en el proceso de transición a la floración, *AGAMOUS* y *SEED STICK*, en la formación de gametofitos (Kooiker et al., 2005; Sieburth & Meyerowitz, 1997; W. Yuan et al., 2016), mientras que para otros

como *AGL6* y *AGL13*, también en la formación de gametofitos, solamente ha sido predicha (Schauer et al., 2009).

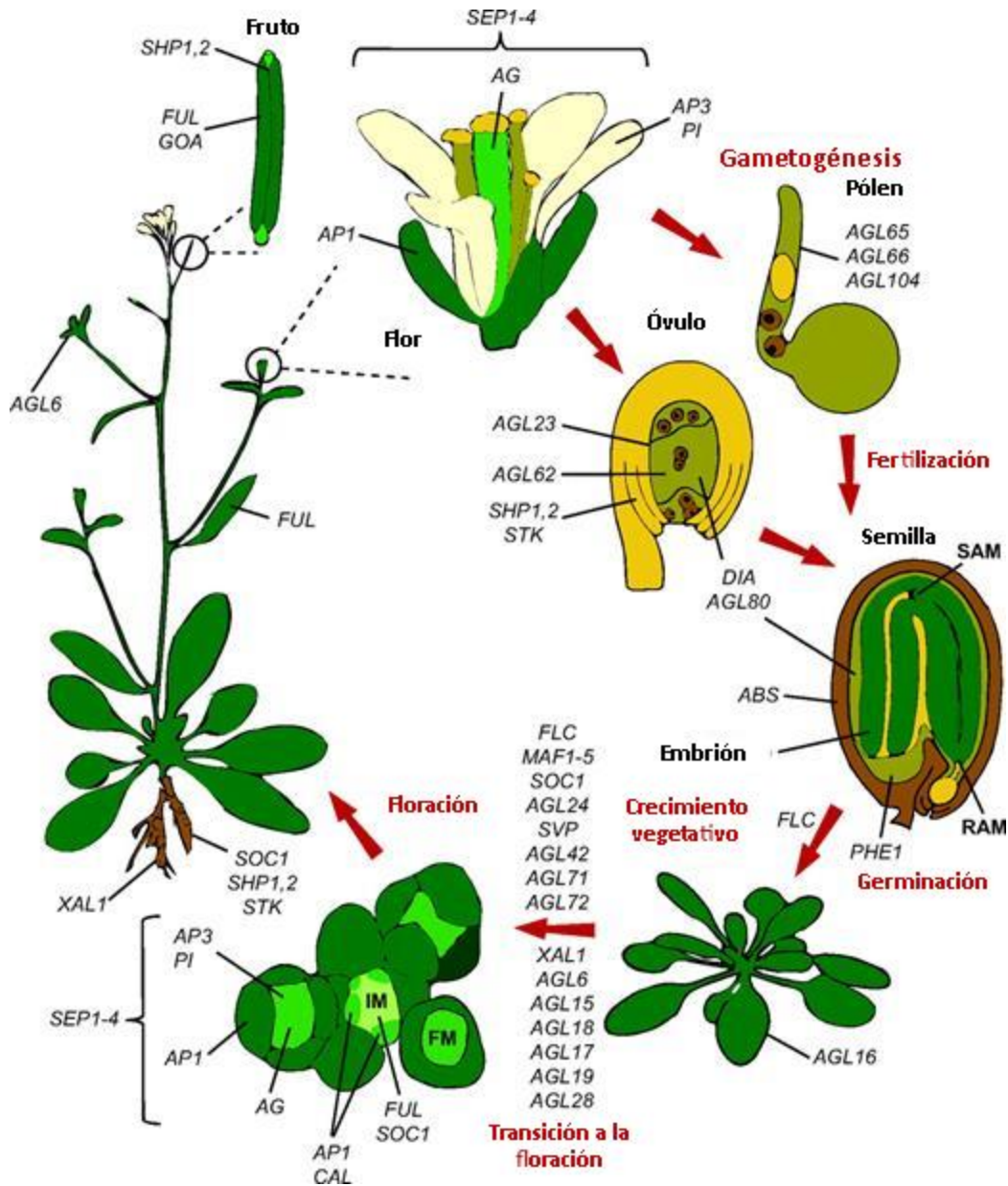


Fig. 5. Esquema modificado de C. Smaczniak et al., (2012), sobre las funciones de genes MADS-box a lo largo del ciclo de vida de *Arabidopsis*.

Los MADS-box reconocen y se unen a motivos de DNA conocidos como cajas CArG (Cezary Smaczniak et al., 2012). La secuencia de la caja consenso canónica es $CC(A/T)_6GG$, sin embargo, existen varios motivos variantes con secuencias como $CC(A/T)_7GG$ o $CC(A/T)_8GG$, los cuales

también son reconocidos por el dominio MADS de estos factores de transcripción (Aerts et al., 2018). Además, se ha observado que la presencia de adeninas seguidas de la secuencia consenso presenta una mayor probabilidad de unión de un MADS-box (Aerts et al., 2018). Se cree que existen propiedades adicionales a las cajas CArG que permiten a estos factores de transcripción unirse en sitios específicos y regular de manera diferencial procesos fisiológicos y del desarrollo (Aerts et al., 2018; Folter & Angenent, 2006).

La distancia entre cajas CArG, así como la posible presencia de motivos híbridos, podrían ser características importantes para su unión. La interacción de los MADS-box formando heterodímeros y heterotetrámeros es común, y el espacio entre las cajas CArG podría explicar que estos se unan en regiones regulatorias en cis distintas dependiendo del acomodo y distancia de estos motivos entre sí, como se ha observado para los genes *SEPALLATA* (Jetha et al., 2014). Más allá, está la posibilidad de la existencia de motivos de unión híbridos, ya sea que esta unión sea cooperativa y se requiera la presencia de ambos factores para una unión óptima, o que se lleve a cabo competencia por el sitio, lo que provocaría una regulación diferencial dependiendo del factor de transcripción unido (Aerts et al., 2018).

Función de los genes MADS-box en el desarrollo

Los MADS-box tienen funciones de genes homeóticos, lo que significa que trabajan como genes de identidad de órganos (Becker & Theißen, 2003; Cezary Smaczniak et al., 2012). Esta función ha sido ampliamente estudiada durante el desarrollo de los órganos florales (Alvarez-Buylla et al., 2010). Además de sus funciones como genes homeóticos, algunos MADS participan dentro de la red de tiempo de floración, donde censan las condiciones ambientales para reprimir o promover la transición de meristemas vegetativos a reproductivos (Becker & Theißen, 2003). También hay MADS-box involucrados en procesos posteriores de formación adecuada de frutos (Gu et al., 1998) y semillas (Becker & Theißen, 2003), en el desarrollo de hojas (Alvarez-Buylla, Pelaz, et al., 2000; Gu et al., 1998) y en el desarrollo de raíz (Alvarez-Buylla et al., 2019)

Papel de los MADS-box en el desarrollo de la raíz

Los MADS-box tipo MIKC participan en todas las etapas del desarrollo de la raíz de *Arabidopsis* (Alvarez-Buylla et al., 2019). Dichos procesos del desarrollo de la raíz pueden dividirse en aquellos relevantes para la arquitectura de la raíz principal, y aquellos importantes para la diferenciación de las raíces laterales (**Fig. 6**) (Alvarez-Buylla et al., 2019). Cerca de la mitad de los genes MADS tipo MIKC son expresados en la raíz (Parenicova et al., 2003), de los cuales una gran parte se desconoce su papel en el desarrollo. En los últimos años se han identificado genes MADS-box de expresión preferencial en raíz pertenecientes a los clados *AGAMOUS* (*AG*), *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1*, *ARABIDOPSIS NITRATE REGULATED1* (*ANR1*) y *AGAMOUS-Like18* (*AGL18*) (Alvarez-Buylla, Liljegren, et al., 2000). Específicamente, *ANR1*, *XAANTAL 1* (*XAL1/AGL12*), *XAANTAL 2* (*XAL2/AGL14*), *AGL17*, *AGL21*, *AGL79*, *AGL42*, *AGL16*, *AGL17* y *AGL18* presentan expresión preferencial en el RAM respecto al SAM (Gan et al., 2005; Nawy et al., 2005).

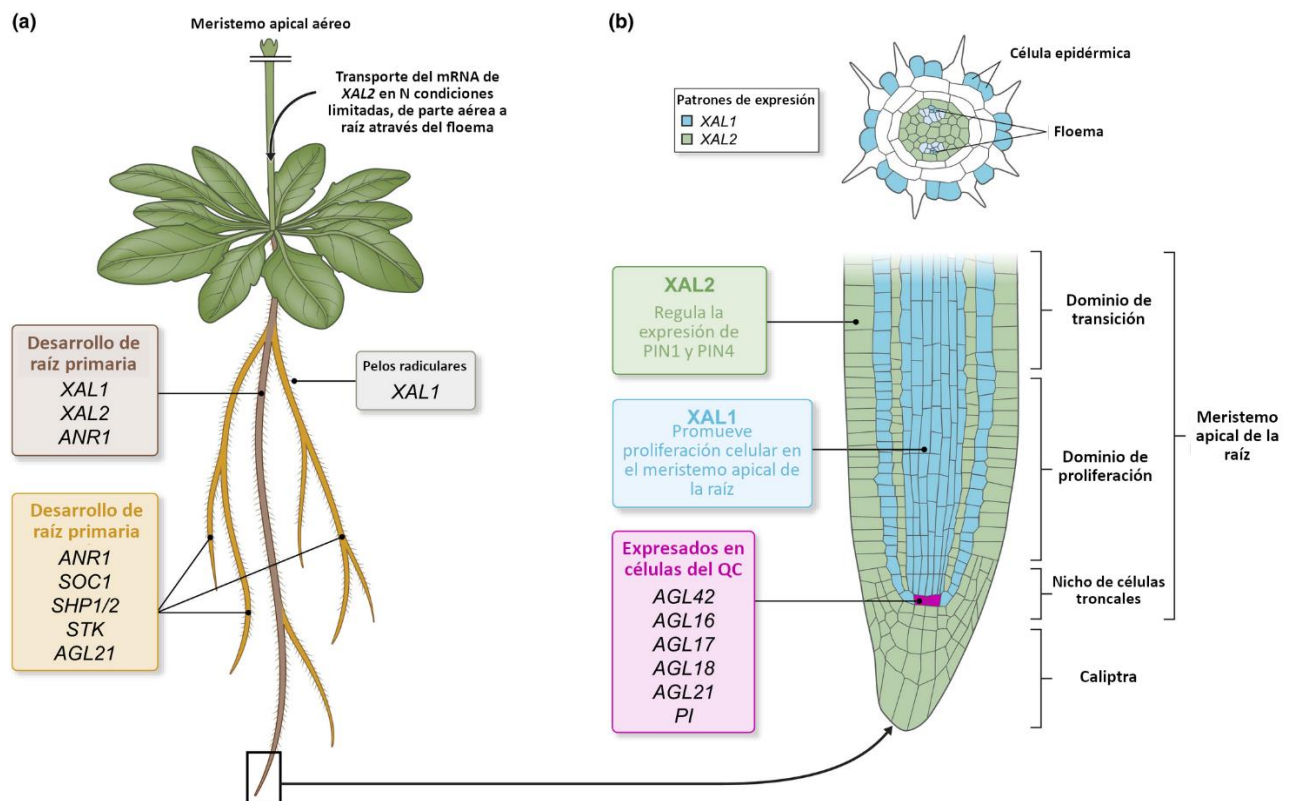


Fig. 6. Esquema modificado de Alvarez-Buylla et al. (2019), sobre los mecanismos de desarrollo y procesos de diferenciación de tejidos de la raíz en los que intervienen los factores de transcripción MADS-box.

Por su parte, en el laboratorio se ha encontrado que XAL1/AGL12 participa durante los eventos que dan lugar al patrón celular característico de la raíz, especialmente en las etapas de proliferación y la transición a la diferenciación (García-Cruz et al., 2016; Tapia-López et al., 2008). Esto mediante la regulación de la tasa de proliferación celular, así como la elongación celular, y modulando la duración del ciclo celular (García-Cruz et al., 2016; Tapia-López et al., 2008). La actividad de XAL2/AGL14 también es necesaria para el desarrollo normal de la raíz y el establecimiento correcto del Centro Quiescente y del nicho de células troncales (Garay-Arroyo et al., 2013). Específicamente, XAL2/AGL14 participa en la regulación del transporte y distribución de auxinas a lo largo de la raíz, a través de la regulación de los transportadores *PIN1* y *PIN4* (Garay-Arroyo et al., 2013).

Importancia del intrón largo en la regulación de un MADS-box *FLC*

El mecanismo de regulación del factor de transcripción *FLC* es un ejemplo de regulación epigenética ampliamente estudiado en *Arabidopsis*, en donde intervienen factores epigenéticos como TrxG, PcG y RNAs largos no codificantes (Whittaker & Dean, 2017). En esta regulación epigenética de *FLC*, su intrón largo de cerca de 3.5 kb, tiene un papel fundamental para su regulación epigenética ya que contiene regiones importantes para llevar a cabo la nucleación de la marca de H3K27me3 (**Fig. 7**) (Qüesta et al., 2016; W. Yuan et al., 2016). Su regulación ha sido principalmente estudiada en un proceso importante para la adecuación de algunas variantes naturales de esta especie, conocido como vernalización. *FLC* es un gen encargado de la represión de la floración durante el desarrollo vegetativo de algunas variantes naturales de *Arabidopsis*. Cuando estas variantes son expuestas a bajas temperaturas por un periodo de tiempo determinado, es decir se vernalizan, *FLC* se reprime mediante mecanismos epigenéticos de tal manera que cuando la temperatura regresa a valores más cálidos se da la floración. Es decir que, para dar paso a la transición a la floración, es necesaria la represión de *FLC* a través de las marcas depositadas por PcG (Berry & Dean, 2015; Coustham et al., 2012). En algunas variantes naturales de *Arabidopsis*, que crecen en zonas con inviernos extremadamente fríos, la regulación de *FLC*

permite que las plantas desarrollen sus órganos vegetativos en otoño y durante el invierno, las temperaturas bajas inducen la represión de *FLC*, de tal manera que cuando llega la primavera puedan desarrollar sus flores y con ello aseguren su descendencia (Berry & Dean, 2015).

Regulación epigenética de *FLC*

La represión de cada copia de *FLC* es llevada a cabo de manera independiente mediante las proteínas PcG, y ocurre en dos pasos principales: 1) nucleación de la marca de H3K27me3 y 2) el esparcimiento de la marca H3K27me3 a lo largo del locus entero de *FLC* (Menon et al., 2021; Yang et al., 2017). En la primera parte de esta regulación resalta el papel que tiene el primer intrón largo de *FLC*, que es el más largo de los MADS-box tipo MIKCC, con cerca de 3.5 kb, y que contiene regiones de gran importancia para su regulación epigenética (Qüesta et al., 2016; W. Yuan et al., 2016).

