

Instituto de Bioquímica Vegetal y Fotosíntesis
Departamento de Bioquímica Vegetal y Biología Molecular
CSIC – Universidad de Sevilla

PAPEL DE LA REGULACIÓN POLYCOMB EN LA PLASTICIDAD CELULAR DE LAS PLANTAS

Trabajo realizado para optar al grado de Doctor por
la Universidad de Sevilla por la Licenciada

Wiam Merini



Directora

Dra. Myriam Calonje Macaya

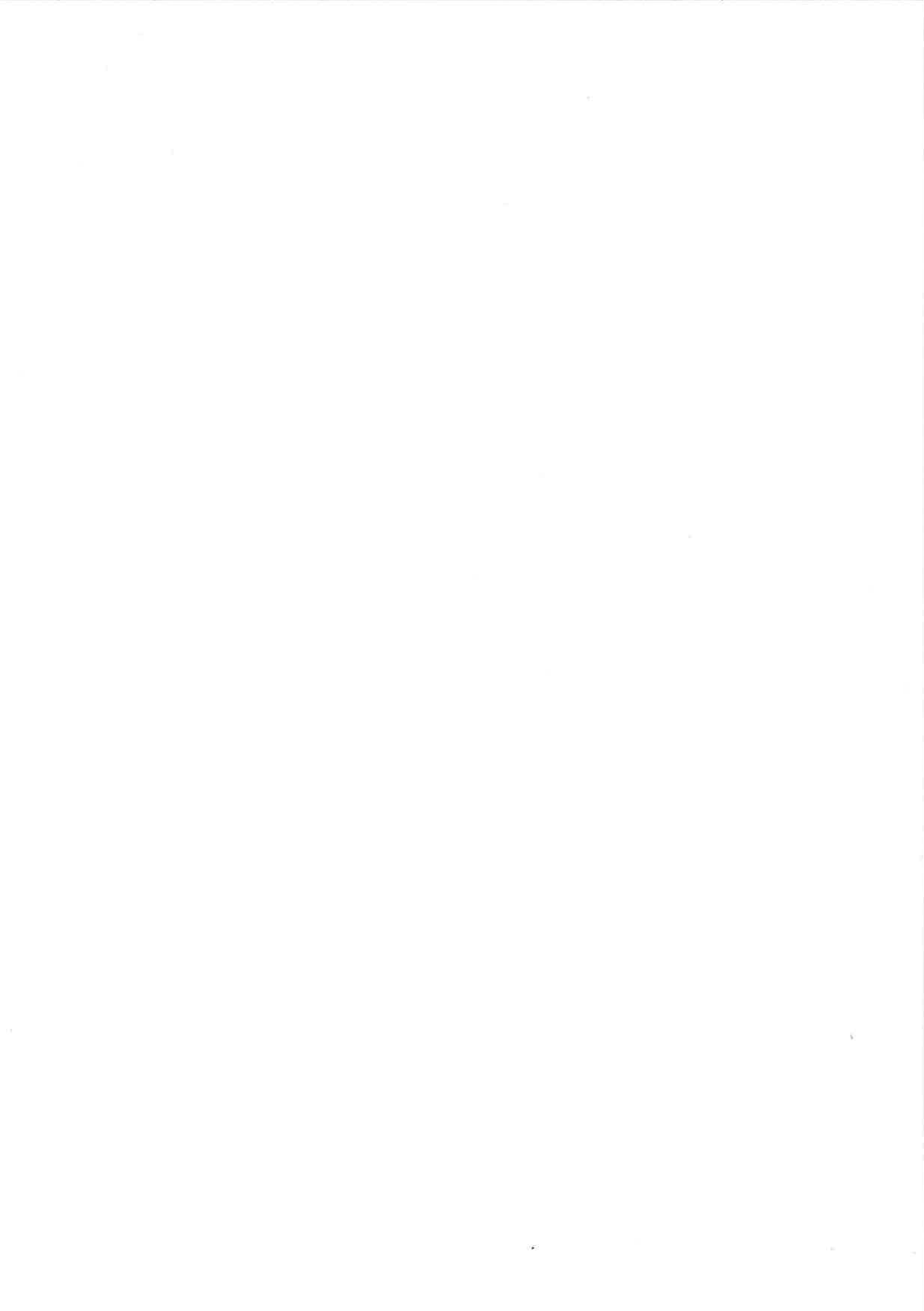
Científica Titular (CSIC)



Tutor

Dr. José María Romero Rodríguez

Catedrático de Bioquímica y Biología Molecular (US)





Este trabajo ha sido realizado en el Instituto de Bioquímica Vegetal y Fotosíntesis de Sevilla (IBVF) y ha sido subvencionado por los proyectos BIO2013-44078-P y BIO2016-76457-P del ministerio de Economía y Competitividad (MINECO) cuya investigadora principal ha sido Dra. Myriam Calonje Macaya.

Wiam Merini fue beneficiaria de un contrato predoctoral para formación de personal investigador (FPI-CSIC) del ministerio de Economía y Competitividad, con referencia **BES2014-069229** adscrito al proyecto BIO2013-44078-P para la realización de esta tesis doctoral.

A mis Padres

AGRADECIMIENTOS

Me gustaría comenzar esta tesis agradeciendo el inmenso apoyo recibido tanto a nivel personal como profesional.

En primer lugar, quiero agradecer a mi directora Dra. Myriam Calonje por haberme dado la oportunidad de realizar mi tesis doctoral en su grupo de investigación. Gracias por su dedicación y su esfuerzo para enseñarme y guiarme en todo el proceso de desarrollo de este trabajo. También a mi tutor Dr. José María Romero por todos sus consejos y su apoyo y a Dr. Francisco Romero por su colaboración en desarrollar los análisis bioinformáticos realizados en esta tesis y por todos los conocimientos que me ha transmitido en este campo.

Gracias a Dra. Cecilia Gotor y Dr. Luis Romero por haberme dado la oportunidad de dar mis primeros pasos en el mundo de la investigación científica. A Ángeles, Irene e Inma por todo lo que he aprendido de vosotras durante mi estancia en el Lab I4 y también durante la tesis, siempre habéis estado dispuestas a echarme una mano cuando lo necesitaba. Gracias a Ana no solo por “introducirme” a la investigación y apoyarme durante los años de la tesis, sino por todos estos años de amistad y los momentos compartidos desde el primer año de la facultad y sobre todo por formar parte muy importante de mi “familia española”. A Juanma, por nuestras conversaciones que siempre me ayudaban a no desanimarme y por tu valiosa amistad.

Gracias a todos los miembros de la unidad de desarrollo de plantas, por la buena acogida en el grupo. A Dr. Federico Valverde por las sugerencias y consejos. A Isa, por tu excelente labor en el laboratorio, que sin ella nuestro trabajo sería mucho más difícil y lento, gracias por tu alegría y tu entrega. A Tere, por estar siempre dispuesta a ayudar. A los alumnos internos que han contribuido al desarrollo de este trabajo: Sara, Mila, Joaquín e Inés. A Eva, la primera fan de mi “chimbo-tarta”. A Sandra, por la gran ayuda para recoger tejido. A Gloria, por esa actitud positiva siempre. A Pedro, por ayudarme a entender estos códigos de bioinformática que al principio me parecían imposibles. A Fernando, por tu ayuda, tu alegría y la pasión que le pones al trabajo, estoy segura que vas a tener mucho éxito. De todos vosotros he aprendido muchísimo.

A todos mis compañeros: Isidro, Lucía, Valle, Víctor, Diego, Manuel, Javi, Laura y Ana por todos los momentos compartidos dentro y fuera del laboratorio.

Para no olvidar a nadie, gracias a todas las personas que he conocido en el IBVF que me ayudaron de alguna forma durante esta etapa.

A mis padres y mi hermano, todo lo que he conseguido ha sido gracias a vosotros. Gracias por el gran esfuerzo y sacrificio que supuso mandarme a estudiar en el extranjero. Por vuestro apoyo incondicional y la fuerza que me transmitís para poder llegar a mis metas. A toda mi familia que siempre me apoyan y me animan, especialmente a mis tíos y mis primos Anas, Safaa, Munaim y Zineb.

ÍNDICE

RESUMEN	5 -
INTRODUCCIÓN GENERAL	9 -
1. ORGANIZACIÓN DE LA CROMATINA DENTRO DEL NÚCLEO	9 -
2. MARCAS EPIGENÉTICAS.....	11 -
3. LAS PROTEÍNAS PCG ESTÁN EVOLUTIVAMENTE CONSERVADAS	13 -
3.1. Componentes PRC2	15 -
3.2. Componentes PRC1	17 -
3.2.1. Actividades del complejo PRC1	20 -
MATERIALES Y MÉTODOS	25 -
1. MATERIAL VEGETAL Y CONDICIONES DE CRECIMIENTO	25 -
1.1. Esterilización de semillas.....	25 -
1.2. Germinación de semillas en cámaras de cultivo <i>in vitro</i>	25 -
1.3. Cultivo de plantas en tierra	25 -
1.4. Cruces de plantas.....	26 -
1.5. Transformación de plantas por inmersión floral	26 -
1.5.1. Generación de las células competentes de <i>A. tumefaciens</i>	26 -
1.5.2. Transformación de <i>A. tumefaciens</i>	26 -
1.5.3. Preparación del medio de <i>A.tumefaciens</i> para la inmersión floral	27 -
1.5.4. Procedimiento de inmersión floral	27 -
1.5.5. Selección de transformantes positivos y genotipado.....	27 -
2. BIOLOGÍA MOLECULAR	28 -
2.1. Extracción de DNA.....	28 -
2.1.1. Método CTAB.....	28 -
2.1.2. Extracción rápida de DNA para genotipado.....	28 -
2.2. Reacción en cadena de la polimerasa (PCR)	29 -
2.2.1. Diseño de oligonucleótidos.....	29 -
2.2.2. PCR utilizando Taq Polimerasa	29 -
2.2.3. PCR utilizando polimerasa de alta fidelidad	29 -
2.3. Electroforesis en geles de agarosa	30 -
2.3.1. Separación de fragmentos de DNA.....	30 -
2.3.2. Aislamiento de los fragmentos de DNA separados.....	30 -

2.4. Extracción de RNA	- 30 -
2.5. Reacción de retrotranscripción (RT)	- 31 -
2.6. RT-PCR cuantitativa (qRT-PCR).....	- 31 -
2.6.1. Diseño de oligonucleótidos.....	- 31 -
2.6.2. Reacciones de qRT-PCR.....	- 31 -
2.6.3. Interpretación de los datos de qRT-PCR	- 32 -
2.7. PCR semicuantitativa (RT-PCR)	- 33 -
2.8. Técnicas de clonaje molecular.....	- 33 -
2.8.1. Clonaje utilizando el sistema de vectores pSPARKII	- 33 -
2.8.2. Clonaje basado en endonucleasas de restricción	- 33 -
2.8.3. Sistema de recapitulación sintética: Construcciones y generación de plantas transgénicas	- 35 -
3. MÉTODOS DE ANÁLISIS Y MANIPULACIÓN DE PROTEÍNAS	- 37 -
3.1. Extracción de proteínas totales.....	- 37 -
3.2. Electroforesis de proteínas en geles de SDS-poliacrilamida.....	- 37 -
3.3. Análisis mediante Western blot.....	- 38 -
4. INMUNOPRECIPITACIÓN DE CROMATINA	- 38 -
4.1. Fijación del material vegetal.....	- 38 -
4.2. Aislamiento de la cromatina.....	- 38 -
4.2.1. Extracción de núcleos.....	- 39 -
4.2.2. Lysis de los núcleos y sonicación de cromatina	- 39 -
4.2.3. Inmunoprecipitación.....	- 39 -
4.2.4. Purificación del DNA: Extracción Fenol-Cloroformo	- 40 -
5. DETERMINACIÓN DE LA ACTIVIDAD B-GLUCURONIDASA (GUS): ENSAYO FLUORIMÉTRICO MUG....	- 41 -
6. ANÁLISIS TRANSCRIPTÓMICO MASIVO BASADO EN DATOS DE RNA-SEQ	- 41 -
6.1. Diseño experimental y secuenciación masiva	- 41 -
6.2. Flujo de trabajo: Protocolo TUXEDO.....	- 42 -
6.3. Análisis de Clustering.....	- 44 -
6.4. Análisis del enriquecimiento de motivos de DNA en los promotores.....	- 44 -
7. ANÁLISIS DE POSICIONAMIENTO Y OCUPACIÓN DE NUCLEOSOMAS BASADO EN ANÁLISIS DE MNASA- SEQ	- 45 -
7.1. Protocolo de digestión con MNasa	- 45 -
7.1.1. Extracción de núcleos	- 45 -

7.1.2. Tratamiento con MNasa	- 45 -
7.1.3. Extracción y precipitación del DNA.....	- 45 -
7.1.4. Purificación del DNA.....	- 45 -
7.2. <i>Preparación de la librería y secuenciación</i>	- 46 -
7.3. <i>Procesamiento de los datos de MNase-seq</i>	- 46 -
CAPÍTULO I: REDES DE REGULACIÓN DE LOS COMPONENTES PRC1.....	- 51 -
1. INTRODUCCIÓN.....	- 51 -
2. OBJETIVO.....	- 56 -
3. RESULTADOS.....	- 57 -
3.1. <i>La acción coordinada de AtBMI1 y EMF1 es necesaria para la transición de la fase juvenil a la fase reproductiva</i>	- 57 -
3.2. <i>Análisis del transcriptoma global de los distintos mutantes atbmi1</i>	- 62 -
3.2.1. Programas de desarrollo desregulados en los mutantes <i>atbmi1</i>	- 72 -
3.2.2. Interacciones reguladoras entre los distintos complejos de cromatina	- 75 -
3.2.3. Reguladores centrales en el programa de transición y desarrollo floral se encuentran inhibidos en el mutante <i>atbmi1a/b/c</i>	- 76 -
3.2.4. VAL1/2 y AtBMI1 co-regulan un conjunto de dianas de AtBMI1	- 77 -
3.2.5. Redes de regulación de AtBMI1, EMF1 y LHP1.....	- 79 -
4. DISCUSIÓN	- 82 -
4.1. <i>Los componentes PRC1 AtBMI1 y EMF1 juegan papeles coordinados en la regulación de la transición de fase juvenil a reproductiva</i>	- 82 -
4.2. <i>El análisis de las redes de regulación de los distintos componentes PRC1 apunta tanto a funciones compartidas entre ellos como a específicas</i>	- 84 -
CAPÍTULO II: EFECTO DE LAS MARCAS H2AK121UB Y H3K27ME3 EN LA ORGANIZACIÓN DE NUCLEOSOMAS EN ARABIDOPSIS	- 93 -
1. INTRODUCCIÓN.....	- 93 -
1.1. <i>Alineamiento de los nucleosomas en el genoma</i>	- 93 -
1.2. <i>La organización de los nucleosomas en las regiones codificantes del genoma</i>	- 95 -
1.3. <i>Determinantes del posicionamiento de los nucleosomas</i>	- 96 -
1.3.1. La secuencia del DNA.....	- 96 -
1.3.2. Transcripción.....	- 96 -
1.3.3. Remodeladores de la cromatina dependientes de ATP	- 97 -
1.3.4. Modificaciones post-traduccionales de histonas (PTM).....	- 99 -
2. OBJETIVO.....	- 100 -

3. RESULTADOS.....	- 102 -
3.1. <i>Obtención de muestras para MNase-seq</i>	- 102 -
3.2. <i>Los distintos grupos de genes diana PcG presentan diferencias en la organización de nucleosomas</i>	- 104 -
3.3. <i>La comparación de perfiles de nucleosomas obtenidos en distintas condiciones de ensayo proporciona información complementaria</i>	- 106 -
3.4. <i>El análisis comparativo de los perfiles de nucleosomas obtenidos en distintas condiciones desvela importantes diferencias entre los grupos de los genes PcG.....</i>	- 108 -
3.5. <i>H2AK121ub afecta al posicionamiento de nucleosomas +1 mientras que H3K27me3 altera la organización de nucleosomas en la región aguas abajo del +1.....</i>	- 109 -
3.6. <i>Los genes marcados con H3K27me3 muestran una accesibilidad reducida a la digestión por MNasa en la región del promotor</i>	- 112 -
3.7. <i>Los perfiles de nucleosomas de los genes only-H2AK121ub o H2AK121ub/H3K27me3 activos desvelan una heterogeneidad transcripcional en los genes de estos conjuntos.....</i>	- 114 -
3.8. <i>Los perfiles de densidad de nucleosomas en atbmi1a/b/c apoyan el papel de las marcas H2AK121ub en determinar el posicionamiento del nucleosoma +1</i>	- 116 -
4. DISCUSIÓN.....	- 120 -
CAPÍTULO III: MECANISMO DE RECLUTAMIENTO DE LA ACTIVIDAD POLYCOMB EN PLANTAS.	- 127 -
1. INTRODUCCIÓN.....	- 127 -
2. OBJETIVO.....	- 132 -
3. RESULTADOS.....	- 133 -
3.1. <i>Las zonas de decaimiento de las marcas H2AK121ub y H3K27me3 se localizan mayoritariamente en regiones promotoras.....</i>	- 133 -
3.2. <i>Identificación de los motivos 6-mer enriquecidos en las regiones del promotor que presentan un decaimiento de las marcas PcG.....</i>	- 137 -
3.3. <i>Recapitulación sintética de los componentes de la maquinaria PcG.....</i>	- 139 -
3.4. <i>La expresión del dominio EAR fusionado a LexABD conduce a la eliminación de la acetilación de H3</i>	- 145 -
4. DISCUSIÓN.....	- 148 -
CONCLUSIONES	- 153 -
ANEXOS	- 155 -
BIBLIOGRAFÍA	- 165 -

RESUMEN

Las proteínas del Grupo Polycomb (PcG) son reguladores epigenéticos que juegan un papel crucial en los procesos de desarrollo, diferenciación y proliferación celular en los organismos eucariotas. Estas proteínas se agrupan en complejos multiproteicos con diferentes actividades dirigidas a la modificación de histonas. En *Arabidopsis* se han identificado dos complejos, el PRC2 (PcG repressive complex 2), que tiene actividad trimethyltransferasa frente a la lisina 27 de la histona H3 (H3K27) y el PRC1 que tiene actividad E3 ubiquitin ligasa frente a la lisina 121 de la histona H2A. La actividad de estos dos complejos modifica la organización de la cromatina dando lugar a la represión estable de los genes diana. En *Arabidopsis*, la composición del PRC2 y su participación, a través de la incorporación de H3K27me3, en el mantenimiento de la represión transcripcional se conoce desde hace tiempo; sin embargo, se pensaba que los componentes del PRC1 no estaban conservados en plantas, por lo que la identificación de un complejo homólogo al PRC1 en animales capaz de monoubiquitinar a la H2A es un hallazgo relativamente reciente. Esta tesis está orientada a conocer el papel que juegan los componentes PRC1 en el mecanismo de represión y su relación con el PRC2.

Los resultados que se presentan muestran que la actividad AtBMI1 es crucial no solo para promover las distintas transiciones de fase a lo largo del desarrollo de la planta, sino también para controlar el desarrollo y crecimiento de los órganos. Nuestros resultados también apuntan a la existencia de diferentes variantes de PRC1 que regulan distintos genes en combinación con el PRC2 para permitir que se den las distintas transiciones de fase. En esta tesis también investigamos el papel que juegan las marcas H2AK121ub y H3K27me3 a nivel de la organización de los nucleosomas para crear una conformación refractaria a la transcripción. Nuestros resultados muestran que estas dos marcas tienen efectos diferentes pero aditivos en la organización de los nucleosomas en los genes diana. Por último, gracias a la construcción de un sistema ideado para recapitular de forma sintética la represión PcG, hemos podido determinar los factores implicados en el mecanismo y los procesos que tienen lugar en paralelo para conseguir una represión efectiva y estable.

En resumen, el trabajo de esta tesis ha aportado información relevante para comprender el funcionamiento de la maquinaria PcG en plantas.

INTRODUCCIÓN GENERAL

INTRODUCCIÓN GENERAL

Las plantas, a diferencia de los animales, no pueden cambiar de ubicación para escapar de condiciones ambientales adversas, como por ejemplo ataques de herbívoros o cambios climáticos (Greb & Lohmann, 2016), por ello han desarrollado mecanismos para prolongar y regular su desarrollo y crecimiento a lo largo de todo su ciclo de vida en función del ambiente, confiriéndoles una plasticidad que asegura su supervivencia (Nicotta et al. 2010). Esto es posible gracias al mantenimiento de un grupo de células madre a lo largo de toda su vida en el ápice del tallo y de la raíz, los meristemos apical y radicular, respectivamente (Morao, Bouyer y Roudier, 2016). El meristemo apical controla la formación de los órganos aéreos, como hojas y flores, y el meristemo radicular el desarrollo de la raíz, en ambos casos influenciados por las condiciones ambientales (Sanchez et al. 2015). Los mecanismos epigenéticos son esenciales tanto para el mantenimiento de la pluripotencia de las células madre como para la diferenciación de los distintos órganos de forma orquestada con las condiciones ambientales (Fletcher, 2017) al jugar un papel importante en la regulación de la expresión génica (Nicotta et al. 2010).

El término “epigenética” se refiere a una serie de procesos moleculares que dan lugar a cambios heredables pero que no conllevan cambios en la secuencia del DNA. En estos procesos participan una amplia variedad de factores con distintas funciones y actividades dirigidas a modificar la organización de la cromatina (Pien y Grossniklaus, 2007).

1. Organización de la cromatina dentro del núcleo

En los organismos eucariotas la información genética se encuentra compactada dentro del núcleo (Petesch y Lis, 2012). El primer nivel de compactación se consigue mediante el enrollamiento de la cadena de DNA en torno a unas partículas protéicas llamadas nucleosomas. Un nucleosoma está constituido por un fragmento de DNA de aproximadamente 147 pb de longitud enrollado alrededor de un octámero de histonas, lo que constituye la unidad de repetición básica de la cromatina. El octámero está formado por dos copias de las histonas H2A, H2B, H3 y H4, bien de las denominadas "canónicas", que se depositan en la cromatina de una manera dependiente de la replicación, o de variantes de estas, que muestran distintas propiedades bioquímicas y son depositadas en la cromatina independientemente de la replicación por maquinaria especializada (Talbert y Henikoff, 2014). Los nucleosomas adyacentes se conectan por un segmento de DNA de una longitud variable llamado DNA espaciador o “DNA linker”,

dando lugar a una estructura de “collar de perlas” (“*Beads on a string*”) [(Eggleston, 2005); Fig. 1].

La siguiente unidad estructural de la cromatina está mediada por la histona H1 (o H5 según especies), que es la histona de unión o “*linker histone*”. Esta histona interacciona con 10 pb de DNA de entrada y de salida del nucleosoma, formando una estructura llamada cromatosoma (Fyodorov *et al.*, 2018). La interacción participa en la condensación de la cromatina modificando la trayectoria del *DNA linker* [(Thoma, Koller y Klug, 1979); Fig. 1]. Tanto las histonas nucleosomales como la histona de unión utilizan los residuos de arginina y lisina cargados positivamente para interaccionar con los fosfatos del esqueleto del DNA (Fyodorov *et al.*, 2018).

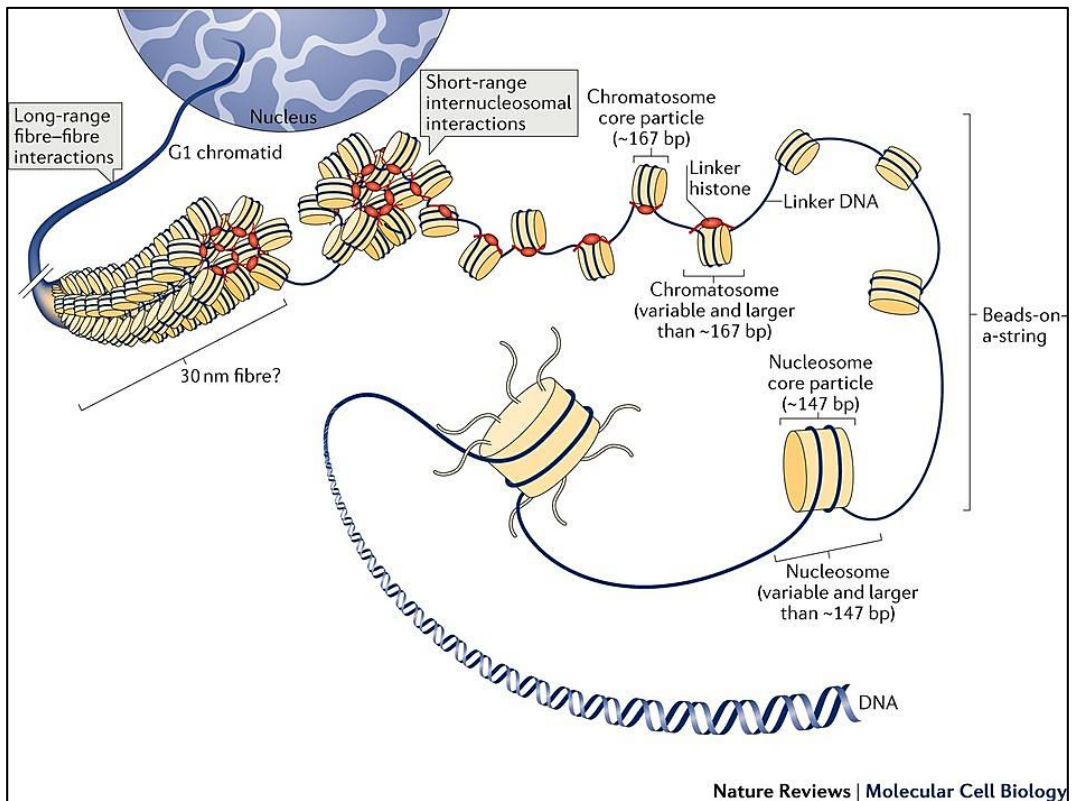


Figura 1: Múltiples niveles de plegamiento de la cromatina. La compactación de la cromatina dentro del núcleo en la interfase se produce a través de interacciones jerárquicas dependientes de histonas, incluida la formación del nucleosoma, la estructura de “collar de perlas”, el cromatosoma y las fibras de 30 nm (cuya existencia es detectable *in vivo* y solo puede ser relevante en tramos cortos de la cromatina) y la asociación de fibras individuales que producen estructuras terciarias (Fyodorov *et al.*, 2018).

La interacción de la histona de unión con el DNA y la interacción de las fibras de cromatina entre sí contribuyen a un mayor grado de compactación de la cromatina (Woodcock y Ghosh, 2010). Además, la longitud del *DNA linker* y las variaciones rotacionales internucleosomales pueden ser determinantes en el grado de plegamiento de la cromatina, dando lugar a estructuras de orden superior [(Grigoryev, 2012); Fig. 1]. El establecimiento y la regulación dinámica de estas estructuras son esenciales para que se den los procesos necesarios a lo largo de la vida de un organismo (Gentry y Hennig, 2014).

La cromatina se divide en dos tipos en función de su estructura y su función; la heterocromatina que está muy condensada, contiene pocos genes y está transcripcionalmente silenciada, y la eucromatina que está enriquecida en genes y menos condensada (Rosa y Shaw, 2013).

La cromatina forma una barrera física frente a los procesos celulares que requieren la accesibilidad al DNA, como por ejemplo la expresión génica, impidiendo el acceso de los factores de transcripción y de la RNA polimerasa II (Petesch y Lis, 2012). Además de los factores de transcripción que se unen al DNA en unas secuencias específicas para activar o reprimir los genes, los factores epigenéticos que modifican y remodelan la estructura de la cromatina son necesarios para establecer y mantener un estado transcripcional específico. Estos factores, en general, forman parte de complejos multiproteicos con actividades dirigidas a modificar de forma reversible el DNA o las histonas nucleosomales. Estas actividades incluyen la incorporación de una amplia variedad de modificaciones covalentes en las histonas o la metilación del DNA, conocidas como modificaciones epigenéticas, la incorporación de variantes de histonas, y actividades dependientes de ATP que alteran las interacciones entre el DNA y las histonas. Las distintas actividades de los factores/complejos epigenéticos colaboran para establecer los cambios requeridos para el mantenimiento de un estado transcripcional determinado.

2. Modificaciones epigenéticas

Una de las modificaciones o marcas epigenéticas mejor estudiadas es la metilación del DNA que consiste en la introducción de un grupo metilo en la posición 5' de la citosina del DNA. La metilación del DNA es un proceso altamente conservado entre animales y plantas y es esencial para la regulación de diferentes procesos biológicos. En plantas, la metilación del DNA puede ocurrir en distintos contextos: CG, CHG y CHH (donde H = A, T ó C) y en cada caso la metilación es catalizada por distintas actividades metiltransferasas. Sin embargo, en animales ocurre principalmente en las secuencias CG (Zhang *et al.*, 2006; J. Wang *et al.*, 2016; Zhang, Lang y Zhu, 2018). Estudios llevados a

cabo para conocer la distribución genómica de la marca de metilación de DNA en *Arabidopsis thaliana* (*Arabidopsis*), han demostrado que la heterocromatina se caracteriza por presentar altos niveles de metilación del DNA y está enriquecida principalmente en los transposones y en secuencias de DNA repetidas, jugando un papel esencial en su silenciamiento (Zhang *et al.*, 2006). Además, un gran porcentaje de genes están metilados, la metilación del DNA en la región codificante de los genes está asociada a genes constitutivamente activos y altamente expresados, mientras que la metilación en los promotores se relaciona con represión transcripcional y con genes específicos de tejido (Zhang *et al.*, 2006).

Por otro lado, existe una gran diversidad de modificaciones que pueden ser incorporadas en las colas de las histonas nucleosomales como son la acetilación, metilación, ubiquitinación y fosforilación, entre otras. Entre las modificaciones más estudiadas en relación con la regulación génica en plantas están la trimetilación de la histona H3 en la lisina 4 (H3K4me3), H3K9me2, H3K9me3, H3K27me1, H3K27me3 y H3K36me3, así como la ubiquitinación de la histona H2A y la histona H2B y la acetilación de la histona H3 en lisina 9 (H3K9ac), H3K14ac y H4K5ac. El efecto de un mismo tipo de modificación depende de la histona en la que se incorpore, el residuo que se modifique, y el grado de modificación, pudiendo así tener un efecto represivo o activador en la expresión génica. Algunas de estas marcas coexisten en estados de cromatina específicos, mientras que otras parecen ser mutuamente excluyentes (Roudier *et al.*, 2011; Sequeira-Mendes *et al.*, 2014). La combinación de estas marcas constituye lo que se conoce como el "código de histonas" (Strahl y Allis, 2000).

La incorporación de cada una de estas modificaciones en las histonas se lleva a cabo a través de actividades específicas. El establecimiento de estas marcas no es algo permanente, sino que puede ser revertido a través de la actividad de factores que eliminan la modificación (Eichten, Schmitz y Springer, 2014). Así, por ejemplo, las distintas marcas de metilación se depositan por diferentes metiltransferasas que contienen el dominio SET. El genoma de *Arabidopsis* contiene al menos 47 metiltransferasas de este grupo (Baumbusch, 2001; Ng *et al.*, 2007), de las cuales hasta diez se han identificado como metiltransferasas de H3K4 (Springer *et al.*, 2003). En *Arabidopsis* también hay 21 desmetilasas que contienen dominios JmJC. Estas se dividen en subclases en función de la presencia de dominios adicionales y tienen actividades específicas de desmetilasa (Lu *et al.*, 2008; Hong *et al.*, 2009). También hay 12 acetiltransferasas (HAT) y 18 desacetilasas de histonas (HDAC), que se subdividen en diferentes clases con actividad específica (Pandey, 2002; Servet, Conde E Silva y Zhou, 2010), y una maquinaria específica de ubiquitinación y desubiquitinación de H2A y H2B, aunque está menos caracterizada (Feng y Shen, 2014).

Además de esta batería de factores “modificadores” y “desmodificadores” de histonas, hay factores que participan en el intercambio de histonas, en el desplazamiento o en el desalajo de nucleosomas que actúan en estrecha colaboración con las marcas (Swygert y Peterson, 2014). Hay cuatro subfamilias principales de remodeladores de cromatina dependientes de ATP (INO80/SWR1, CHD, ISWI y SWI/SNF) con diversos papeles remodeladores (Han *et al.*, 2015). Aunque no vamos a introducir en detalle la información disponible acerca de todas estas marcas y los factores asociados a su incorporación/eliminación, sí queremos reflejar la gran complejidad que envuelve el código de histonas. Este trabajo se centra en conocer el papel que juegan los factores modificadores del Grupo Polycomb (PcG) y sus marcas en plantas. Estos factores se agrupan en dos tipos de complejos, uno con actividad H3K27me3 y otro con actividad H2AK121ub. Estas actividades son necesarias para el mantenimiento estable de la represión de los genes diana en *Arabidopsis*.

3. Las proteínas PcG están evolutivamente conservadas

Las proteínas PcG fueron identificadas por primera vez en *Drosophila melanogaster* (*Drosophila*) como represores de los genes homeóticos implicados en el desarrollo embrionario (Lewis, 1978). Hoy en día se sabe que estas proteínas regulan un gran número de genes implicados en procesos de diferenciación y desarrollo en eucariotas, operando tanto en la mayoría de animales multicelulares como en plantas (Schwartz y Pirrotta, 2013). Se han identificado ortólogos de uno de los complejos PcG, el Polycomb Repressive Complex 2 (PRC2) en varios organismos unicelulares y primitivos como los ciliados (*Tetrahymena thermophila*), diatomeas (*Phaeodactylum tricornutum*), hongos filamentosos (*Neurospora crassa*), algas verdes (*Chlamydomonas reinhardtii*) y algas rojas (*Cyanidioschyzon merolae*). Sin embargo, ortólogos obvios de otro de los complejos PcG, el PRC1, no están presentes en los genomas de estos organismos, pero sí lo están en animales superiores (*Bilateria*) y en plantas terrestres como *Marchantia polymorpha* y *Physcomitrella patens*, lo que sugiere una función de PRC1 relacionada con el control de la multicelularidad u organización tisular compleja o ambas; sin embargo, no todos sus componentes se han conservado (Schubert, 2019).

La complejidad de la maquinaria PcG se ha expandido durante la evolución, encontrándose un número mayor de parálogos de los distintos componentes en organismos superiores (Sowpati, Ramamoorthy y Mishra, 2015; Huang *et al.*, 2019), lo que indica una diversificación de funciones de los complejos por combinación de distintos parálogos, a pesar de que puedan seguir mostrando algún grado de redundancia (Sowpati, Ramamoorthy y Mishra, 2015; Huang *et al.*, 2019).

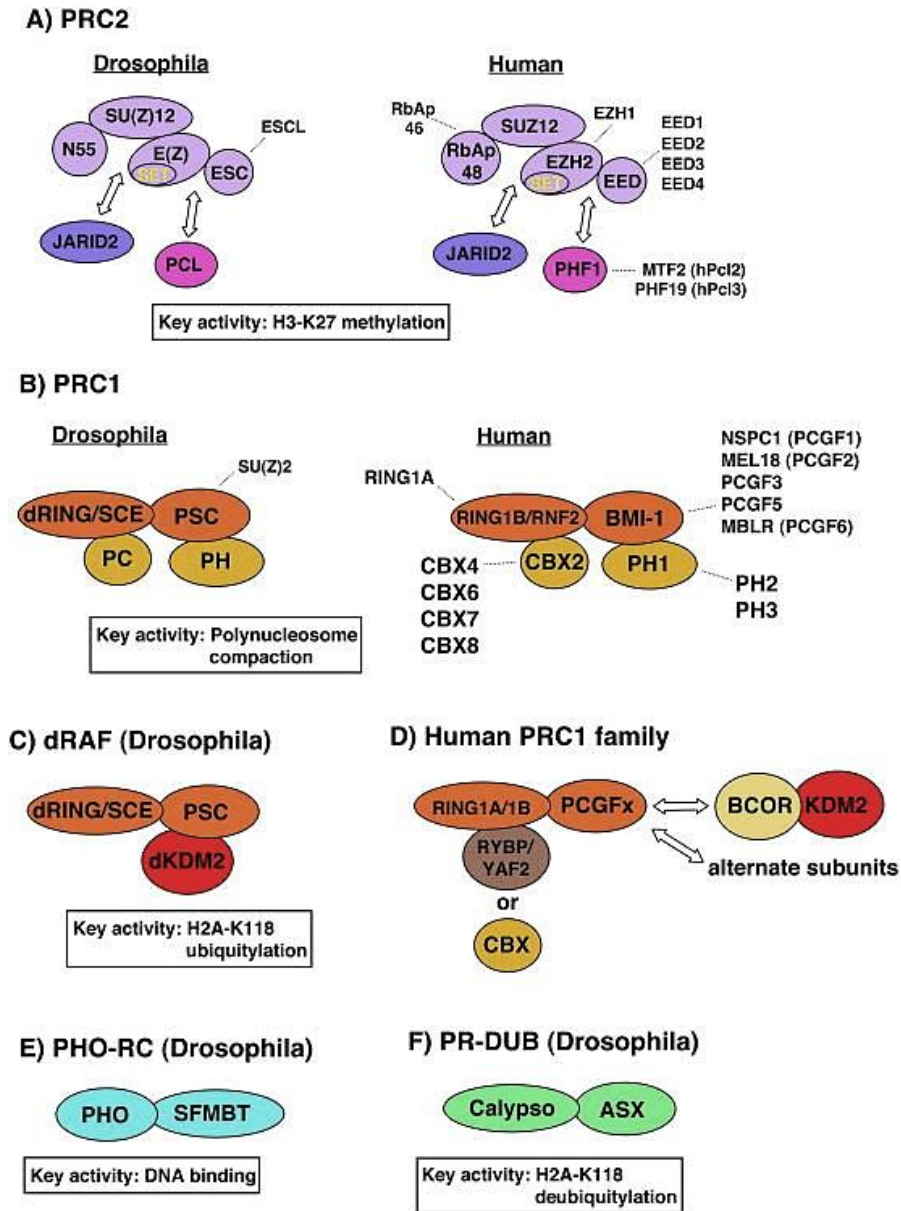


Figura 2: Composición y actividades de los complejos PcG. A) complejos de la familia PRC2. Las subunidades centrales están en color lavanda y las flechas indican asociación de subunidades opcionales. Las líneas discontinuas indican subunidades alternativas derivadas de múltiples copias de genes o variantes de proteínas de un solo gen. B) y C) Complejos de la familia PRC1. B) PRC1 canónico con 4 subunidades centrales incluido un homólogo de Pc (CBX en humanos). C) y D) Variantes de PRC1 que contienen las subunidades KDM2 y/o RYBP. En humanos, en los complejos PRC1 las subunidades RYBP y CBX son mutuamente excluyentes. La ubiquitinación ocurre en H2A-K119 en humanos, que corresponde a K118 en Drosophila. E) PHO-RC y F) PR-DUB de Drosophila. En humanos existen homólogos de PHO (YY1), SFMBT, ASX y Calypso (BAP1). (Jeffrey A. Simon y Kingston, 2013).

En la actualidad, en *Drosophila* han sido caracterizados cinco complejos multiproteicos diferentes necesarios para el establecimiento y mantenimiento de la represión: PRC2 con actividad metiltransferasa que metila la histona H3 en la lisina 27 (H3K27me3), PRC1 con actividad E3 ligasa que cataliza la monoubiquitinación de la histona H2A (H2Aub) y también tiene otra actividad relacionada con la compactación de la cromatina, Pc-repressive deubiquitinase complex (PR-DUB) que es una deubiquitinasa de la H2Aub (Cao *et al.*, 2002; Czermin *et al.*, 2002; Kuzmichev, 2002; Müller *et al.*, 2002; Wang *et al.*, 2004; Lagarou *et al.*, 2008; Scheuermann *et al.*, 2010), el complejo dRing-associated factors (dRAF) que además de monoubiquitinar a la H2A desmetila H3K36me2 y finalmente, Pho-repressive complex (Pho_RC) que reconoce los elementos “polycomb response elements” (PREs) a través de la subunidad Pho y participa en el reclutamiento de los complejos PRC1 y PRC2 a sus genes diana [(Müller y Kassis, 2006; Ringrose y Paro, 2007); Fig. 2]. En vertebrados se han identificado complejos homólogos a los de *Drosophila* aunque la existencia de subunidades alternativas ha dado lugar a múltiples complejos de la misma familia, especialmente del PRC1 [(Fig. 2); (Jeffrey A Simon y Kingston, 2013)].

En plantas, los dos únicos complejos PcG identificados hasta la fecha son PRC1 y PRC2. Estos complejos tienen la misma actividad enzimática que sus equivalentes en animales, aunque en el caso del PRC1 algunos de los componentes no están conservados y han sido re-emplazados por proteínas específicas de plantas (J. A. A. Simon y Kingston, 2013; Schwartz y Pirrotta, 2013; Calonje, 2014).

3.1. Componentes PRC2

En *Drosophila*, el PRC2 se compone de 4 subunidades de copia única: Enhancer of zeste (E(z)), Extra sex combs (Esc), Suppressor of zeste 12 (Su(z)12) y el Nucleosome-remodeling factor (Nurf 55) (Czermin *et al.*, 2002; Müller *et al.*, 2002). Una fracción de PRC2 está unida a una proteína Polycomb-like (Pcl) que contiene un dominio PLANT HOMEODOMAIN (PHD) y que es necesaria para incorporar altos niveles de las marcas H3K27me3 en los genes diana [(Nekrasov *et al.*, 2007), (Fig. 2)]. En vertebrados, existen diferentes homólogos de los componentes de PRC2: E(z) tiene dos homólogos EZH1 y EZH2 (Sarma *et al.*, 2008). Embryonic Ectoderm Development (EED) es el homólogo de Esc, y tiene 4 isoformas que se generan de los diferentes sitios de inicio de transcripción del mismo RNAm (Kuzmichev *et al.*, 2005). El homólogo de Su(z)12 es SUZ12 y el de Nurf55 es el Retinoblastoma-associated protein 46/48 (RbAp46/48) (Kuzmichev, 2002). El componente Pcl conserva su dominio PHD en vertebrados y tiene tres homólogos PCL1, PCL2 y PCL3 [(Sarma *et al.*, 2008), (Fig. 2)].

Las cuatro subunidades de PRC2 de *Drosophila* están conservadas en *Arabidopsis*. Existen diferentes parálogos de cada componente de PRC2 en *Arabidopsis*, la combinación de estos parálogos da lugar a diferentes complejos PRC2 (Fig. 3). Los homólogos de EZH2 son MEDEA (MEA) (Grossniklaus *et al.*, 1998), CURLY LEAF (CLF) (Goodrich *et al.*, 1997) y SWINGER (SWN) (Chanvivattana, 2004). EMBRYONIC FLOWER2 (EMF2), FERTILISATION INDEPENDENT SEED2 (FIS2) y VERNALIZATION2 (VRN2) son los homólogos de SU(Z)12 (Yoshida *et al.*, 2001). FERTILIZATION INDEPENDENT ENDOSPERM (FIE) es el homólogo de EED (Ohad *et al.*, 1999) y MULTICOPY SUPPRESSOR OF IRA1–5 (MSI1–5) son los cinco homólogos de RbAp46/48; sin embargo, solo MSI1 forma parte de los complejos PRC2 (Köhler *et al.*, 2003; De Lucia *et al.*, 2008; Derkacheva *et al.*, 2013). Las subunidades responsables de la actividad trimetil-transferasa del PRC2 son CLF y SWN durante el desarrollo del esporofito y MEA en el endospermo (Chanvivattana *et al.*, 2004). Mientras que MSI1 y FIE se presentan en los diferentes PRC2 (Derkacheva y Hennig, 2014), las proteínas SU(Z)12-like (FIS2, EMF2 y VRN2) son las que determinan la especificidad de los diferentes complejos PRC2 (Chanvivattana *et al.*, 2004).

De acuerdo con esto, en *Arabidopsis* se han identificado tres complejos PRC2 que actúan en diferentes fases del desarrollo (Derkacheva y Hennig, 2014) y se clasifican según la subunidad SU(Z)12-like que contienen (Mozgova, Köhler y Hennig, 2015): el complejo EMF2-PRC2 (formado por EMF2, FIE, CLF, o SWN y MSI1) que controla el desarrollo vegetativo de la planta y regula la expresión de los genes responsables de la transición de la fase vegetativa a la fase reproductiva (Kinoshita *et al.*, 2001; Derkacheva *et al.*, 2013), el complejo VRN2-PRC2 (VRN2, FIE, CLF, o SWN y MSI1) que está implicado en la respuesta a la vernalización (De Lucia *et al.*, 2008) y por último, FIS2-PRC2 que previene el desarrollo del endospermo en ausencia de la fertilización [(Kohler, 2003; Wang *et al.*, 2006), (Fig. 3)].

La actividad de PRC2 puede ser modulada por diferentes proteínas que interaccionan con el complejo. Por ejemplo, existen proteínas con dominio PHD, como VERNALIZATION INSENSITIVE3 (VIN3) (Sung y Amasino, 2004; Wood *et al.*, 2006), VERNALIZATION5 (VRN5 o VIL1) (Sung, Schmitz y Amasino, 2006; Greb *et al.*, 2007), y VIN3-LIKE 2 (VIL2) (Kim y Sung, 2010) que se ha demostrado que co-purifican con el complejo VRN2-PRC2. Estas proteínas son necesarias para mantener el silenciamiento epigenético del represor floral *FLOWERING LOCUS C (FLC)* durante la vernalización, contribuyendo así al enriquecimiento en las marcas H3K27me3 en este *locus* (Kim y Sung, 2013).

(PCGFs); BMI1, NSPC1, MEL18, PCGF3, PCGF5 y MBLR [(Z. Gao et al. 2012; Calonje 2014); (Fig.2 y 4)].




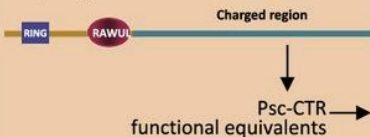








Drosophila	Arabidopsis	Vertebrates
<p>dRING (Sce)</p> 	<p>AtRING1A, AtRING1B</p> 	<p>RING1A, RING1B</p> 
<p>Psc, Su(z)2</p> 	<p>AtBMI1A, AtBMI1B, AtBMI1C</p> 	<p>PCGF (Nspc1, Mel18, Pcgf3, Bmi1, Pcgf5, Mblr)</p> 
<p>Ph-d, Ph-p</p> 	<p>Ph</p> 	<p>PH1, PH2, PH3</p> 
<p>Pc</p> 	<p>LHP1</p> 	<p>Cbx2, Cbx4, Cbx6, Cbx7, Cbx8</p> 

Figura 4: Representación esquemática de la estructura de dominios de los componentes de PRC1 en Drosophila, Arabidopsis y vertebrados (Merini y Calonje, 2015) .

En plantas, hasta no hace mucho tiempo se pensaba que PRC1 estaba ausente y por tanto que la monoubiquitinación de la H2A no era necesaria para la represión PcG en estos organismos. En su lugar, se identificaron unos factores específicos de plantas que actúan junto con el PRC2 para mantener la represión. Se identificó LHP1, que se asocia a las marcas H3K27me3 (Turck *et al.*, 2007); EMBRYONIC FLOWER1 (EMF1), un posible equivalente funcional a Psc de Drosophila, que reprime a los genes homeóticos de floración, como es el caso de *AGAMOUS (AG)*, durante el desarrollo vegetativo (Calonje *et al.*, 2008; Kim, Zhu y Sung, 2010; Kim *et al.*, 2012); y VRN1, que actúa con el complejo VRN2-PRC2 para mantener la represión de *FLC* [(Hennig y Derkacheva, 2009), (Fig. 4)]. Estos factores parece que desempeñan funciones PRC1 al reconocer las marcas H3K27me3 y participar en el mantenimiento de la represión. En ese momento, la ausencia de homólogos a los componentes PRC1 y la participación de proteínas específicas de plantas se asociaban con la existencia de un mecanismo de represión más flexible que les permite adaptarse a estímulos exteriores y modular su desarrollo en función de estos, algo que no ocurre en animales (Pien y Grossniklaus, 2007). Sin embargo, estudios posteriores basados en la arquitectura de dominios desvelaron la presencia de homólogos de las proteínas RING1 y PCGF [(Sanchez-Pulido *et al.*, 2008), Fig 4]. Mediante este estudio se identificaron proteínas que conservan el dominio RING finger en su región N-terminal y en su región C-terminal presentan un nuevo dominio Ubiquitin-like llamado RAWUL (*Ring-finger And WD40 associated Ubiquitin-Like*) (Sanchez-Pulido *et al.*, 2008). La

presencia de estos dos dominios son la característica particular de las proteínas RING finger de PRC1 que las distingue de otras proteínas (Sanchez-Pulido *et al.*, 2008). Se identificaron 2 homólogos de las proteínas RING en Arabidopsis, AtRING1A y AtRING1B y tres de las proteínas PCGF, AtBMI1A, AtBMI1B y AtBMI1C (Sanchez-Pulido *et al.*, 2008)

LHP1 fue identificado como subunidad de PRC1 específica de plantas que tiene un papel similar a la subunidad Pc de Drosophila, que se une a la marca H3K27me3 tanto *in vivo* como *in vitro* mediante su cromodominio localizado en su extremo N-terminal (Exner *et al.* 2009). LHP1 a nivel de secuencia es el homólogo de la subunidad HP1 [(Heterochromatin Protein1), (Fig. 4)] que se une a la marca H3K9me3 en Drosophila. Sin embargo, a diferencia de HP1 que se localiza en la heterocromatina, LHP1 se localiza en la eucromatina, indicando que juega un papel en la regulación de la expresión génica (Turck *et al.* 2007, Derkacheva *et al.* 2013). Varios estudios han demostrado que LHP1 interacciona físicamente con los componentes del complejo PRC1, AtRING1A/B y AtBMI1A/B/C *in vitro* (Bratzel *et al.*, 2010). Además, los análisis de localización a nivel genómico global indican que LHP1 se localiza en las regiones del genoma enriquecidas en H3K27me3. Estos resultados sugerían que LHP1 es un componente de PRC1 que reconoce y se une a la marca H3K27me3 y que forma un complejo PRC1-like junto con AtRING1A/B y AtBMI1A/B/C (Turck *et al.*, 2007; Calonje, 2014). Sin embargo, estudios posteriores demostraron que LHP1 co-purifica con PRC2 y juega un papel en el establecimiento de las marcas H3K27me3 en los genes diana (Derkacheva *et al.*, 2013; H. Wang *et al.*, 2016).

El mutante *lhp1* presenta un fenotipo de floración temprana y hojas rizadas similar al fenotipo del mutante de PRC2 *clf*, pero diferente al fenotipo de los mutantes de pérdida de función de AtRING y AtBMI1 (Gaudin *et al.*, 2001; Xu y Shen, 2008; Bratzel *et al.*, 2010). Además, LHP1 ha sido purificado con el complejo PRC2 y se ha demostrado que co-inmunoprecipita con los componentes del complejo PRC2, MSI1 y EMF2 *in vivo* e interacciona *in vitro* con MSI1 (Feng y Lu, 2017). Más importante, LHP1 afecta a los niveles de H3K27me3 en los genes diana del PRC2, incluidos *FLC*, *SEPALLATA3 (SEP3)* y *AG* cuyos niveles de H3K27me3 están significativamente reducidos en los mutantes *lhp1* (Derkacheva *et al.*, 2013).

La interacción de LHP1 con PRC1 y PRC2 indica que podría actuar como una conexión entre los dos complejos (Feng y Lu, 2017). También, se ha demostrado mediante un ensayo de inmunoprecipitación que LHP1 forma parte del complejo PRC2 y es independiente del PRC1 (H. Wang *et al.*, 2016).

El componente EMBRYONIC FLOWER 1 (EMF1) es un regulador transcripcional específico de plantas, que se ha demostrado que interacciona con los componentes del

PRC1, AtRING1A/B y AtBMI1A/B/C (Bratzel *et al.*, 2010, 2012). EMF1 contiene la secuencia de la región del C-terminal de Psc que de forma similar a Psc es capaz de inhibir la transcripción y participar en la compactación de la cromatina [(Calonje *et al.*, 2008; Beh, Colwell y Francis, 2012), Fig. 4]. Los análisis de localización a nivel genómico muestran un enriquecimiento de EMF1 en las regiones ocupadas por las marcas H3K27me3 sugiriendo que actúa junto con el PRC2 para establecer las marcas H3K27me3 (Kim *et al.*, 2012). Además, se ha demostrado que EMF1 forma junto con LHP1 y una desmetilasa de H3K4me3 llamada JUMONJ14 (PKDM7B/JMJ14) un complejo PRC1-like que tiene un papel crucial en la represión del flórigeno *FLOWERING LOCUS T (FT)* y la regulación de la transición floral en Arabidopsis (Wang *et al.*, 2014).

Aunque LHP1 y EMF1 pueden interactuar tanto con AtRING1 como AtBMI1 (Bratzel *et al.*, 2010; Chen *et al.*, 2010), resultados recientes mostraron que también copurifican con los componentes de PRC2 (Derkacheva *et al.*, 2013; Liang *et al.*, 2015); por lo tanto, no está claro en qué contexto llevan a cabo su función.

Otro componente relacionado con la función de PRC1, es la proteína VRN1 que tiene un dominio B3 de unión a DNA y parece que forma parte del complejo PRC1. Esta proteína es esencial para la represión de *FLC* durante la vernalización. Hay evidencias de que la función de VRN1 esté asociada a las marcas de represión H3K9me2 y que actúa aguas abajo del PRC2 (Bastow *et al.*, 2004). Por lo tanto, VRN1 se propuso como componente de PRC1, aunque su función no se conoce bien todavía (Calonje *et al.*, 2008; Holec y Berger, 2012). Más recientemente se han identificado otros factores asociados a la función PRC1 como las proteínas VP1/ABI3-Like 1/2/3 (VAL1/2/3) (Yang *et al.*, 2013; Qüesta *et al.*, 2016; Yuan *et al.*, 2016; D. H. Chen *et al.*, 2018), Alfin-like 1–7 (AL1–7) (Molitor *et al.*, 2014; Peng *et al.*, 2018), Early Bolting in Short Days (EBS) y Short Life (SHL) (Li *et al.*, 2018). Se ha propuesto que estos factores participan en el reclutamiento de PRC1 a sus genes diana (algunos de estos factores serán introducidos en mayor profundidad en los capítulos de esta tesis); sin embargo, no se han identificado en plantas homólogos del componente de PRC1 Ph de animales. Es muy probable que exista una proteína específica de plantas, todavía sin identificar, que lleve a cabo la función de Ph.

3.2.1. Actividades del complejo PRC1

La existencia de componentes PRC1 evolutivamente conservados en plantas, como se ha mencionado, es un hallazgo relativamente reciente, pero quedaba por determinar si estos componentes eran capaces de mediar la monoubiquitinación de la H2A. Mediante ensayos *in vitro* e *in vivo* se demostró que efectivamente, las proteínas AtBMI1 estaban implicadas en el establecimiento de esta marca, concretamente en la

lisina 121 (Bratzel *et al.*, 2010). Los mutantes *atbmi1a/b* presentan una disminución global de las marcas H2AK121ub que se refleja en la disminución de las marcas en la región próxima al sitio de inicio de la transcripción (Transcription Start Site: TSS) (Yang *et al.*, 2013). La actividad de monoubiquitinación de H2A de AtRING1 *in vivo* no ha sido comprobada todavía, por lo tanto, no se sabe cómo afecta a los niveles de H2AK121ub. Es posible que las proteínas AtBMI1 lleven a cabo la actividad del complejo y que AtRING1 module esta actividad o viceversa, pero también puede que las dos proteínas tengan actividad E3 ubiquitin ligasa *in vivo*. Los mutantes de pérdida de función de AtRING1 presentan un fenotipo más atenuado que los mutantes *atbmi1a/b* y las diferencias en el fenotipo de ambos mutantes (Xu y Shen, 2008; Bratzel *et al.*, 2010) sugieren que podrían tener alguna función de forma independiente aunque todavía no hay evidencias que apoyen esta suposición.

Curiosamente, en Arabidopsis las actividades de Psc se llevan a cabo aparentemente por diferentes proteínas: AtBMI1 que están implicadas en la monoubiquitinación de H2A (Sanchez-Pulido *et al.*, 2008; Bratzel *et al.*, 2010) y EMF1 que tiene la función de Psc-CTR (Calonje *et al.*, 2008; Beh, Colwell y Francis, 2012). De acuerdo con el papel de AtBMI1 en la incorporación de H2AK121ub, los mutantes *atbmi1* presentan una disminución de los niveles de H2AK121ub en los genes de maduración de semilla; sin embargo, en los mutantes *emf1-2* no se observa una alteración significativa de los niveles de H2AK121ub en estos genes (Yang *et al.*, 2013). Además, genes diana de EMF1, como por ejemplo el gen homeótico de floración *AG*, no tienen las marcas H2AK121ub (Yang *et al.*, 2013). De acuerdo con estos resultados, los genes de maduración de semilla no se sobre-expresan en los mutantes *emf1-2* (Yang *et al.*, 2013) y los mutantes *atbmi1* o *atring1* no sobre-expresan los genes de floración (Xu y Shen, 2008; Bratzel *et al.*, 2010). Por lo tanto, parece que existen dos tipos de represión mediada por el PRC1 en Arabidopsis. Puesto que en Arabidopsis las actividades de Psc no se asocian en la misma proteína, es probable que estas proteínas formen complejos PRC1 diferentes que llevan a cabo diferentes tipos de represión. Sin embargo, también hay evidencias que indican que pueden coexistir en el mismo complejo; la represión de varios genes diana requiere la participación de ambas proteínas RING finger y EMF1, como es el caso de *FLC* y *MAF4/5* (Kim, Zhu y Sung, 2010; Molitor *et al.*, 2014). Además, EMF1 interacciona con las proteínas AtBMI1 (Bratzel *et al.*, 2010, 2012) y se une a los genes regulados por AtBMI1 *in vivo* (Kim *et al.*, 2012). Todos estos datos sugerían que la asociación de las distintas subunidades variaba en función de los genes diana.