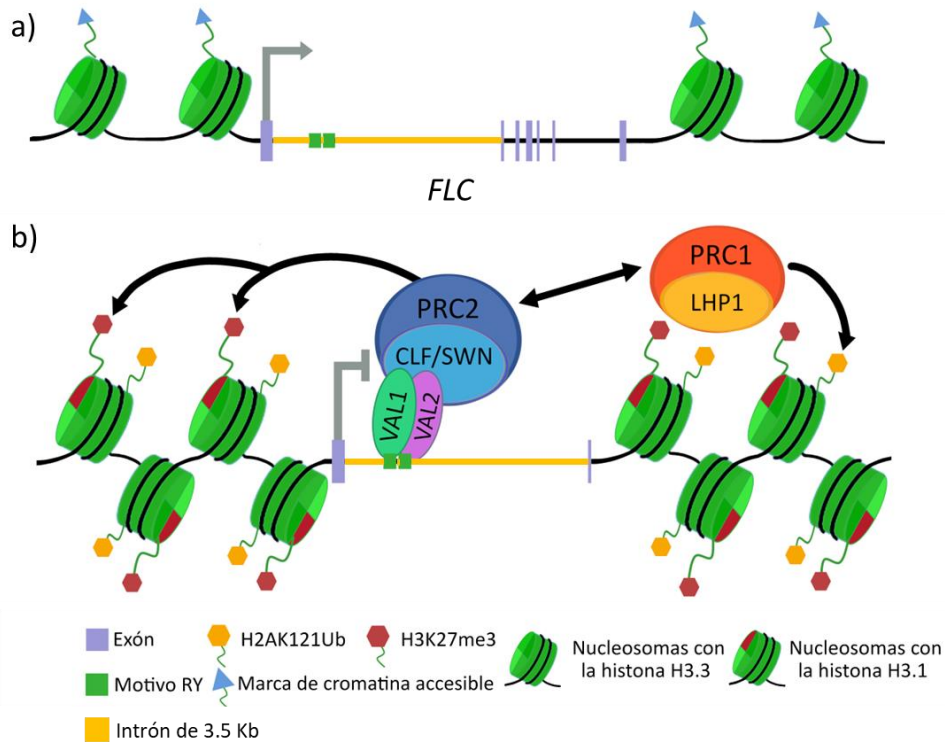


Fig. 7. Esquema con la estructura del gen *FLC* y la regulación negativa donde intervienen los factores de transcripción VAL1 y VAL2. En a) se representa la estructura de la cromatina de *FLC* cuando se encuentra transcripcionalmente activo. En b) está representado el reclutamiento de las maquinarias PRC2 y PRC1 por parte de los factores de transcripción VAL1 y VAL2.

La región de nucleación de la marca H3K27me3 abarca aproximadamente tres nucleosomas, sobre el exón uno y el inicio del primero intrón. Los procesos de nucleación y cambio de un estado de activación a uno de represión son de naturaleza estocástica, lo que significa que diferentes copias del gen *FLC* cambian a distintos tiempos, y que la porción de copias del gen que son nucleadas por PRC2 incrementan durante la exposición al frío (Menon et al., 2021). Variaciones en cis de esta región, en diferentes accesiones de *Arabidopsis*, modulan el requerimiento del tiempo de vernalización necesario para llevar a cabo el silenciamiento del gen (Coustham et al., 2012). Dentro de la porción de la región de nucleación encontrada en el primer intrón de *FLC*, se ha descrito la presencia de dos motivos tipo Sph/Ry (5'-TGCATG-3') (**Fig. 7**), los cuales son reconocidos por dominios de unión a DNA tipo B3 (Qüesta et al., 2016; W. Yuan et al., 2016). A este elemento se le ha otorgado el nombre de *Cold Memory Element* (CME), debido a que es requerido para el mantenimiento estable del silenciamiento del gen, funcionando como memoria celular de exposición previa a condiciones de frío (W. Yuan et al., 2016). Se ha probado la interacción directa de VIVIPAROUS1/ABI3-LIKE 1 y 2 (VAL1 y VAL2), dos factores de transcripción pertenecientes a la familia LAV B3, con los motivos RY (**Fig. 7**) (Wu et al., 2018), de los cuales previamente se conocía su función como represores de genes de maduración de la semilla durante etapas del desarrollo de plántulas de *Arabidopsis* (Tsukagoshi et al., 2007).

A diferencia de *Drosophila* o mamíferos, donde se conoce más sobre la manera en que la maquinaria de PcG reconoce a sus loci blancos, en plantas todavía no se tiene claro un mecanismo de reconocimiento y que éste sea generalizado. Sin embargo, las proteínas VAL1 y VAL2 podrían ejemplificar un mecanismo de reconocimiento y reclutamiento de PcG a sus genes blanco (L. Yuan et al., 2021). Las mutantes de pérdida de función simples y dobles *val1* y *val2* presentan niveles bajos de la marca H3K27me3 sobre *FLC* durante y posterior a la vernalización (Sasnauskas et al., 2018; L. Yuan et al., 2021). De la misma manera, la deposición y mantenimiento de la marca H3K27me3 se ve reducida sobre el locus de *FLC* al realizar mutaciones simples sobre la secuencia de los motivos RY, puesto que los dominios B3 en ambas proteínas son altamente específicos en el reconocimiento de sus motivos (Sasnauskas et al., 2018). Conjuntamente al dominio B3, de unión a DNA, VAL1 VAL2 también cuentan con un dominio EAR de represión transcripcional, y dos dominios de unión a histonas: un dominio CW y un dominio

PHD-L (Wu et al., 2018). El dominio EAR es capaz de interactuar físicamente con la proteína AtSAP18, quién regula el reclutamiento de la deacetilasa de histonas HDA19, además de ser una subunidad del complejo ASAP (Qüesta et al., 2016). El dominio PHD-L reconoce el estado de metilación de la histona H3 a través de interacciones con las marcas de H3K27me2 y H3K27me3 (W. Yuan et al., 2016). Además, el homólogo de la proteína RING de PRC1, AtBMI1A, también se comprobó que interactúa *in vivo* con VAL1 y VAL2 (**Fig. 7**) (Qüesta et al., 2016). Asimismo, se ha probado la interacción de VAL1 y VAL2 con el componente de Polycomb LIKE HETEROCHROMATIN PROTEIN 1 (LHP1), y las subunidades catalíticas de PRC2 SWINGER (SWN) y CURLY LEAF (CLF) (**Fig. 7**) (Qüesta et al., 2016; L. Yuan et al., 2021).

Las múltiples interacciones en las que participan los factores de transcripción VAL1 y VAL2 permiten proponer mecanismos por los que se lleva a cabo el silenciamiento transcripcional de *FLC* por parte de PcG. Se propone que, una vez unidas a los motivos RY en el CME de *FLC*, VAL1 y VAL2 reclutan a LHP1. LHP1 a su vez es capaz de interactuar con MSI1 (Derkacheva et al., 2013), componente de PRC2, reclutando a PRC2 para llevar a cabo la deposición de la H3K27me3. En conjunto con los complejos SWN/CLF-PRC2, VAL1 y VAL2 se pueden asociar a PRC1 y/o a complejos de desacetilasas de histonas para facilitar el esparcimiento de la marca por PRC2 a lo largo del gen (Qüesta et al., 2016; Sasnauskas et al., 2018; Wu et al., 2018; L. Yuan et al., 2021; W. Yuan et al., 2016). Además, el proceso de mantenimiento del silenciamiento transcripcional mediante la H3K27me3 requiere el recambio de la variante H3.3 por la variante H3.1. Como se mencionó anteriormente, la variante H3.1 es depositada durante la replicación de DNA, y es necesaria debido a la monometilación preferencial en la K27 (H3K27me1) que llevan a cabo ATRX5/6, proveyendo un sustrato que puede ser metilado rápidamente por PRC2 para mantener los niveles necesarios de H3K27me3 (**Fig. 7**) (Jiang & Berger, 2017a).

La conservación de una estructura génica similar a lo largo de los factores de transcripción MADS-box tipo MIKC (Parenicova et al., 2003), así como la importancia de su correcta regulación transcripcional para llevar a cabo sus funciones durante el desarrollo de Arabidopsis (Ng & Yanofsky, 2001), podría indicar que también se conserva la regulación transcripcional en cis en la cual, se ve involucrada la presencia de motivos RY, en esta subfamilia de factores de transcripción. Datos previos obtenidos en el laboratorio (Alejo-Vinogradova, 2020; Vega-León,

2018) han demostrado la regulación mediante la maquinaria de PcG de una serie de MADS-box en la raíz de Arabidopsis. La mutante de pérdida de función de la histona metiltransferasa CLF (*clf-29*) muestran un aumento en la expresión de los MADS-box *AGL17*, *AGL19*, *FLC* y *XAL2/AGL14* (Vega-León, 2018), lo que concuerda con la actividad represora de PcG. Al mismo tiempo, *AGL17*, *AGL19*, *AGL42*, y *XAL2/AGL14* aumentan su expresión en la mutante de pérdida de función de LHP1 (*lhp1-6*), la lectora de la H3K27me3, debido a que son blancos directos de la maquinaria de PRC1 (Alejo-Vinogradova, 2020). Considerando estos datos previos, este trabajo pretende ahondar en la posible conservación de un patrón de regulación en cis de factores de transcripción MADS-box relevantes para el desarrollo de la raíz de Arabidopsis, a través de los posibles motivos RY encontrados en sus intrones largos.

Hipótesis:

Dentro de los intrones largos de los genes MADS-box tipo MIKC existe un patrón de represión transcripcional generalizado por parte de la maquinaria del grupo PcG, en el cual están involucrados los motivos RY.

Objetivos:

General:

- Estudiar el papel de los intrones largos de los factores de transcripción MADS-box tipo MIKC, en su regulación epigenética.

Específicos:

- Realizar análisis *in silico* de expresión de los 45 genes MADS-box tipo MIKC durante el desarrollo de *Arabidopsis*.
- Identificar secuencias consenso o motivos compartidos en los intrones largos de dichos genes.
- Analizar las implicaciones de los motivos encontrados en la regulación transcripcional de los genes MADS-box.
- Identificar patrones de marcas epigenéticas de activación (H3K4me3) y represión (H3K27me3), y su asociación con las variantes de histonas H3.1 y H3.3, en los genes MADS-box.

Materiales y métodos

Filogenia de los genes MADS-box

La filogenia utilizada en este trabajo fue obtenida a partir del artículo de Martínez-Castilla & Alvarez-Buylla (2003). Esta filogenia fue construida utilizando datos de las regiones codificantes de los genes y no de regiones no codificantes. La filogenia está enraizada por el MADS-box tipo I *AGL103*.

Análisis de expresión *in silico*

Los análisis de expresión de los MADS-box tipo MIKC fueron realizados utilizando el programa Genevestigator (Hruz et al., 2008), el cual permite analizar la regulación transcripcional de genes entre una amplia cantidad de condiciones experimentales al mismo tiempo (NEBION, 2021). Específicamente para este proyecto, se utilizó un *Compendium-wide análisis* con la condición de búsqueda de desarrollo, el cual resume los cambios en niveles de expresión, representado por el porcentaje de potencial de expresión, de los 45 genes MADS-box tipo MIKC a lo largo de 10 etapas distintas del desarrollo de Arabidopsis: estadio 0 – germinación de la semilla, emergencia de hipocótilo y cotiledones (1-5.9 días); estadio 1 – primera etapa del desarrollo foliar (6-13.9 días); estadio 2 – segunda etapa del desarrollo foliar (14-17.9 días); estadio 3 – tercera etapa del desarrollo foliar (18-20.9 días); estadio 4 – cuarta etapa del desarrollo foliar (21-24.9 días); estadio 5 – emergencia del escapo floral (25-28.9 días); estadio 6 – formación de flores (29-35.9 días); estadio 7 – apertura de flores y formación de silicuas (36-44.9 días); estadio 8 – apertura de silicuas y senescencia de la planta (45-50 días); estadio 9 - desarrollo de la semilla.

A su vez, se realizó un *Compendium-wide analysis* con la condición de búsqueda de anatomía, específicamente en los tejidos de la raíz de Arabidopsis (raíz principal, punta de la raíz, RAM, zona de elongación, zona de maduración, estele, periciclo y raíz lateral). La plataforma de datos seleccionada fue *Affymetrix Arabidopsis ATH1 Genome Array*. En dicha base de datos del programa no se encontraron datos para los genes *AGL70* y *AGL63*, por lo que no fueron

considerados para estos análisis. El porcentaje de potencial de expresión es un indicador del nivel de expresión máxima de un gen. Representa el percentil más alto de todos los valores de expresión de este gen (NEBION, 2021).

La selección de los genes más expresados en raíz fue realizada a través de la normalización de los datos de potencial de expresión de cada uno de los 45 genes MADS-box tipo MIKC contra un gen que codifica para una helicasa de RNA (AT1G58050), debido a su expresión constitutiva y no variable en raíz (Czechowski et al., 2005).

Búsqueda *in silico* de motivos RY y cajas CArG

Para realizar la búsqueda de motivos RY se utilizó la herramienta *AthaMap* (Steffens et al., 2004). Esta herramienta muestra un mapa de posibles sitios de unión de factores de transcripción a lo largo del genoma entero de Arabidopsis, basándose en predicciones realizadas con matrices de pesos posicionales (Bülow et al., 2009). La búsqueda de los motivos se llevó a cabo a lo largo de los intrones largos de 19 de los 22 MADS-box tipo MIKC de mayor expresión en la raíz. Para confirmar la presencia de estos motivos, se utilizó la herramienta *FIMO* de *MEME Suite* (Grant et al., 2011), la cual es capaz de escanear secuencias dadas en búsqueda de motivos especificados por el usuario. Por último, los motivos se confirmaron manualmente en la secuencia de cada gen utilizando *Serial Cloner* (Perez, 2004), un programa para la visualización y análisis de secuencias.

Gráficos de distribución de marcas y variantes de histonas

Los datos de deposición de las variantes de histonas H3.1 y H3.3 fueron tomados del artículo de Stroud et al. (2012). Estos datos son resultado de ChIP-seq utilizando plántulas completas de 10 días. Los datos de la marca de histonas H3K27me3 provienen del artículo de Zhang et al. (2007). Estos datos son resultado de ensayos de ChIP utilizando plántulas completas de 10-14 días. Los datos de la marca de histonas H3K4me3 fueron tomados de Zhang et al. (2009). Para estos datos de ChIP se utilizó la parte aérea de plantas de 21 días.

El análisis de los datos y la obtención de las gráficas de distribución de variantes de histonas y marcas de histonas, se llevó a cabo a través del programa *Integrated Genome Browser* (IGB) (Nicol et al., 2009). Se utilizó como referencia la versión TAIR10 del genoma de Arabidopsis.

Para las gráficas de distribución de marcas de histonas, se tomó como umbral un valor de 0.8. Por otro lado, los datos de variantes de histonas se obtuvieron ya procesados por Stroud *et al.*, (2012), por lo que no fue necesario asignarles un valor umbral.

Como validación del procesamiento de los datos de la histona H3.1 y la variante H3.3, se realizó un gráfico de distribución para cada cromosoma y se comparó con los datos publicados por Stroud *et al.* (2012) (**Fig. A1** en [Anexos](#)).

Metaplots y Mapas de calor

Los datos utilizados para la elaboración de estos gráficos se obtuvieron de las publicaciones de You *et al.*, (2017) (E-MTAB-4680 y E-MTAB-4684) y Stroud *et al.*, (2012) (ID: 200036629).

Se realizaron archivos con extensión BED que contenían las coordenadas de los intrones largos de los MADS-box de mayor expresión en raíz, usando como referencia la versión del TAIR10 del genoma de Arabidopsis. Se tomó en cuenta la información de todos los genes MADS-box tipo MIKC que contienen un intrón largo en su secuencia, que son 38 de los 45 totales. El gen *XAL1/AGL12* contiene dos intrones largos, por lo que ambos fueron considerados para los análisis.

A continuación, se enlistan los paquetes y herramientas utilizados para el procesamiento de los datos:

- *Bowtie2 v.2.3.4.3*: alineamiento de lecturas del input de la marca al genoma (Langmead & Salzberg, 2012)
- *Samtools v.1.9*: generación del índice de lecturas alineadas al genoma para input y marca (Li, *et al.*, 2009).
- *bamCompare v.3.5.1*: cálculo de matriz para representación del metaplot (Ramírez, *et al.*, 2016).
- *plotHeatmap v.3.5.1*: representación del metaplot, a partir de la matriz generada anteriormente, en formato de gráfico junto con el mapa de calor (Ramírez, *et al.*, 2016).

Resultados:

La expresión de los genes MADS-box tipo MIKC tienen un patrón espacial y temporal diferencial en los diferentes tejidos de Arabidopsis.