Los resultados de esta tesis doctoral se recogen en las publicaciones:

- **Picó S*, Ortiz-Marchena MI*, Merini W*, Calonje M**, 2015. « Deciphering the Role of POLYCOMB REPRESSIVE COMPLEX1 Variants in Regulating the Acquisition of Flowering Competence in Arabidopsis», *Plant Physiology* 128:6-97. doi: 10.1104/pp.15.00073.
 - **Merini W*, Romero-Campero FJ*, Gomez-Zambrano A, Zhou Y, Turck F, Calonje M**, 2017 «The Arabidopsis Polycomb Repressive Complex 1 (PRC1) Components AtBMI1A, B, and C Impact Gene Networks throughout All Stages of Plant Development», *Plant Physiology*. American Society of Plant Biologists, 173(1), pp. 627-641. doi: 10.1104/pp.16.01259.
- (*) indica igual contribución al trabajo.

Se prevee que este trabajo de lugar a otras dos publicaciones que en la actualidad están en preparación.

La doctoranda además ha participado activamente en la revisión:

- **Merini W, Calonje M** (2015) «PRC1 is taking the lead in PcG repression», *Plant Journal*, 83(1). doi: 10.1111/tpj.12818.

y en otra publicación relacionada con el tema:

- **Gómez-Zambrano A., Merini W.y Calonje M.** (2019) «The repressive role of Arabidopsis H2A.Z in transcriptional regulation depends on AtBMI1 activity», *Nature Communications*. Nature Publishing Group, 10(1). doi: 10.1038/s41467-019-10773-1.

MATERIALES Y MÉTODOS

MATERIALES Y MÉTODOS

1. MATERIAL VEGETAL Y CONDICIONES DE CRECIMIENTO

1.1. Esterilización de semillas

Antes de germinar las semillas de Arabidopsis, se procedió a su esterilización con etanol y lejía. Todo el procedimiento se realizó en campana de esterilidad. Las semillas se colocaron en un tubo Eppendorf y se lavaron con una solución de etanol al 70% durante un minuto en agitación. Después, se retiró el etanol y se realizó una incubación con una solución de lejía al 10% durante 20 minutos en agitación a temperatura ambiente, utilizando un agitador rotatorio Multi Bio RS-24 (Biosan). Por último, las semillas se lavaron 3 veces con agua estéril.

1.2. Germinación de semillas en cámaras de cultivo *in vitro*

Las semillas estériles se sembraron en medio Murashige y Skoog (MS) de Duchefa Biochemie (Murashige y Skoog, 1962) a pH 5,7 ajustado con NaOH 1N y suplementado con sacarosa 1 % (p/v) y Agar 0,8% (p/v).

Con el fin de sincronizar la germinación de las semillas, las placas se mantuvieron a 4°C durante 1-4 días, después de los cuales se colocaron en cámaras de cultivo *in vitro* SG-1400 (Radiber SA, Spain) en condiciones de día largo (16 horas de luz y 8 horas de oscuridad a una temperatura de 22°C durante el día y 18°C durante la noche y una intensidad lumínica entre 100 y 120 $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$).

Para la selección de las plantas transgénicas con una resistencia a un agente químico, se utilizó medio MS suplementado con el compuesto de selección. Las plantas se seleccionaron en medio MS suplementado con Kanamicina (50 $\mu\text{g}/\mu\text{l}$) o Higromicina (25 $\mu\text{g}/\mu\text{l}$), dependiendo de la resistencia del plásmido usado para la transformación. Los antibióticos se esterilizaron por filtración utilizando filtros Millex® GP (MILLIPORE) de 0,2 μm .

1.3. Cultivo de plantas en tierra

Después de aproximadamente 10 días de crecimiento en placas de MS en las cámaras de cultivo *in vitro*, se procedió a trasplantar las plántulas a tierra en semilleros (Arasystem, Bélgica) introducidos en bandejas de cultivo. La tierra (turba vegetal tipo Floraska, Alemania) se trató previamente con el antifúngico Terraclos SuperX (Etridiazol 6 % + Quintoceno 24 %, Dow Agoscience Ibérica, S.A. España) al 0,5 % y se suplementó con

10% de Vermiculita. La tierra fue proporcionada por el Servicio de Cultivos Biológicos del Centro de Investigaciones Científicas Isla de la Cartuja.

Las plantas se cultivaron en las cámaras de cultivo (Aralab, Climatic Chamber FITOCLIMA 20000EH), en condiciones de día largo (16 horas de luz y 8 horas de oscuridad) a una temperatura constante de 21°C, o en invernadero utilizando lámparas de vapor de sodio (400 W) para alcanzar las 16 horas de luz.

1.4. Cruces de plantas

Para cruzar plantas de *Arabidopsis*, se procedió a abrir los capullos cerrados de las flores de la planta madre; utilizando una pinza de punta fina se quitaron cuidadosamente los sépalos, los pétalos y las anteras. Después, se separaron las anteras de la planta paterna para poner en contacto el polen con el estigma de la planta madre.

1.5. Transformación de plantas por inmersión floral

La transformación de *Arabidopsis* se llevó a cabo según el método estándar de inmersión floral (Bechtold y Pelletier, 1998).

1.5.1. Generación de las células competentes de *A. tumefaciens*

Para generar células de *Agrobacterium tumefaciens* (*A. tumefaciens*) transformables por el método de choque térmico, se utilizó el método de cloruro cálcico. Una colonia de *Agrobacterium tumefaciens* (cepa C58C1) se inoculó en 3 ml de medio LB con el antibiótico rifampicina (25 µg/mL) y se incubó a 30°C en agitación durante toda la noche. Al día siguiente, se inocularon 50 ml de medio LB con 2 ml del preinóculo en un matraz de 250 ml y se incubó a 30°C durante aproximadamente 2 horas hasta que el cultivo alcanzó una densidad óptica DO_{600nm} entre 0,5 y 1. Después, el cultivo se enfrió en hielo durante 10 minutos y se centrifugó a 5000 g durante 15 minutos a 4°C en tubos de centrifuga estériles y enfriados previamente. Seguidamente, se descartó el sobrenadante y se resuspendieron las células en 1 ml de cloruro cálcico $CaCl_2$ 20 mM frío. Finalmente, se dispensaron 100 µl de la suspensión en tubos de 1,5 ml enfriados previamente y se congelaron en nitrógeno líquido inmediatamente. Los tubos se almacenaron a -80°C.

1.5.2. Transformación de *A. tumefaciens*

Las construcciones generadas en los vectores binarios pCAMBIA 1305 y pPZP211 fueron comprobadas por genotipado y secuenciación. Para su transformación en planta,

se introdujeron previamente en células de *A. tumefaciens* (Cepa C58C1) por el método de choque térmico.

La transformación de las células competentes se llevó a cabo añadiendo 5 µl del plásmido de interés a las células competentes congeladas. Las células se dejaron 30 minutos en hielo y se sometieron después a choque térmico incubándolas 5 minutos a 37°C. Seguidamente, se añadieron 800 µl de medio LB a las células y se incubaron durante dos horas a 28°C en agitación (300 RPM). Finalmente, las células se transfirieron a una placa de medio LB suplementado con el antibiótico de resistencia del plásmido y rifampicina (25 µg/ml) y se incubaron a 28°C durante dos días.

1.5.3. Preparación del medio de *A. tumefaciens* para la inmersión floral

Un clon de *A. tumefaciens* portador del plásmido de interés se inoculó en 5 ml de medio LB con el antibiótico de selección del plásmido y rifampicina (25 µg/mL) y se incubó a 28°C en agitación durante toda la noche. Al día siguiente, se añadieron 200 µl del preinóculo a 200 ml de medio LB suplementado con el antibiótico de selección en un matraz de 1 L y se incubó a 28°C durante toda la noche. Cuando el cultivo alcanzó una densidad óptica DO_{600nm} entre 0,7 y 1, se centrifugó a 5000 g durante 20 minutos a temperatura ambiente y se resuspendieron las células en 200 ml del medio de infiltración (sacarosa al 5%) conteniendo 40 µl de SILWET® L-77 (Lehle Seeds) que actúa como surfactante, el cual se añadió pocos minutos antes de la inmersión floral.

1.5.4. Procedimiento de inmersión floral

Para la transformación mediante inmersión floral se utilizó el ecotipo silvestre de *Arabidopsis thaliana* Columbia-0 (WT-Col0). Para aumentar el número de ramas de la planta, se cortó el tallo principal inmediatamente después de su aparición. Después de aproximadamente 7-10 días las plantas alcanzaron el estado óptimo para la inmersión. Para aumentar la eficiencia de la transformación, se cortaron todas las flores abiertas y las silicuas. Las inflorescencias se sumergieron en la suspensión de *A. tumefaciens* previamente preparada, durante 20 segundos con movimiento continuo. Las plantas sumergidas, se mantuvieron tumbadas en una bandeja de plástico cubierta de papel de aluminio durante toda la noche. A la mañana siguiente se descubrieron y se pusieron en posición vertical.

1.5.5. Selección de transformantes positivos y genotipado

Para la selección de las plantas transformadas estables, las semillas T1 fueron germinadas en placas de MS suplementado con el antibiótico de selección. Las plantas resistentes fueron transferidas a tierra y genotipadas mediante PCR utilizando

oligonucleótidos específicos. Además, se analizó la segregación de estas plantas para seleccionar las líneas heterocigotas que portaban una sola copia del transgen, ajustándose a la segregación mendeliana 3:1 (tres resistentes frente a una sensible al antibiótico).

Cuando era requerido, se comprobó la expresión del gen mediante RT-PCR (Ver sección 2.6) y/o la proteína mediante Western blot (Ver sección 3.3)

2. BIOLOGÍA MOLECULAR

2.1. Extracción de DNA

Se utilizaron dos métodos diferentes de extracción de DNA genómico (DNAg). Para obtener DNAg de alta calidad y con alto rendimiento, se utilizó el método del bromuro de cetiltrimetilamonio (CTAB) (Clarke, 2009), mientras que se llevó a cabo un método de extracción rápida de DNAg para proceder al genotipado de líneas mutantes a larga escala (Kasajima *et al.*, 2004).

2.1.1. Método CTAB

La extracción de DNAg se llevó a cabo a partir de una plántula o una hoja de roseta. Para ello, se introdujo el tejido en un tubo Eppendorf y se congeló inmediatamente en nitrógeno líquido. Posteriormente, se trituró el tejido utilizando una varilla de plástico especial para tubos Eppendorf. A continuación, se añadieron 800 µl de la solución de CTAB (CTAB 2%, Tris-HCl 100 mM pH 8.0, EDTA 20mM, NaCl 1.4 M) conteniendo 0,2 % de β-Mercaptoetanol y se incubó a 65°C durante 30 minutos con una agitación suave (400 RPM). Después de la incubación, se añadió el mismo volumen de una solución de cloroformo:alcohol isoamílico (24:1) a la muestra, se mezcló y se incubó en hielo durante 15 minutos. La muestra se centrifugó a 16000 g durante 5 minutos a 4°C. A continuación, se transfirieron 600 µl de la fase acuosa a un nuevo tubo. Para precipitar el DNAg, se añadieron 400 µl de isopropanol frío y se incubó a -20°C durante 20 minutos. Tras la incubación, se centrifugó a 16000 g durante 10 minutos. El precipitado se lavó dos veces con etanol al 70 %, se secó en un desecador al vacío y el DNAg se resuspendió en 50 µl de tampón Tris-EDTA (TE) pH 8.0 suplementado con RNase A. Para controlar su calidad se cargaron 2 µl del DNAg extraído en un gel de agarosa y se midió su concentración utilizando Espectrofotómetro NanoDrop™ modelo ND-1000 (Thermo Fisher Scientific).

2.1.2. Extracción rápida de DNA para genotipado

Esta extracción se llevó a cabo a partir de un cotiledón de una plántula de 10-14 días de edad. El tejido se introdujo en un tubo Eppendorf conteniendo 100 µl de tampón

de extracción rápida (Tris-HCL 200 mM (pH 7,5), NaCl 250 mM, EDTA 25 mM, SDS 0,5%) y se homogeneizó con una varilla de plástico especial para tubos Eppendorf. Se añadieron 300 µl del tampón de extracción rápida y se centrifugó durante 3 minutos a 20000 g. El sobrenadante (350 µl), se transfirió a un tubo nuevo y se mezcló con 350 µl de isopropanol. La muestra se incubó durante tres minutos a temperatura ambiente y después se centrifugó a 20000 g durante 5 minutos. Se retiró el sobrenadante, y el precipitado se lavó dos veces con 400 µl de etanol al 70 % y se secó al aire. Finalmente, el DNAg se resuspendió en 50 µl de tampón TE.

2.2. Reacción en cadena de la polimerasa (PCR)

2.2.1. Diseño de oligonucleótidos

Para el diseño de los oligonucleótidos y el cálculo de la temperatura de hibridación se utilizó la herramienta *OligoAnalyzer 3.1* de “Integrated DNA Technology” (IDT).

2.2.2. PCR utilizando Taq Polimerasa

Las reacciones de PCR para las comprobaciones rutinarias y genotipados se llevaron a cabo utilizando una polimerasa convencional (MyTaq DNA Polymerase, Bionline). Las reacciones se llevaron a cabo en un volumen final de 25 µl. La mezcla de la reacción y las condiciones de los ciclos de amplificación se realizaron según las recomendaciones del fabricante.

Las reacciones de PCR se realizaron en el termociclador “peqSTAR 96X Universal Gradient” (PEQLAB). Los fragmentos amplificados se visualizaron en un gel de agarosa al 1% preteñido con SYBR™ Safe (Invitrogen).

2.2.3. PCR utilizando polimerasa de alta fidelidad

Se utilizó una polimerasa de alta fidelidad (Velocity DNA polimerase, Bionline) con el fin de reducir los posibles errores al amplificar fragmentos de DNAg o cDNA para su posterior clonaje.

Las reacciones de PCR se llevaron a cabo siguiendo el protocolo descrito en el manual.

2.3. Electroforesis en geles de agarosa

2.3.1. Separación de fragmentos de DNA

Para la separación de los fragmentos de DNA por tamaño, se utilizaron geles de agarosa estándares. En la mayoría de los casos al 1% (p/v), pero para la separación de fragmentos con un tamaño inferior a 200 pb, la concentración se aumentó a 1,5% (p/v).

La agarosa se disolvió en un tampón Tris, Borato y EDTA (TBE) estándar y se suplementó con SYBR Safe a una concentración final de 1X para la posterior visualización del DNA a la luz UV. Los geles se colocaron en una cubeta de electroforesis y se cubrieron con tampón TBE al 1X. Las muestras de DNA se mezclaron con un tampón de carga 1X para su posterior carga en los pocillos del gel. Para determinar el tamaño de los fragmentos de DNA obtenidos, se cargaron en un pocillo adyacente 3 µl de un indicador comercial de peso molecular (HyperLadder 1kb o HyperLadder 50 bp (Bioline), dependiendo del tamaño de DNA esperado). Los geles se visualizaron en un transiluminador UNIVERSAL HOOD II (BioRad).

2.3.2. Aislamiento de los fragmentos de DNA separados

Los fragmentos de DNA separados en el gel de agarosa se purificaron utilizando el kit *ISOLATE II PCR and Gel* (Bioline) siguiendo el protocolo descrito en el manual. Brevemente, la banda conteniendo el fragmento de DNA de interés se cortó del gel de agarosa y se pasó a un tubo Eppendorf. Se añadió un volumen de 200 µl de tampón de extracción por cada 100 mg de gel y se incubó a 55°C hasta la completa disolución de la agarosa. La solución se cargó en la columna *ISOLATE II PCR and Gel Kit Column*, se centrifugó 30 segundos a 11000 g y se lavó dos veces con el tampón de lavado *Wash Buffer CW*. El DNA se eluyó con 15-30 µl de agua Milli-Q. La calidad de la purificación y la concentración del DNA fueron comprobadas mediante electroforesis en gel de agarosa.

2.4. Extracción de RNA

La extracción de RNA total se llevó a cabo utilizando el kit *ISOLATE II RNA Plant* (Bioline) siguiendo el protocolo descrito en el manual.

Se recogieron aproximadamente 20-50 mg de tejido fresco y se congelaron inmediatamente en nitrógeno líquido. El tejido se trituró en un tubo *Eppendorf* de 1,5 ml manteniéndolo en nitrógeno líquido utilizando una varilla de plástico especial para estos tubos. Después, se añadieron 350 µl de tampón de lisis RLY (conteniendo 3,5 µl de β-mercaptoetanol) a la muestra para llevar a cabo la lisis celular. El lisado se cargó en una

columna *ISOLATE II Filter* colocada en un tubo de 2 ml y se centrifugó a 11000 g durante 1 min. El filtrado se transfirió a otro tubo y se le añadió un volumen de 350 µl de etanol al 70 %. La mezcla se agitó con vortex dos veces durante 5 segundos. Posteriormente, se pasó la mezcla por la columna *ISOLATE II RNA Mini* con una membrana que retiene el RNA. Con el fin de quitar las sales retenidas en la membrana junto al RNA, se añadieron 350 µl del tampón MEM y se centrifugó durante 1 min a 11000 g. Se añadieron 95 µl de tampón RDN conteniendo DNasa I a la membrana y se incubó 15 min a temperatura ambiente. Seguidamente, se lavó la membrana con 200 µl de tampón RW1 y dos veces con 600 µl de tampón RW2. Después de secar completamente la membrana, el RNA se eluyó con 50 µl de agua Milli-Q libre de RNasa. La calidad del RNA se comprobó cargando 2 µl del RNA extraído en un gel de agarosa al 1% y la cantidad se midió en un Espectrofotómetro NanoDrop™ modelo ND-1000 a 260 nm. Los cocientes $Abs_{260/280}$ y $Abs_{260/230}$ que estiman la proporción de ácidos nucleicos frente a proteínas y frente a polisacáridos, respectivamente, nos permitieron determinar la integridad del RNA extraído; el valor de ambos índices tiene que estar entre 1,8 y 2.

2.5. Reacción de retrotranscripción (RT)

Para obtener el DNA complementario (cDNA) a partir del RNA extraído, se utilizó el kit *QuantiTect Reverse Transcription* (Qiagen) y se procedió de acuerdo con el protocolo descrito por el fabricante, utilizando una concentración final de 300-500 ng de RNA en 10 µl de volumen de reacción.

2.6. RT-PCR cuantitativa (qRT-PCR)

2.6.1. Diseño de oligonucleótidos

El diseño de los oligonucleótidos de qRT-PCR se llevó a cabo utilizando la herramienta *PrimerQuest* facilitada por “Integrated DNA Technologies” (IDT), teniendo en cuenta los requisitos siguientes: tamaño del producto de PCR obtenido entre 80 y 120 pb, temperatura de hibridación de 60°C y tamaño del oligonucleótido de 25 pb, además de asegurar que los oligonucleótidos hibridan en dos exones separados para poder distinguir entre el cDNA y la posible contaminación de DNA genómico (DNAg).

2.6.2. Reacciones de qRT-PCR

Para analizar el nivel de expresión génica de forma cuantitativa, se llevó a cabo una RT-PCR cuantitativa en el sistema Mutilicolor *Real-Time PCR Detection System iQ™5* (BioRad).

Para la cuantificación se generó una curva estándar para cada pareja de oligonucleótidos utilizada, mediante unas diluciones seriadas de un cDNA de control (1, 1/10 y 1/100). El gen *ACTIN2* fue utilizado como control interno para la normalización de la cantidad de cDNA. Todas las muestras se cuantificaron por triplicado.

Las reacciones de PCR se llevaron a cabo utilizando el kit *SensiFAST™ SYBR® & Fluorescein* (Bioline). La mezcla de la reacción y las condiciones de amplificación se ajustaron según las indicaciones del fabricante. El programa constaba de tres pasos: i) 1 ciclo (95 °C, 2 min); ii) 40 ciclos (95°C, 5 s; 60 °C, 5 s y 72 °C, 6 s) iii) 1 ciclo (72 °C, 6 s).

Para comprobar que no se generaba un segundo producto de la PCR inespecífico, se realizó un análisis de desnaturalización después de los ciclos de amplificación. Para ello, se aplicó una escala creciente de 0,1°C de temperatura entre 70 y 90°C. En todas las reacciones, se confirmó la existencia de un solo producto de PCR que contribuía a la señal de fluorescencia.

2.6.3. Interpretación de los datos de qRT-PCR

➤ Cálculo de curva estandar

Los datos de la qRT-PCR fueron interpretados en base al resultado del sistema *iQ5 Optical System Software* (BioRad). Para generar la curva estándar, el programa genera unas curvas de amplificación representando los valores de la emisión de fluorescencia frente al número de ciclos. Para asegurar una estimación óptima, se consideraron los siguientes factores: eliminar el nivel de fluorescencia intrínseca del fondo, la fluorescencia tenía que estar dentro de la fase exponencial de la reacción de PCR y la eficiencia de la reacción tenía que ser muy próxima a 1. Basado en esto, se define el valor de Ct como el número del ciclo al que la fluorescencia alcanza el valor del umbral fijado. Considerando estos parámetros, el programa calcula la curva estándar que permite estimar los valores de la concentración relativa del cDNA (Calc. Conc.), basado en la ecuación siguiente: $Ct = M * \log(\text{Calc. Conc.}) + B$ (donde Ct = ciclo umbral, M = pendiente de la recta, B = intersección con el eje y, Calc. Conc. = concentración relativa del cDNA).

➤ Cuantificación del cDNA

Para cuantificar los cDNA de interés, se utilizaron los valores de la concentración calculada (Calc. Conc.) determinados por el programa de acuerdo con la ecuación indicada anteriormente. Después de calcular la media de los triplicados, se normalizaron los valores con aquellos obtenidos para el gen *ACTIN2*. Los resultados se representaron mediante un diagrama de barras incluyendo las barras de error de los triplicados utilizando Excel (Microsoft).

2.7. PCR semicuantitativa (RT-PCR)

En las reacciones de PCR semicuantitativa se utilizó cDNA pero se llevó a cabo una PCR convencional utilizando la polimerasa MyTaq siguiendo las indicaciones del fabricante. En este caso, como se utilizaban los oligonucleótidos de la qRT-PCR, la temperatura de hibridación fue 60°C y el número de ciclos fue reducido dependiendo del gen de interés, para el gen control *ACTIN2* el número de ciclos fue 23 y para los genes de interés, en general fue 35.

2.8. Técnicas de clonaje molecular

2.8.1. Clonaje utilizando el sistema de vectores pSPARKII

Los fragmentos amplificados a partir de DNA genómico o a partir de cDNA fueron clonados primero en el vector pSPARKII GQ489185 (Canvax). Este sistema permite clonar fragmentos de PCR con extremos romos amplificados con una Taq-polimerasa de alta fidelidad.

pSPARKII tiene el sitio de clonaje insertado dentro de la región codificante del péptido alpha de la enzima beta-galactosidasa. La inserción de un fragmento de DNA en este sitio inactiva el péptido alpha, por lo tanto, los clones recombinantes se pueden identificar directamente mediante la selección azul/blanco en placas de 5-Bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-Gal) e Isopropyl β -D-1-thiogalactopyranoside (IPTG), de tal forma que los clones positivos se seleccionaron en placas de medio LB suplementado con ampicilina (50 mg/ml), cuyo gen de resistencia está contenido en el vector, X-GAL (2% p/v) e IPTG (100 mM).

2.8.2. Clonaje basado en endonucleasas de restricción

Las construcciones se llevaron a cabo en los vectores binarios pCAMBIA 1305 y pPZP211 mediante la manipulación de estos plásmidos usando enzimas de restricción. A continuación, se explica de forma general cómo se llevaron a cabo las digestiones con las endonucleasas de los fragmentos y los vectores, así como la ligación de ellos.

➤ Diseño de oligonucleótidos

Para amplificar los fragmentos de DNA a partir de DNAg o de cDNA, se diseñaron oligonucleótidos en los extremos del fragmento conteniendo los sitios de corte de las enzimas seleccionadas para el clonaje.

➤ Aislamiento de los fragmentos de interés

Los plásmidos pCAMBIA 1305 y pPZP211 constan de sitio de clonaje múltiple (MCS) lo que permite el clonaje de fragmentos con diferentes enzimas de restricción. El plásmido pSPAPRKII en el que previamente se había clonado el fragmento de interés y el plásmido de destino (pCAMBIA 1305 o pPZP211) fueron tratados con las mismas enzimas. Las enzimas de restricción utilizadas fueron de New England (NEB, BioLabs) o Fermentas. La mezcla de reacción de un volumen final de 20 µl contenía 1 unidad de la enzima de restricción, 0,5 – 1 µg de plásmido, 2 µl de tampón específico de la enzima (10X). Las reacciones se incubaron durante 1-2 horas a 37°C. Tras la digestión, los fragmentos de DNA se separaron en un gel de agarosa y se purificaron utilizando el kit *ISOLATE II PCR and Gel* (Bioline). Se cargaron en un gel 2 µl de los fragmentos purificados para evaluar sus cantidades relativas.

➤ Reacción de ligación

Para la reacción de ligación, se utilizó la T4 DNA-Ligase (Canvax). La cantidad del inserto fue aproximadamente tres veces mayor que la del vector linealizado. La reacción se llevó a cabo en un volumen final de 10 µl y se incubó a temperatura ambiente durante toda la noche. Al día siguiente, la reacción de ligación se usó para la transformación de las células competentes *E.Coli* DH5α por choque térmico. Las células transformadas se transfirieron a una placa de medio LB suplementado con el antibiótico de selección. Se comprobaron 10 clones resistentes por PCR de colonias, y se purificaron los plásmidos de dos clones positivos que a su vez se comprobaron por PCR y por enzimas de restricción. Los clones fueron secuenciados por el servicio de secuenciación de la compañía STABVIDA.

➤ Células competentes *E.coli* y procedimiento de transformación

Las células competentes *Escherichia.coli* (*E.coli*) de la estirpe DH5α fueron proporcionadas por el servicio de cultivo del Instituto de Biología Vegetal y Fotosíntesis (IBVF).

La transformación de las células competentes se llevó a cabo añadiendo 10 µl de la reacción de la ligación a las células competentes congeladas. Las células se dejaron 30 minutos en hielo y se sometieron después a choque térmico incubándolas 45 segundos a 42°C y pasándolas a hielo durante 2 minutos. Seguidamente, se añadieron 800 µl de medio LB a las células y se incubaron durante una hora a 37°C en agitación (300 RPM). Finalmente, las células se transfirieron a una placa de medio LB suplementado con el antibiótico de resistencia y se incubaron a 37°C durante toda la noche.

➤ Selección de transformantes positivos y aislamiento de plásmidos

Los clones de bacterias seleccionados previamente en placas de LB con el antibiótico de selección, se cultivaron en medio líquido para aislar el plásmido. Para ello, se seleccionó una colonia y se inoculó en 3 ml de medio LB líquido suplementado con el antibiótico de selección. El cultivo se incubó toda la noche a 37°C en agitación.

Para aislar el plásmido, se utilizó el kit de miniprep *ISOLATE II Plasmid Mini* (Bioline) siguiendo el protocolo descrito por el fabricante. Los plásmidos se eluyeron con 50 µl de tampón de elución del kit.

➤ Secuenciación del DNA

Los insertos de interés fueron secuenciados por el servicio de secuenciación de la compañía STABVIDA. La secuenciación se llevó a cabo por la técnica de Sanger (Sanger et al. 1977).

Para comprobar la calidad de la secuenciación, se visualizó el fichero del electroferograma mediante Chromas Light 2.1.1 (Technenlysium). Se utilizó la herramienta *ClustalW2* (EMBL-EBI) para realizar los alineamientos de las secuencias.

2.8.3. Sistema de recapitulación sintética: Construcciones y generación de plantas transgénicas

Para la construcción del sistema, se generaron dos construcciones de forma separada. La primera es la construcción que contiene el promotor sintético que controla la expresión del gen *GUS*. Para su construcción se utilizó el promotor constitutivo cauliflower mosaic virus (*CaMV 35S*). Para posibilitar la unión de las proteínas de interés, se insertaron 4 copias de los elementos de unión a DNA (operador) de la proteína LexA (*LexAO*). La inserción del *LexAO* no interfiere con la expresión constitutiva del gen reportero *GUS* en ausencia de posibles represores sintéticos. Para generar la construcción, se reemplazó el promotor *CaMV 35S* que dirige la expresión del gen reportero *GUS* en el vector binario *pCAMBIA 1305.1* por el promotor *CaMV 35S* con la secuencia de *LexAO (LexAO-35S)*, insertada justo aguas arriba de la caja TATA. Esta construcción se transformó en plantas WT-Col0 para generar las plantas transgénicas *WT/pLexAO-35s:GUS (LexAO:GUS)*.

También se generaron construcciones del promotor sintético *pLexAO-35s:GUS* en el que se introdujo junto al operador LexAO una secuencia que contiene un motivo *GAGA* y dos motivos *TELOBOX* obtenida a partir del gen *ABI3* (a la que denominamos PRE) (Xiao

et al., 2017). Se generaron dos construcciones, una con la secuencia PRE insertada justo aguas abajo de LexAO (*pLexAO-PRE-35s:GUS*) y otra con la secuencia PRE insertada justo aguas arriba del LexAO (*pPRE-LexAO-35s-GUS*) (Fig. 44). Estas construcciones se introdujeron en plantas WT-Col0 obteniéndose las plantas transgénicas *LexAO-PRE:GUS* y *PRE-LexAO:GUS*.

La expresión de un transgen puede verse afectada, dependiendo del lugar de su inserción en el genoma. Por lo tanto, las líneas transgénicas fueron seleccionadas en base a una adecuada expresión del gen reportero *GUS*. Para ello, se midió la actividad *GUS* en plántulas T2 de diferentes líneas transgénicas T1 de *LexAO:GUS*, *PRE-LexAO:GUS* y *PRE-LexAO:GUS*. Todas las líneas obtenidas presentaban altos niveles de expresión de *GUS*, por lo tanto, se eligió una de las que tenían mayor expresión para cada línea transgénica.

La segunda construcción consiste en la fusión entre una secuencia de 252 pb correspondiente al dominio de unión a DNA de la proteína LexA bacteriana (LexABD) (Schaumberg et al., 2015) y el cDNA del gen de interés. Las construcciones se generaron bajo el control del promotor constitutivo *G10-90*. Se introdujo el promotor *G10-90*, la fusión *LexABD-cDNA* y el terminador *OCTOPINE SYNTHASE (OCS)* en este orden, en el vector binario pPZP211. Se generaron las construcciones conteniendo una batería de diferentes componentes PcG y factores de interés. Esta tesis se centra en el análisis de las proteínas de fusión LexABD y AtBMI1A, AtRING1B, VAL1 y el dominio EAR fusionado a la señal de localización nuclear (NLS) de la proteína VAL1.

Estas construcciones se introdujeron en plantas WT-Col0 obteniéndose las plantas transgénicas WT/LexABD-AtBMI1A (BD-AtBMI1A), WT/LexABD-AtRING1B (BD-AtRING1B), WT/LexABD-VAL1 (BD-VAL1), y WT/LexABD-EAR (BD-EAR). Se seleccionaron aquellas plantas que expresan la proteína de fusión mediante análisis de Western blot utilizando el anticuerpo anti-LexABD.

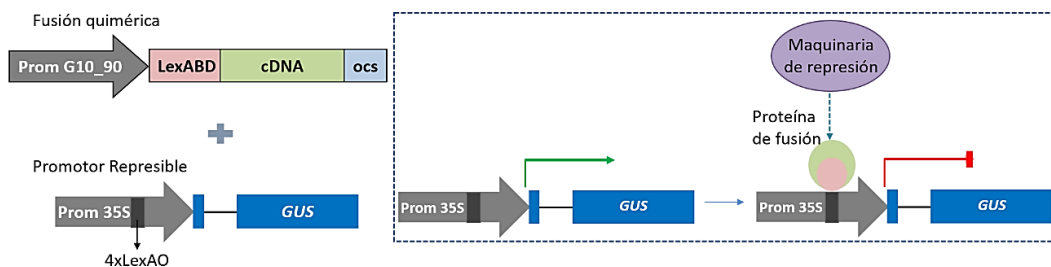


Figura 5: Esquema representativo del sistema de recapitulación sintética. (a) Diagrama de la construcción de la fusión quimérica y la construcción del promotor sintético represible conteniendo 4 sitios de unión de LexA (operadores, *LexAO*) insertados en el promotor CaMV 35S. **(b)** Representación esquemática del sistema de recapitulación sintética. Las proteínas de fusión LexABD-PcG se unen a los sitios de unión LexAO y reclutan la maquinaria PcG.

Para tener los insertos del promotor sintético y la proteína de fusión en la misma planta, se cruzaron las distintas plantas obteniendo las líneas transgénicas: *LexAO/BD-AtBMI1A*, *LexAO/BD-AtRING1B*, *LexAO/BD-VAL1* y *LexAO/BD-EAR*.

Además, se cruzaron las plantas que contienen el promotor represible *PRE-LexAO:GUS* y *LexAO-PRE:GUS* con las plantas *BD-AtBMI1A*, *BD-AtRING1B* y *BD-VAL1* para obtener las plantas *PRE-LexAO/BD-AtBMI1A*, *PRE-LexAO/BD-AtRING1B*, *PRE-LexAO/BD-VAL1*, *LexAO-PRE/BD-AtBMI1A*, *LexAO-PRE/BD-AtRING1B*, *LexAO-PRE/BD-VAL1*.

3. MÉTODOS DE ANÁLISIS Y MANIPULACIÓN DE PROTEÍNAS

3.1. Extracción de proteínas totales

Con el fin de comprobar si las plantas transgénicas contenían las construcciones de proteínas de fusión a LexABD, se realizó una extracción de proteínas totales para su posterior análisis mediante Western blot.

Para la extracción de las proteínas se recogieron 100 mg de tejido y se congelaron en nitrógeno líquido inmediatamente. El tejido se maceró en nitrógeno líquido y se le añadió un volumen de 200 µl de tampón de extracción de proteínas (50 mM Tris-HCl pH 8,0, 150 mM NaCl, 10 mM EDTA, 1% NPO₄, 0.45% deoxicolato de sodio 1% SDS) suplementado con un cóctel de inhibidores de proteasas (PIC, Roche) al 1X y PMSF (1 mM). Después, se centrifugó durante 20 min a 20000 g a 4°C. Se recogió el sobrenadante y se le añadió tampón de carga de proteínas (LAEMMLI, 1970). Las muestras se cargaron en un gel de acrilamida después de hervirlas durante 10 min a 98°C.

3.2. Electroforesis de proteínas en geles de SDS-poliacrilamida

Los extractos de proteínas se analizaron mediante la separación de las proteínas por electroforesis en geles de poliacrilamida en condiciones desnaturizantes (SDS-PAGE), siguiendo el protocolo descrito por Laemmli (1970). Se prepararon las dos partes del gel: el gel separador (10 % (v/v) de acrilamida:bis-acrilamida 29:1, DCS al 0,1 % (p/v) en tampón Tris-HCl 0,375 M pH 8,8) y el concentrador (4 % de acrilamida, DCS al 0,1 % (p/v) y Tris-HCl 0,125 M pH 6,8). Ambos geles se polimerizaron con persulfato de amonio (APS) al 0,5 % (p/v) y TEMED 6,6 M.

Para cargar las muestras de proteínas, se mezclaron con tampón de carga (LAEMMLI, 1970) y se hirvieron durante 10 minutos a 98°C para desnaturizar las proteínas. Después, las muestras se cargaron en los pocillos del gel. En un pocillo adicional, se cargó un marcador de peso molecular de proteínas (bioBLU Prestained Protein Ladder, gTPbio). La electroforesis se realizó a un amperaje constante de 20 mA

por gel, utilizando un tampón de electroforesis Tris-Glicina (Tris-HCl 25 mM pH 8,3, glicina 0,19 M y SDS 0,1 % (p/v)).

3.3. Análisis mediante Western blot

Para la detección de las proteínas utilizando anticuerpos, las proteínas separadas se transfirieron a una membrana de polivinilideno fluoride Immobilon-P, PVDF (Millipore) en un sistema de transferencia semi-seco (Trans-Blot® Turbo™ Transfer System, Biorad) según las recomendaciones del fabricante. Antes de la transferencia, la membrana se activó sumergiéndola en metanol al 100 % durante 45 segundos, se lavó con agua destilada y se sumergió en el tampón de transferencia (Tris-HCl 25 mM (pH 8,3), glicina 0,19 M, metanol 20% (v/v)). La transferencia se realizó en las condiciones siguientes: 25 V, 1,3 mA, durante 10 minutos. Después, la membrana se bloqueó sumergiéndola en metanol al 100 % y se dejó secar durante 30 minutos a 37°C o durante toda la noche a temperatura ambiente. La membrana se incubó durante 2 horas con el anticuerpo primario anti-lexABD a una dilución 1:1000 preparado en una solución de TBS (Tris-HCl 20 mM (pH 7,5), NaCl 150 mM) con 5 % (p/v) de leche en polvo. Después se realizaron 3 lavados de 5 minutos con una solución de TBSt (TBS y Tween 0,1 % (v/v)), para eliminar los anticuerpos unidos de forma inespecífica. A continuación, se incubó con el anticuerpo secundario preparado en una solución de TBS con 5 % (p/v) de leche en polvo, durante una hora en agitación. Finalmente, se realizaron 3 lavados de 5 minutos con una solución de TBSt (TBS + Tween 20%). El anticuerpo secundario está unido a la enzima peroxidasa que nos permite detectar el complejo antígeno-anticuerpo. El revelado se realizó mediante una reacción quimioluminiscente utilizando el reactivo de detección comercial ECL Select™ Amersham™ según las indicaciones del fabricante. Las membranas se visualizaron en el sistema *ChemiDoc™ Imaging System (BioRad)*

4. Inmunoprecipitación de cromatina

El protocolo de la inmunoprecipitación de cromatina está basado en el protocolo publicado previamente (Yang *et al.*, 2013).

4.1. Fijación del material vegetal

Se recolectó 1 g de plántulas de 7 días. El tejido fue fijado en una solución de formaldehído al 1% por infiltración al vacío durante 20 minutos. La fijación se paró añadiendo una solución de glicina a una concentración final de 0,125 M y aplicando vacío durante 10 min. Tras esto, el tejido fue lavado tres veces con agua destilada, y tras eliminar el exceso de agua, se congeló y se almacenó a -80°C.

4.2. Aislamiento de la cromatina

4.2.1. Extracción de núcleos

El tejido fijado se maceró en nitrógeno líquido hasta obtener un polvo fino utilizando un mortero. Seguidamente, el polvo se resuspendió en 25 ml de tampón de extracción 1 (Tris-HCl 10 mM (pH 8), sacarosa 0,4 M, MgCl₂ 10 mM, β-Mercaptoetanol 5 mM, PMSF 0,1 mM) manteniéndolo en hielo hasta su completa descongelación. Después, la suspensión se filtró pasándola por doble capa de Miracloth (Merck) y se centrifugó a 2100 g durante 20 minutos a 4°C. El sobrenadante se desechó y el precipitado se resuspendió en el tampón de extracción 2 (Tris-HCl 10 mM (pH 8), sacarosa 0,25 M, MgCl₂ 10 mM, Triton X-100 1% (v/v), β-Mercaptoetanol 5 mM, PMSF 0,1 mM), incubándose 10 minutos en hielo. Seguidamente, se centrifugó a 5000 g durante 10 minutos a 4°C. Se repitió una vez más el lavado con el tampón de extracción 2.

4.2.2. Lysis de los núcleos y sonicación de cromatina

El precipitado se resuspendió en 600 µl de tampón de lisis (Tris-HCl 20 mM (pH 8), NaCl 150 mM, SDS 0,1% (v/v), Triton X-100 1% (v/v), EDTA 2 mM). Para fragmentar la cromatina y obtener fragmentos de DNA entre 0,5-0,2 Kb, la solución de cromatina se sometió a 11 ciclos de sonicación de 30 segundos ON/OFF en el Bioruptor® Pico (Diagenode). Después de la sonicación, la muestra se centrifugó a 16000 g durante 10 minutos y se recogió el sobrenadante (cromatina sonicada).

4.2.3. Inmunoprecipitación

➤ Unión del anticuerpo

La cromatina sonicada se dividió en dos alícuotas de 300 µl y una de 30 µl que se reservó como *Input*. Las alícuotas de 300 µl se utilizaron para la inmunoprecipitación (muestras IP), para ello, se añadió 1 µl del anticuerpo correspondiente a cada muestra y se incubó a 4°C en continua agitación durante 16 horas.

➤ Bloqueo de las perlas Protein A Sepharose (GE Healthcare)

Las perlas de Protein A Sepharose tras ser hidratadas en tampón de lisis se incubaron con Seroalbúmina bovina (BSA) a una concentración final de 10 mg/ml y DNA fragmentado (Salmon Sperm DNA sheared, Invitrogen) a 10 mg/ml a 4°C durante toda la noche para bloquear posibles uniones inespecíficas. Después, se lavaron 3 veces con el tampón de lisis y se resuspendieron en un volumen del mismo tampón para obtener una suspensión de perlas al 50 %.

➤ Captura de los inmunocomplejos

Para recuperar los inmunocomplejos formados, se añadieron 60 µl de la suspensión de perlas a las muestras IP y se incubaron 2 horas a 4°C con agitación. Después, se lavaron con 1 ml de los siguientes tampones en el siguiente orden:

1- Tampón de lisis (con baja concentración de sales) (Tris-HCl 20 mM (pH 8), NaCl 150 mM, SDS 0,1% (v/v), Triton X-100 1% (v/v), EDTA 2 mM).

2- Tampón con alta concentración de sales ((Tris-HCl 20 mM (pH 8), NaCl 500 mM, SDS 0,1% (v/v), Triton X-100 1% (v/v), EDTA 2 mM).

3- Tampón LiCl (Tris-HCl 10 mM (pH 8), NaCl 150 mM, NP-40 1% (v/v), Desoxicolato de sodio 1% (p/v), EDTA 1 mM).

4- Tampón TE (Tris-HCl 10 mM (pH 8), EDTA 1 mM).

Después de cada lavado se centrifugó a 800 g durante 1 minuto y se retiró el tampón con cuidado de no coger las perlas.

➤ Elución y reversión de la fijación de la cromatina

Para liberar los inmunocomplejos, se añadió a las perlas un volumen de 250 µl de tampón de elución (SDS 1% (v/v), NaHCO₃ 0,1 M) recién preparado y se incubó a 65°C durante 30 minutos con agitación (800 RPM) utilizando el incubador Eppendorf Thermomixer. Tras la incubación, se llevó a cabo una centrifugación de 5 minutos a 3000 g para recoger el sobrenadante y transferirlo a un tubo nuevo. Al *Input* de cromatina se le añadió un volumen de 220 µl de tampón de elución. Tanto a las muestras IP como al *Input* se les añadió un volumen de 10 µl de NaCl 5 M y se incubaron a 65°C durante toda la noche para revertir la fijación.

4.2.4. Purificación del DNA: Extracción Fenol-Cloroformo

Tras revertir la fijación, cada muestra se trató con 10 µg de RNase A durante 60 min a 37°C y seguidamente con 20 µg de Proteinasa K durante 1 hora a 45°C para limpiar el DNA inmunoprecipitado de proteínas y RNA. Seguidamente, se añadió el mismo volumen (250 µl) de una solución Fenol/cloroformo/isoamilo a (24:1:1) se agitó con vortex y se centrifugó a alta velocidad durante 15 minutos. La fase acuosa se pasó a un tubo nuevo y se le añadieron 20 µg de glucógeno que actuará como agente coprecipitante en la posterior precipitación del DNA. La precipitación del DNA se llevó a cabo añadiendo 2,5

volumenes (625 μ l) de etanol al 100% frío a cada muestra e incubando 2 horas a -80°C . Después, las muestras se centrifugaron a alta velocidad a 4°C durante 30 minutos para recoger el DNA precipitado. El DNA se lavó dos veces con etanol al 70% y se secó utilizando el desecador al vacío. Finalmente, se añadieron 50 μ l de TE para resuspender el DNA.

5. Determinación de la actividad β -glucuronidasa (GUS): Ensayo fluorimétrico MUG

La actividad de la β -glucuronidasa (GUS) se cuantificó utilizando el reactivo 4-methylumbelliferyl β -D-glucuronide (4-MUG) como sustrato que tras su hidrólisis por la actividad de GUS genera un compuesto fluorescente; el 4-metil-unbeliferona (4-MU).

Para realizar el ensayo de cuantificación de actividad GUS a larga escala, se optimizó el método descrito previamente (Blazquez, 2007) disminuyendo el tiempo de ensayo y de manipulación.

En cada pocillo de una placa de 96 de fondo negro, se sumergió una plántula (recogida 7 días después de la germinación, 7 DAG) en 150 μ l de tampón de lisis (tampón fosfato 50 mM (pH 7), EDTA 10 mM, Triton X-100 0,1 % (v/v)) conteniendo 4-MUG 1 mM. Después de incubar durante 90 minutos a 37°C , se añadieron 50 μ l de una solución de carbonato sódico (Na_2CO_3) 1 M para parar la reacción. Se cuantificó la fluorescencia del 4-MU producido midiendo directamente en un lector de microplacas *Varioskan LUX Multimode Microplate Reader* (Thermo Fisher), utilizando una excitación a 365 nm y midiendo la emisión a 455 nm.

6. Análisis transcriptómico masivo basado en datos de RNA-seq

6.1. Diseño experimental y secuenciación masiva

El diseño experimental de nuestro análisis consistió en dos réplicas biológicas independientes de cada genotipo (WT-Col0, *atbmi1a*, *atbmi1b*, *atbmi1a/b* y *atbmi1a/b/c*) ya que en análisis previos se ha comprobado que este número de réplicas biológicas son suficientes para capturar la variabilidad en muestras de RNA-seq (Conesa *et al.*, 2016). Para la extracción del RNA se utilizó el kit *ISOLATE II RNA Plant* (Bioline), siguiendo las instrucciones del fabricante. La calidad y la concentración del RNA se comprobaron mediante la cuantificación fotométrica utilizando el Espectrofotómetro NanoDrop™. La preparación de la librería se llevó a cabo según las instrucciones del fabricante (TruSeq RNA Sample Prep Kit v2, Illumina). La secuenciación de las librerías se realizó con el secuenciador Illumina HiSeq 2000, produciendo aproximadamente una media de 15 millones de lecturas emparejadas (paired-end) de 100 pb para cada muestra. Estos datos

están disponibles en la base de datos GEO (Gene Expression Omnibus) bajo el número de acceso GSE83568.

6.2. Flujo de trabajo: Protocolo TUXEDO

El flujo de trabajo seguido para analizar los datos masivos secuenciados se basa en el protocolo TUXEDO (Trapnell *et al.*, 2012) implementado en *scripts* para computación de altas prestaciones usando el sistema de gestión de colas en clústeres de ordenadores SGE (Sun Grid Engine). Para el control de calidad de las lecturas, se utilizó la herramienta *FastQC* (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) y se determinó la calidad por pares de bases. Todas las muestras analizadas fueron de alta calidad, y no fue necesario pre-procesar las lecturas para eliminar las lecturas de baja calidad (Fig. 13).

El genoma de referencia de *Arabidopsis thaliana* Col-0 y la anotación de los transcritos fueron descargados de la base de datos Phytozome (TAIR10) en formato fastq y GTF [(Gene Transfer Format, (Goodstein *et al.*, 2012)], respectivamente. La construcción del índice del genoma de referencia para facilitar la fase de alineamiento de lecturas cortas se llevó a cabo con la herramienta *Bowtie* (Langmead *et al.*, 2009). El alineamiento de las lecturas al genoma de referencia, el ensamblado de transcritos y la estimación de la expresión génica se realizaron utilizando las herramientas *TopHat*, *Cufflinks* y *Cuffdiff* (Trapnell *et al.*, 2012) usando los parámetros por defecto. Se obtuvo un alto porcentaje de alineamientos concordantes entre lecturas cortas emparejadas (Tabla 3). Las lecturas fueron alineadas al genoma de referencia utilizando *TOPHAT* que a su vez hace uso del alineador ultra-rápido de lecturas cortas *BOWTIE* posibilitando la existencia de los huecos en el alineamiento causados por intrones. Con las secuencias alineadas, se realizó un ensamblado de transcritos para cada muestra utilizando *CUFFLINKS* que también cuantifica los niveles de expresión de cada transcrito en la muestra. Para poder estimar la expresión diferencial de los transcritos, todos los transcritos ensamblados en las diferentes muestras se combinaron integrándolos y unificándolos en el mismo fichero GTF con la herramienta *CUFFMERGE*. Para calcular la expresión diferencial y la significancia estadística comparando las diferentes condiciones se utilizó la herramienta *CUFFDIFF*. Los niveles de expresión génica se estimaron en fragmentos por kilobase de exon por millón de lecturas mapeadas (FPKM) (Mortazavi *et al.*, 2008).

Se usó *CummeRbund* (Goff, Trapnell y Kelley, 2014), un paquete de R de Bioconductor (<http://www.bioconductor.org/>) para el análisis posterior de los datos y la representación gráfica de los resultados. Debido al efecto masivo en algunos de los mutantes analizados y con el objetivo de descartar posibles efectos secundarios, los genes expresados de forma diferencial fueron seleccionados como aquellos genes que presentan una expresión diferencial con un factor de cambio (del inglés *fold change*, FC)

mayor de 4 cuando se compara con WT y un q-valor < 0,05 (p-valor corregido según Bejamini-Hochberg). Los diagramas de Venn que comparan los diferentes grupos de genes que se expresan de forma diferencial se generaron con la herramienta *Venny* (<http://bioinfogp.cnb.csic.es/tools/venny/index.html>) y la significancia estadística de las intersecciones con los genes marcados con H3K27me3 se determinaron usando el test exacto de Fisher.

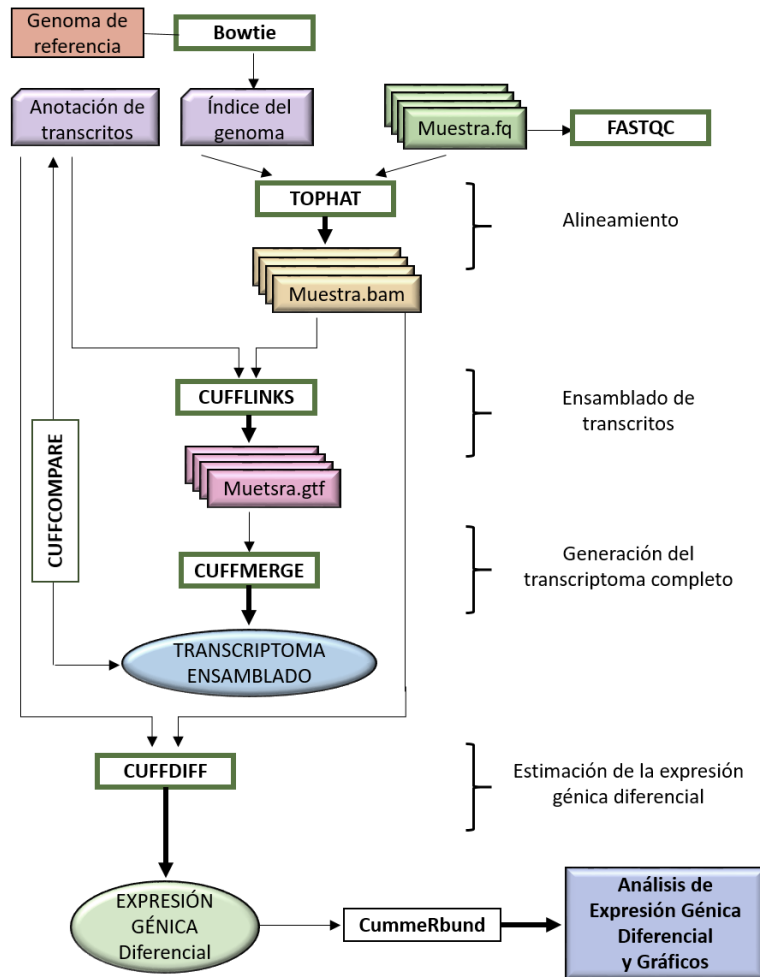


Figura 6: Flujo de trabajo del protocolo TUXEDO. Representación esquematizada de los pasos seguidos para el análisis de datos de RNA-seq. El genoma de referencia se obtuvo de la base de datos Phytozome (TAIR 10). El análisis de la expresión génica medida por FPKM se realizó mediante el alineamiento de las secuencias cortas al genoma de referencia utilizando las herramientas *Bowtie* y *Tophat*. El ensamblado de los transcritos y la generación del transcriptoma completo se llevaron a cabo mediante *Cufflinks* y *Cuffdiff*. El análisis de la expresión diferencial y las representaciones gráficas se realizaron con el paquete de R *CummeRbund*.

El enriquecimiento funcional en términos de ontología de los diferentes conjuntos de genes expresados de forma diferencial se realizó utilizando las herramientas *AgriGO* (Du *et al.*, 2010) y el paquete de bioconductor de R *ClusterProfiler* (Yu *et al.*, 2012). La redundancia en los enriquecimientos funcionales anteriores se eliminó y visualizó usando *ReViGO* (Supek *et al.*, 2011).

Se realizó un análisis de componentes principales (PCA) para comprobar que los datos eran de buena calidad y las muestras se agrupaban según los distintos genotipos. Para ello, se usó la función básica del paquete base de R *prcomp* realizando un escalado previo de los datos para obtener datos normalizados con media cero y desviación típica 1.

6.3. Análisis de *Clustering*

El análisis de agrupamiento o *clustering* (Anexos 1 y 2) se realizó usando el algoritmo de *clustering* jerárquico implementado en la función *hclust* del paquete *cluster* de R tomando como datos de entrada los datos normalizados de la expresión génica medidos en FPKM. La visualización de los grupos de genes (*clústeres*) se realizó usando la función *heatmap* del paquete base de R.

6.4. Análisis del enriquecimiento de motivos de DNA en los promotores

Se utilizó la herramienta software *HOMER* [(Hypergeometric Optimization of Motif EnRichment; (Heinz *et al.*, 2010)] para la detección de motivos de DNA significativamente enriquecidos en los promotores de genes de interés. Específicamente se utilizó la función *findMotifs.pl* usando las bases de datos de motivos de unión de DNA para plantas *AGRIS* (Davuluri *et al.*, 2003), *JASPAR* (Sandelin, 2004), y *AthaMap* (Steffens, 2004). El script *findMotifs.pl* se utilizó con los parámetros por defecto para detectar los motivos de 6,7,8 y 9 pb de longitud. En todos los análisis se usó como fondo para determinar el enriquecimiento, el conjunto total de los promotores de todos los genes anotados de *Arabidopsis thaliana* (TAIR10). Se definieron los promotores de los genes como la secuencia de DNA de 1000 pb que se encuentra aguas abajo del codón de inicio del gen correspondiente.

Para determinar la proximidad de los motivos *ABRE* y *GT-box*, primero se identificaron las localizaciones de los motivos *ABRE* en los promotores y después se extrajeron las secuencias de DNA de 100 pb aguas arriba y aguas abajo del centro del motivo *ABRE*. Se llevó a cabo un análisis de enriquecimiento del motivo *GT-box* en estas secuencias de DNA utilizando la función *findMotifsGenome.pl* utilizando los parámetros por defecto. Las significancias del solapamiento entre los motivos se llevaron a cabo como

un análisis de enriquecimiento de la secuencia de DNA que resulta de la combinación de los dos motivos. Las secuencias de DNA utilizadas en este análisis fueron descargadas utilizando la funcionalidad *BioMart* asociada a Phytozome (Goodstein *et al.*, 2012).

7. Análisis de posicionamiento y ocupación de nucleosomas basado en análisis de MNasa-seq

7.1. Protocolo de digestión con MNasa

7.1.1. Extracción de núcleos

Se recolectó 1 g de plántulas de 7 DAG de edad y se congeló inmediatamente en nitrógeno líquido. Para extraer los núcleos se siguió el protocolo descrito en la sección 4.2.1.

7.1.2. Tratamiento con MNasa

El extracto de núcleos se resuspendió en el tampón de digestión (Tris-HCl 5 mM pH 8, CaCl₂ 0,5 mM, NaCl 5 mM, BSA 1X, 2 µl de RNase A, PIC 1X) y se digirió con 1 unidad/µl de nucleasa micrococcal MNase (NEB) durante 10 minutos a 37 °C. Transcurrido este tiempo, se añadieron EDTA y SDS para una concentración final de 20 mM y 1% (v/v), respectivamente, para parar la digestión.

7.1.3. Extracción y precipitación del DNA

La muestra se trató con 10 µg de RNase A durante 1 hora a 37 °C y seguidamente con 20 µg de Proteinasa K durante 1 hora a 65 °C. Después, se añadió el mismo volumen de una solución Fenol/cloroformo/isoamilo a (24:1:1), se agitó con vortex y se centrifugó a alta velocidad durante 5 minutos a temperatura ambiente. La fase acuosa se pasó a un tubo nuevo, se le añadieron 60 µl de acetato sódico 3 M y 2,5 volúmenes de etanol 100% y se incubó durante 2 horas a -20 °C. Después, la muestra se centrifugó a alta velocidad durante 30 minutos y el precipitado obtenido se limpió dos veces con etanol al 70 %. Se eliminó todo el etanol de la muestra en un desecador al vacío y finalmente, el DNA se resuspendió en 50 µl de agua Milli-Q.