Como primer análisis para este proyecto, se llevaron a cabo análisis *in silico* sobre la expresión de los MADS-box tipo MIKC a lo largo del desarrollo de Arabidopsis. Para ello, se utilizó el programa de análisis de microarreglos Genevestigator ([Materiales y métodos](#)). En la **Fig. 8**, se puede observar que la expresión de los factores de transcripción es altamente dependiente de la etapa del ciclo de vida de la planta y, por lo tanto, su regulación transcripcional y sus funciones como reguladores de otros genes también son dependientes de la etapa del desarrollo en que se encuentre la planta.

A partir de este análisis, se buscaron posibles relaciones entre clados de genes y patrones de expresión. Esto siguiendo la idea de que genes filogenéticamente cercanos entre sí podrían llevar a cabo funciones similares y además sufrir patrones de regulación también similares. Para ello se utilizó la filogenia de los MADS-box publicada por Martínez-Castilla & Alvarez-Buylla (2003), la cual incluye a los 45 genes MADS-box tipo MIKC, y está enraizada por el MADS-box tipo I *AGL103*. Únicamente clados como el de *FLC-like* (*AGL31*, *MAF1*, *AGL68*, *FLC2* y *FLC*), y el clado *A-related* (*SEP1*, *SEP2*, *SEP3*, *AGL3*, *AGL13*, *AGL6*, *CAL*, *AP1*, *FUL*, y *AGL79*) presentan patrones de expresión similares, entre los genes que los conforman, en el desarrollo de la semilla y en el desarrollo de flores y frutos, respectivamente. Cabe mencionar que la filogenia utilizada en este trabajo se construyó utilizando datos de regiones codificantes de los genes (Martínez-Castilla & Alvarez-Buylla, 2003), dejando de lado posible información relevante encontrada en regiones no codificantes.

A partir de un análisis de expresión en los tejidos de raíz ([Materiales y métodos](#)), se seleccionaron los genes MADS-box tipo MIKC con mayor expresión en tejidos de la raíz (**Fig. 9**). De los 22 genes seleccionados, dos de ellos pertenecen al grupo de los MIKC*, *AGL30* y *AGL65*, que junto con *AGL15* no contienen intrones largos (>500 pb) en su secuencia (**Fig. 10**), por lo que

no fueron considerados para los análisis posteriores. En los 19 genes restantes, el intrón largo se encuentra en la primera posición, es decir, corresponde al primer intrón, a excepción de *XAL1/AGL12*, *AGL5/SHP2*, *AGL44/ANR1*, *AGL19*, *XAL2/AGL14* y *AGL20/SOC1*, que se encuentra en la segunda posición. De este grupo de genes, *FLC* es el gen que tiene el intrón más largo de 3,493 pb, mientras que el más corto pertenece a *AGL18* con 520 pb.

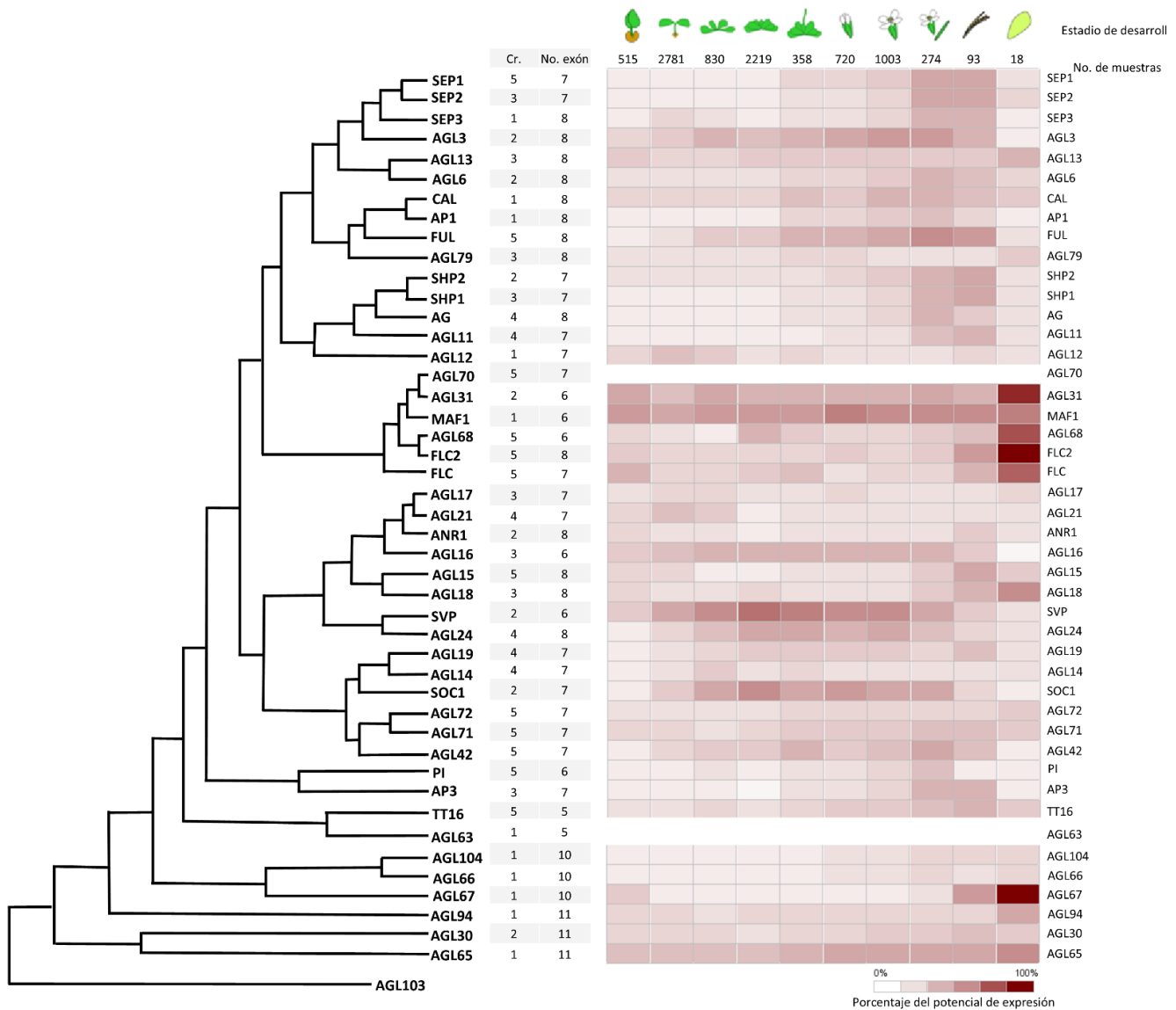


Fig. 8. Árbol filogenético de los MADS-box tipo MIKC y la expresión diferencial de los 45 genes a lo largo de 10 etapas del desarrollo de Arabidopsis. La primera columna (Cr.) indica el número de cromosoma en el que se encuentra el gen, la segunda columna (No. exón) el número de exones que tiene el gen. Los genes *AGL70* y *AGL63* no presentan datos puesto que estos no se encontraban en las bases de datos utilizadas en el programa Geneinvestigator. El árbol fue enraizado al gen *AGL103*, un MADS-box tipo I. La intensidad del color rojo indica el nivel de expresión en el tejido correspondiente, donde un tono intenso indica mayor expresión que un tono claro. La expresión se encuentra cuantificada por el porcentaje del potencial de expresión.

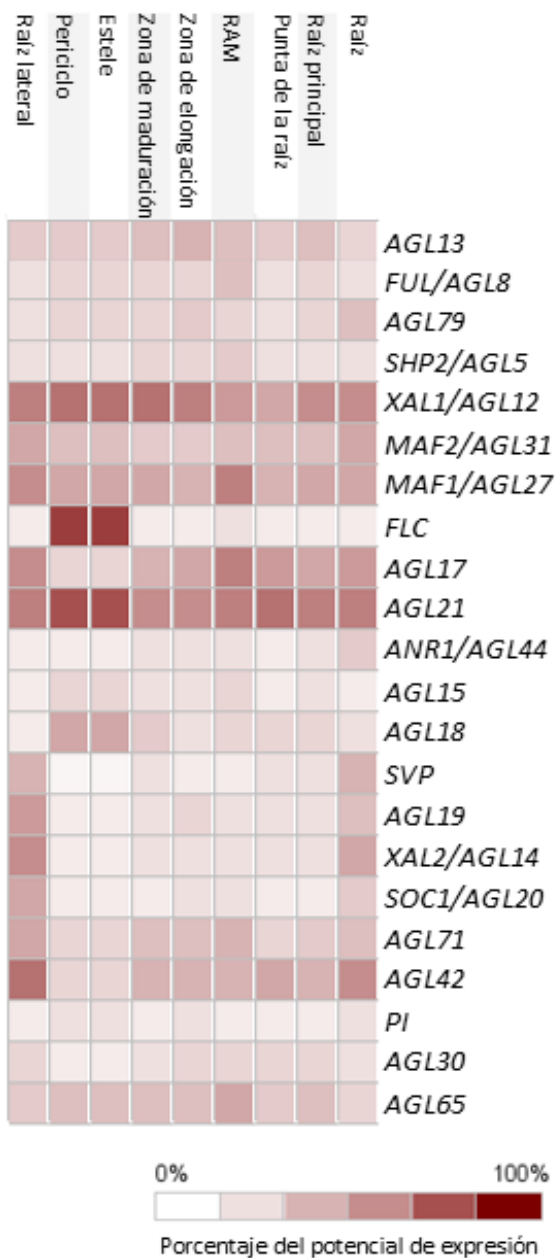


Fig. 9. Mapa de calor de los 22 MADS-box tipo MIKC de mayor expresión en los tejidos de raíz. La intensidad del color rojo indica el nivel de expresión en el tejido correspondiente, donde un tono intenso indica mayor expresión que un tono claro. La expresión se encuentra cuantificada por el porcentaje del potencial de expresión.

Los motivos RY están conservados en la mayoría de los intrones largos de los MADS-box expresados en raíz.

A partir de los MADS-box seleccionados por presentar los niveles de expresión más altos en tejido de raíz, se llevó a cabo la búsqueda *in silico* de motivos RY en sus intrones largos, tal como ocurre en *FLC*.

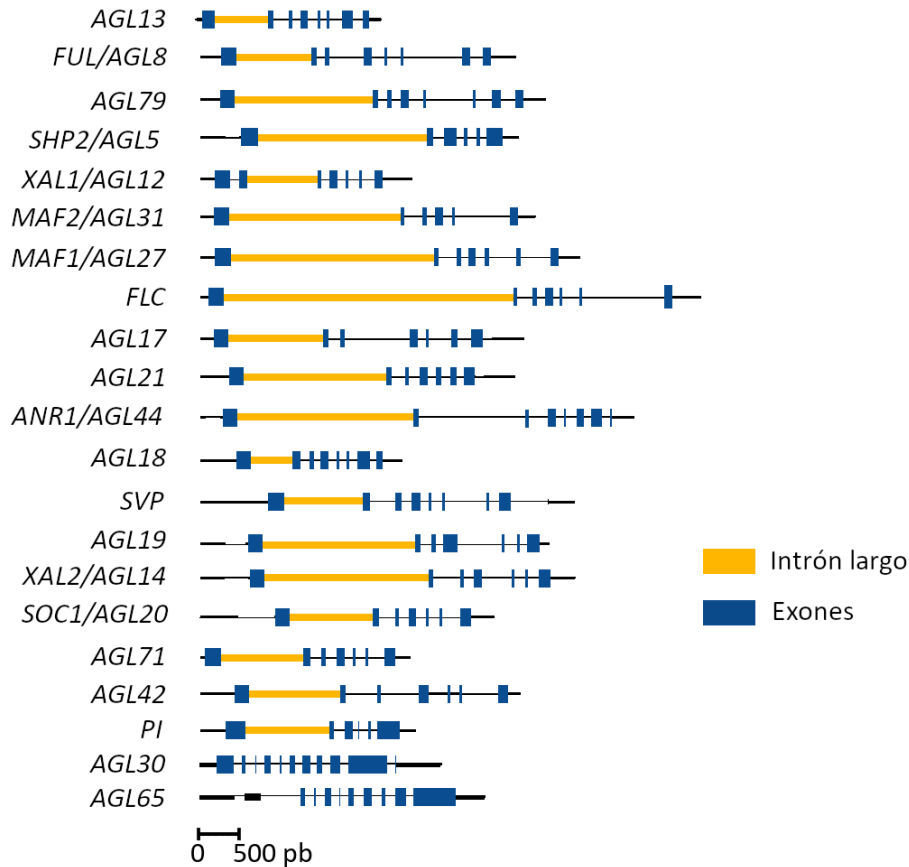


Fig. 10. Estructura genética de 21 de los MADS-box tipo MIKC más expresados en raíz. Los rectángulos amarillos indican el tamaño y presencia de los intrones largos, mientras que los rectángulos azules representan a los exones.

Los genes fueron divididos en tres grupos dependiendo de la cantidad de motivos RY que presentaron: 1) aquellos genes que no presentaron ningún motivo (*FUL*, *AGL17*, *AGL18*, *SVP*, *SOC1* y *AGL42*); 2) aquellos genes que presentaron al menos un motivo RY (*XAL1/AGL12*, *AGL31* y *AGL71*); y 3) aquellos que presentaron dos o más motivos RY (*AGL79*, *SHP2*, *MAF1*, *FLC*, *AGL21*, *ANR1*, *AGL19*, *XAL2/AGL14*, *AGL13*, y *PI*) (**Fig. 11**).

Puesto que previamente se ha observado que los factores de transcripción MADS-box son regulados por elementos dentro de la misma familia (Jetha et al., 2014), y además que es importante la disposición de los mismos para tener la capacidad de unirse en sitios específicos y regular de manera diferencial procesos fisiológicos y del desarrollo (Aerts et al., 2018; Folter & Angenent, 2006), se llevó a cabo la búsqueda de cajas CArG dentro de los intrones (**Fig. 11**). En este caso, todos los genes presentaron por lo menos una de estas cajas, (*FUL*, *SVP*, *AGL71* y *PI*), y hasta 13 en *MAF1*. De manera interesante, podemos observar que predomina la presencia de

cajas CARg no canónicas, donde únicamente se encontró una caja CARg canónica dentro del intrón largo de *AGL19* (Fig. 11).