7.1.4. Purificación del DNA

El DNA digerido fue separado en un gel de agarosa *Certified Low Melt Agarose* (BioRad) al 2% preteñido con SYBR-safe. Después, la banda conteniendo los fragmentos correspondientes a la fracción mononucleosomal (aproximadamente 150 pb) se cortó del

gel y se purificó utilizando *ISOLATE II PCR and Gel Kit* (Bioline) siguiendo el protocolo descrito en el manual.

7.2. Preparación de la librería y secuenciación

La librería se preparó utilizando el kit *TruSeq DNA Sample Prep Kit* siguiendo el protocolo descrito por el fabricante (Illumina).

La secuenciación de las librerías de dos réplicas biológicas de cada genotipo se llevó a cabo mediante el secuenciador Illumina HiSeq 4000 de. La tabla 1 muestra el número de lecturas emparejadas de 50 pb que se obtuvieron en cada muestra.

Tabla 1: Número de lecturas obtenidas en cada réplica.

Muestra	Número de lecturas
WT rep1	44456648
WT rep2	50458445
<i>atbmi1 a/b/c</i> rep1	61043019
<i>atbmi1 a/b/c</i> rep2	50011049

7.3. Procesamiento de los datos de MNase-seq

El control de calidad de los datos de secuenciación recogidos en ficheros fastq se llevó a cabo utilizando la herramienta software *FastQC* (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).

Las lecturas fueron alineadas al genoma de referencia TAIR10 (Lamesch *et al.*, 2012) utilizando *BOWTIE* (Langmead *et al.*, 2009) permitiendo un error de 2 nucleótidos, produciendo un alto porcentaje de alineamientos en todas las muestras como viene detallado en la tabla 2. De esta forma se generaron ficheros en formato BAM.

Por una parte, estos ficheros BAM se usaron para la determinación de los picos correspondientes a la localización de nucleosomas en los dos genotipos. Para ello, se utilizó la función *Dpos* de la herramienta *DANPOS2* (Chen *et al.*, 2015) eliminando los duplicados con el parámetro *clonalcut*. La representación de la comparativa de la ocupación de los nucleosomas detectados se realizó a través de gráficos de nubes de puntos (scatterplots) con un *script* de R desarrollado para tal efecto.

Tabla 2: Porcentaje de alineamiento de lecturas cortas al genoma de referencia TAIR10

Muestra	Porcentaje de alineamiento
WT rep1	81,72%
WT rep2	81,06%
<i>atbmi1 a/b/c</i> rep1	79,06%
<i>atbmi1 a/b/c</i> rep2	77,99%

Por otra parte, los ficheros BAM alineados se limpiaron eliminando los duplicados utilizando la herramienta *samtools rmdup* (Li *et al.*, 2009), los ficheros BAM resultantes se convirtieron en ficheros WIG utilizando el paquete de R *nucleR* (Flores y Orozco, 2011), disponible en Bioconductor. Para convertir los ficheros BAM a ficheros WIG, primero se importaron los ficheros BAM en R, las lecturas se procesaron utilizando la función *processReads* recortando cada lectura a 40 pb alrededor de su díada utilizando el parámetro *trim=40*. A continuación, se utilizó la función *filterFFT* del paquete *nucleR* para eliminar el ruido y suavizar la curva que representa el conteo de lecturas en cada posición a lo largo de los cromosomas con el parámetro *pcKeepComp=0.01*. Finalmente, los resultados se exportaron en ficheros WIG utilizando la función *export.wig*.

Para generar los perfiles de ocupación medios de los nucleosomas en los distintos conjuntos de genes, se utilizó la instrucción *Dprofile* de la herramienta *DANPOS2* utilizando como datos de entrada los ficheros WIG generados en el paso anterior y los conjuntos de genes correspondientes. La información necesaria sobre estos conjuntos de genes se extrajo utilizando el paquete de R *TxDb.Athaliana.BioMart.plantsmart28* de Bioconductor. La representación gráfica final de los perfiles medios de ocupación se generó utilizando un *script* de R desarrollado de forma específica para este fin. Además, la ocupación media se visualizó utilizando mapas de calor usando la función *heatmap* del paquete *base* de R.

La determinación de la posición ocupada con mayor probabilidad por el nucleosoma +1 se llevó a cabo identificando el punto donde se alcanzaba el máximo de la ocupación dentro del pico observado para cada conjunto de genes. Las distribuciones globales del posicionamiento del nucleosoma +1 se visualizaron mediante gráficos de cajas y bigotes usando la función *boxplot* del paquete *base* de R. La normalidad de estos datos se estudió usando el test de Shapiro-Wilk implementado en la función *shapiro.test*. Puesto que los datos fueron no-normales, se aplicó el test de Mann-Whitney-Wilcoxon usando la función *wilcox.test* para determinar la significancia de las diferencias observadas entre los diferentes grupos de genes.

La ocupación de la región libre de nucleosomas se determinó para cada conjunto de genes calculando el área bajo la curva del correspondiente perfil de ocupación usando la función *auc* del paquete de R *AUC*. De forma similar a lo anteriormente descrito, se representó la distribución global y se analizó la significancia de las diferencias observadas.

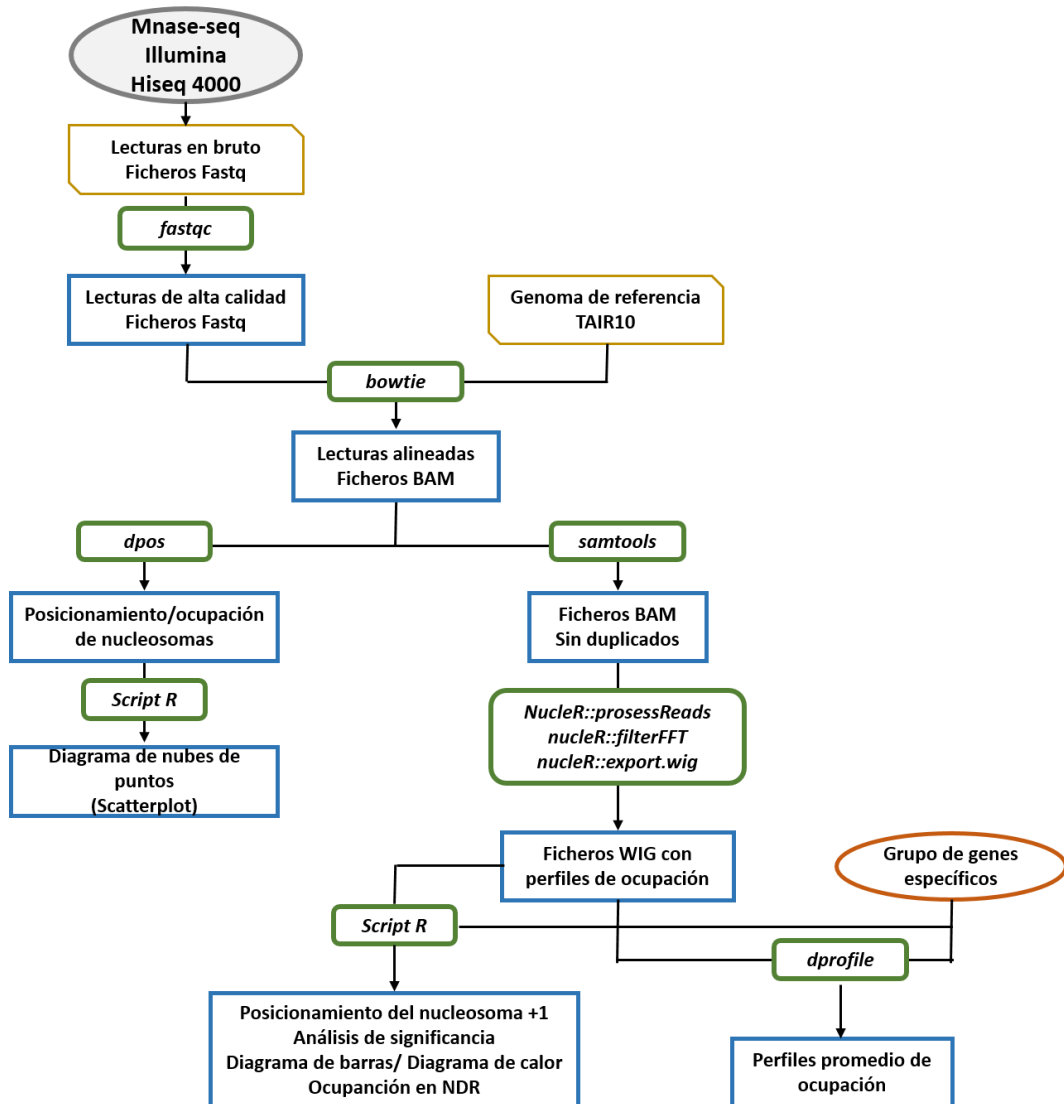


Figura 7: Flujo de trabajo del análisis de MNase-seq

CAPÍTULO I

Redes de regulación de los
componentes PRC1

CAPÍTULO I: Redes de regulación de los componentes PRC1

1. Introducción

Las proteínas PcG son necesarias para mantener la identidad celular y controlar las transiciones de fase durante el desarrollo en *Arabidopsis* (Holec y Berger, 2012). Prueba de ello son las anomalías fenotípicas severas que presentan los mutantes en las diferentes subunidades de los complejos PRC1 y PRC2 (Chanvivattana, 2004; Aichinger *et al.*, 2009; Bratzel *et al.*, 2010; Chen *et al.*, 2010; Bouyer *et al.*, 2011; Yang *et al.*, 2013). Estos complejos, al igual que en animales, incorporan modificaciones en las histonas para regular la expresión génica.

La trimetilación de la histona H3K27 (H3K27me3) es mediada por los componentes del complejo PRC2, CURLY LEAF (CLF), SWINGER (SWN) o MEDEA (MEA), que constituyen distintos complejos PRC2. Mientras que CLF y SWN forman parte de los complejos PRC2 durante el desarrollo del esporofito, MEA confiere la actividad enzimática al complejo durante el desarrollo del gametofito y la formación inicial de la semilla. Las proteínas CLF y SWN muestran un papel redundante durante el desarrollo, de tal forma que el doble mutante *clf/swn* tiene un fenotipo más severo que los mutantes simples; sin embargo, esta redundancia es parcial puesto que SWN no puede rescatar el fenotipo de floración temprana y las hojas curvadas del mutante *clf* (Chanvivattana *et al.*, 2004). La pérdida de función de CLF y SWN lleva a una de-diferenciación celular dando lugar a unas estructuras similares a los callos (Aichinger *et al.*, 2009).

Los distintos complejos PRC2 actúan en diferentes procesos y la especificidad de cada uno viene determinada por las proteínas FERTILIZATION INDEPENDENT SEED 2 (FIS2), VERNATIZATION 2 (VRN2) y EMBRYONIC FLOWER 2 (EMF2). El complejo FIS2-PRC2 actúa específicamente durante el desarrollo del gametofito femenino y de la semilla. En los mutantes *fis2*, *mea* y *fie* el endospermo se desarrolla en ausencia de la fertilización, lo que da lugar a que el desarrollo del embrión se detenga, y por tanto no se producen semillas viables (Ingouff, Haseloff y Berger, 2005; Rodrigues *et al.*, 2010; Hehenberger *et al.*, 2012; Schmidt *et al.*, 2013; Hands, Rabiger y Koltunow, 2016). Estos mutantes sobre-expresan los genes con el dominio MADS-box *PHERES 1* (*PHE1*) y *AGL62* implicados en el desarrollo de semilla (Köhler *et al.*, 2003; Kradolfer, Hennig y Köhler, 2013; Förderer, Zhou y Turck, 2016).

A diferencia de FIS2-PRC2, los complejos VRN2-PRC2 y EMF2-PRC2 participan en la regulación del desarrollo del esporofito (Gendall *et al.*, 2001; Yoshida *et al.*, 2001; Chanvivattana *et al.*, 2004; Schonrock *et al.*, 2006; Wood *et al.*, 2006; De Lucia *et al.*, 2008;

Jiang *et al.*, 2008; Derkacheva *et al.*, 2013; Li *et al.*, 2016). El complejo VRN2-PRC2 regula el tiempo de floración mediante el mantenimiento de la represión estable de *FLOWERING LOCUS C (FLC)* después de la vernalización (Gendall *et al.*, 2001; Chanvivattana *et al.*, 2004; Butenko y Ohad, 2011), consistente con ello, en los mutantes *vrn2* después del tratamiento con frío los niveles de expresión de *FLC* están elevados (De Lucia *et al.*, 2008; Förderer, Zhou y Turck, 2016). Por el contrario, los mutantes severos *emf2* forman una flor terminal inmediatamente después de la germinación sin pasar por la fase vegetativa del desarrollo. Este fenotipo es causado por la activación de genes de floración como *APETALA 1 (AP1)*, *AP3*, *PISTILLATA (PI)* y *AG* (Sung *et al.*, 1992; Yang, Chen y Sung, 1995; Kinoshita *et al.*, 2001; Chanvivattana *et al.*, 2004; Calonje *et al.*, 2008; Bouyer *et al.*, 2011).

Las proteínas AtBMI1 (A, B o C) y AtRING1 (A o B) constituyen el módulo del complejo PRC1 que cataliza la monoubiquitinación de la histona H2A en la lisina 121 (H2AK121ub). Los mutantes simples *atbmi1a* o *b* y *atring1a* o *b* presentan un fenotipo similar al silvestre (WT-like), sin embargo, los mutantes dobles *atbmi1a/b* y *atring a/b* presentan alteraciones fenotípicas importantes, lo que indica un cierto grado de redundancia entre los diferentes parálogos; sin embargo, no se sabe si, como en el caso de los componentes de PRC2, los distintos parálogos muestran una cierta especificidad por un conjunto de dianas (Bratzel *et al.*, 2010; Chen *et al.*, 2010; Yang *et al.*, 2013).

Mientras que los genes *AtBMI1A* y *B* se expresan en todos los órganos de la planta, *AtBMI1C* se expresa únicamente durante el desarrollo del endospermo, estambre y raíz (Li *et al.*, 2011; Bratzel *et al.*, 2012). Curiosamente, la expresión de *AtBMI1C* en el endospermo tiene lugar a partir del alelo materno, mientras que en el estambre y raíz se expresa bialelicamente (Bratzel *et al.*, 2012). La descendencia de las plantas *atbmi1a^{-/-}b^{-/-}c^{+/-}* presenta semillas abortadas y plantas con un fenotipo similar a *atbmi1a/b* en las partes aéreas, pero con defectos más severos en la raíz (Yang *et al.*, 2013). Sin embargo, la sobre-expresión de *AtBMI1C* es capaz de recuperar parcialmente los defectos de la raíz en los mutantes *atbmi1a/b* (Bratzel *et al.*, 2010; Yang *et al.*, 2013).

La pérdida de función de AtBMI1 conduce a la desrepresión de caracteres embrionarios tras la germinación como consecuencia de la sobre-expresión de los genes de maduración de semilla (e.g. *FUSCA 3 (FUS3)*, *LEAFYCOTYLEDON 1 (LEC1)*, *ABSCISIC ACID INSENSITIVE 3 (ABI3)*). También se produce una expresión ectópica de los reguladores de las células madre (e.g. *WUSCHEL (WUS)*, *WUSCHEL-related homeo- box 5 (WOX5)*) (Bratzel *et al.*, 2010). Curiosamente, estos mutantes muestran un fenotipo similar a aquel presentado por los mutantes de PRC2 (Chanvivattana *et al.*, 2004; Aichinger *et al.*, 2009). Los mutantes *atring1 a/b* también presentan una desregulación del programa

embrionario (Chen *et al.*, 2010), indicando que AtBMI1, AtRING1 y PRC2 regulan un subgrupo común de genes diana.

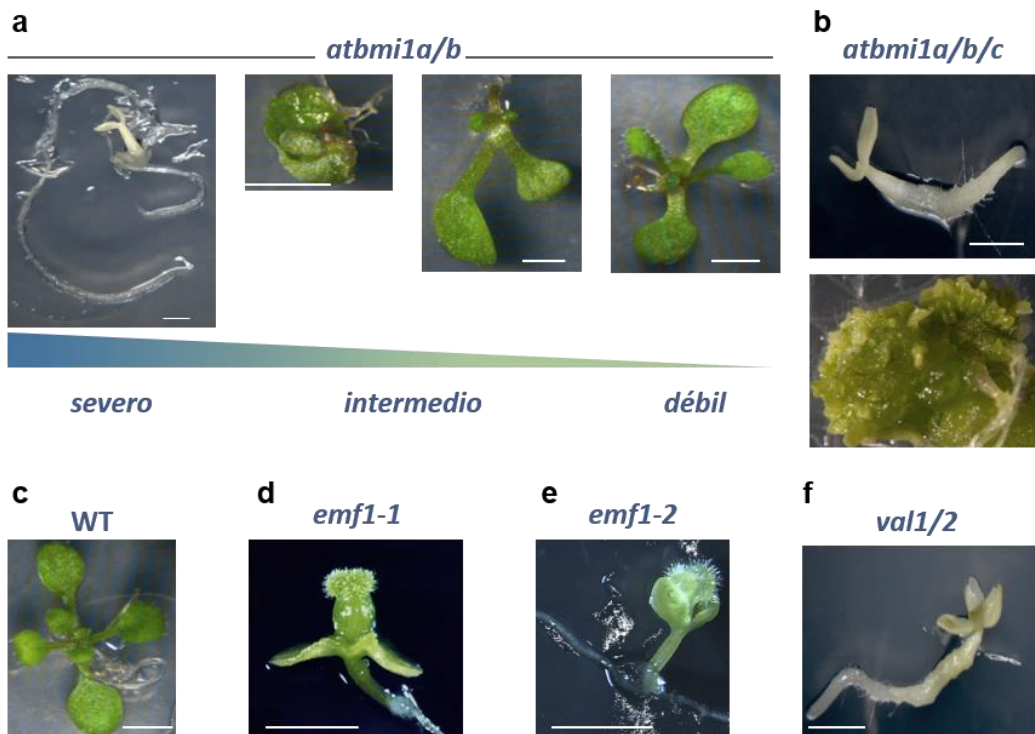


Figura 8: Fenotipos de los mutantes PRC1. (a) Fenotipos de los mutantes dobles *atbmi1a/b* a los 10 días después de la germinación (DAG). Los mutantes *atbmi1a/b* presentan un gradiente de fenotipos que va desde el severo al más débil. (b) Fenotipo del triple mutante *atbmi1a/b/c* a los 10 DAG (panel superior) y a los 30 DAG (panel inferior). (c) Fenotipo de WT-Col0 a los 10 DAG. (d) Fenotipo del mutante *emf1-2* a los 10 DAG. (e) Fenotipo del mutante *val1/2* a los 10 DAG. Barras = 2 mm.

Los mutantes *atbmi1a/b* presentan un gradiente de fenotipos que van desde plántulas con cotiledones blanquecinos cuyo desarrollo se detiene (mutante severo, Fig. 8a), pasando por plántulas con cotiledones verdes pero retorcidos que no pueden generar hojas verdaderas más adelante en el desarrollo (mutante intermedio, Fig. 8a), a plántulas con cotiledones verdes pero retorcidos que generan hojas retorcidas pero que pueden completar el desarrollo (mutantes débiles, Fig. 8a). Los mutantes severos e intermedios se quedan detenidos en la fase embrionaria formando estructuras similares a los callos y generando embriones somáticos (Bratzel *et al.*, 2010; Chen *et al.*, 2010). El hecho de que exista este gradiente de severidad del fenotipo es debido a la distinta penetrancia del alelo *AtBMI1B* (Bratzel *et al.*, 2010). En cualquier caso, el fenotipo de estos mutantes intermedios sugeriría que estas proteínas deben de regular, además de la transición del

desarrollo embrionario al vegetativo, otros estadios y transiciones a lo largo del desarrollo.

De acuerdo con esto, distintos datos apuntan a que las proteínas AtBMI1 pueden estar involucradas en la regulación del tiempo de floración. La sobre-expresión del gen *AtBMI1C* bajo el control del promotor CaMV35S produce un incremento de los niveles de monoubiquitinación de la histona H2A que resulta en un fenotipo de floración temprana. Por consiguiente, las líneas sobre-expresoras presentan unos niveles bajos del represor de floración *FLC* y una activación del gen promotor de la floración *FT*, lo que sugiere la participación de las proteínas AtBMI1 en el control del tiempo de floración (Li *et al.*, 2011). Además, se ha demostrado que la pérdida de función de las proteínas AtRING1A da lugar a un fenotipo de floración tardía (Shen *et al.*, 2014).

La proteína EMF1 participa en la represión del programa reproductivo en la fase vegetativa. Los mutantes *emf1* se saltan la fase vegetativa del desarrollo y florecen justo después de la germinación (Sung *et al.*, 1992; Yang, Chen y Sung, 1995; Kinoshita *et al.*, 2001). Los mutantes débiles *emf1-1* producen unas inflorescencias pequeñas después de la germinación y algunas flores estériles (Fig. 8d). Todas las hojas son pequeñas sin pecíolo y presentan características de hojas caulinares. En el mutante severo *emf1-2* los órganos laterales se convierten en carpelos [(Fig. 8e), (Sung *et al.*, 1992; Yang, Chen y Sung, 1995)]. El análisis estructural de los motivos de la proteína EMF1 sugiere que podría tener la función de regulador transcripcional (Moon *et al.*, 2003). Los mutantes *emf1* presentan expresión ectópica de varios genes, entre ellos los genes homeóticos de floración como *AP1*, *AP3*, *PISTILLATA (PI)* y *AG* (Moon *et al.*, 2003). Curiosamente, los mutantes *emf1-1* tienen exactamente el mismo fenotipo que los mutantes *emf2*, además, hay un alto grado de solapamiento de los genes desregulados en ambos mutantes (Moon *et al.*, 2003), sugiriendo una cooperación de EMF1 y EMF2 en la represión de los genes homeóticos de floración durante el desarrollo vegetativo (Chen *et al.*, 1997). Por otro lado, EMF1 interacciona *in vitro* con las proteínas AtBMI1A y AtBMI1B (Bratzel *et al.*, 2010).

LHP1 participa en la represión de la transición de la fase vegetativa a la fase reproductiva, y en la vernalización. El mutante simple *lhp1^{-/-}* presenta un fenotipo pleiotrópico que incluye hojas de roseta más pequeñas, estrechas y curvadas hacia arriba, un peso reducido de la planta (más de un 50% de reducción), floración temprana, una estructura de flor terminal y silicuas con tamaño más pequeño (Larsson, Landberg y Meeks-Wagner, 1998; Gaudin *et al.*, 2001; Kotake *et al.*, 2003). Distintos ensayos han demostrado que LHP1 interacciona con los componentes del PRC1, AtRING1 y AtBMI1 (Xu y Shen, 2008; Bratzel *et al.*, 2010; Chen *et al.*, 2010) y también con EMF1 (Wang *et al.*, 2014), indicando que LHP1 podría formar parte del complejo PRC1 en Arabidopsis. Sin embargo, LHP1 también co-purifica con el complejo PRC2 (Mylne *et al.*, 2006; H. Wang

et al., 2016). Además, los genes diana de LHP1 responsables del fenotipo del mutante, son también dianas de los complejos PRC2, incluidos *FT*, *FLC* y *AG* (Kotake *et al.*, 2003; Mylne *et al.*, 2006; Sung, Schmitz y Amasino, 2006). Por lo tanto, LHP1 podría actuar de puente entre PRC1 y PRC2 puesto que puede interactuar con los componentes de ambos complejos (Feng y Lu, 2017).

Otras proteínas relacionadas con las proteínas PcG son las proteínas VAL (VP1/ABI3-LIKE) 1/2/3 que pertenecen a una subfamilia de proteínas específicas de plantas que contienen un dominio B3 (Swaminathan, Peterson y Jack, 2008), que se unen a los elementos *LEC2/ABI3/VP1* también conocidos como elementos *RY* (*CATGCA*) (Suzuki, Wang y McCarty, 2007). Se ha demostrado que las proteínas VAL1/2 y AtBMI1 son necesarias para la represión del programa de maduración de semilla tras la germinación (Suzuki y McCarty, 2008; Guerriero *et al.*, 2009; Bratzel *et al.*, 2010). De acuerdo con esto, los mutantes *atbmi1a/b/c* y *val1/2* presentan el mismo fenotipo (Yang *et al.*, 2013). Además, estudios previos del grupo han demostrado que VAL1/2 interactúa con AtBMI1A/B *in vitro* y que son necesarias para la monoubiquitinación de H2A en los genes de maduración de semilla, indicando que VAL1/2 actúa aguas arriba de AtBMI1 en la represión (Yang *et al.*, 2013). Dado que los promotores de los genes de maduración de semilla contienen los elementos *RY/Sph*, se ha propuesto que las proteínas VAL están involucradas en el reclutamiento específico de PRC1 a estos genes (Yang *et al.*, 2013). De acuerdo con estos resultados, estudios recientes han demostrado que una mutación puntual en una secuencia *RY* localizada en el primer intrón de *FLC* previene su silenciamiento epigenético durante la vernalización (Qüesta *et al.*, 2016; Yuan *et al.*, 2016).

Por tanto, no estaba clara la identidad del PRC1. De acuerdo con estos datos, es probable que existan distintos complejos dependiendo de la combinación de distintas subunidades. Tampoco estaba claro si los distintos parálogos de AtBMI1, a pesar de que el fenotipo de los mutantes dobles y triples sugería que actúan de forma redundante, podrían desempeñar funciones específicas.

2. Objetivo

Nos propusimos investigar el papel que desempeñan las distintas proteínas AtBMI1 a lo largo del desarrollo, su relación con otros componentes de PRC1 y su conexión con el complejo PRC2.

Distintos resultados apoyan la participación de las proteínas AtBMI1 en la regulación de la transición del desarrollo embrionario al vegetativo mediante la represión de los genes de maduración de semilla durante la germinación. Sin embargo, no se había investigado en profundidad su posible implicación en otros procesos o etapas de desarrollo. Tampoco estaba claro si las distintas proteínas AtBMI1 regulan distintos genes de forma específica. Por otro lado, a pesar de que EMF1 y LHP1 participan en la represión del desarrollo reproductivo durante la fase vegetativa, su papel en la regulación de otros procesos tampoco estaba claro. Es más, a pesar de que estas proteínas interactúan con las proteínas AtBMI1, las grandes diferencias que muestran los fenotipos de los mutantes *atbmi1* y *emf1* a priori sugieren que estas proteínas regulan distintos procesos de desarrollo. Tampoco se conocía si las proteínas AtBMI1 actúan siempre en combinación con los complejos PRC2 o podrían actuar de forma independiente. Más aún, a pesar de que la pérdida de función de las proteínas VAL da lugar al mismo fenotipo que el del mutante *atbmi1a/b/c*, y el mutante *val1val2* también sobre-expresa los genes de maduración de semilla, no se sabe si los factores VAL y AtBMI1 operan juntos en la regulación de otros genes.

Con el fin de contestar estas preguntas, se llevó a cabo por un lado la caracterización de los mutantes débiles *atbmi1a/b*, y por otro el análisis transcriptómico de los diferentes mutantes *atbmi1*, así como el de mutantes en diferentes proteínas PcG tanto de PRC1 como de PRC2.

3. Resultados

3.1. La acción coordinada de AtBMI1 y EMF1 es necesaria para la transición de la fase juvenil a la fase reproductiva

El fenotipo que presentan los mutantes *atbmi1a/b* y *atbmi1a/b/c* severos, consistente en plantas que tras germinar continúan en estado embrionario, permitió intuir que las proteínas AtBMI1 jugaban un papel importante en la transición de la fase embrionaria a la vegetativa a través de la represión de los genes embrionarios, pero no permitía vislumbrar otras posibles funciones. Sin embargo, el hecho de que el alelo *atbmi1b* manifieste una penetrancia variable llegándose a expresar en cierta medida en algunas plantas de forma estocástica (Bratzel *et al.*, 2010), hace que estos mutantes presenten un gradiente de fenotipos. Tal es el caso que, algunos de estos mutantes, los más débiles, son capaces de florecer y generar semillas viables (Bratzel *et al.*, 2010), lo que ha resultado ser una ventaja para poder estudiar otros procesos de desarrollo. Cabe destacar que los mutantes *atbmi1a/b* débiles no presentan un fenotipo de floración temprana como se ha observado en otros mutantes PcG como *emf1* o *emf2*, lo que es sorprendente dado que se supone que PRC1 y PRC2 regulan conjuntamente los genes diana.

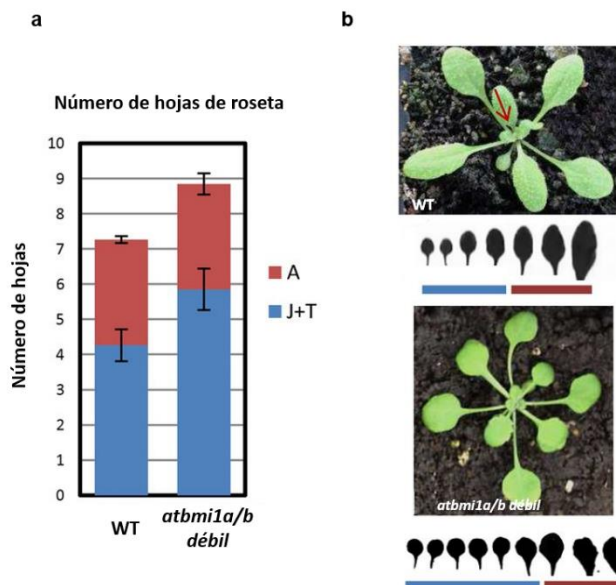


Figura 9: (a) Tiempo de floración de plantas silvestres WT-Colo y *atbmi1a/b* debil. El tiempo se midió en número de hojas de roseta producidas a partir del meristemo apical antes de la floración; se calificaron 16 a 20 plantas de cada línea. Las barras de error indican la desviación estándar (SD). **(b)** Se diferenciaron las hojas juveniles (J) y de transición (T) de las hojas adultas (A) por su forma.

Para estudiar si las proteínas AtBMI1 juegan algún papel en la regulación del proceso de floración, primero analizamos si el tiempo de floración estaba afectado en los mutantes *atbmi1a/b* débiles. Se comparó el tiempo de floración en días y en número de hojas de roseta antes de generar el botón floral entre las plantas mutantes *atbmi1a/b* débiles y WT-Col0 bajo condiciones de día largo (DL) (Fig. 9a).

Se observó que los mutantes *atbmi1a/b* débiles florecían 3 días más tarde que las plantas WT-Col0 y presentaban dos hojas de más antes de desarrollar el botón floral (Fig. 9a). Estas hojas presentaban una forma redonda y un peciolo largo (Fig. 9b), que son características de plantas juveniles (Wu *et al.*, 2009), lo que sugería que estos mutantes presentaban una fase juvenil prolongada.

La sobre-expresión de *miR156* prolonga la fase vegetativa y retrasa la transición floral (Hwan Lee, Joon Kim y Ahn, 2012). *miR156* está codificado por ocho genes en Arabidopsis [*MIR156A* a *MIR156H*; (Reinhart *et al.*, 2002)]. Dos de estos genes, *MIR156A* y *MIR156C* han sido identificados como dianas directas del gen de maduración de semilla *FUSCA3* (*FUS3*). *FUS3* activa la expresión de *MIR156A/C* durante el desarrollo de la semilla, y esta expresión es importante después de la germinación para retrasar la transición de la fase juvenil a la fase adulta (Wang y Perry, 2013). Además, *MIR156A* y *MIR156C* contienen elementos *RY* en su extremo 5' y en el gen, que son los elementos de DNA reconocidos específicamente por los dominios B3 de unión al DNA (B3 DNA-binding domain) de la proteína *FUS3* (Wang y Perry, 2013).

Puesto que *FUS3* está sobre-expresado en los mutantes *atbmi1* y *clf-28/swn-7* pero no en *emf1* o *emf2* [(Yang *et al.*, 2013); Fig. 10a], se analizaron los niveles de los transcritos de *pri-MIR156A/C* en estos mutantes (Fig. 10a). En los tres mutantes *atbmi1a/b*, los niveles de *pri-MIR156A/C* aumentaban de forma drástica a los 14 DAG, especialmente en los fenotipos intermedio y severo, y en *clf-28/swn-7* (Fig. 10a), pero no se encontraron alterados en *emf1-2* y *emf2-2* (Fig. 10a). Además, los niveles de los *pri-MIR156* eran similares en los mutantes *atring1a/b* y *atbmi1a/b* débiles (Fig. 10b), lo que sugiere que la expresión de *miR156* está regulada tanto por AtBMI1 como AtRING1. Estos resultados indican que la prolongación de la fase juvenil en los mutantes *atbmi1a/b* débiles es consecuencia de la sobre-expresión de *miR156*. Sin embargo, esto podría ser un efecto indirecto de la expresión ectópica de *FUS3* que muestran estos mutantes.

Para determinar si las proteínas AtBMI1 juegan un papel directo en la regulación de la expresión de *pri-miR156A/C*, se compararon los niveles de las marcas H2AK121ub en la región del sitio de inicio de la transcripción (TSS) de *MIR156A* y *MIR156C* en plantas WT-Col0 y *atbmi1a/b* débil a los 10 DAG. Los resultados mostraron que los niveles de estas

marcas en *MIR156A* estaban disminuidos en los mutantes *atbmi1a/b* y que los niveles en *MIR156C* aparentemente disminuían, aunque la variación experimental fue grande (Fig. 11a). Puesto que la actividad AtBMI1 se requiere para la deposición de las marcas H3K27me3 mediada por PRC2 en varios genes diana (Yang *et al.*, 2013), se analizaron los niveles de las marcas H3K27me3 en el TSS de estos genes (Fig. 11b), encontrándose que los niveles estaban reducidos en el mutante *atbmi1a/b* débil en ambos casos (Fig. 11b); más aún, las marcas H3K27me3 desaparecían de estos genes en el triple mutante *atbmi1a/b/c* (Fig. 11b), lo que indica que *MIR156A* y *MIR156C* están regulados directamente por la maquinaria PcG.

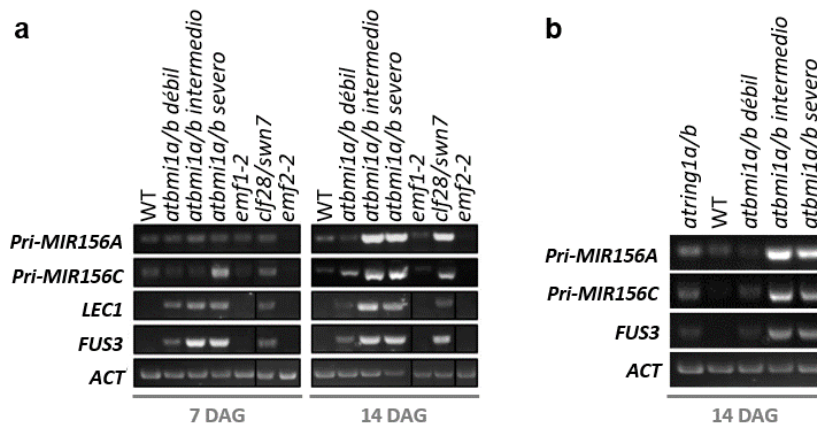


Figura 10: Los mutantes *atbmi1a/b* sobre-expresan *MIR156A* y *MIR156C*. (a) Niveles de expresión de *pri-MIR156A*, *pri-MIR156C*, y los genes de maduración de semilla *LEAFY COTYLEDON1(LEC1)* y *FUS3* en los diferentes mutantes a los 7 y 14 DAG crecidos en condiciones de DL a ZT1. (b) Niveles de expresión de *pri-MIR156A*, *pri-MIR156C* y *FUS3* en WT y mutantes *ating1a/b* y *atbmi1a/b* a los 14 DAG crecidos en condiciones de DL a ZT1. Se utilizó *ACT2* como control interno. *FUS3* y *pri-MIR156A/C* se sobre-expresan en los mutantes *ating1a/b* al mismo nivel que en *atbmi1* débiles.

En la transición de la fase juvenil a la adulta, las plantas adquieren la competencia para florecer. En condiciones normales, conforme avanza la edad de la planta, los niveles de *miR156* se reducen y aumenta la expresión de *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL)*. SPL9 activa la expresión de *pri-MIR172b* que a su vez reprime a los represores florales *AP2-like* que inhiben la expresión de *FT* (Wang *et al.*, 2014). También, SPL3 regula directamente la expresión de *FT* (Kim *et al.*, 2012). Además, los altos niveles de *miR156* reducen la habilidad de FT para inducir la floración mediante la represión de la actividad de SPL en el meristemo apical (SAM) (Wang, Czech y Weigel, 2009). Por lo tanto, la acción de los SPL y *miR172* contribuye al establecimiento del umbral de FT necesario para la floración y además a preparar el SAM para que responda a la señal de floración.

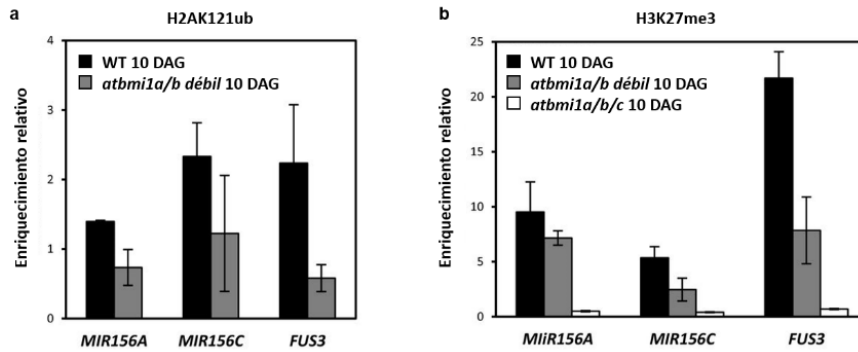


Figura 11: MIR156A y MIR156C son dianas directas de AtBMI1. (a) Análisis ChIP de los niveles de H2AK121ub en el TSS de MIR156A y MIR156C en plantas WT y *atbmi1a/b* débil a los 10 DAG. Se utilizó FUS3 como control positivo. (b) Análisis ChIP de los niveles de H3K27me3 en el TSS de MIR156A y MIR156C en plantas WT, *atbmi1a/b* débil y *atbmi1a/b/c* a los 10 DAG. Se utilizó FUS3 como control positivo. Los DNA inmunoprecipitados se cuantificaron y se normalizaron frente a ACT7. Las barras de error indican la desviación estándar SD de al menos dos réplicas biológicas.

Para determinar si los niveles de expresión de *pri-MIR156A/C* en los mutantes *atbmi1* se correlacionaban con los niveles de *SPL3*, *SPL9*, *pri-MIR172b* y *FT*, y si los patrones de expresión de estos genes podían explicar el tiempo de floración, se comparó la expresión de todos ellos en los mutantes y WT-Col0 de 10 días de edad (Fig. 12). Los resultados muestran que de acuerdo con los niveles de *pri-MIR156A/C* en los mutantes *atbmi1a/b*, los niveles de *SPL3*, *SPL9* y *pri-MIR172b* son más bajos en los mutantes confirmando su estado juvenil. En consecuencia, los niveles de *FT* son más bajos en estos mutantes, lo que resulta en un retraso de la floración en los mutantes *atbmi1a/b* débiles. En el caso de los mutantes intermedios y severos no se observa nunca floración, esto no solo debe ser consecuencia de los niveles bajos de estos genes sino también a causa de su desarrollo detenido en la fase embrionaria.

Por otro lado, los niveles de expresión de *SPL3*, *SPL9* y *pri-MIR172b* estaban elevados en los mutantes *emf1-2* (Fig. 12). Estos resultados apoyan datos anteriores en los que se ha demostrado que *SPL9* es una diana de EMF1 (Kim *et al.*, 2012); por lo tanto, la desrepresión de *SPL9* puede dar lugar a la activación de *pri-MIR172b* en los mutantes *emf1-2*. Además, se ha visto que *SPL3* está sobre-expresado en las plantas transgénicas que expresan un cDNA de EMF1 antisentido bajo el control del promotor del gen de identidad del meristemo floral *LEAFY* (*LFY:asEMF1*; (Pu *et al.*, 2013)). También, se ha demostrado que varios genes *MIR172* son dianas directas de EMF1 (Kim *et al.*, 2012). Los niveles de expresión de los genes *SPL3*, *SPL9* y *pri-MIR172b* aumentan también en el mutante *emf2-2*, aunque no tanto como en *emf1-2*, probablemente debido a la redundancia de EMF2 y VRN2 en la regulación de estos genes, ya que ambos co-regulan un conjunto de genes diana (Lafos *et al.*, 2011). Estos resultados indican que EMF1 y EMF2

regulan directa o indirectamente los niveles de *miR172*. Notablemente, los niveles de *pri-MIR156*, *SPLs* y *pri-MIR172b* en *emf1-2* y *emf2-2* pueden explicar la expresión de *FT* independiente de *CONSTANS* (*CO*) y la adquisición de la competencia para la floración tan temprana de estos mutantes.

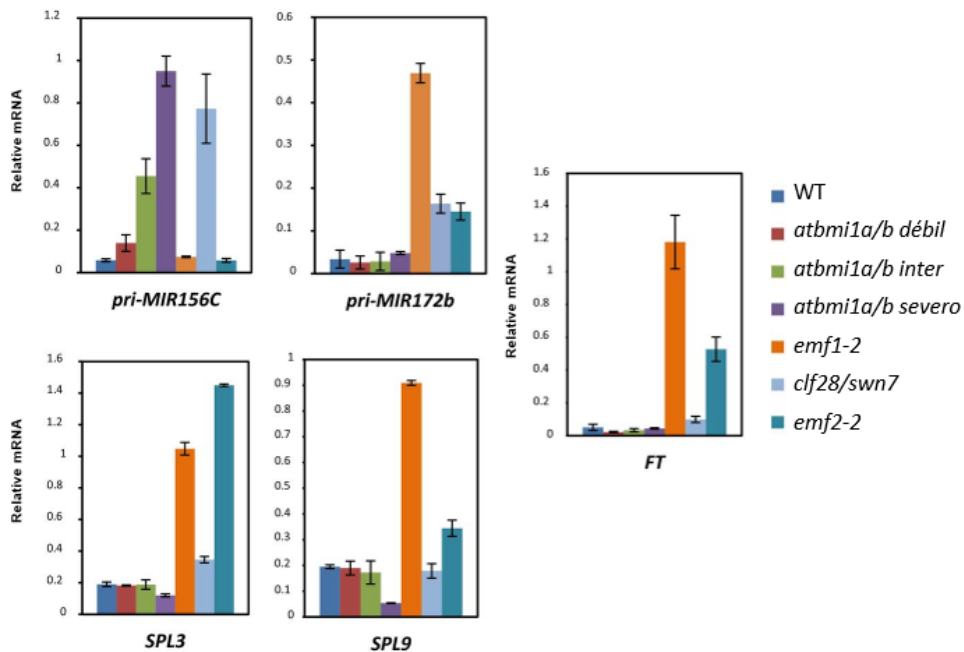


Figura 12: Regulación de *miR156* y *miR172* mediada por AtBMI1-PRC1. Se muestran los niveles de expresión de *pri-MIR156C*, *pri-MIR172b*, *SPL3*, *SPL9* y *FT* en plantas WT y mutantes a los 10 DAG. Las cuantificaciones fueron normalizadas frente a *ACT2*. Las barras de error indican la desviación estándar (SD) de dos réplicas biológicas.

En los mutantes de pérdida de función de PRC2 *clf-28/swn-7*, los niveles de *SPL3* y *SPL9* eran solo algo más altos que en WT-Col0 (Fig. 12), y la expresión de *pri-MIR172b* no llegaba a los niveles encontrados en el mutante *emf1-2*. Es muy probable que los altos niveles de *pri-MIR156A/C* en estos mutantes afecten a la expresión de *pri-MIR172b* mediante la reducción de los niveles de *SPL*. De acuerdo con estos datos, *clf-28/swn-7* no presenta altos niveles de expresión de *FT*.

Estos resultados ponen de manifiesto las funciones coordinadas de las proteínas AtBMI1 y EMF1 en la regulación de la transición de la fase juvenil a la adulta mediante la represión de *miR156* y *miR172*, respectivamente.

3.2. Análisis del transcriptoma global de los distintos mutantes *atbmi1*

De acuerdo con estos y otros resultados anteriores (Bratzel *et al.*, 2010; Yang *et al.*, 2013), parece que las proteínas AtBMI1 están implicadas en la regulación de la transición de la fase embrionaria a la vegetativa y en la transición de la fase juvenil a la adulta; sin embargo, no se sabía si jugaban algún papel en la regulación de otros procesos de desarrollo. Por otra parte, tampoco estaba claro si las tres proteínas AtBMI1 regulan los mismos genes de forma redundante o por el contrario presentan diversidad de funciones.

Estudios previos mostraron que los genes *AtBMI1A* y *AtBMI1B* se expresan en todos los tejidos de la planta y actúan principalmente de forma redundante a lo largo del desarrollo (Bratzel *et al.*, 2010). Sin embargo, *AtBMI1C* solo se expresa en raíces, endospermo y estambre. Además, su sobre-expresión no logra rescatar los defectos del doble mutante *atbmi1a/b* en la parte aérea y rescata el fenotipo de la raíz solo parcialmente (Yang *et al.*, 2013; Merini y Calonje, 2015). Esto sugiere un proceso de divergencia funcional de *AtBMI1C* (Yang *et al.*, 2013; Merini y Calonje, 2015). No obstante, los mutantes dobles *atbmi1a/c* y *atbmi1b/c* no presentan alteraciones fenotípicas (Yang *et al.*, 2013), lo que indica que la pérdida de la función de *AtBMI1C* es compensada por las otras proteínas AtBMI1.

Por tanto, para estudiar el papel de las diferentes proteínas AtBMI1 en la regulación de la expresión génica, se llevaron a cabo estudios transcriptómicos utilizando la técnica de secuenciación masiva del RNA (RNA-seq) de los distintos mutantes *atbmi1*. Nuestro diseño experimental consistió en dos réplicas biológicas recogidas 10 días después de la germinación de los siguientes genotipos: silvestre Col-0 (WT), *atbmi1a*, *atbmi1b*, *atbmi1a/b* y *atbmi1a/b/c*. Los mutantes dobles *atbmi1a/b* presentan un amplio rango de variación fenotípica debido a la distinta penetrancia del alelo *atbmi1b* (Bratzel *et al.*, 2010). Por lo que para el análisis se seleccionó el mutante *atbmi1a/b* con el fenotipo más severo, el cual se diferencia del triple mutante *atbmi1a/b/c* principalmente en que son capaces de desarrollar raíz [(Yang *et al.*, 2013); (Fig. 8a, b)].

El análisis computacional de los datos de RNA-seq se realizó usando el sistema de colas SGE (del inglés *Sun Grid Engine*), siguiendo el protocolo Tuxedo (Trapnell *et al.*, 2012) para el ensamblado de los transcritos y el análisis de la expresión génica diferencial. Las muestras de secuenciación fueron de alta calidad (Fig. 13; Tabla 3). Los genes expresados diferencialmente o DEGs (del inglés *Differential Expression Genes*) fueron determinados utilizando un criterio restrictivo que consiste en la combinación de un factor de proporción o FC (del inglés *fold change*) mayor que 4 y un q-valor inferior a 0,05 ya que se observó un efecto severo en el transcriptoma de los mutantes con respecto al tipo silvestre. El número de genes detectados como presentes o expresados en al menos una de las muestras fue de 24503, representando el 72,96% del transcriptoma completo

de Arabidopsis. Para la preparación de la librería se realizó una extracción polyA para centrarnos en mRNA, lo que provoca que se pierdan los RNA no codificantes lo cual explica el porcentaje de detección.

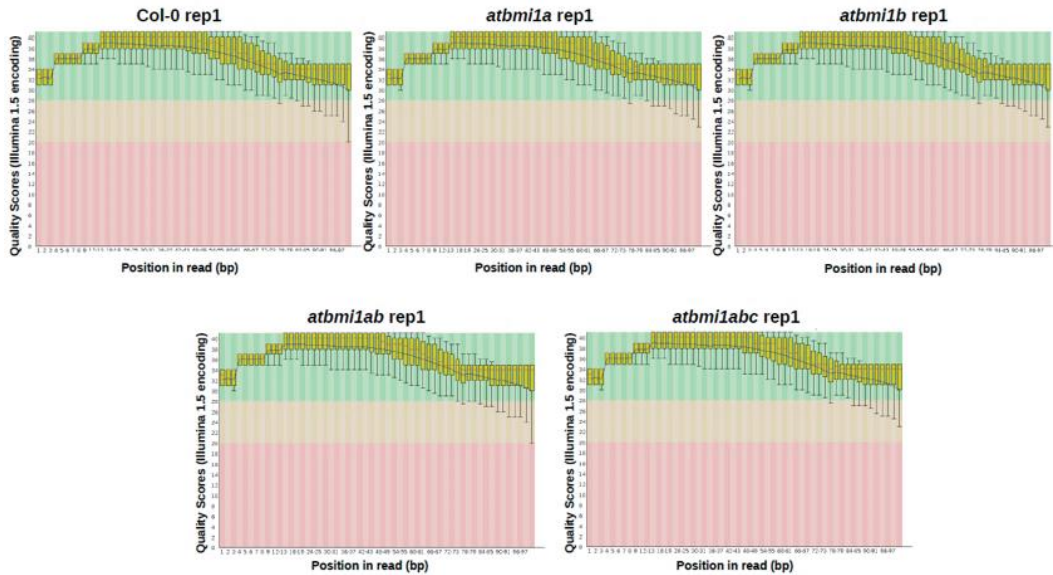


Figura 13: Análisis de la calidad de secuenciación. Diagrama de cajas representando el índice de calidad de cada base en función de su posición (Illumina 1.5 encoding), en la primera réplica de cada una de las muestras. En todas las muestras el índice de calidad se encuentra en el área verde indicando una alta calidad de secuenciación. Se observa una disminución de la calidad al final de las lecturas; sin embargo, nunca entran en el área roja.

Tabla 3: Número de lecturas y porcentaje de alineamiento con el genoma de referencia.

Muestra	Número de lecturas	Porcentaje de alineamientos concordantes
WT-Col0 rep1	14578745	96,5%
WT-Col0 rep2	16253159	96,6%
<i>atbmi1a</i> rep1	14982986	94,5%
<i>atbmi1a</i> rep2	18324516	96,1%
<i>atbmi1b</i> rep1	15714678	95,8%
<i>atbmi1b</i> rep2	18412180	95,3%
<i>atbmi1a/b</i> rep1	14885215	94,9%
<i>atbmi1a/b</i> rep2	14547885	95,1%
<i>atbmi1a/b/c</i> rep1	14832480	94,5%
<i>atbmi1a/b/c</i> rep2	12349849	94,9%

De media, el número de lecturas en cada muestra es aproximadamente 15 millones y la media de la tasa de alineamiento de pares concurrentes es mayor que el 95%. Esto indica una alta calidad de secuenciación y la ausencia de contaminación de las muestras.

La comparación de los diferentes transcriptomas de los mutantes con el del tipo silvestre WT mostró que los mutantes simples tenían alterado un 3-4% del transcriptoma mientras que los mutantes dobles *atbmi1a/b* y triples *atbmi1a/b/c* tenían alterado el 15% y 20%, respectivamente (Fig. 14a). Al llevar a cabo un análisis de componentes principales (PCA), encontramos que los transcriptomas de WT y los mutantes simples *atbmi1a* y *atbmi1b* formaban un mismo grupo (*cluster*), mientras que los transcriptomas de los mutantes *atbmi1a/b* y *atbmi1a/b/c* constituían dos grupos distintos y distantes tanto al *cluster* constituido por WT y los mutantes simples como entre sí, lo que indica que estos mutantes presentan diferencias importantes (Fig. 14b).

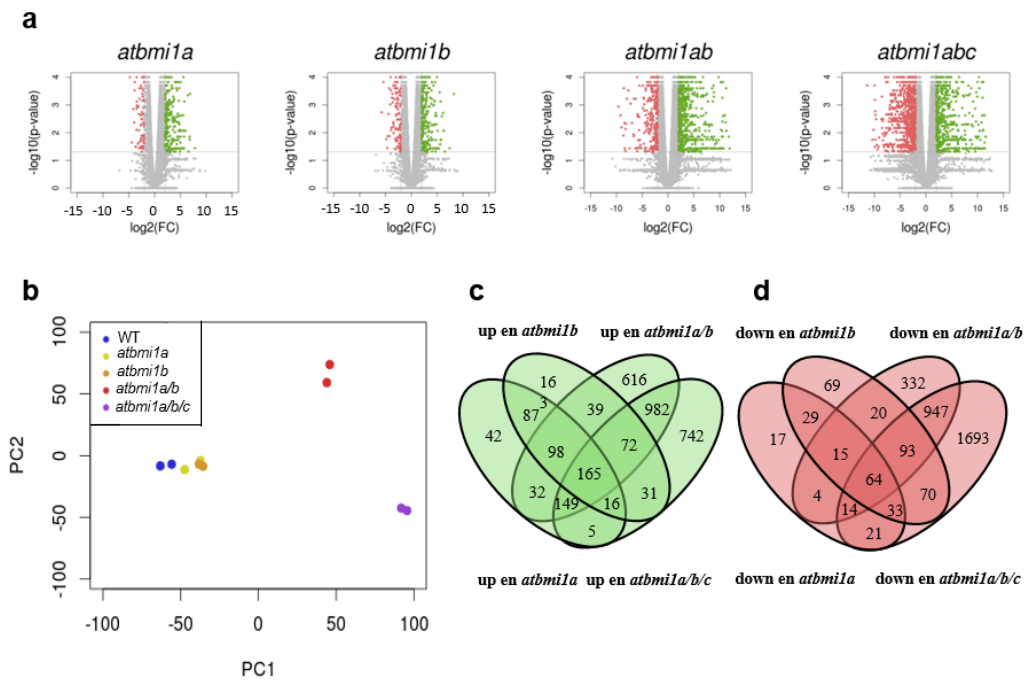


Figura 14: Análisis transcriptómico de WT y los diferentes mutantes *atbmi1* seleccionados a los 10 DAG. (a) Diagramas de dispersión (Volcano plots) representando los genes que se expresan de forma diferencial (DEGs) en los mutantes *atbmi1* comparados con el WT con un FC 4 y q-valor < 0,05. Lo genes activados en los mutantes de forma diferencial se representan en color verde y los reprimidos en color rojo. **(b)** Análisis de componentes principales (PCA) que muestra que WT, *atbmi1a* y *atbmi1b* forman grupos (*cluster*) similares, mientras que *atbmi1a/b* y *atbmi1a/b/c* forman dos grupos distintos. **(c, d)** Alteración de la expresión génica en los mutantes *atbmi1*. Diagramas de Venn con los genes activados (up) (c) y los reprimidos (down) (d) en los diferentes mutantes.

Seguidamente, llevamos a cabo el análisis de los genes que presentaban expresión diferencial en cada tipo de mutante. En términos generales encontramos que los mutantes *atbmi1a*, *atbmi1b* y *atbmi1a/b* presentaban un mayor número de genes activados que de genes reprimidos (Tabla 4), lo que apoya el papel de las proteínas AtBMI1 en la represión transcripcional; sin embargo, en los mutantes *atbmi1a/b/c* el número de genes reprimidos era mayor que el número de genes activados lo cual a priori puede resultar sorprendente pero, teniendo en cuenta que el desarrollo de estos mutantes se detiene en la fase de maduración de la semilla y no desarrollan ningún órgano vegetativo, la activación de determinados genes en este contexto puede tener un fuerte impacto negativo en la expresión génica. Por otro lado, la falta de activación en el mutante triple de muchos de los genes activados en el mutante doble puede deberse a la ausencia de factores específicos de tejido requeridos para su activación, al carecer los mutantes *atbmi1a/b/c* de tejidos diferenciados. La comparación de los genes expresados de forma diferencial en los distintos mutantes puso de manifiesto que la mayoría de los genes con

expresión alterada en los mutantes simples estaban también alterados en los mutantes doble y triple (Fig. 14c y d).

Tabla 4: Número de genes que se expresan de forma diferencial (activados y reprimidos) en los diferentes mutantes.

Expresión génica	<i>atbmi1a</i>	<i>atbmi1b</i>	<i>atbmi1a/b severo</i>	<i>atbmi1a/b/c</i>
Genes activados	594	671	2153	2162
Genes reprimidos	197	393	1489	2935

A continuación, llevamos a cabo un estudio del enriquecimiento funcional en términos de Ontología génica (GO) de los genes activados en los mutantes severos *atbmi1a/b* y *atbmi1a/b/c* (Anexos 1 y 2). Encontramos que estos genes están principalmente asociados con la respuesta a diferentes estímulos (e.g. estrés hídrico, temperatura, hormonas) y con el metabolismo de lípidos (e.g. transporte, biosíntesis, almacenamiento). El mismo estudio con los genes reprimidos mostró un enriquecimiento en los términos GO relacionados con la fotosíntesis y los procesos metabólicos (Anexos 1 y 2). Esto encaja bien con el estado de desarrollo de los mutantes que se quedan atrapados en la fase de maduración de semilla (Yang *et al.*, 2013). Durante esta fase, las semillas adquieren tolerancia a la desecación y acumulan productos de reserva que prevalecen en forma de lípidos (Vicente-Carbajosa y Carbonero, 2005), mientras que la estructura de los cloroplastos se desorganiza (Delmas *et al.*, 2013).