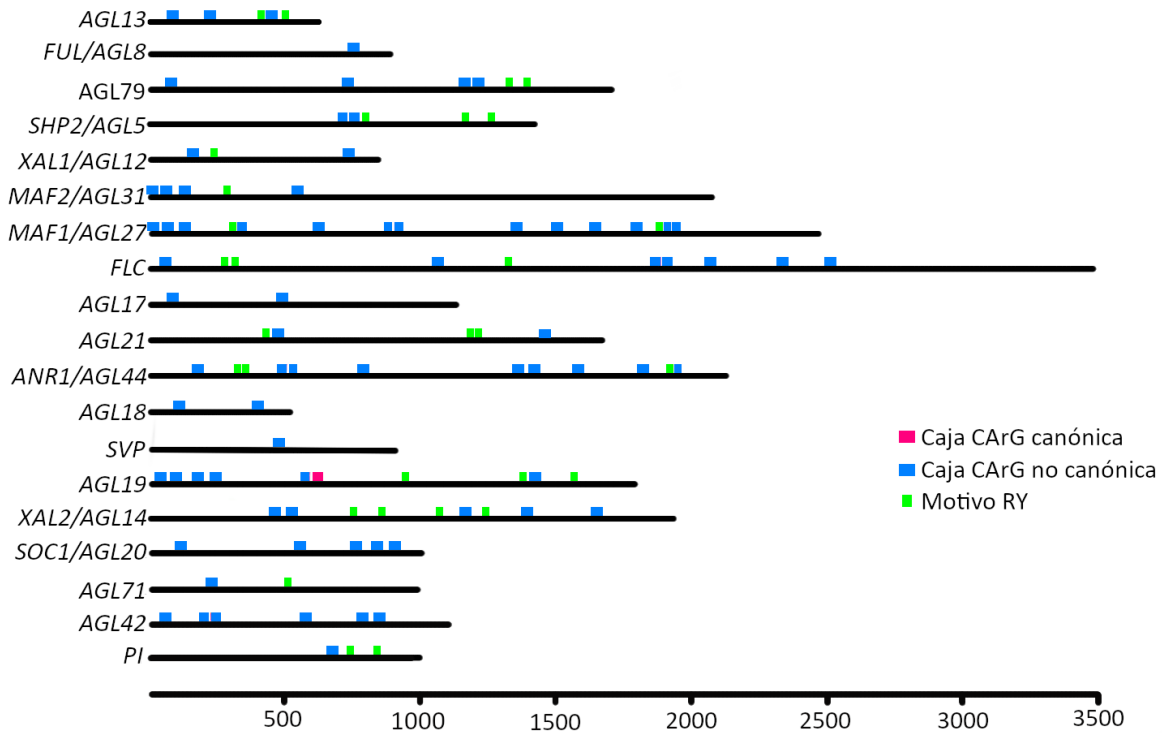
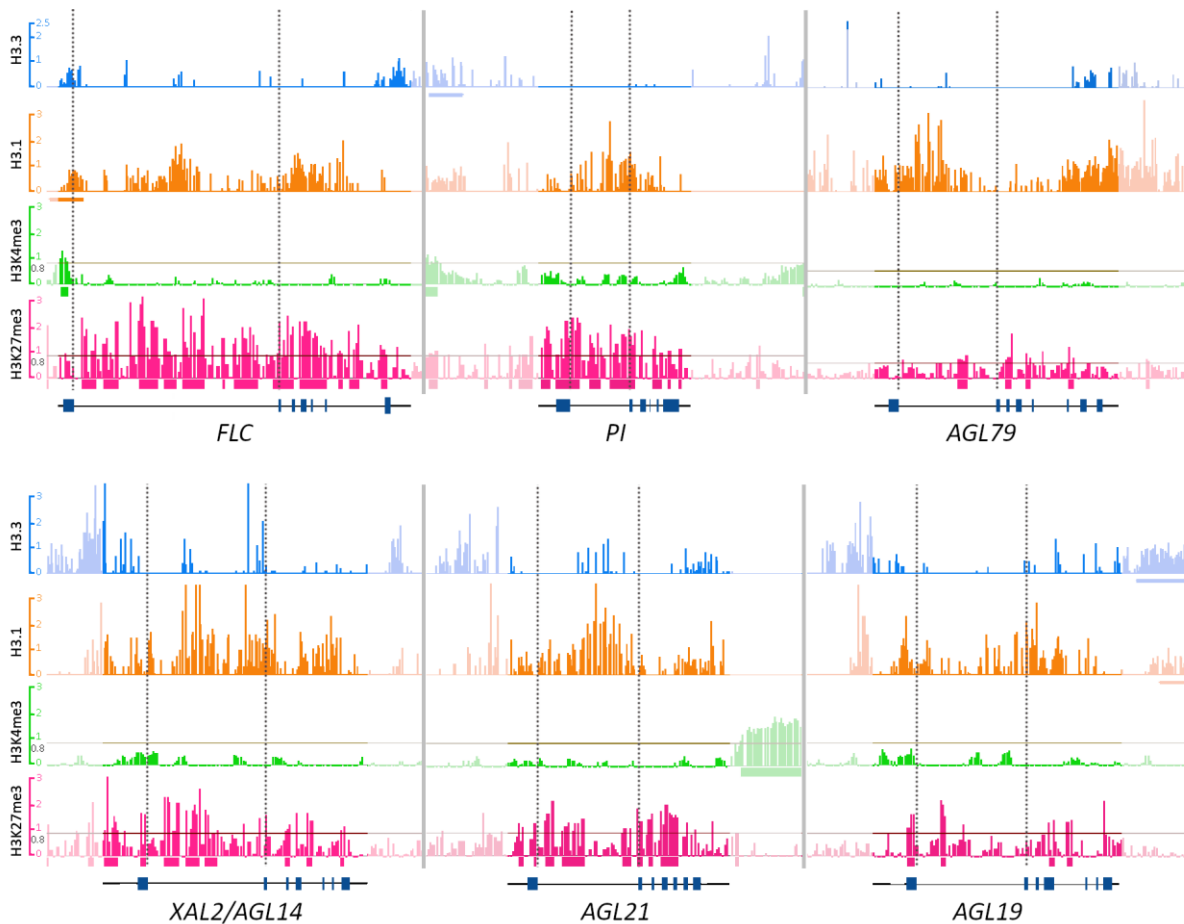


Fig. 11. Motivos RY y cajas CARg situadas dentro de los intrones largos de los MADS-box tipo MIKC más expresados en raíz. Los recuadros color verde indican el lugar donde se sitúan los motivos RY, en azul las cajas CARg no canónicas y en rosa las cajas CARg canónicas. En la parte inferior de la figura se encuentra una escala que indica las pares de bases que mide cada intrón.

Una aparente relación se encuentra entre la cantidad de motivos RY dentro del intrón largo, la deposición diferencial de variantes y las marcas de histonas.

Con el objetivo de identificar patrones de distribución de las marcas de H3K4me3 y de H3K27me3, así como su asociación con las variantes de histonas H3.1 y H3.3, en los 19 genes MADS-box seleccionados, se llevaron a cabo análisis *in silico* ([Materiales y métodos](#)) de distribución de estas variantes y marcas de histonas. En **Fig. 12-14** se muestran los gráficos de distribución realizados con el programa IGB a partir de los datos de Stroud *et al.* (2012), Zhang *et al.* (2007) y Zhang *et al.* (2009) ([Materiales y métodos](#)).

En la **Fig. 12** se observan las gráficas de distribución para los 10 genes MADS-box tipo MIKC de mayor expresión en raíz que contienen dos o más motivos RY dentro de su intrón largo (*AGL79*, *SHP2*, *MAF1*, *FLC*, *AGL21*, *ANR1*, *AGL19*, *AGL14/XAL2*, *AGL13*, y *PI*). A excepción de *AGL13* y *SHP2*, ninguno de los genes presenta una distribución significativa de la marca de H3K4me3. Por otro lado, los genes *FLC*, *PI*, *AGL79*, *AGL14*, *AGL21*, y *AGL19* presentan una distribución significativa de la marca de H3K27me3, y a su vez presentan niveles significativos de la variante de histonas H3.1. *ANR1*, *AGL13* y *SHP2* presentan niveles bajos de la marca de H3K27me3, y picos significativos de ambas variantes de histonas a lo largo del gen.



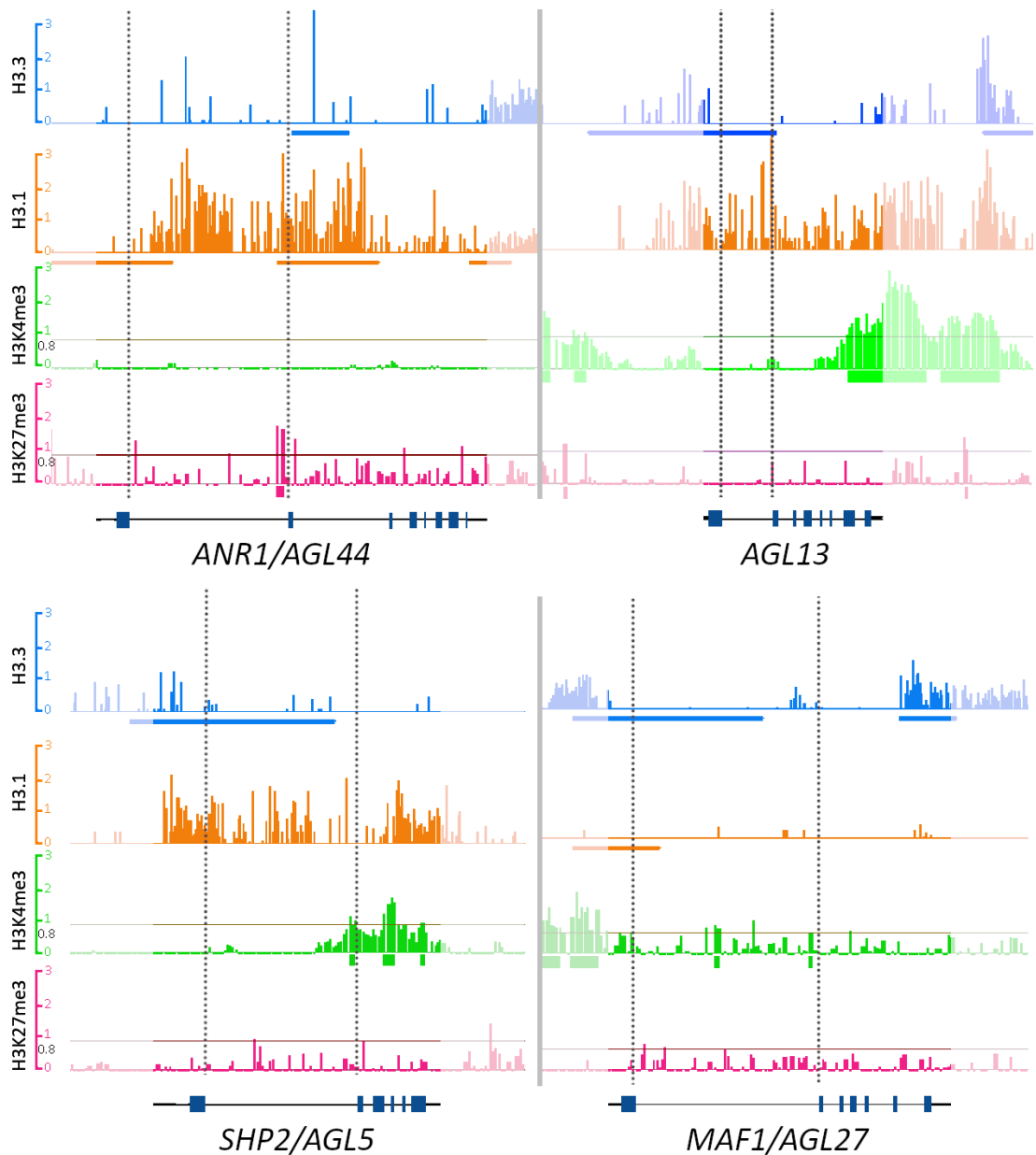


Fig. 12. Gráficos de distribución de variantes de histonas H3.1 y H3.3, así como de las marcas de histonas de H3K4me3 y H3K27me3 en los genes MADS-box tipo MIKC de mayor expresión en raíz de *Arabidopsis* que presentan dos o más motivos RY dentro de sus intrones largos. En color azul se muestran los datos de distribución de la variante H3.3, en naranja de la H3.1, en verde la marca de H3K4me3 y en rosa la marca de H3K27me3. Debajo de los gráficos de distribución se encuentra la estructura del gen con los exones resaltados en color azul, así como el nombre de cada uno. Las líneas puntadas de color gris indican la posición del intrón largo. Para las gráficas de distribución de marcas de histonas, se tomó como umbral un valor de 0.8.

Por último, *MAF1* no presenta niveles significativos de distribución para ninguna de las marcas analizadas ni para las variantes de histonas.

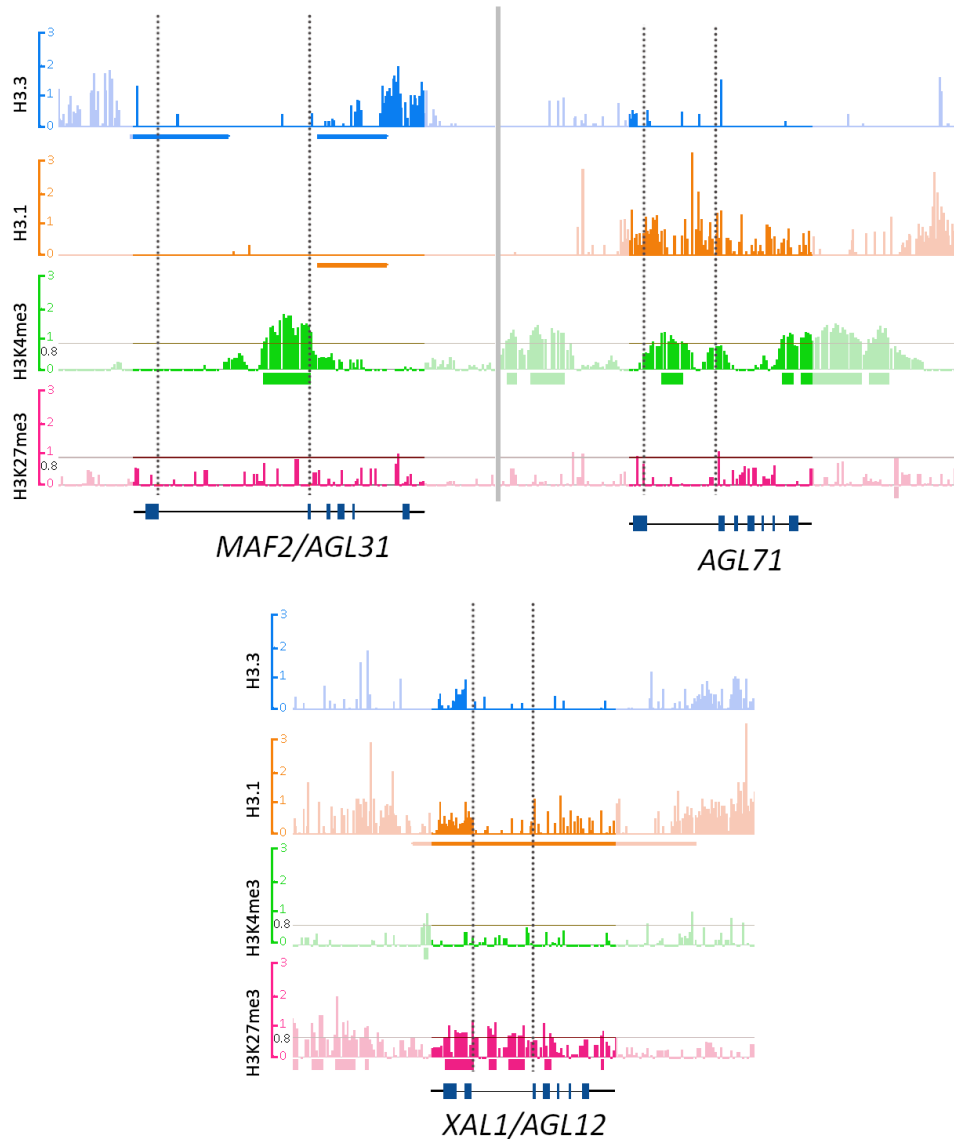


Fig. 13. Gráficos de distribución de variantes de histonas H3.1 y H3.3, así como de las marcas de histonas de H3K4me3 y H3K27me3 en los genes MADS-box tipo MIKC de mayor expresión en raíz de *Arabidopsis* que presentan un único motivo RY dentro de sus intrones largos. En color azul se muestran los datos de distribución de la variante H3.3, en naranja de la H3.1, en verde la marca de H3K4me3 y en rosa la marca de H3K27me3. Debajo de los gráficos de distribución se encuentra la estructura del gen con los exones resaltados en color azul, así como el nombre de cada uno. Las líneas puntadas de color gris indican la posición del intrón largo. Para las gráficas de distribución de marcas de histonas, se tomó como umbral un valor de 0.8.