Las proteínas PcG están involucradas en la represión de genes reguladores centrales (Xiao y Wagner, 2015), por lo tanto, el alto grado de desregulación génica encontrada en los diferentes mutantes *atbmi1* puede ser consecuencia directa o indirecta de la pérdida de función de AtBMI1, o la combinación de las dos. También hay que tener en cuenta que no todos los genes directamente regulados por las proteínas AtBMI1 tienen que presentar una expresión alterada ya que puede darse el caso de que no estén presentes sus activadores transcripcionales, como ha sido demostrado en otros mutantes de pérdida de función de PcG (Bouyer *et al.*, 2011; Kim *et al.*, 2012; Derkacheva *et al.*, 2013). Por lo tanto, para focalizar el estudio en los genes más probables de ser dianas directas de la regulación PcG se consideraron únicamente los genes activados, al estar estas proteínas implicadas en la represión de la transcripción, y que portaban las marcas H3K27me3 establecidas por el PRC2, ya que las actividades de PRC1 y PRC2 se requieren para mantener la represión estable de un gran número de genes en otros organismos (Schwartz y Pirrotta, 2013; Comet y Helin, 2014). Para identificar estos genes activados y marcados con H3K27me3 (up_K27), se intersecaron los genes activados en los diferentes

mutantes con el conjunto de 5360 genes marcados con H3K27me3 en WT, previamente identificados en dos análisis independientes en plantas del mismo estado de desarrollo (Bouyer *et al.*, 2011; Kim *et al.*, 2012). El número de genes en cada genotipo se indica en la tabla 5. Se encontró un solapamiento significativo entre los genes marcados con H3K27me3 y los genes activados en los diferentes mutantes excepto en *atbmi1b* (Fig. 15a), probablemente porque la expresión del gen está reducida (“*knock-down*”), no completamente eliminada (Bratzel *et al.*, 2010), y por tanto, el número de genes desregulados no es tan alto como en los otros casos. El mismo análisis utilizando los genes reprimidos mostró un solapamiento no-significativo en todos los casos excepto en *atbmi1a/b/c* debido al alto número de genes reprimidos en este mutante (Fig. 15b) probablemente debido a efectos secundarios.

Tabla 5: Número de genes activados en los diferentes mutantes y marcados con H3K27me3 en plántulas WT de la misma edad (up_K27)

Expresión génica	<i>atbmi1a</i>	<i>atbmi1b</i>	<i>atbmi1a/b</i> <i>severo</i>	<i>atbmi1a/b/c</i>
Genes Up_K27	196	145	766	765

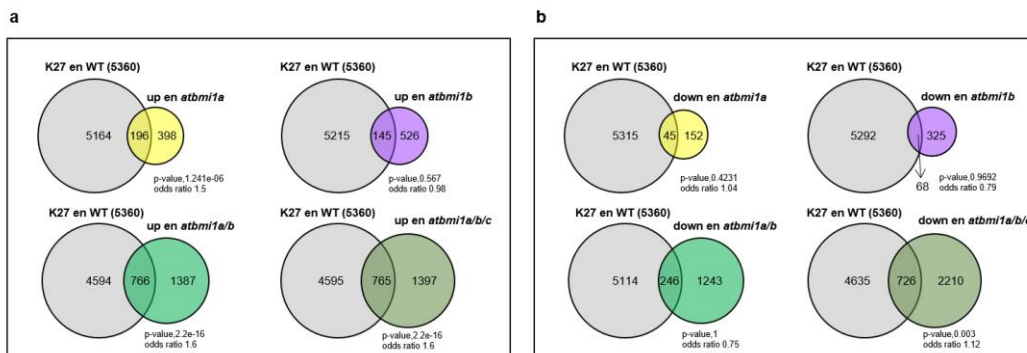


Figura 15: Posibles dianas directas de AtBMI1. (a) Representación mediante diagramas de Venn de los genes activados (up) en los diferentes mutantes y marcados con H3K27me3 (K27) en plántulas WT de la misma edad. Todos estos solapamientos son significativos con un p-valor menor que $1,2 \times 10^{-6}$ y odds ratios mayor que 1,5 de acuerdo con el test exacto de Fisher excepto en el caso del mutante *atbmi1b* probablemente porque es un mutante *knock-down*. **(b)** Representación mediante diagramas de Venn de los genes activados (up) en los diferentes mutantes y marcados con H3K27me3 (K27) en plántulas WT de la misma edad. Todos estos solapamientos no son significativos, el p-valor es mayor que 0,4231 y odds ratios menor que 1,044 de acuerdo con el test exacto de Fisher excepto en el caso del mutante *atbmi1a/b/c*, debido probablemente a su estado de desarrollo.

Para investigar la existencia de genes diana específicos de AtBMI1A y AtBMI1B, se compararon los genes up_K27 en los mutantes simples y dobles (Fig. 16a). El número de genes en el mutante doble es considerablemente más alto que en los mutantes simples, lo que indica un alto grado de redundancia funcional entre AtBMI1A y AtBMI1B. Además, la mayoría de los genes up_K27 en los mutantes simples están incluidos en el grupo de los genes up_K27 en el mutante doble; sin embargo, hay un grupo de genes que se activan exclusivamente en *atbmi1a* y *atbmi1a/b* o en *atbmi1b* y *atbmi1a/b* (104 y 27 genes, respectivamente).

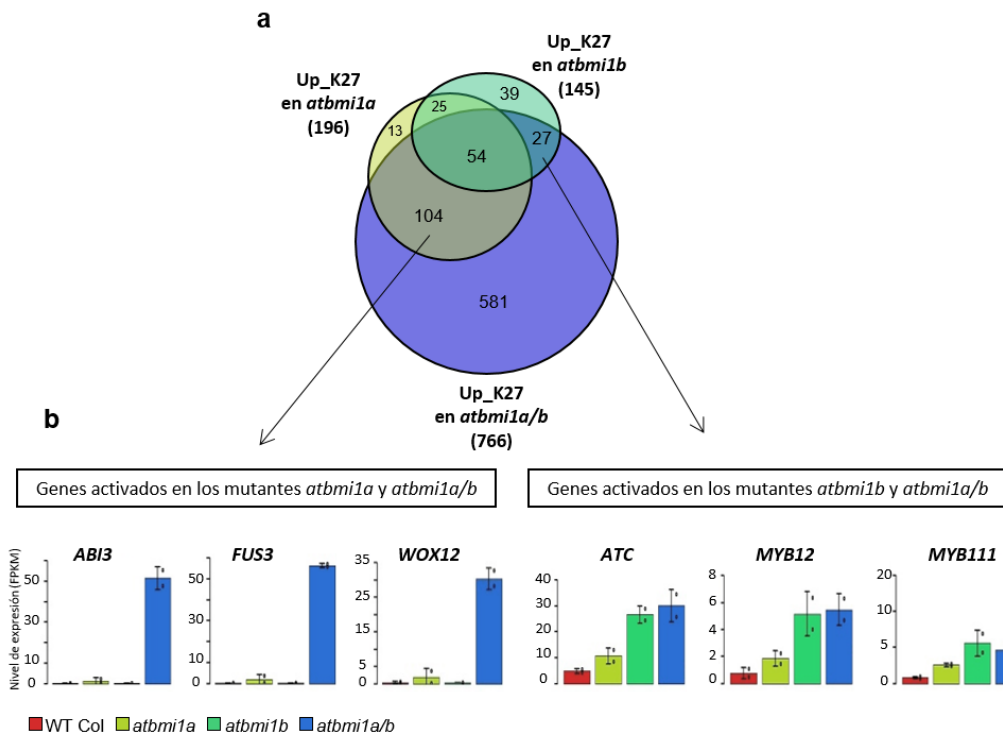


Figura 16: Genes regulados por AtBMI1A y AtBMI1B. (a) Diagrama de Venn que muestra el número de los genes up_K27 que solapan entre los mutantes *atbmi1a*, *atbmi1b* y *atbmi1a/b*. Todos los solapamientos son significativos con un p-valor inferior a $2,2 \times 10^{-16}$ y odds ratios mayor que 17 de acuerdo con el test exacto de Fisher (b) Niveles de expresión de los genes que están aparentemente activados exclusivamente en los mutantes *atbmi1a* o *atbmi1b* en los diferentes genotipos.

Los genes up_K27 en *atbmi1a* y *atbmi1a/b* parece que podrían ser dianas específicas de AtBMI1A; sin embargo, el análisis de la expresión en los diferentes mutantes mostró que su nivel de expresión era muy bajo en comparación con el mutante doble (Fig. 16b), lo que indica que AtBMI1B es capaz de compensar la pérdida de función de AtBMI1A, casi en su totalidad, en la regulación de estos genes. Por otro lado, el hecho de que el mutante *atbmi1b* produzca bajos niveles de transcrito de *AtBMI1B* puede

explicar que la expresión de alguno de estos genes esté más afectada en el mutante simple *atbmi1a* (Bratzel *et al.*, 2010). No obstante, se encontró un número de genes específicamente sensibles a la pérdida de función de AtBMI1B que su expresión no incrementaba en los mutantes dobles (Fig. 16b), indicando una cierta especificidad en la regulación de estos genes.

Para investigar el grado de redundancia entre AtBMI1A/B y AtBMI1C, se compararon los genes up_K27 en *atbmi1a/b* y *atbmi1a/b/c* encontrándose tres grupos de genes. Un primer grupo (Cluster I) de genes activados en los dos mutantes (aproximadamente 2/3 de los genes de cada genotipo), un segundo grupo (Cluster II) de genes activados exclusivamente en *atbmi1a/b/c* (1/3 de los genes activados en este mutante), y un tercer grupo (Cluster III) de genes únicamente activados en *atbmi1a/b* (1/3 de los genes activados en este mutante) (Fig. 17a). Seguidamente, se analizó el nivel de expresión de los genes de los distintos grupos en los dos mutantes y en el WT representándose en un mapa de calor (Fig. 17b).

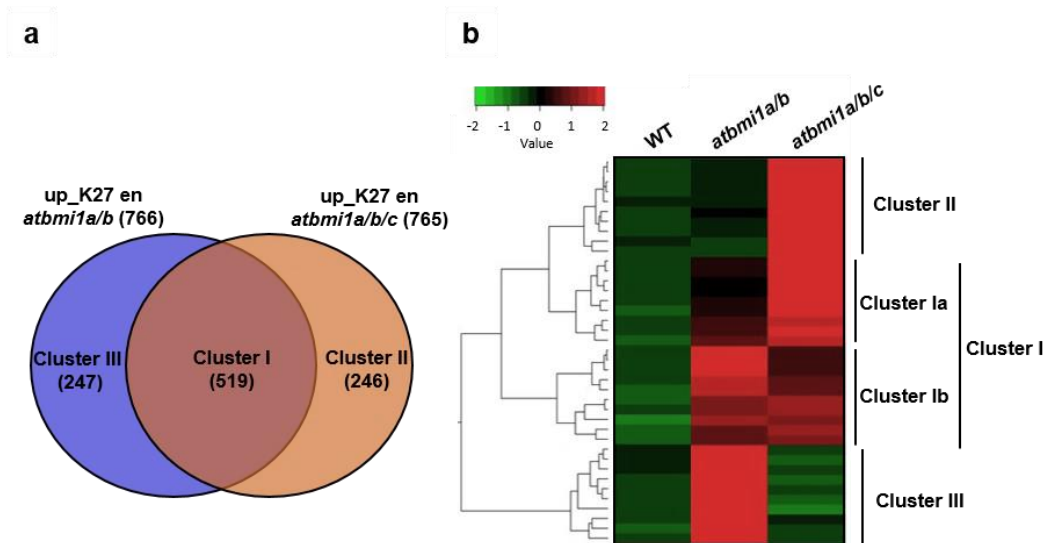


Figura 17: Redundancia funcional entre AtBMI1A/B y AtBMI1C. (a) Análisis de *Clustering* de los genes up_K27 en los mutantes *atbmi1a/b* y *atbmi1a/b/c*. El solapamiento es significativo con un p-valor inferior a $2,2 \times 10^{-16}$ y un odds ratio mayor que 21 según el test exacto de Fisher. **(b)** Niveles de expresión en WT, *atbmi1a/b* y *atbmi1a/b/c* de los genes de los diferentes *clusters*. El código de color representa los valores de expresión normalizados medidos en FPKM.

Los genes en el cluster I se clasificaron en dos sub-grupos de acuerdo a su expresión. El sub-grupo Ia (Cluster Ia) incluye los genes que presentan un aumento gradual de la expresión en los mutantes dobles y triples, sugiriendo una regulación redundante por AtBMI1A/B y por AtBMI1C (Fig. 17b y Fig. 18a). El sub-grupo Ib (Cluster Ib) contiene genes cuya regulación parece depender exclusivamente de AtBMI1A/B, puesto que la pérdida de función de AtBMI1C no afecta sustancialmente a sus niveles de expresión, en general (Fig. 17b y Fig. 18a). El cluster II incluye genes activados exclusivamente en *atbmi1a/b/c*, lo que podría indicar que estos genes son dianas específicas de AtBMI1C, o alternativamente, que AtBMI1C compensa completamente la pérdida de función de AtBMI1A/B en la regulación de estos genes (Fig. 17b y Fig. 18a). Para discernir entre estas dos posibilidades, se midieron los niveles de expresión de los genes del cluster II en WT, el mutante simple *atbmi1c* y el mutante triple *atbmi1a/b/c*, en plantas completas y en raíces a los 10 días después de la germinación (10 DAG) mediante RT-PCR cuantitativa (qRT-PCR). Puesto que estos genes no se activan en el mutante simple *atbmi1c* (Fig. 18b), se concluyó que AtBMI1C compensa la pérdida de la función de AtBMI1A/B en la regulación de estos genes. Finalmente, los genes del cluster III estaban exclusivamente activados en los mutantes *atbmi1a/b*, (Fig. 17b y Fig. 18a). Aunque a priori este resultado era inesperado, podría ser explicado si la activación de estos genes requiriese un estado de desarrollo que no se alcanza en *atbmi1a/b/c*.

En resumen, todos estos datos indican que AtBMI1A y B regulan genes predominantemente de forma redundante, mientras que AtBMI1C afecta solamente a un sub-grupo de posibles dianas de AtBMI1A/B.

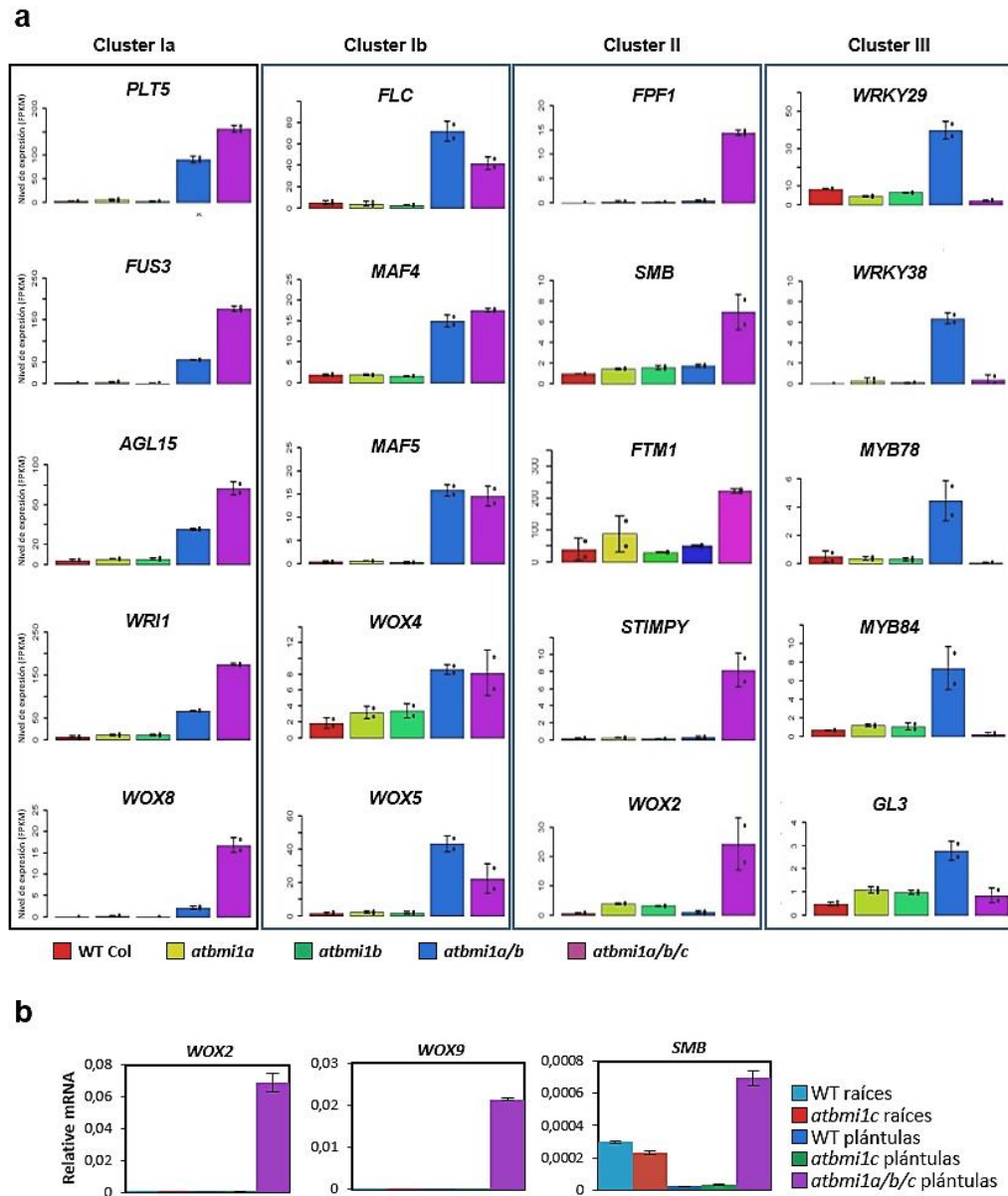


Figura 18: Genes expresados diferencialmente en *atbmi1a/b* y *atbmi1a/b/c*. (a) Niveles de expresión de varios genes de los diferentes clusters en WT y los mutantes. (b) Análisis mediante qRT-PCR de los niveles de expresión de *WOX2*, *WOX9* y *SMB* en toda la plántula y en las raíces en WT y en los mutantes *atbmi1c* y *atbmi1a/b/c*. Las cuantificaciones son relativas a los niveles de *ACT2*. Se indican las barras de error de tres réplicas independientes.

3.2.1. Programas de desarrollo desregulados en los mutantes *atbmi1*

En estudios anteriores se describió el papel de las proteínas AtBMI1 en la regulación de varios genes de maduración de semilla (Bratzel *et al.*, 2010; Chen *et al.*, 2010; Yang *et al.*, 2013) y los genes relacionados con la germinación (Molitor *et al.*, 2014). Además, un análisis transcriptómico previo del mutante *atbmi1a/b* de fenotipo débil (*atbmi1a/b* débil) confirmó la implicación de AtBMI1 en la regulación de los genes de desarrollo de semilla (H. Wang *et al.*, 2016). La comparación de los genes marcados con H3K27me3 y activados en *atbmi1a/b* débil (con FC ≥ 2 , de acuerdo con Wang *et al.* 2016) con aquellos genes en el mutante *atbmi1a/b* con fenotipo severo, mostró que hay un número significativo de genes que se activan solo en el mutante severo (Fig. 19a).

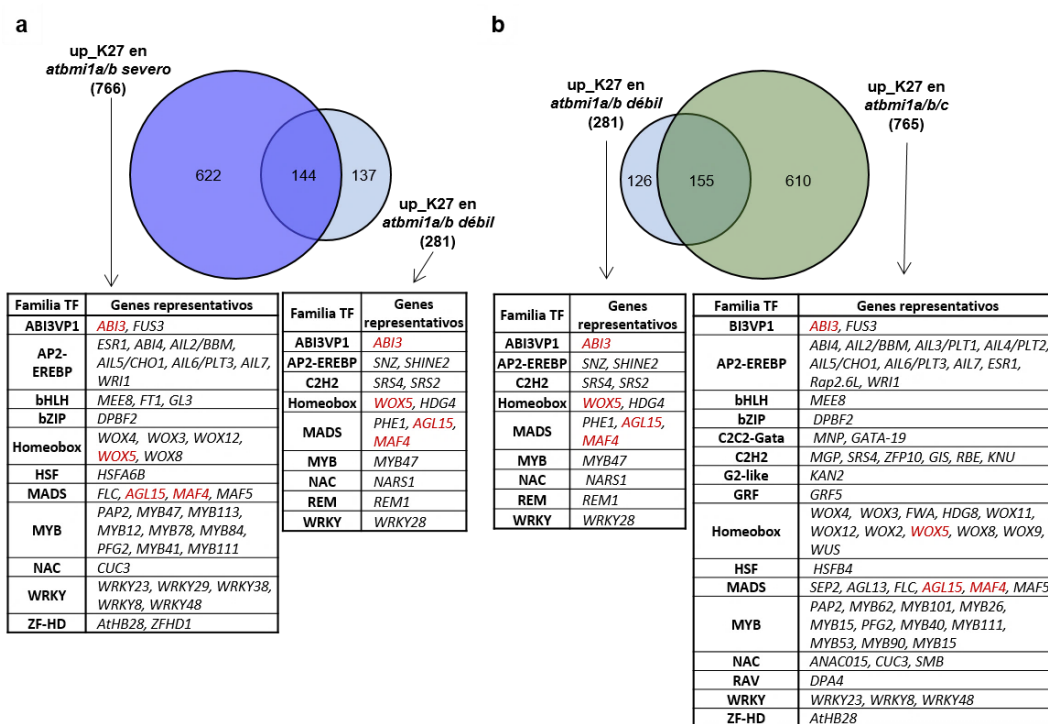


Figura 19: Diferentes patrones de expresión génica en los mutantes *atbmi1a/b* débiles y severos. (a) Diagrama de Venn que muestra el solapamiento entre los genes up_K27 en los mutantes *atbmi1a/b* débiles y severos. El solapamiento es significativo con un p-valor inferior a $2,2 \times 10^{-16}$ y un odds ratio mayor que 15 de acuerdo con el test exacto de Fisher. Se indican algunos factores de transcripción representativos (TF) de cada conjunto. Los TF encontrados en los dos conjuntos están marcados en rojo. **(b)** Diagrama de Venn que muestra el solapamiento entre los genes up_K27 en los mutantes *atbmi1a/b* débil y *atbmi1a/b/c*. El solapamiento es significativo con un p-valor inferior a $2,2 \times 10^{-16}$ y un odds ratio mayor que 17 según el test exacto de Fisher. Se indican algunos factores de transcripción representativos de cada conjunto. Los TFs encontrados en los dos conjuntos están marcados en rojo.

Entre los genes activados en ambos mutantes, había genes que habían sido identificados previamente como dianas de AtBMI1, como son *ABSCISIC ACID INSENSITIVE3 (ABI3)*, y *DELAY OF GERMINATION 1 (DOG1)*; sin embargo, otras dianas bien conocidas de AtBMI1, como *FUSCA 3 (FUS3)* o *BABYBOOM (BBM)* (Yang *et al.*, 2013), estaban activadas solo en *atbmi1a/b* severo. Se obtuvieron resultados similares comparando los datos de *atbmi1a/b* débil y *atbmi1a/b/c* (Fig. 19b). Por lo tanto, para obtener una imagen más completa de los procesos de desarrollo regulados por las proteínas AtBMI1 y, puesto que *atbmi1a/b/c* presenta las alteraciones más fuertes de desarrollo, se estudiaron las funciones de desarrollo anotadas para los genes up_K27 en estos mutantes.

➤ **Maduración de semilla y dormición**

El análisis de los genes up_K27 en el triple mutante *atbmi1a/b/c* desveló nuevos genes diana implicados en la maduración de semilla y la respuesta a ácido abscísico (ABA), como *FUS3* y *ABI4*, y en la dormición de semilla, como *SOMNUS (SOM)*. También encontramos genes implicados en la regulación del metabolismo de carbohidratos y lípidos, como *WRINKLED 1 (WRI1)* (Fig. 19b). La mayoría de estos genes se inactivan después de la germinación en el WT; sin embargo, los ABI se requieren para la respuesta de la planta a varios estreses bióticos y abióticos (Cutler *et al.*, 2010), lo que sugiere la participación de las proteínas AtBMI1 en la regulación de respuestas a las condiciones ambientales.

➤ **Genes específicos del endospermo**

Además de los genes de maduración de semilla, encontramos que en los mutantes *atbmi1a/b/c* también estaban activados genes que en condiciones normales se expresan de forma predominante en el endospermo, pero no en la cubierta (testa) de la semilla o en tejidos vegetativos (Wolff *et al.*, 2011). Muchos de ellos son genes que se expresan preferentemente del alelo materno en el endospermo [*FLOWERING WAGENINGEN (FWA)*, *HOMEODOMAIN GLABROUS 8 (HDG8)*, y *AtBMI1C*] o del paterno [*PICKLE RELATED 2 (PKR2)*, *VARIANT IN METHYLATION 5 (VIM5)*, AT2G21930 y AT3G49770] (Fig. 19b).

➤ **Genes relacionados con el mantenimiento del meristemo y la proliferación celular**

atbmi1a/b/c también presentó activación de los genes implicados en el mantenimiento del meristemo y la proliferación celular a lo largo de la vida de la planta. Curiosamente, se activan dos familias de genes con un papel crucial en estos procesos. La

primera engloba los genes *PLETHORA (PLT)* o *AINTEGUMENTA-LIKE (AIL)*. Seis de los ocho miembros de esta familia estaban activados y marcados con K27 (up_K27) en los mutantes *atbmi1a/b/c* (*PLT1/2/3/5/7* y *BBM*; Fig. 19b y Fig. 20). Algunos de estos genes *PLT* tienen papeles superpuestos en la regulación de la embriogénesis, el mantenimiento del meristemo apical de la raíz y del tallo y en la iniciación del desarrollo de órganos (Horstman *et al.*, 2014). La segunda es la familia de los genes *WUSCHEL homeobox-containing (WOX)*, que contiene 14 miembros (van der Graaff, Laux y Rensing, 2009), de los cuales *WUS* y *WOX2/3/4/5/8/9/11/12* estaban activados en los mutantes *atbmi1a/b/c* (Fig. 19b y Fig. 20). Estos factores promueven la división celular y previenen la diferenciación celular prematura, procesos que son cruciales para el mantenimiento de las células madre y la formación de los órganos. Asimismo, otros genes con funciones relacionadas se encontraron activados en los mutantes, por ejemplo, *CUP SHAPED COTYLEDON 3 (CUC3)*, *ENHANCER OF SHOOT REGENERATION 1 (ESR1)* y *GROWTH REGULATING FACTOR 5 (GRF5)* (Fig. 19b).

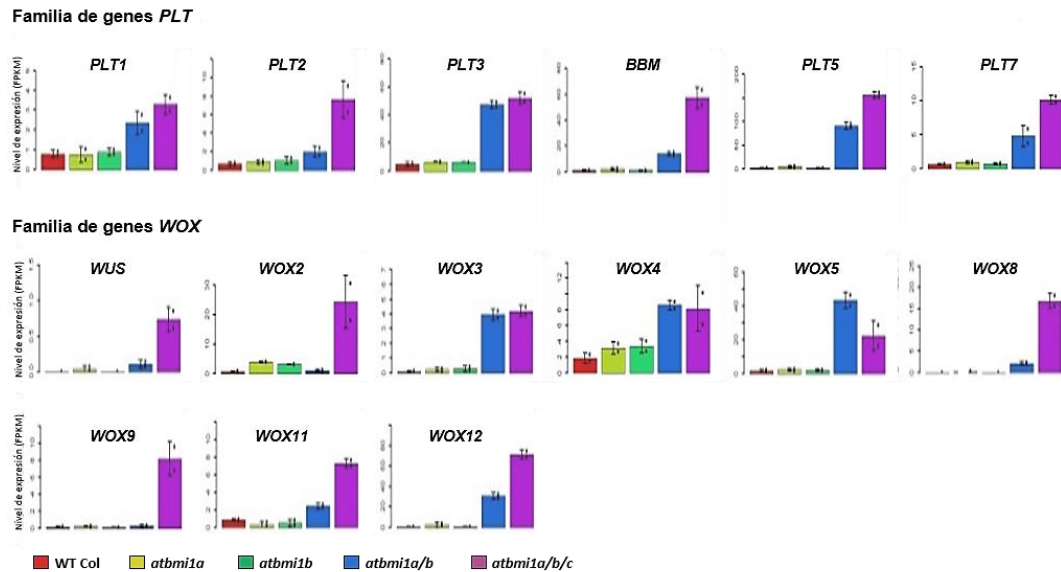


Figura 20: Niveles de expresión de los genes de las familias *PLT* y *WOX* en WT y los mutantes *atbmi1*.