Por su parte, aquellos genes que presentan únicamente un motivo RY dentro de su intrón largo (*MAF2*, *AGL71* y *XAL1/AGL12*) presentan patrones de distribución más heterogéneos entre sí (**Fig. 13**). *XAL1/AGL12* presenta una distribución significativa tanto de la marca de H3K27me3 como de la variante H3.1 a lo largo del gen, mientras que presenta niveles bajos de la marca de H3K4me3 y de la variante H3.3. Por otro lado, *MAF2/AGL31* y *AGL71* presentan niveles bajos de la marca de H3K27me3 y niveles significativos de la marca de H3K4me3. *AGL71* no tiene niveles significativos de ninguna de las dos variantes de histonas analizadas, mientras que *MAF2/AGL31* presenta un mayor nivel de la variante H3.3.

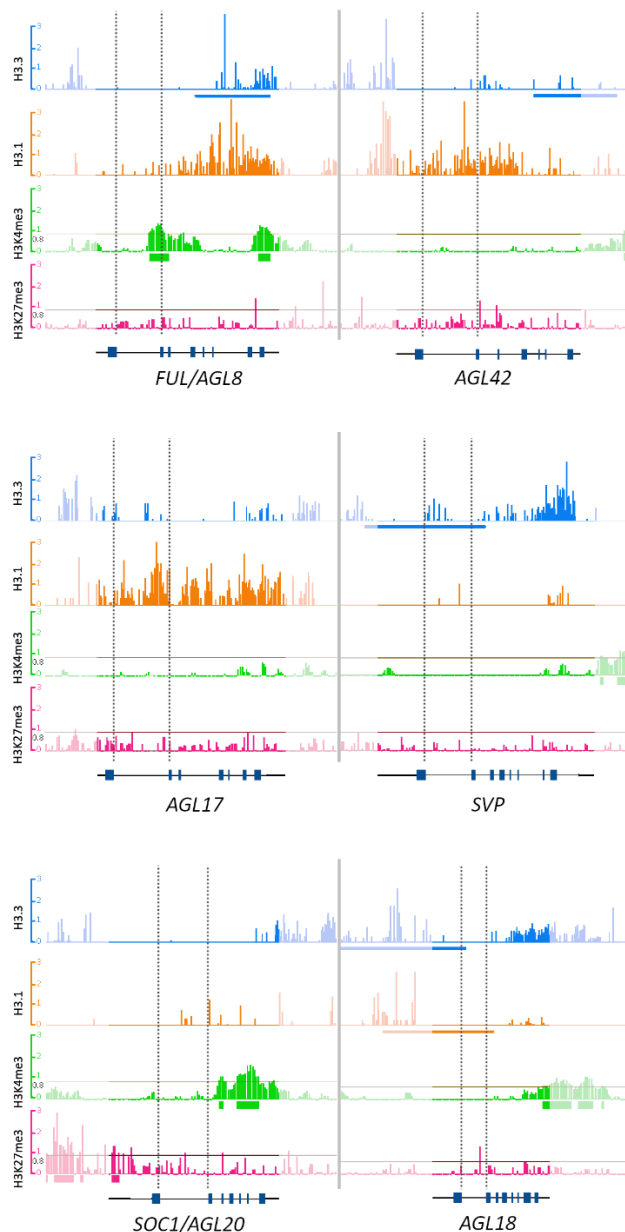
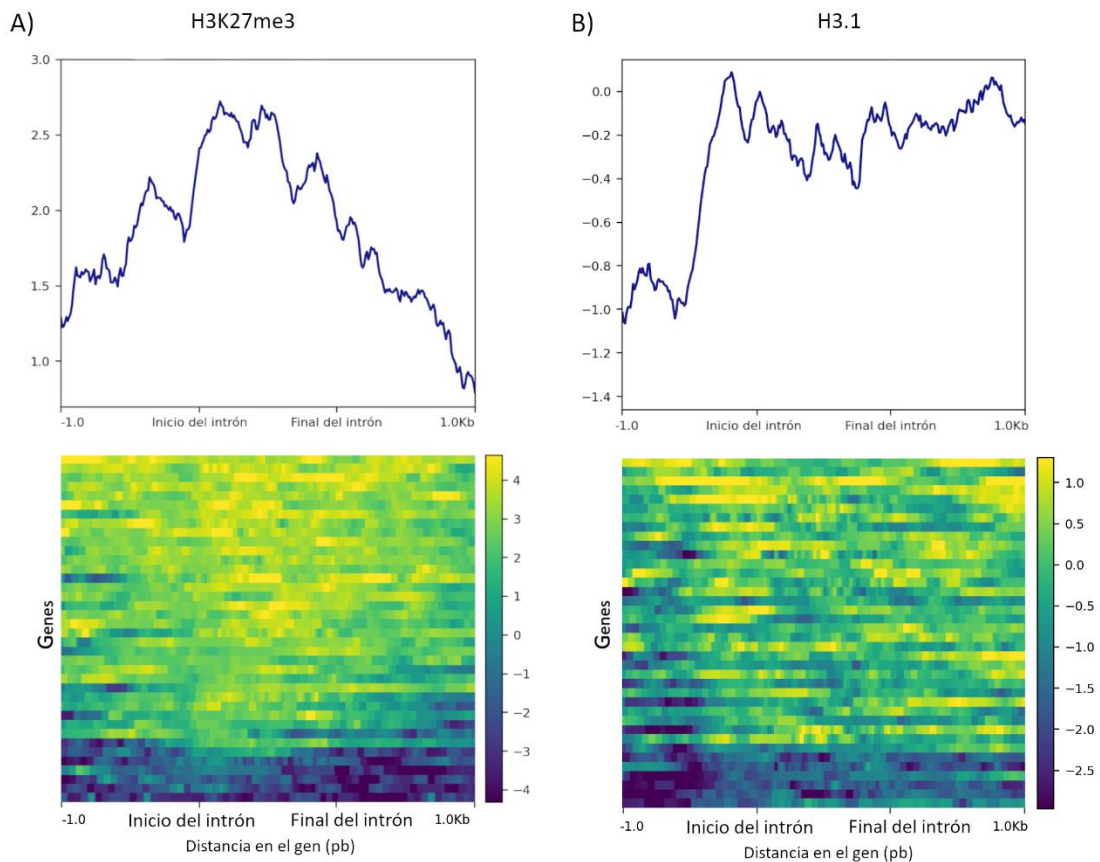


Fig. 14. Gráficos de distribución de variantes de histonas H3.1 y H3.3, así como de las marcas de histonas de H3K4me3 y H3K27me3 en los genes MADS-box tipo MIKC de mayor expresión en raíz de *Arabidopsis* que no presentan ningún motivo RY dentro de sus intrones largos. En color azul se muestran los datos de distribución de la variante H3.3, en naranja de la H3.1, en verde la marca de H3K4me3 y en rosa la marca de H3K27me3. Debajo de los gráficos de distribución se encuentra la estructura del gen con los exones resaltados en color azul, así como el nombre de cada uno. Las líneas puntadas de color gris indican la posición del intrón largo. Para las gráficas de distribución de marcas de histonas, se tomó como umbral un valor de 0.8.

Ninguno de los genes que carecían de motivos RY en sus intrones largos (*FUL/AGL8*, *AGL17*, *AGL18*, *SVP*, *SOC1/AGL20* y *AGL42*) presenta niveles significativos de la marca de H3K27me3 (**Fig. 14**). *FUL/AGL8*, *SOC1/AGL20* y *AGL18* presentan niveles significativos de la marca de H3K4me3 solo en unas pequeñas porciones del gen. De la misma manera, *FUL/AGL8*, *AGL42*, *SVP* y *AGL18* presentan niveles significativos de la variante H3.3 en algunas porciones del gen. En la **Tabla 1** se resumen los datos de distribución de las variantes H3.1 y H3.3 como de las marcas de histonas de H3K27me3 y H3K4me3 en los 19 genes analizados.



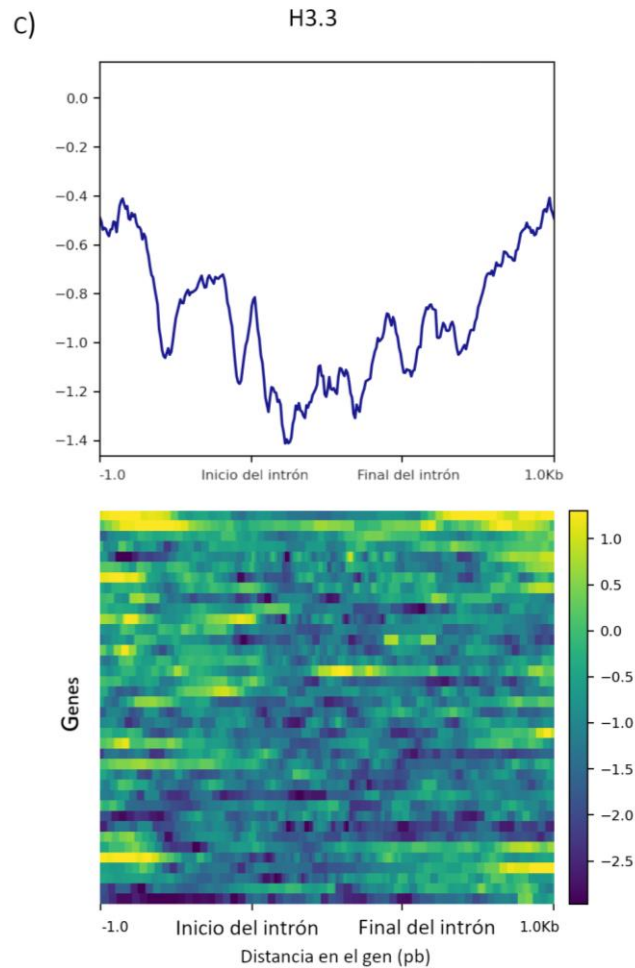


Fig. 15. Metaplots y mapas de calor de la distribución de las variantes H3.1, H3.3 y la marca de histonas H3K27me3 en el intrón largo de los MADS-box tipo MIKC. Se tomó en cuenta la información de todos los genes MADS-box tipo MIKC que contienen un intrón largo en su secuencia (38/45). El gen *XAL1/AGL12* contiene dos intrones largos, por lo que ambos fueron considerados para los análisis. Los datos utilizados para la elaboración de estos gráficos se obtuvieron de las publicaciones de You *et al.*, (2017) y Stroud *et al.*, (2012). Se usó como referencia la versión del TAIR10 del genoma de *Arabidopsis*.

Estos resultados nos llevaron a suponer que, en los intrones largos de los genes MADS-box tipo MIKC, existe una correlación entre la deposición de la marca de H3K27me3 y la presencia de la variante de histonas H3.1. Por esta razón, se llevó a cabo un análisis bioinformático que tomara en cuenta las coordenadas precisas de los intrones largos para alinearlas con el input de la variante o marca de histonas escogida.

En la **Fig. 15**, se encuentran los gráficos con los alineamientos de lecturas del input de la marca de H3K27me3, y las variantes de histonas H3.1 y H3.3 (**Fig. 15 A, B y C**, respectivamente) al genoma, utilizando los datos de You *et al.*, (2017) y Stroud *et al.*, (2012), respectivamente. Se tomó en cuenta la información de todos los genes MADS-box tipo MIKC que contienen un intrón largo en su secuencia (37/45, el gen *AGL11* y *ANR1/AGL44* contienen dos intrones largos, por lo que ambos fueron considerados).

Los gráficos de la marca de H3K27me3 (**Fig. 15 A**) muestran que la distribución de la marca tiene un pico en la porción inicial y media del intrón, y esta misma disminuye hacia la parte anterior y posterior al intrón. De manera similar, la variante de histonas H3.1 (**Fig. 15 B**) presenta un pico en la porción anterior inmediata al intrón, así como al inicio de este. La marca después disminuye en la región posterior al intrón.

Por otro lado, la variante H3.3 (**Fig. 15 C**) se ve agotada especialmente en la región inicial del intrón, aunque presenta niveles bajos en toda la región intrónica y en las porciones de una 1kb aledañas al intrón largo.

En resumen, los resultados de este trabajo indican la existencia de un patrón de regulación, donde, los intrones largos de los factores de transcripción MADS-box de mayor expresión en raíz, que contienen en su secuencia uno o más motivos RY, están relacionados con la deposición significativa de la marca de H3K27me3, así como de la variante H3.1.

Tabla 1: Resumen de distribución de las variantes de histonas H3.1, H3.3, distribución de las marcas de histonas de H3K4me3 y de H3K27me3, así como la interacción física de los factores de transcripción VAL1 y VAL2 con los 22 genes MAD5-box tipo MIKC de mayor expresión en la raíz de *Arabidopsis*. Datos de interacción fueron tomados de L. Yuan *et al.*, 2021. * representa que la interacción fue probada. + representa distribución significativa de las variantes y marcas en el gen.

ATG	Nombre del gen	Motivos RY	VAL1 target	VAL1-2 targets	VAL2 targets	Histona H3.1	Histona H3.3	H3K27me3	H3K4me3
At3g61120	AGL13	2 o +	*	*	*	+			+
At5g60910	FU1/AGL8	0	*	*	*	+	+		+
At3g30260	AGL79	2 o +				+		+	
At2g42830	SHP2/AGL5	2 o +				+	+		
At1g71692	XAL1/AGL12	1	*	*	*	+		+	
At5g65050	MAF2/AGL31	1					+		+
At1g77080	MAF1/AGL27	2 o +							
At5g10140	FLC	2 o +	*			+		+	
At2g22630	AGL17	0				+			
At4g37940	AGL21	2 o +				+			+
At2g14210	ANR1/AGL44	2 o +	*	*	*	+			
At5g13790	AGL15	0	*	*	*				
At3g57390	AGL18	0	*	*	*				
At2g22540	SVP	0	*	*	*		+		
At4g22950	AGL19	2 o +				+		+	
At4g11880	XAL2/AGL14	2 o +				+			
At2g45660	SOC1/AGL20	0	*	*	*				+
At5g51870	AGL71	1							+
At5g62165	AGL42	0			*	+			
At5g20240	PI	2 o +	*			+		+	

Discusión:

Una primera aproximación para encontrar a aquellos genes MADS-box tipo MIKC que se regulan a través de motivos en cis dentro de su secuencia, fue buscar una posible relación entre la cercanía filogenética de los genes y la expresión en el desarrollo de *Arabidopsis* (**Fig. 3**). Sin embargo, únicamente el clado de *FLC-like* (*AGL31*, *MAF1*, *AGL68*, *FLC2* y *FLC*), así como el clado A-related (*SEP1*, *SEP2*, *SEP3*, *AGL3*, *AGL13*, *AGL6*, *CAL*, *AP1*, *FUL*, y *AGL79*) presentan patrones de expresión similares en el desarrollo de la semilla y de flores y frutos, respectivamente. La filogenia utilizada en este trabajo (Martínez-Castilla & Alvarez-Buylla, 2003), así como la gran mayoría de filogenias que se han construido para los MADS-box (Becker & Theißen, 2003; Parenicova et al., 2003; Pelucchi et al., 2002), utilizan datos de regiones codificantes de los genes y excluyen aquellos de regiones no codificantes. La realización futura de filogenias que también consideren a las regiones no codificantes de genes será importante, puesto que cada vez vemos más trabajos en los que estas regiones son también importantes para que se lleve a cabo la función adecuada de los genes.

Como se mencionó anteriormente en este trabajo, la expresión de estos factores de transcripción es altamente dependiente de la etapa del ciclo de vida de la planta y, por lo tanto, su regulación transcripcional y sus funciones como reguladores de otros genes también son dependientes de la etapa del desarrollo en que se encuentre la planta (e.g. *XAL2/AGL14* (Garay-Arroyo et al., 2013; Pérez-Ruiz et al., 2015)). Además, como observamos, y coincidiendo con la literatura, un mismo factor de transcripción puede tener diferente expresión en los distintos órganos de la planta, y no funcionar como un gen de identidad de un órgano único o un tipo celular. Se ha reportado la expresión preferencial de factores de transcripción MADS-box tipo MIKC en tejidos de raíz respecto a la parte aérea (Alvarez-Buylla, Liljegren, et al., 2000), más no se han descrito los mecanismos epigenéticos por los cuales su expresión es regulada. Debido a esto, se decidió utilizar a aquellos MADS-box tipo MIKC de mayor expresión en raíz para realizar los análisis siguientes.

De los 22 genes MADS-box tipo MIKC seleccionados por presentar una mayor expresión en tejidos de la raíz, 19 genes conservan la presencia de un intrón largo (<500 pb). Dicho intrón se

encuentra ya sea en la primera o segunda posición dentro de los genes (**Fig. 4**). Cabe mencionar que aquellos genes que no conservan un intrón largo se encuentran en la base del árbol filogenético. En un principio, la presencia del intrón largo es interesante puesto que permite interpretar que este ha sido conservado a lo largo de la evolución por su posible papel en la regulación de los genes. Aunado a esto, una búsqueda de motivos RY dentro de los intrones largos, mostró que la mayoría (13/19 genes revisados) presenta por lo menos uno de los motivos, donde el gen *XAL2/AGL14* presenta cuatro de estos motivos. Además, se ha demostrado la regulación de algunos factores de transcripción MADS-box por miembros de la misma familia, a través de motivos conocidos como cajas CARG (Jetha et al., 2014). Incluso, algunos de ellos pueden autorregularse (De Folter et al., 2005; Tilly et al., 1998; Zhu & Perry, 2005). Se ha observado que, para que una caja CARG sea funcional es importante el contexto en el que se encuentra dentro de la secuencia de DNA. Por ejemplo, la distancia entre distintas cajas CARG, así como la presencia de motivos de unión de otros factores de transcripción cercanos a las cajas CARG (Aerts et al., 2018; Folter & Angenent, 2006). Una búsqueda de cajas CARG dentro de los intrones largos demostró que estos se encuentran conservados en todos los genes analizados (**Fig. 6**). Todos los genes presentaron por lo menos una de estas cajas, aún en aquellos genes que no presentan motivos RY. Entre las cajas CARG encontradas predomina la presencia de cajas no canónicas, y únicamente se encontró una caja CARG canónica dentro del intrón largo de *AGL19*. Cabe mencionar que tanto las cajas canónicas como las no canónicas son funcionales e importantes para la regulación transcripcional (Aerts et al., 2018; Folter & Angenent, 2006), además, se ha demostrado su importancia en la regulación por la maquinaria de Polycomb. Se han encontrado ejemplos en los que la unión de un factor de transcripción tipo MADS cercano a la región de nucleación de Polycomb, causa que las proteínas de PCG se separen del locus y lleven a la pérdida de la marca de H3K27me3 (Sun et al., 2014). Por lo tanto, determinar la funcionalidad de las cajas CARG encontradas dentro de los intrones largos de genes MADS-box, sería interesante para describir su posible papel en la regulación por Polycomb.

Por otra parte, los factores de transcripción de la familia B3 (e.g. LEC2, FUS3, ABI3) también presentan requerimientos similares para que ocurra su unión a los motivos RY (5'-CATGCA-3') (Baud et al., 2016; Mönke et al., 2012). Se ha observado que en los factores de transcripción LEC2,

FUS y ABI3, la unión a sus genes blanco es dependiente de la presencia de diferentes motivos con sustituciones en los sitios que rodean al *core* del motivo RY (5'-CATG-3'), nombrados motivos RY-Like (Baud et al., 2016). En el caso de los factores de transcripción VAL1 y VAL2, se ha probado *in vivo* que su unión es a la secuencia consenso 5'-CATGCA-3' (L. Yuan et al., 2021), aunque no se ha indagado sobre otros requerimientos para que estos motivos sean funcionales.

Con base en la regulación en cis reportada para *FLC* (Qüesta et al., 2016; Wu et al., 2018; L. Yuan et al., 2021; W. Yuan et al., 2016), buscamos marcas y variantes de histonas que pudieran permitirnos predecir la función de motivos RY dentro de los intrones largos de otros MADS-box tipo MIKC. Para *FLC*, se han propuesto mecanismos distintos mediante los cuales se lleva a cabo su silenciamiento transcripcional por parte de la maquinaria de PcG. Los factores de transcripción VAL1 y VAL2 son capaces de interactuar con proteínas de diferentes complejos que actúan durante el silenciamiento transcripcional, tanto en la deposición de la marca de H3K27me3 (PRC1 a través de LHP1, y PRC2 a través de MSI1, SWN y CLF), como aquellos que facilitan el esparcimiento de la marca a lo largo del gen (Qüesta et al., 2016; Sasnauskas et al., 2018; Wu et al., 2018; L. Yuan et al., 2021; W. Yuan et al., 2016). Debido a esto, se eligió revisar la distribución de la marca de H3K27me3 a lo largo de los 19 genes seleccionados, así como la marca de H3K4me3, una marca antagónica de PcG, que participa en la activación transcripcional. Por otro lado, en el proceso de represión también se ha visto involucrado el recambio de la variante de histona H3.3 por la variante H3.1, recambio que genera un sustrato que puede ser metilado rápidamente por PRC2 para mantener los niveles necesarios de H3K27me3 (Jiang & Berger, 2017a). A pesar de presentar heterogeneidad, el análisis *in silico* de estas marcas y variantes de histonas apunta a una correlación entre la presencia de motivos RY y niveles significativos de deposición de la marca de H3K27me3 y la variante H3.1 (**Fig. 7-9, Tabla 1**). La mitad de los genes que presentan uno o más motivos RY dentro de su intrón largo (*FLC*, *PI*, *AGL79*, *AGL14*, *AGL21*, *AGL19* y *XAL1/AGL12*) presentan una distribución significativa de la marca de H3K27me3, a la vez que presentan un enriquecimiento significativo de la variante de histonas H3.1. Mientras que ninguno de los genes que carecen de motivos RY en sus intrones largos (*FUL/AGL8*, *AGL17*, *AGL18*, *SVP*, *SOC1/AGL20* y *AGL42*) presenta niveles significativos de la marca de H3K27me3 y, en cambio presentan una distribución significativa de la marca de H3K4me3 o de la variante H3.3.

Para poder distinguir de mejor manera un patrón en la distribución de marcas y el enriquecimiento de las variantes de histonas, se llevó a cabo un análisis bioinformático que tomó en cuenta las coordenadas precisas de los intrones largos para alinearlas con la señal de la variante o marca de histonas escogida. Este análisis (**Fig. 10**), que contiene la información para todos aquellos MADS-box tipo MIKC que presentan un intrón largo, apoya la correlación observada entre la marca de H3K27me3 y la variante H3.1. El enriquecimiento en ambos casos tiene un pico en la porción inicial y media del intrón, mientras que la deposición se ve disminuida en las regiones de 1 Kb aledañas al intrón largo. Por otro lado, la variante H3.3 muestra una disminución importante especialmente en la región inicial del intrón, aunque presenta niveles bajos en toda la región del intrón. Esta distribución antagónica de la histona H3.1 y su variante H3.3, va en sentido con la sustitución de la Ala³¹ en H3.1 (Jiang & Berger, 2017b), puesto que se favorece la deposición de la marca de H3K27me3 y, por lo tanto, la represión de los factores de transcripción.

En la mayoría de los eucariontes se ha descrito una tendencia por los primeros intrones en las región 5' de los genes a tener una mayor longitud que los consecuentes hacia la región 3' (Bradnam & Korf, 2008). La razón por la cual es importante crear una distinción entre aquellos intrones cercanos al extremo 5' del resto es debido a su relevancia en la regulación de la transcripción de genes. En *Arabidopsis* y en arroz, por ejemplo, se han encontrado motivos responsables del incremento en transcripción mediado por intrones (*Intron-Mediated Enhancement*) proximales a la región promotora (Back & Walther, 2021; Rose et al., 2008). Más aún, se han descrito ejemplos de genes en *Caenorhabditis elegans* (Okkema et al., 1993), así como en *Arabidopsis* (Gallegos & Rose, 2017), donde el primer intrón tiene la flexibilidad de fungir como sitio de inicio de la transcripción ante la delección del promotor. Estudios de Yuan, et al., (2021), encontraron que a nivel genómico la distribución de VAL1 y VAL2 se encuentra flanqueando regiones intergénicas, con un pico de señal alto en la región cercana al sitio de inicio de la transcripción. Lo cual tiene sentido con los primeros intrones, siendo sitios importantes de regulación. A su vez, esto lo podemos ver con el pico de distribución de la marca de H3K27me3 que encontramos en los genes MADS-box tipo MIKC analizados en este trabajo.

Además de los genes MADS-box, en *Arabidopsis* también se ha encontrado la regulación por motivos RY en el intrón largo del gen *FLOWERING LOCUS T (FT)*, importante para la transición a la floración (Jing et al., 2019). Los motivos RY en *FT* se encuentran en el segundo intrón, y se encuentran conservados en sus homólogos de las especies cercanas *Arabidopsis lyrata*, *Brassica oleracea* y *Brassica rapa* (Jing et al., 2019). De manera similar al mecanismo descrito para *FLC* (Qüesta et al., 2016; W. Yuan et al., 2016), VAL1 participa en la represión de *FT* por PcG mediante la unión a los motivos RY y la subsecuente interacción con complejos de PRC1 y PRC2 (Jing et al., 2019). Patrones de regulación similares parecen conservarse evolutivamente, puesto que se han observado en otras especies de plantas. En la especie *Oryza sativa* (arroz), también se ha observado que la presencia de un motivo RY en el intrón largo (7.8 Kb) del gen *EUI1* es necesaria para reclutar a los factores de transcripción tipo B3 OsVAL2 y OsGD1 (Xie et al., 2018). *EUI1* codifica para una enzima que desactiva al ácido giberélico (GA), una fitohormona importante en el proceso de desarrollo de plantas, específicamente en la determinación de su longitud. Las plantas que presentan la delección de dicho motivo RY, resultan ser de menor tamaño, debido a que la represión de la expresión de *EUI1* no puede llevarse a cabo de manera correcta (no es reclutada la maquinaria de PcG ni el complejo de desacetilazas de histonas), lo cual genera que disminuyan los niveles de GA disponibles (Xie et al., 2018). Por otro lado, el contexto de la región génica en la que se localizan los motivos RY también resulta relevante para el tipo de regulación en el que participan. Alteraciones a motivos RY en el promotor de ciertos genes en las especies *Vicia faba*, *Arachis hypogaea* y *Glycine max* (haba, cacahuate y soya, respectivamente), han demostrado ser importantes para llevar a cabo la función óptima del promotor pero, además, causan que se pierda la expresión tejido específico de los mismos (Bäumlein et al., 1991; Fujiwara & Beachy, 1994; Neelakandan et al., 2022).

En su trabajo, Yuan *et al.*, (2021) realizaron experimentos donde demuestran la interacción de los factores de transcripción VAL1 y VAL2 con genes a lo largo del genoma. En la **Tabla 1** se encuentran resumidos los resultados de sus experimentos para los 19 genes de mayor expresión en raíz que fueron revisados en este trabajo. Al comparar las predicciones de los motivos RY realizadas en este trabajo, contra la unión reportada de VAL1 y VAL2 (Yuan *et al.*, 2021), encontramos algunas discrepancias. *AGL15*, *FUL/AGL8*, *AGL30*, *SOC1/AGL20*, *AGL18*, *AGL42* y

SVP, por ejemplo, no contienen ningún motivo RY dentro de su intrón largo, sin embargo, los datos de Yuan *et al.*, (2021) muestran que VAL1 y/o VAL2 sí presentan unión a dichos genes. Por otro lado, genes como *AGL79*, *AGL19*, *XAL2/AGL14*, y *AGL21*, a pesar de presentar dos o más motivos RY dentro de su intrón largo, no presentan la unión esperada de VAL1 y VAL2. Estas discrepancias podrían deberse al hecho de que se trata de genes que dan identidad a órganos y tejidos de las plantas, por lo que es importante tomar en cuenta que los resultados pueden variar dependiendo de la etapa de desarrollo en la que se encuentre la planta y el tejido u órganos utilizados. En sus experimentos, Yuan *et al.*, (2021) utilizaron plántulas completas, por lo que no es posible hacer conclusiones concretas de genes que tienen expresión en tiempos y espacios específicos (e.g. en raíz).

Conocer la regulación transcripcional llevada a cabo por elementos en cis dentro de intrones largos nos coloca ante una nueva visión sobre la manera en la que estudiamos la regulación génica en la actualidad. Estos trabajos nos hacen reconsiderar la importancia del modelo clásico de regulación transcripcional en el que los factores de transcripción se unen a secuencias conservadas de DNA dentro de los promotores. Aunque este modelo funciona para la mayoría de los genes, genera dificultades al momento de entender el comportamiento de casos especiales. Hasta ahora, los mecanismos de regulación transcripcional que más se han estudiado son aquellos en los que los intrones intervienen de manera positiva. Sin embargo, cada vez caemos más en cuenta de la importancia de los intrones en la represión de la transcripción.

Las predicciones realizadas en este trabajo sientan las bases moleculares para una amplia variedad de investigaciones futuras en genes particulares. Sería interesante confirmar experimentalmente la acción de los factores de transcripción VAL1 y VAL2 en la regulación negativa de aquellos genes con motivos RY dentro de sus intrones largos. Por ejemplo, podría determinarse la relevancia de una mayor cantidad de motivos RY respecto a aquellos que solamente tienen uno o ningún motivo. Así como el contexto en la secuencia de DNA que propicie la función de dichos motivos en la regulación negativa por la maquinaria de PcG.

Conclusiones

A partir de los datos analizados *in silico*, podemos llegar a las siguientes conclusiones:

- La expresión de los MADS-box tipo MIKC es dependiente de la etapa del desarrollo y del tejido que se estudie.
- Dentro de los intrones largos de la mayoría de los factores de transcripción MADS-box tipo MIKC se encuentran conservados motivos RY.
- Los intrones largos de los genes analizados en este estudio contienen cajas CArG dentro de su secuencia.
- Existe una correlación entre la presencia de motivos RY y el enriquecimiento de la marca de H3K27me3 y la variante de histona H3.1.
- La ausencia de motivos RY está relacionada negativamente con la deposición de la marca de H3K27me3 y la variante H3.1

Anexos

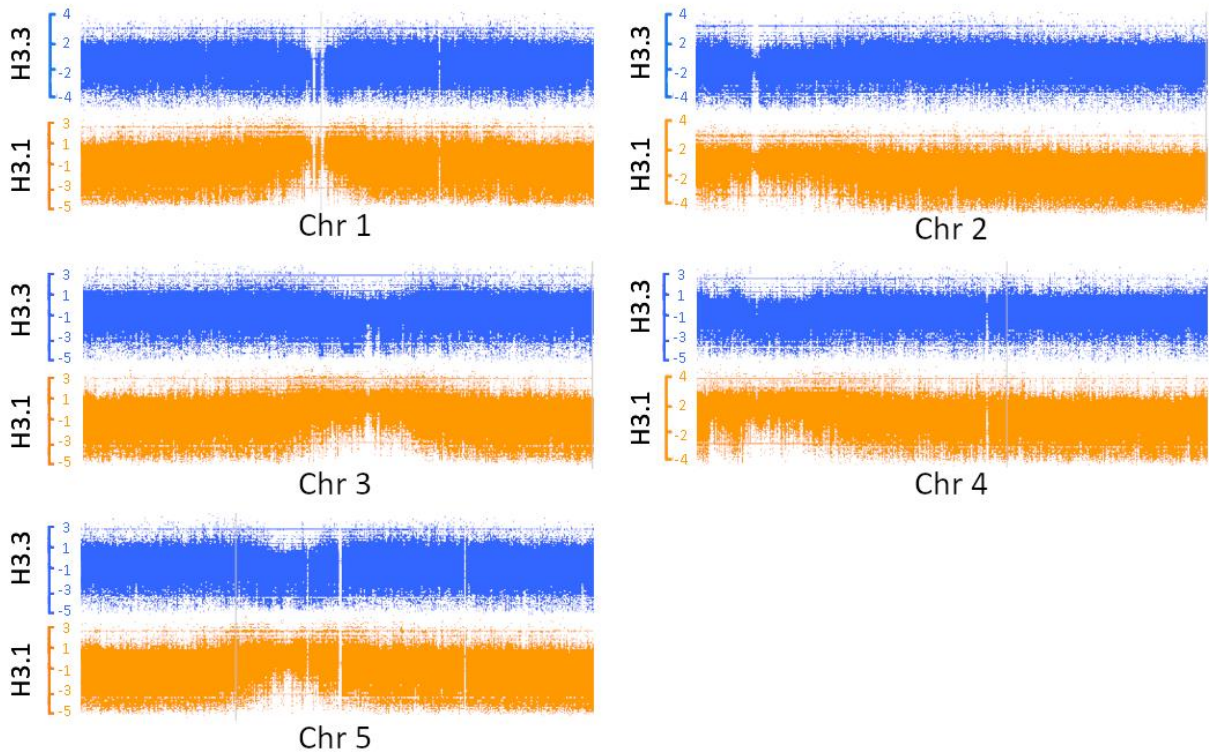


Fig. S1. Gráficos de distribución de variantes de histonas H3.1 y H3.3, en los cinco cromosomas de *Arabidopsis*. En color azul se muestran los datos de distribución de la variante H3.3, en naranja de la H3.1.

Referencias:

- Aerts, N., de Bruijn, S., van Mourik, H., Angenent, G. C., & van Dijk, A. D. J. (2018). *Comparative analysis of binding patterns of MADS-domain proteins in Arabidopsis thaliana*. BMC Plant Biology, 18(1), 1–16.
- Alejo Vinogradova, M. (2020). *El desarrollo de la raíz de Arabidopsis descrito a través de LHP1* (Maestría). Universidad Nacional Autónoma de México.
- Alvarez-Buylla, E. R., Benítez, M., Corvera-Poiré, A., Chaos Cador, Á., de Folter, S., Gamboa de Buen, A., Garay-Arroyo, A., García-Ponce, B., Jaimes-Miranda, F., Pérez-Ruiz, R. V., Piñeyro-Nelson, A., & Sánchez-Corrales, Y. E. (2010). *Flower Development*. The Arabidopsis Book, 8(1), e0127.
- Alvarez-Buylla, E. R., García-Ponce, B., Sánchez, M. de la P., Espinosa-Soto, C., García-Gómez, M. L., Piñeyro-Nelson, A., & Garay-Arroyo, A. (2019). *MADS-box genes underground becoming mainstream: plant root developmental mechanisms*. New Phytologist, 223(3), 1143–1158.
- Alvarez-Buylla, E. R., Liljegren, S. J., Pelaz, S., Gold, S. E., Burgeff, C., Ditta, G. S., Vergara-Silva, F., & Yanofsky, M. F. (2000). *MADS-box gene evolution beyond flowers: Expression in pollen, endosperm, guard cells, roots and trichomes*. Plant Journal, 24(4), 457–466.
- Alvarez-Buylla, E. R., Pelaz, S., Liljegren, S. J., Gold, S. E., Burgeff, C., Ditta, G. S., De Pouplana, L. R., Martínez-Castilla, L., & Yanofsky, M. F. (2000). *An ancestral MADS-box gene duplication occurred before the divergence of plants and animals*. Proceedings of the National Academy of Sciences of the United States of America, 97(10), 5328–5333.
- Alvarez-Venegas, R., Pien, S., Sadler, M., Witmer, X., Grossniklaus, U., & Avramova, Z. (2003). *ATX-1, an Arabidopsis homolog of trithorax, activates flower homeotic genes*. Current Biology, 13(8), 627–637.
- Andrés, F., & Coupland, G. (2012). *The genetic basis of flowering responses to seasonal cues*. Nature Reviews Genetics, 13(9), 627–639.
- Arteaga-Vazquez, M. A., & Chandler, V. L. (2010). *Paramutation in maize: RNA mediated trans-generational gene silencing*. Current Opinion in Genetics and Development, 20(2), 156–163.
- Arya, G., & Schlick, T. (2009). *A tale of tails: how histone tails mediate chromatin compaction in different salt and linker histone environments*. Journal of Physical Chemistry A, 113(16) 4045–4059.
- Back, G., & Walther, D. (2021). *Identification of cis-regulatory motifs in first introns and the prediction of intron-mediated enhancement of gene expression in Arabidopsis thaliana*. BMC Genomics, 22(1), 1–24.
- Baile, F., Gómez-Zambrano, Á., & Calonje, M. (2022). *Roles of Polycomb complexes in regulating gene expression and chromatin structure in plants*. Plant Communications, 3(1), 1–17.
- Baud, S., Kelemen, Z., Thévenin, J., Boulard, C., Blanchet, S., To, A., Payre, M., Berger, N., Effroy-Cuzzi, D., Franco-Zorrilla, J. M., Godoy, M., Solano, R., Thevenon, E., Parcy, F., Lepiniec, L., & Dubreucq, B. (2016). *Deciphering the molecular mechanisms underpinning the transcriptional control of gene expression by master transcriptional regulators in arabidopsis seed*. Plant Physiology, 171(2), 1099–1112.
- Baulcombe, D. C., & Dean, C. (2014). *Epigenetic regulation in plant responses to the environment*. Cold Spring Harbor Perspectives in Biology, 6(9), 1–20.
- Bäumlein, H., Nagy, I., Villarroel, R., Inzé, D., & Wobus, U. (1991). *Cis-analysis of a seed protein gene promoter: the conservative RY repeat CATGCATG within the legumin box is essential for tissue-specific expression of a legumin box*. The Plant Journal, 2(1 992), 233–239.
- Becker, A., & Theissen, G. (2003). *The major clades of MADS-box genes and their role in the development and evolution of flowering plants*. Molecular Phylogenetics and Evolution, 29(3), 464–489.
- Berry, S., & Dean, C. (2015). *Environmental perception and epigenetic memory: Mechanistic insight through FLC*. Plant Journal, 83(1), 133–148.
- Bevan, M., & Walsh, S. (2005). *The Arabidopsis genome: A foundation for plant research*. Genome Research, 15(12), 1632–

- Bowman, J. (Ed.). (2012). *Arabidopsis: an atlas of morphology and development*. Springer Science & Business Media.
- Bradnam, K. R., & Korf, I. (2008). *Longer first introns are a general property of eukaryotic gene structure*. *PLoS ONE*, 3(8).
- Bülow, L., Engelmann, S., Schindler, M., & Hehl, R. (2009). *AthaMap, integrating transcriptional and post-transcriptional data*. *Nucleic Acids Research*, 37(SUPPL. 1), 983–986.
- Carles, C. C., & Fletcher, J. C. (2009). *The SAND domain protein ULTRAPETALA1 acts as a trithorax group factor to regulate cell fate in plants*. *Genes and Development*, 23(23), 2723–2728.
- Coustham, V., Li, P., Strange, A., Lister, C., Song, J., & Dean, C. (2012). *Quantitative modulation of polycomb silencing underlies natural variation in vernalization*. *Science*, 337(6094), 584–587.
- Cui, J., Zhang, Z., Shao, Y., Zhang, K., Leng, P., & Liang, Z. (2015). *Genome-wide identification, evolutionary, and expression analyses of histone H3 variants in plants*. *BioMed Research International*, 2015.
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K., & Scheible, W. R. (2005). *Genome-wide identification and testing of superior reference genes for transcript normalization in arabidopsis*. *Plant Physiology*, 139(1), 5–17.
- De Folter, S., Immink, R. G. H., Kieffer, M., Pařenicová, L., Henz, S. R., Weigel, D., Busscher, M., Kooiker, M., Colombo, L., Kater, M. M., Davies, B., & Angenent, G. C. (2005). *Comprehensive interaction map of the Arabidopsis MADS box transcription factors*. *Plant Cell*, 17(5), 1424–1433.
- Della Pina, S., Souer, E., & Koes, R. (2014). *Arguments in the evo-devo debate: Say it with flowers!* *Journal of Experimental Botany*, 65(9), 2231–2242.
- Derkacheva, M., Steinbach, Y., Wildhaber, T., Mozgová, I., Mahrez, W., Nanni, P., Bischof, S., Gruissem, W., & Hennig, L. (2013). *Arabidopsis MSI1 connects LHP1 to PRC2 complexes*. *EMBO Journal*, 32(14), 2073–2085.
- Doğan, E. S., & Liu, C. (2018). *Three-dimensional chromatin packing and positioning of plant genomes*. *Nature Plants*, 4(8), 521–529.
- Fischle, W., Wang, Y., & Allis, C. D. (2003). *Histone and chromatin cross-talk*. *Current Opinion in Cell Biology*, 15(2), 172–183.
- Fletcher, J. C. (2017). *State of the art: TrxG factor regulation of post-embryonic plant development*. *Frontiers in Plant Science*, 8(November), 1–8.
- Folter, S. de, & Angenent, G. C. (2006). *Trans meets cis in MADS science*. *Trends in Plant Science*, 11(5), 224–231.
- Foroozani, M., Holder, D. H., & Deal, R. B. (2022). *Histone Variants in the Specialization of Plant Chromatin*. *Annual Review of Plant Biology*, 73, 149–172.
- Fujiwara, T., & Beachy, R. N. (1994). *Tissue-specific and temporal regulation of a β -conglycinin gene: roles of the RY repeat and other cis-acting elements*. *Plant Molecular Biology*, 24(2), 261–272.
- Gallegos, J. E., & Rose, A. B. (2017). *Intron DNA sequences can be more important than the proximal promoter in determining the site of transcript initiation*. *Plant Cell*, 29(4), 843–853.
- Gan, Y., Filleur, S., Rahman, A., Gotensparre, S., & Forde, B. G. (2005). *Nutritional regulation of ANR1 and other root-expressed MADS-box genes in Arabidopsis thaliana*. *Planta*, 222(4), 730–742.
- Garay-Arroyo, A., Ortiz-Moreno, E., De La Paz Sánchez, M., Murphy, A. S., García-Ponce, B., Marsch-Martínez, N., De Folter, S., Corvera-Poiré, A., Jaimes-Miranda, F., Pacheco-Escobedo, M. A., Dubrovsky, J. G., Pelaz, S., & Álvarez-Buylla, E. R. (2013). *The MADS transcription factor XAL2/AGL14 modulates auxin transport during Arabidopsis root development by regulating PIN expression*. *EMBO Journal*, 32(21), 2884–2895.
- García-Cruz, K. V., García-Ponce, B., Garay-Arroyo, A., De La Paz Sanchez, M., Ugartechea-Chirino, Y., Desvoyes, B., Pacheco-Escobedo, M. A., Tapia-López, R., Ransom-Rodríguez, I., Gutierrez, C., & Alvarez-Buylla, E. R. (2016). *The MADS-box XAANTAL1 increases proliferation at the Arabidopsis root stem-cell niche and participates in transition to differentiation by regulating cell-cycle components*. *Annals of Botany*, 118(4), 787–796.
- Gayon, J. (2016). *De Mendel à l'épigénétique : histoire de la génétique*. *Comptes Rendus - Biologies*, 339(7–8), 225–230.

- Gonzalez, D. H. (Ed.). (2015). *Plant transcription factors: evolutionary, structural and functional aspects*. Academic Press.
- Gramzow, L., & Theissen, G. (2010). *A hitchhiker's guide to the MADS world of plants*. *Genome Biology*, 11(6), 1–11.
- Grant, C. E., Bailey, T. L., & Noble, W. S. (2011). *FIMO: Scanning for occurrences of a given motif*. *Bioinformatics*, 27(7), 1017–1018.
- Gu, Q., Ferrándiz, C., Yanofsky, M. F., & Martienssen, R. (1998). *The FRUITFULL MADS-box gene mediates cell differentiation during Arabidopsis fruit development*. *Development*, 125(8), 1509–1517.
- Guo, L., Yu, Y., Law, J. A., & Zhang, X. (2010). *Set domain group2 is the major histone H3 lysine 4 trimethyltransferase in Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, 107(43), 18557–18562.
- Heitz, E. (1929). Heterochromatin, *Chromocentren, Chromomeren*. *Berichte Der Deutschen Botanischen Gesellschaft*, 47(4), 274–284.
- Hennig, L., Bouveret, R., & Gruissem, W. (2005). *MSI1-like proteins: An escort service for chromatin assembly and remodeling complexes*. *Trends in Cell Biology*, 15(6), 295–302.
- Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W., & Zimmermann, P. (2008). *Genevestigator V3: A Reference Expression Database for the Meta-Analysis of Transcriptomes*. *Advances in Bioinformatics*, 2008, 1–5.
- Jetha, K., Theißen, G., & Melzer, R. (2014). *Arabidopsis SEPALLATA proteins differ in cooperative DNA-binding during the formation of floral quartet-like complexes*. *Nucleic Acids Research*, 42(17), 10927–10942.
- Jiang, D., & Berger, F. (2017a). *DNA replication-coupled histone modification maintains Polycomb gene silencing in plants*. *Science*, 357(6356), 1146–1149.
- Jiang, D., & Berger, F. (2017b). *Histone variants in plant transcriptional regulation*. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1860(1), 123–130.
- Jing, Y., Guo, Q., & Lin, R. (2019). *The B3-domain transcription factor VAL1 regulates the floral transition by repressing flowering locus T*. *Plant Physiology*, 181(1), 236–248.
- Kim, J. M., Sasaki, T., Ueda, M., Sako, K., & Seki, M. (2015). *Chromatin changes in response to drought, salinity, heat, and cold stresses in plants*. *Frontiers in Plant Science*, 6(MAR), 1–12.
- Köhler, C., & Hennig, L. (2010). *Regulation of cell identity by plant Polycomb and trithorax group proteins*. *Current Opinion in Genetics and Development*, 20(5), 541–547.
- Köhler, C., Hennig, L., Bouveret, R., Gheyselink, J., Grossniklaus, U., & Gruissem, W. (2003). *Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development*. *EMBO Journal*, 22(18), 4804–4814.
- Köhler, C., Hennig, L., Spillane, C., Pien, S., Gruissem, W., & Grossniklaus, U. (2003). *The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1*. *Genes and Development*, 17(12), 1540–1553.
- Kooiker, M., Airoldi, C. A., Losa, A., Manzotti, P. S., Finzi, L., Kater, M. M., & Colombo, L. (2005). *Basic Pentacysteine1, a GA binding protein that induces conformational changes in the regulatory region of the homeotic Arabidopsis gene Seedstick*. *Plant Cell*, 17(3), 722–729.
- Kralemann, L. E. M., Liu, S., Trejo-Arellano, M. S., Muñoz-Viana, R., Köhler, C., & Hennig, L. (2020). *Removal of H2Aub1 by ubiquitin-specific proteases 12 and 13 is required for stable Polycomb-mediated gene repression in Arabidopsis*. *Genome Biology*, 21(1), 1–19.
- Lafos, M., Kröll, P., Hohenstatt, M. L., Thorpe, F. L., Clarenz, O., & Schubert, D. (2011). *Dynamic regulation of H3K27 trimethylation during Arabidopsis differentiation*. *PLoS Genetics*, 7(4).
- Langmead, B. & Salzberg, S.L. *Fast gapped-read alignment with Bowtie 2*. *Nat. Methods* 9, 357–359 (2012).
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... & Durbin, R. (2009). *The sequence alignment/map format and SAMtools*. *Bioinformatics*, 25(16), 2078–2079.

- Loidl, P. (2004). *A plant dialect of the histone language*. Trends in Plant Science, 9(2), 84–90.
- Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. (1997). *Crystal structure of the nucleosome core particle at 2.8 Å resolution*. Nature, 389(6648), 251–260.
- Lyle, A. (2014). Epigenetics (Primera ed). Garland Science.
- Margueron, R., & Reinberg, D. (2010). *Chromatin structure and the inheritance of epigenetic information*. Nature Reviews Genetics, 11(4), 285–296.
- Martínez-Castilla, L. P., & Alvarez-Buylla, E. R. (2003). *Adaptive evolution in the Arabidopsis MADS-box gene family inferred from its complete resolved phylogeny*. Proceedings of the National Academy of Sciences of the United States of America, 100(23), 13407–13412.
- McGarry, R. C., & Ayre, B. G. (2012). *Manipulating plant architecture with members of the CETS gene family*. Plant Science, 188–189, 71–81.
- Menon, G., Schulten, A., Dean, C., & Howard, M. (2021). *Digital paradigm for Polycomb epigenetic switching and memory*. Current Opinion in Plant Biology, 61, 1–9.
- Meyerowitz, E. M. (1987). *Arabidopsis thaliana*. Annual Review of Genetics, 21(1), 93–111.
- Meyerowitz, E. M. (1994). *Plant developmental biology: green genes for the 21st century*. BioEssays, 16(9), 621–625.
- Meyerowitz, E. M. (2001). *Prehistory and history of Arabidopsis research*. Plant Physiology, 125(1), 15–19.
- Molitor, A., & Shen, W. H. (2013). *The Polycomb Complex PRC1: Composition and Function in Plants*. Journal of Genetics and Genomics, 40(5), 231–238.
- Mönke, G., Seifert, M., Keilwagen, J., Mohr, M., Grosse, I., Hähnel, U., Junker, A., Weisshaar, B., Conrad, U., Bäumlein, H., & Altschmied, L. (2012). *Toward the identification and regulation of the Arabidopsis thaliana ABI3 regulon*. Nucleic Acids Research, 40(17), 8240–8254.
- Nam, J., DePamphilis, C. W., Ma, H., & Nei, M. (2003). *Antiquity and evolution of the MADS-box gene family controlling flower development in plants*. Molecular Biology and Evolution, 20(9), 1435–1447.
- Nawy, T., Lee, J. Y., Colinas, J., Wang, J. Y., Thongrod, S. C., Malamy, J. E., Birnbaum, K., & Benfey, P. N. (2005). *Transcriptional profile of the arabidopsis root quiescent center*. Plant Cell, 17(7), 1908–1925.
- NEBION. (2021). *Genevestigator*. Manual de usuario. Extraído de:
https://genevestigator.com/userdocs/manual/GENEVESTIGATOR_UserManual.pdf
- Neelakandan, A. K., Wright, D. A., Traore, S. M., Chen, X., Spalding, M. H., & He, G. (2022). *CRISPR/Cas9 Based Site-Specific Modification of FAD2 cis-Regulatory Motifs in Peanut (Arachis hypogaea L)*. Frontiers in Genetics, 13(April), 1–10.
- Ng, M., & Yanofsky, M. F. (2001). *Function and evolution of the plant MADS-box gene family*. Nature Reviews Genetics, 2(3), 186–195.
- Nicol, J. W., Helt, G. A., Blanchard, S. G., Raja, A., & Loraine, A. E. (2009). *The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets*. BMC Bioinformatics, 25(20), 2730–2731.
- Nie, X., Wang, H., Li, J., Holec, S., & Berger, F. (2014). *The HIRA complex that deposits the histone H3.3 is conserved in Arabidopsis and facilitates transcriptional dynamics*. Biology Open, 3(9), 794–802.
- Okkema, P. G., Harrison, S. W., Plunger, V., Aryana, A., & Fire, A. (1993). *Sequence requirements for myosin gene expression and regulation in Caenorhabditis elegans*. Genetics, 135(2), 385–404.
- Ornelas-ayala, D., Garay-arroyo, A., García-ponce, B., Álvarez-buylla, E. R., De, M., & Sanchez, P. (2021). *The Epigenetic Faces of*. Frontiers in Plant Science, 12(February), 1–8.
- Ornelas-Ayala, D., Vega-León, R., Petrone-Mendoza, E., Garay-Arroyo, A., García-Ponce, B., Álvarez-Buylla, E. R., & Sanchez, M. de la P. (2020). *ULTRAPETALA1 maintains Arabidopsis root stem cell niche independently of ARABIDOPSIS TRITHORAX1*. New Phytologist, 225(3), 1261–1272.

- Otero, S., Desvoyes, B., Peiró, R., & Gutierrez, C. (2016). *Histone H3 dynamics reveal domains with distinct proliferation potential in the arabidopsis root*. *Plant Cell*, 28(6), 1361–1371.
- Parenicova, L., de Folter, S., Kieffer, M., Horner, D. S., Favalli, C., Busscher, J., ... & Colombo, L. (2003). Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world. *The Plant Cell*, 15(7), 1538-1551. Pelucchi, N., Fornara, F., Favalli, C., Masiero, S., Lago, C., Pè, E. M., Colombo, L., & Kater, M. M. (2002). *Comparative analysis of rice MADS-box genes expressed during flower development*. *Sexual Plant Reproduction*, 15(3), 113–122.
- Pérez-Ruiz, R. V., García-Ponce, B., Marsch-Martínez, N., Ugartechea-Chirino, Y., Villajuana-Bonequi, M., De Folter, S., Azpeitia, E., Dávila-Velderrain, J., Cruz-Sánchez, D., Garay-Arroyo, A., Sánchez, M. D. L. P., Estévez-Palmas, J. M., & Álvarez-Buylla, E. R. (2015). *XAANTAL2 (AGL14) is an important component of the complex gene regulatory network that underlies arabidopsis shoot apical meristem transitions*. *Molecular Plant*, 8(5), 796–813.
- Pikaard, C. S., & Scheid, O. M. (2014). *Epigenetic regulation in plants*. *Cold Spring Harbor Perspectives in Biology*, 6(12), 1–32.
- Perez, F. (2004). *Serial Cloner v. 2.6. 0. Software SerialBasics*.
- Poethig, R. S. (2003). *Phase Change and the Regulation of Developmental Timing in Plants*. 301(July), 334–337.
- Probst, A. V., Desvoyes, B., & Gutierrez, C. (2020). *Similar yet critically different: The distribution, dynamics and function of histone variants*. *Journal of Experimental Botany* (Vol. 71, Issue 17, pp. 5191–5204).
- Qüesta, J. I., Song, J., Geraldo, N., An, H., & Dean, C. (2016). *Arabidopsis transcriptional repressor VAL1 triggers Polycomb silencing at FLC during vernalization*. *Science*, 353(6298), 485–488.
- Ramirez-Parra, E., & Gutierrez, C. (2007). *The many faces of chromatin assembly factor 1*. In *Trends in Plant Science*, 12(12), 570-576.
- Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F., & Manke, T. (2016). *deepTools2: a next generation web server for deep-sequencing data analysis*. *Nucleic acids research*, 44(W1), W160-W165.
- Ravi, M., Shibata, F., Ramahi, J. S., Nagaki, K., Chen, C., Murata, M., & Chan, S. W. L. (2011). *Meiosis-Specific loading of the Centromere-Specific histone CENH3 in Arabidopsis thaliana*. *PLoS Genetics*, 7(6), e1002121.
- Riechmann, J. L., & Ratcliffe, O. J. (2000). *A genomic perspective on plant transcription factors*. *Current Opinion in Plant Biology*, 3(5), 423–434.
- Rose, A. B., Elfers, T., Parra, G., & Korf, I. (2008). *Promoter-proximal introns in Arabidopsis thaliana are enriched in dispersed signals that elevate gene expression*. *Plant Cell*, 20(3), 543–551.
- Sanchez, M., Aceves-García, P., Petrone, E., Steckenborn, S., Vega-León, R., Álvarez-Buylla, E. R., Garay-Arroyo, A., & García-Ponce, B. (2015). *The impact of Polycomb group (PcG) and Trithorax group (TrxG) epigenetic factors in plant plasticity*. *New Phytologist*, 208(3), 684–694.
- Sasnauskas, G., Kauneckaitė, K., & Siksnys, V. (2018). *Structural basis of DNA target recognition by the B3 domain of Arabidopsis epigenome reader VAL1*. *Nucleic Acids Research*, 46(8), 4316–4324.
- Schauer, S. E., Schlüter, P. M., Baskar, R., Gheyselinck, J., Bolaños, A., Curtis, M. D., & Grossniklaus, U. (2009). *Intronic regulatory elements determine the divergent expression patterns of agamous-like6 subfamily members in arabidopsis*. *Plant Journal*, 59(6), 987–1000.
- Scheres, B., Benfey, P., & Dolan, L. (2002). *Root Development*. *The Arabidopsis Book*, 1, e0101.
- Schönrock, N., Bouveret, R., Leroy, O., Borghi, L., Köhler, C., Gruissem, W., & Hennig, L. (2006). *Polycomb-group proteins repress the floral activator AGL19 in the FLC-independent vernalization pathway*. *Genes and Development*, 20(12), 1667–1678.
- Schwartz, Y. B., & Pirrotta, V. (2007). *Polycomb silencing mechanisms and the management of genomic programmes*. *Nature Reviews Genetics*, 8(1), 9–22.
- Sieburth, L. E., & Meyerowitz, E. M. (1997). *Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically*. *Plant Cell*, 9(3), 355–365.

- Smaczniak, C., Immink, R. G. H., Angenent, G. C., & Kaufmann, K. (2012). *Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies*. *Development*, 139(17), 3081–3098.
- Steffens, N. O., Galuschka, C., Schindler, M., Bülow, L., & Hehl, R. (2004). *AthaMap: An online resource for in silico transcription factor binding sites in the Arabidopsis thaliana genome*. *Nucleic Acids Research*, 32(DATABASE ISS.), 368–372.
- Stroud, H., Otero, S., Desvoyes, B., Ramírez-parra, E., Jacobsen, S. E., & Gutierrez, C. (2012). *Genome-wide analysis of histone H3.1 and H3.3 variants in Arabidopsis thaliana*. *PNAS*, 109(14), 5370–5375.
- Sun, B., Looi, L. S., Guo, S., He, Z., Gan, E. S., Huang, J., Xu, Y., Wee, W. Y., & Ito, T. (2014). *Timing mechanism dependent on cell division is invoked by Polycomb eviction in plant stem cells*. *Science*, 343(6170).
- Tapia-López, R., García-Ponce, B., Dubrovsky, J. G., Garay-Arroyo, A., Pérez-Ruiz, R. V., Kim, S. H., Acevedo, F., Pelaz, S., & Alvarez-Buylla, E. R. (2008). *An AGAMOUS-related MADS-box gene, XAL1 (AGL12), regulates root meristem cell proliferation and flowering transition in Arabidopsis*. *Plant Physiology*, 146(3), 1182–1192.
- Tilly, J. J., Allen, D. W., & Jack, T. (1998). *The CARG boxes in the promoter of the Arabidopsis floral organ identity gene APETALA3 mediate diverse regulatory effects*. *Development*, 125(9), 1647–1657.
- Truskina, J., Han, J., Chrysanthou, E., Galvan-Ampudia, C. S., Lainé, S., Brunoud, G., Macé, J., Bellows, S., Legrand, J., Bågman, A. M., Smit, M. E., Smetana, O., Stigliani, A., Porco, S., Bennett, M. J., Mähönen, A. P., Parcy, F., Farcot, E., Roudier, F., ... Vernoux, T. (2021). *A network of transcriptional repressors modulates auxin responses*. *Nature*, 589(7840), 116–119.
- Tsakagoshi, H., Morikami, A., & Nakamura, K. (2007). *Two B3 domain transcriptional repressors prevent sugar-inducible expression of seed maturation genes in Arabidopsis seedlings*. *Proceedings of the National Academy of Sciences of the United States of America*, 104(7).
- Vega-León, M. (2018). *Función de los factores epigenéticos ATX1 y CLF en la regulación de genes de la familia MADS-box en Arabidopsis thaliana* (Maestría). Universidad Nacional Autónoma de México.
- Wang, Q., & Shen, W. H. (2018). *Chromatin modulation and gene regulation in plants: Insight about prc1 function*. *Biochemical Society Transactions*, 46(4), 957–966.
- Whittaker, C., & Dean, C. (2017). *The FLC locus: A platform for discoveries in epigenetics and adaptation*. *Annual Review of Cell and Developmental Biology*, 33(June), 555–575.
- Wollmann, H., Holec, S., Alden, K., Clarke, N. D., Jacques, P. É., & Berger, F. (2012). *Dynamic deposition of histone variant H3.3 accompanies developmental remodeling of the Arabidopsis transcriptome*. *PLoS Genetics*, 8(5), 28–31.
- Wu, B., Zhang, M., Su, S., Liu, H., Gan, J., & Ma, J. (2018). *Structural insight into the role of VAL1 B3 domain for targeting to FLC locus in Arabidopsis thaliana*. *Biochemical and Biophysical Research Communications*, 501(2), 415–422.
- Xie, Y., Zhang, Y., Han, J., Luo, J., Li, G., Huang, J., Wu, H., Tian, Q., Zhu, Q., Chen, Y., Kawano, Y., Liu, Y. G., & Chen, L. (2018). *The Intronic cis Element SE1 Recruits trans-Acting Repressor Complexes to Repress the Expression of ELONGATED UPPERMOST INTERNODE1 in Rice*. *Molecular Plant*, 11(5), 720–735.
- Xu, F., Kuo, T., Rosli, Y., Liu, M. Sen, Wu, L., Chen, L. F. O., Fletcher, J. C., Sung, Z. R., & Pu, L. (2018). *Trithorax Group Proteins Act Together with a Polycomb Group Protein to Maintain Chromatin Integrity for Epigenetic Silencing during Seed Germination in Arabidopsis*. *Molecular Plant*, 11(5), 659–677.
- Yang, H., Berry, S., Olsson, T. S. G., Hartley, M., Howard, M., & Dean, C. (2017). *Distinct phases of Polycomb silencing to hold epigenetic memory of cold in Arabidopsis*. *Science*, 357(6356), 1142–1145.
- Yao, X., Feng, H., Yu, Y., Dong, A., & Shen, W. H. (2013). *SDG2-Mediated H3K4 Methylation Is Required for Proper Arabidopsis Root Growth and Development*. *PLoS ONE*.
- You, Y., Sawikowska, A., Neumann, M., Posé, D., Capovilla, G., Langenecker, T., Neher, R. A., Krajewski, P., & Schmid, M. (2017). *Temporal dynamics of gene expression and histone marks at the Arabidopsis shoot meristem during flowering*. *Nature Communications*, 8(May 2016).
- Yuan, L., Song, X., Zhang, L., Yu, Y., Liang, Z., Lei, Y., Ruan, J., Tan, B., Liu, J., & Li, C. (2021). *The transcriptional repressors VAL1 and VAL2 recruit PRC2 for genome-wide Polycomb silencing in Arabidopsis*. *Nucleic Acids Research*, 49(1), 98–113.
- Yuan, W., Luo, X., Li, Z., Yang, W., Wang, Y., Liu, R., Du, J., & He, Y. (2016). *A cis cold memory element and a trans epigenome*

reader mediate Polycomb silencing of FLC by vernalization in Arabidopsis. Nature Genetics, 48(12), 1527–1534.

Zhang, X., Bernatavichute, Y. V., Cokus, S., Pellegrini, M., & Jacobsen, S. E. (2009). *Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in Arabidopsis thaliana.* Genome Biology, 10(6), 1–14.

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y. V., Pellegrini, M., Goodrich, J., & Jacobsen, S. E. (2007). *Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis.* PLoS Biology.

Zhao, T., Holmer, R., De Bruijn, S., Angenent, G. C., Van Den Burg, H. A., & Eric Schranz, M. (2017). *Phylogenomic synteny network analysis of MADS-box transcription factor genes reveals lineage-specific transpositions, ancient tandem duplications, and deep positional conservation.* Plant Cell, 29(6), 1278–1292.

Zhu, C., & Perry, S. E. (2005). *Control of expression and autoregulation of AGL15, a member of the MADS-box family.* Plant Journal, 41(4), 583–594.