➤ **Genes específicos de desarrollo de la raíz**

Además de los genes implicados en el mantenimiento del meristemo de la raíz, el mutante *atbmi1a/b/c* presentaba una activación de los genes que juegan un papel crucial en el desarrollo post-embriionario de la raíz, como *CEGENDUO (CEG)*, *MAGPIE (MGP)*, *INDOLE-3-ACETIC ACID INDUCIBLE 30 (IAA30)*, *ROOT MERISTEM GROWTH FACTOR 2 (RGF2)*, y el factor de transcripción de la clase IIB NAC *SOMBRERO (SMB)*, confirmando la importancia de la función de AtBMI1 en el desarrollo de la raíz (Fig. 19b).

➤ **Otros genes de desarrollo**

Entre los genes up_K27 en los mutantes *atbmi1a/b/c* había también genes implicados en la regulación de otros procesos de desarrollo, como el desarrollo del gametofito, el desarrollo de la hoja y la transición floral [e.j. *KANADI 2 (KAN2)*, *KNUCKLES (KNU)*, *DEVELOPMENT-RELATED PcG TARGET IN THE APEX 4 (DPA4)*, *SEPALLATA 2 (SEP2)*, *FLOWERING LOCUS C (FLC)*, *MADS AFFECTING FLOWERING 4 (MAF4)*, *MAF5* y *FACTOR PROMOTING FLOWERING 1 (FPF1)*] (Fig. 19b).

➤ **Procesos de metabolismo secundario**

En los mutantes *atbmi1a/b/c*, también se activaban genes implicados en procesos del metabolismo secundario como aquellos que participan en el metabolismo de los fenilpropanoides, como *CHALCONE SYNTHASE (CHS)*, *TRANSPARENT TESTA 4 (TT4)*, *CHALCONE ISOMERASE (CHI, TT5)*, *FLAVONOID 3'-HYDROXYLASE (F3'H, TT7)*, *DIHYDROFLAVONOL 4-REDUCTASE (DFR)*, y factores de transcripción como *AtMYB90*, *PRODUCTION OF ANTHOCYANIN PIGMENT 2 (PAP2)*, *AtMYB111* y *AtMYB101*; (Fig. 19b).

Todos estos datos indican que la función de AtBMI1 en *Arabidopsis* es necesaria para regular una diversidad de procesos de desarrollo mucho mayor de los que anteriormente habían sido caracterizados.

3.2.2. Interacciones reguladoras entre los distintos complejos de cromatina

Los mutantes *atbmi1a/b/c* presentaban activación de algunos genes PcG o relacionados con la función PcG, como *AtRING1A*, *AtRING1B*, *VAL1*, *VAL2*, y *VIN3*, lo que sugiere que la maquinaria PcG regula su propia expresión; sin embargo, la expresión de otros de estos genes no cambiaba de forma significativa, como es el caso de *CLF*, *SWN*, *MEA*, *EMF2*, *VRN2*, *FIS2*, *MSI1*, *FIE*, *EMF1* y *LHP1* (Fig. 21). Por otro lado, los genes del grupo Trithorax (TrxG) *ULTRAPETALA 1 (ULT1)*, *ULT2* y *PKR2* que antagonizan la acción de los complejos PcG estaban también activados en estos mutantes (Fig. 21). La desregulación de algunos de estos factores probablemente contribuya a la severa alteración de los patrones de expresión que muestra este mutante.

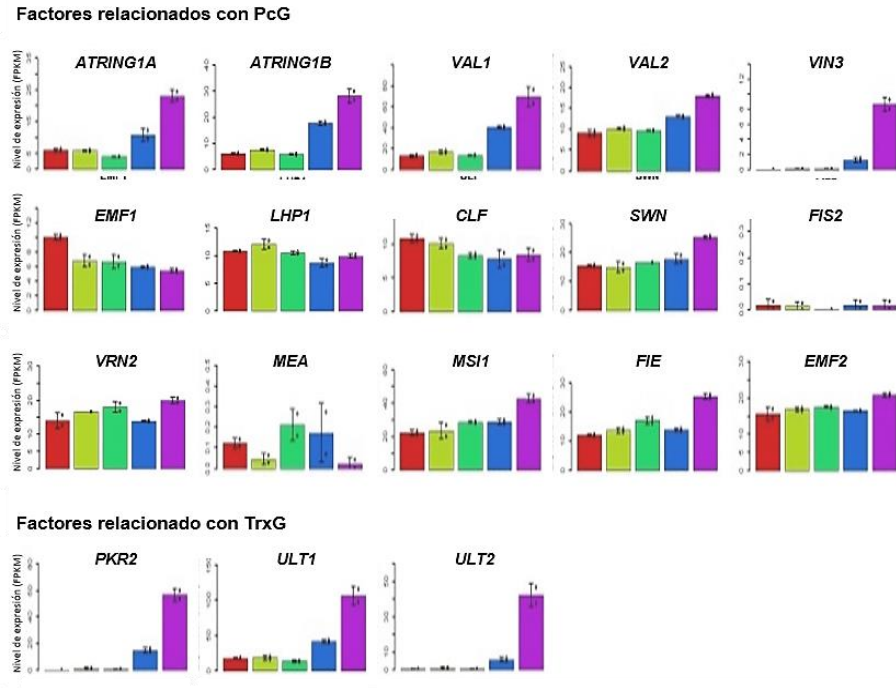


Figura 21: Niveles de expresión de los factores relacionados con cromatina que pertenecen a las familias PcG y TrxG en WT y los mutantes *atbmi1*.

3.2.3. Reguladores centrales en el programa de transición y desarrollo floral se encuentran inhibidos en el mutante *atbmi1a/b/c*

Sorprendentemente, varios factores de transcripción MADS-box necesarios para especificar la identidad del meristemo floral o involucrados en el desarrollo del órgano floral estaban reprimidos en los mutantes *atbmi1a/b/c* (Fig. 22) [e.g. *AGL42*, *SUPPRESSOR OF CONSTANS 1 (SOC1)*, *SEP3*, *SEP4*, *AGL24*, *SHORT VEGETATIVE PHASE (SVP)*]; pero también lo estaban genes claves para la regulación de la floración, como *TEMPRANILLO 1 (TEM1)* y varios *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPLs)* (e.g. *SPL2,3,4,8,12*). Más aún, encontramos que los genes de floración, los cuales muestran niveles de expresión basales en plantas WT de 10 DAG, tenían niveles de expresión aún más bajos en *atbmi1a/b/c* [e.g. *AG*, *APETALA 3 (AP3)*, *FT*; Fig. 22], lo que indica que la función de AtBMI1 es necesaria para la regulación adecuada del desarrollo floral.

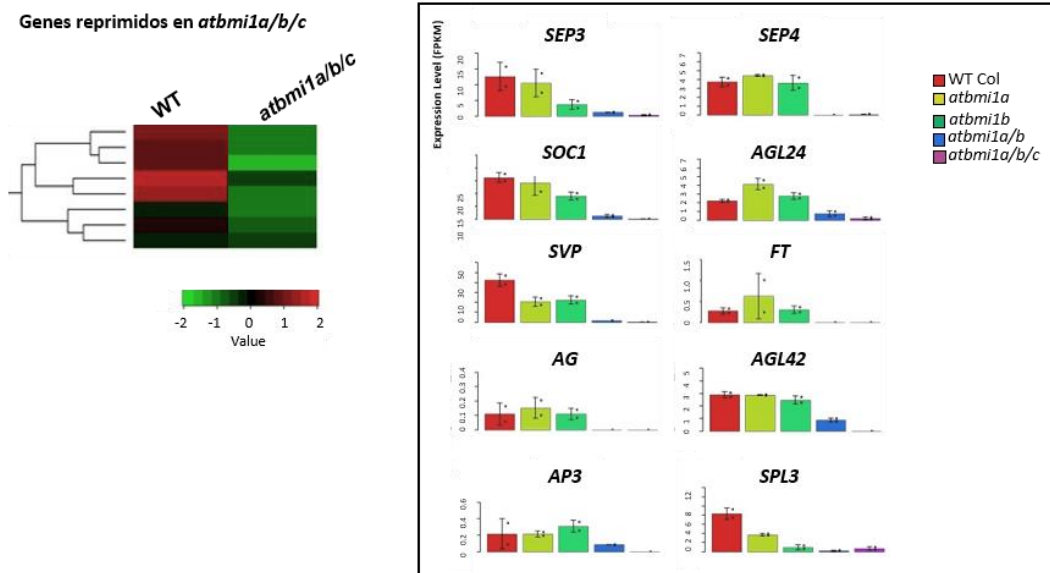


Figura 22: Expresión génica de los genes reprimidos en los mutantes *atbmi1a/b/c*. Los genes de floración están reprimidos en los mutantes *atbmi1a/b/c*.

3.2.4. VAL1/2 y AtBMI1 co-regulan un conjunto de dianas de AtBMI1

A pesar de la similitud del fenotipo de los mutantes *val1/2* y *atbmi1a/b/c* y de que ambos sobre-expresan un conjunto común de genes, parece que hay otros genes regulados por AtBMI1 independientemente de VAL. Por ejemplo, *WUS*, está regulado por AtBMI1, pero no por VAL1/2, lo que indica que existen también diferencias en las dianas de estas proteínas (Yang *et al.*, 2013). Para determinar hasta qué punto las proteínas VAL1/2 y AtBMI1 actúan conjuntamente en la regulación de la expresión génica, se compararon los genes activados en el mutante *val1/2* (Suzuki, Wang y McCarty, 2007) y marcados con H3K27me3 en WT de acuerdo con nuestros datos (up_K27 en *val1/2*) con los genes up_K27 en *atbmi1a/b/c* (Fig. 23a). El 70% de los genes up_K27 en *val1/2* están incluidos en los genes up_K27 en *atbmi1a/b/c*; estos genes representan un tercio de los genes up_K27 en *atbmi1a/b/c*, lo que indica que estas proteínas co-regulan un gran número de genes; sin embargo, AtBMI1 tiene otras funciones independientes de VAL1/2. Se identificaron los motivos de DNA reguladores en *cis* enriquecidos en los promotores de los genes co-regulados por AtBMI1 y VAL1/2, y aquellos enriquecidos en los genes únicamente regulados por AtBMI1. En los genes co-regulados por las dos proteínas, encontramos, en efecto, un enriquecimiento en los motivos *RY* y también en los elementos *ABRE* (*ABA Responsive Elements*) [ACGT o *G-box* (Choi *et al.*, 2000)] (Fig. 23a).

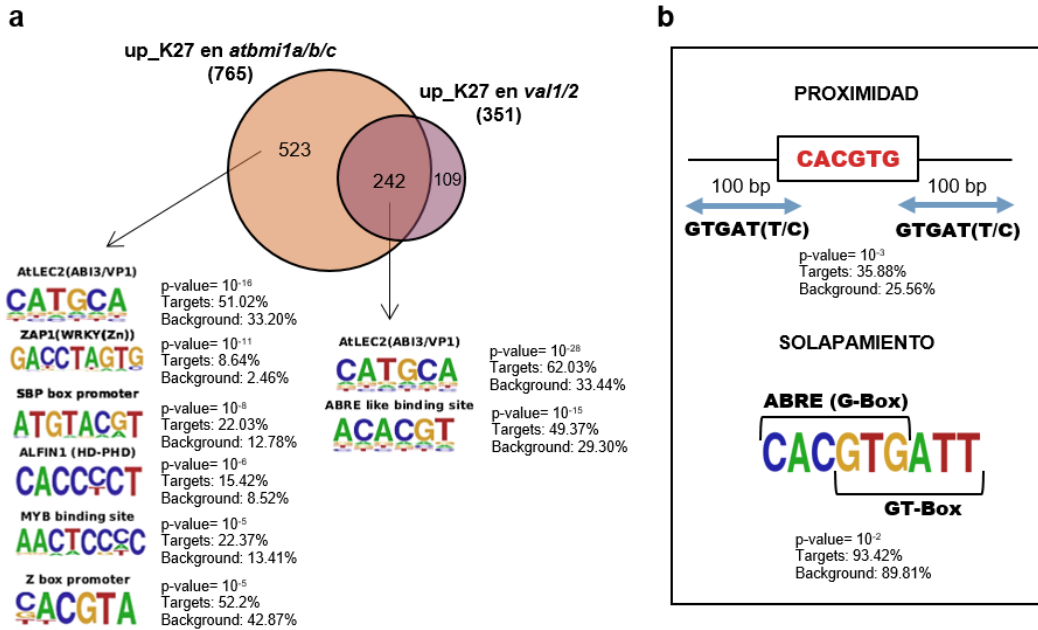


Figura 23: Interacción entre las proteínas AtBMI1 con las proteínas VAL1/2. (a) Diagrama de Venn que muestra el solapamiento entre los genes up_K27 en los mutantes *atbmi1a/b/c* y *val1/2*. Secuencias LOGO de los elementos reguladores en *cis* enriquecidos solo en los genes up_K27 en *atbmi1a/b/c* y los comunes entre *atbmi1a/b/c* y *val1/2*. (b) Proximidad y solapamiento de los motivos ABRE/G-box y GT-box en el promotor de los genes co-regulados por AtBMI1 y VAL1/2. Se indica el p-valor y el porcentaje en las dianas.

Los elementos ABRE/G-box son motivos de unión de los factores de transcripción bZIP como ABI5 (Carles *et al.*, 2002). Los elementos RY y ABRE se localizan en las regiones 5' aguas arriba de los genes regulados por los factores ABI3/VP1 y ABA (Suzuki, Ketterling y McCarty, 2005), y son necesarios para el control de la expresión de los genes de maduración de semilla (Santos-Mendoza *et al.*, 2008). Por otro lado, los elementos GT-box (GTGATT y sus variaciones) (Gao *et al.*, 2009) son reconocidos por una proteína de tipo *trihelix* específica de planta, ARABIDOPSIS 6B-INTERACTING PROTEIN 1-LIKE 1 (ASIL1). Esta proteína es necesaria para la represión de los genes de maduración de semilla después de la germinación. Estos elementos están significativamente próximos a los elementos ABRE/G-box y RY en el promotor de varios genes de maduración de semilla. Además, los elementos GT-box solapan frecuentemente con los elementos ABRE/G-box, lo que sugiere que ASIL1 podría competir con la unión de los factores de transcripción activadores para la represión de los genes embrionarios (Gao *et al.*, 2009). Así, se estudió la proximidad entre ambos elementos en el promotor de los genes co-regulados por AtBMI1 y VAL1/2 y se encontró que estos elementos estaban significativamente próximos (Fig. 23b) y, además, ambos elementos se solapan de forma significativa en el promotor de estos genes (Fig. 23b).

Por otro lado, los promotores de los genes up_K27 exclusivamente en *atbmi1a/b/c* también presentaban un alto enriquecimiento en los elementos *RY*, lo que sugiere que la represión de estos genes puede estar funcionalmente relacionada con otros factores de transcripción con dominio B3. En este grupo de genes los elementos *RY* y *ABRE/GT-box* no estaban significativamente próximos. El grupo de genes up_K27 únicamente en *atbmi1a/b/c* presentaba un enriquecimiento en otros motivos, como los sitios de unión *SQUAMOSA BINDING PROTEIN (SBP)*, *ZAP1 (WRKY)*, *ALFIN1* y *MYB* y un motivo *Z-box* que es el sitio de unión de una nueva clase de factores de transcripción, los *Z-box BINDING FACTORS (ZBF)*, cuyas funciones en la regulación del desarrollo está comenzando a ser desvelada [(Gangappa *et al.*, 2013), (Fig. 23a)]. Los elementos *ALFIN1* son sitios de unión de unas proteínas específicas de plantas, las *ALFIN1-like [AL1-7]*; (Lee *et al.*, 2009)], que actúan como represores transcripcionales (Wei *et al.*, 2015) e interactúan con *AtRING1* y *AtBMI1* (Molitor *et al.*, 2014), lo que confirma la existencia de módulos mixtos conteniendo un factor con dominio B3 y otros factores que colaboran en la represión de los genes mediada por *AtBMI1*.

3.2.5. Redes de regulación de *AtBMI1*, *EMF1* y *LHP1*

Para estudiar la relación entre las proteínas *AtBMI1* y *EMF1*, se compararon los genes diana de *EMF1*, determinados anteriormente en un análisis de inmunoprecipitación de cromatina seguido de hibridación de arrays (ChIP-chip) (Kim *et al.*, 2012), y marcados con H3K27me3 en WT (*K27_EMF1*), con los genes desregulados (activados y reprimidos) en los mutantes *atbmi1a/b/c* (Fig. 24a). El análisis de *Clustering* mostró que hay un número significativo (786) de dianas directas de *EMF1* que se desregulan en los mutantes *atbmi1a/b/c*. La mitad de los genes up_K27 en los mutantes *atbmi1a/b/c* son dianas directas de *EMF1* (Fig. 25a), lo que sugiere que las proteínas *EMF1* y *AtBMI1* co-regulan un número considerable de genes.

Sin embargo, encontramos pocos genes solapantes al comparar los genes up_K27 en *atbmi1a/b/c* y *emf1-2* [Fig. 25b; (Kim, Zhu y Sung, 2010)]. Además, la mayoría de los genes diana de *EMF1* que estaban up_K27 en *atbmi1a/b/c* no se encontraron activados en los mutantes *emf1-2*, lo que concuerda con datos anteriores que muestran que solo un pequeño porcentaje de los genes diana de *EMF1* están activados en los mutantes *emf1-2* (Kim *et al.*, 2012).

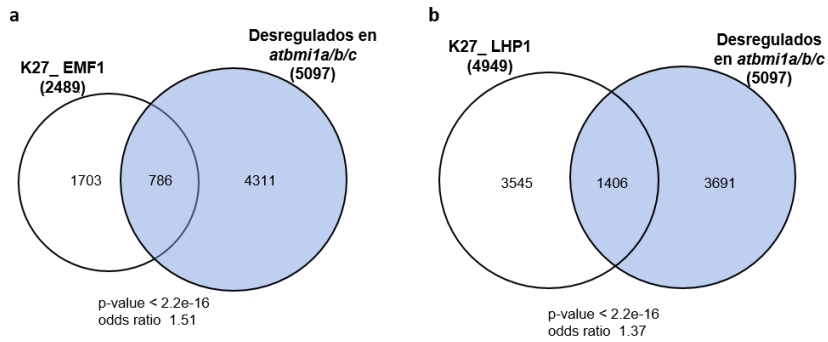


Figura 24: Relación funcional entre AtBMI1, EMF1 y LHP1. (a) Análisis de *Clustering* de los genes desregulados (activados y reprimidos) en *atbmi1a/b/c* y los genes marcados con H3K27me3 a los cuales se une EMF1. (b) Análisis de *Clustering* de los genes desregulados (activados y reprimidos) en *atbmi1a/b/c* y los genes marcados con H3K27me3 a los cuales se une LHP1. Estos solapamientos son significativos (se indican los resultados de p-valor y el test exacto de Fisher).

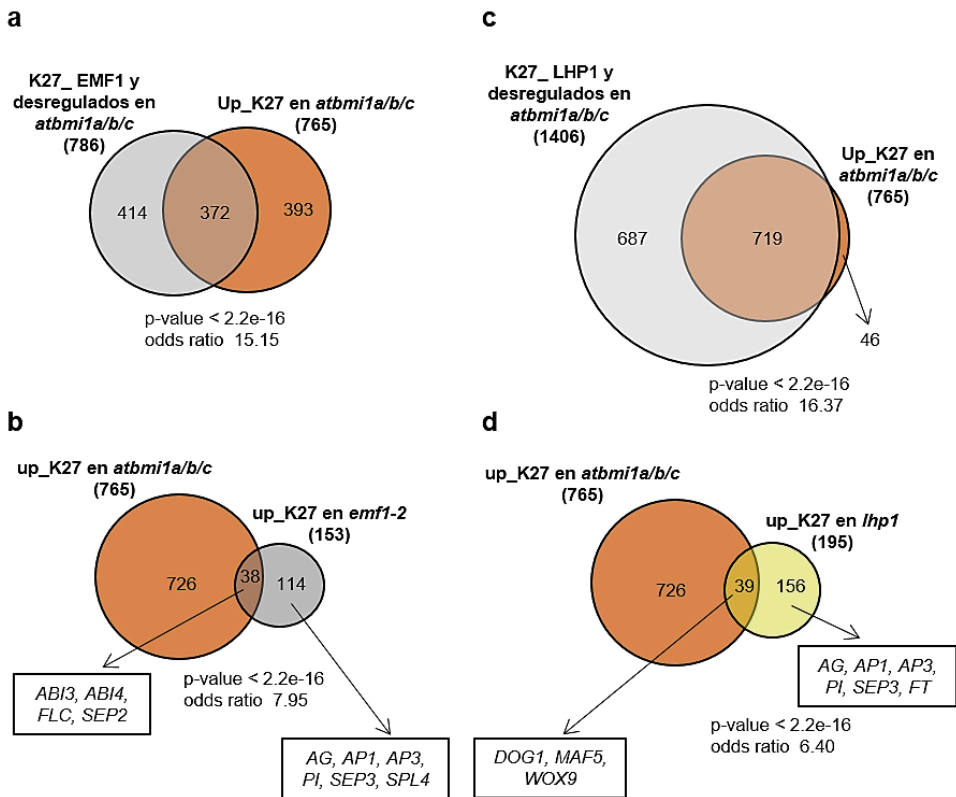


Figura 25: Redes de regulación de AtBMI1, EMF1 y LHP1. (a) Comparación de los genes marcados con H3K27me3 a los cuales se une EMF1 y se activan en *atbmi1a/b/c* con los genes up_K27 en *atbmi1a/b/c*. (b) Diagrama de Venn de los genes up_K27 en *atbmi1a/b/c* y *emf1-2*. (c) Comparación de los genes marcados con H3K27me3 a los cuales se une LHP1 y se activan en *atbmi1a/b/c* con los genes up_K27 en *atbmi1a/b/c*. (d) representación mediante diagrama de Venn de los genes up_K27 en *atbmi1a/b/c* y *lhp1*. Se indican algunos genes representativos de los genes que se solapan y los que no. Todos los solapamientos son significativos (se indican los resultados del p-valor y el test exacto de Fisher).

Por otro lado, varios resultados han mostrado que la localización de LHP1 coincide con el 85-90% de posiciones marcadas con H3K27me3 en el genoma de *Arabidopsis* (Turck *et al.*, 2007; Zhang *et al.*, 2007; Engelhorn *et al.*, 2012); de acuerdo con esto, el 92,3% de los genes marcados con H3K27me3 (4949 de 5360) están ocupados por LHP1 (Veluchamy *et al.*, 2016). Encontramos que, entre estos genes, 1406 estaban desregulados (activados y reprimidos) en los mutantes *atbmi1a/b/c* (Fig. 24b). Más aún, el 93,9% de los genes up_K27 en *atbmi1a/b/c* eran dianas de LHP1 (Fig. 25c), lo que indica que AtBMI1 y LHP1 co-regulan un gran número de genes. Sin embargo, como ocurría en el caso de EMF1, encontramos un número pequeño de genes solapantes entre los genes marcados con H3K27me3 y activados en los mutantes *lhp1* (up_K27 en *lhp1*) (FC ≥ 2 , según Wang *et al.* 2016) y los genes up_K27 en los mutantes *atbmi1a/b/c* (Fig. 25d), lo que indica que la pérdida de función de LHP1 por sí misma no tiene mucho impacto en la expresión de los genes regulados por AtBMI1. Es interesante resaltar que la pérdida de función de LHP1 y EMF1 afecta significativamente a la expresión de los genes del desarrollo reproductivo. Estos genes no se activan en los mutantes *atbmi1a/b/c* y algunos de ellos están incluso reprimidos, lo que podría indicar que LHP1 y EMF1 juegan papeles diferentes en distintos grupos de genes. Por tanto, a pesar de que EMF1, LHP1 y AtBMI1 co-localizan en un gran número de genes, la pérdida de función de estas tres actividades PcG no conlleva al mismo resultado en términos de desregulación de los genes diana.

4. Discusión

En Arabidopsis, la regulación PcG se lleva a cabo mediante la actividad de tres diferentes complejos PRC2, que regulan diferentes fases del desarrollo y presentan una especificidad parcial por los genes diana específicos, y el complejo PRC1. Estudios sobre el mecanismo PcG han demostrado que el PRC1 se requiere para establecer las marcas H3K27me3 en algunos genes diana tanto en Arabidopsis (Yang *et al.*, 2013; Calonje, 2014) como en animales (Schwartz y Pirrotta, 2013; Comet y Helin, 2014), situando al PRC1 en una posición decisiva para la represión de algunos genes. A pesar de la identificación de varias posibles subunidades de PRC1 (Merini y Calonje, 2015) y de algunas evidencias que sugieren la existencia de diferentes variantes funcionales de este complejo (Yang *et al.*, 2013; Calonje, 2014; Wang *et al.*, 2014; Merini y Calonje, 2015), no está claro si se requiere una combinación de diferentes componentes de PRC1 para ejercer diferentes mecanismos.

Resultados previos demostraron que las proteínas RING finger AtBMI1 y AtRING1 son necesarias para la represión de los genes de maduración de semilla después de la germinación, mientras que EMF1 se requiere para la represión del programa floral durante la fase vegetativa del desarrollo (Moon *et al.*, 2003; Calonje *et al.*, 2008; Bratzel *et al.*, 2010; Chen *et al.*, 2010), indicando que diferentes componentes del PRC1 están implicados en la regulación de diferentes grupos de genes diana. Por otro lado, otros estudios sugieren que todos estos componentes se requieren para la regulación de otro grupo de genes diferente. Por ejemplo, se ha demostrado que AtRING1A participa en la represión de *FLC*, *MAF4* y *MAF5* (Shen *et al.*, 2014) y EMF1 en la represión de *FLC* (Kim, Zhu y Sung, 2010).

En este trabajo, hemos explorado otras posibles funciones de las proteínas AtBMI1 durante el desarrollo de plantas aparte de su función en la represión de los genes de maduración de semilla después de la germinación.

4.1. Los componentes PRC1 AtBMI1 y EMF1 juegan papeles coordinados en la regulación de la transición de fase juvenil a reproductiva

Nuestros resultados han revelado que las proteínas AtBMI1 y EMF1 juegan un papel crucial en la regulación de la transición de la fase juvenil a la adulta. Más importante aún, nuestros datos apuntan a un modelo en el que dos variantes funcionales de PRC1, AtBMI1-PRC1 y EMF1-PRC1, coordinan la adquisición de la competencia para florecer y contribuyen a alcanzar el umbral de FT necesario para la floración, a través de la

regulación de los niveles de dos componentes claves en la transición de la fase juvenil a la adulta (Huijser y Schmid, 2011), *miR156* y *miR172*, respectivamente (Fig. 26).

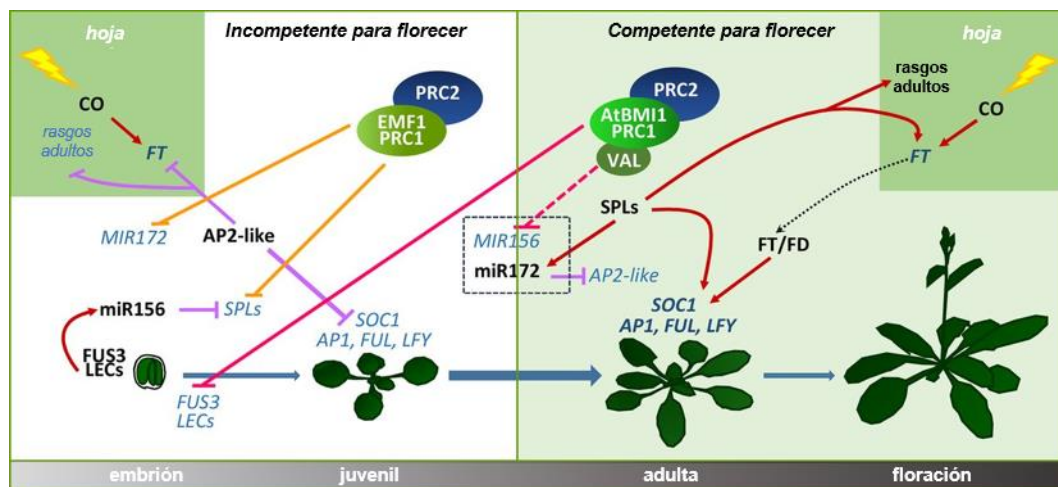


Figura 26: Modelo de las funciones de las variantes AtBMI1-PRC1 y EMF1-PRC1 en la regulación de la transición de la fase juvenil a la adulta mediante la represión de *miR156* y *miR172*. EMF1-PRC1 reprime *MIR172* y los SPL para mantener la fase juvenil. A medida que la planta crece, los niveles de *miR156* bajan mediante la represión mediada por AtBMI1-PRC1, lo que permite el desarrollo de los rasgos adultos y la adquisición de la competencia para florecer. Las líneas rectas moradas con barras indican regulación negativa; las líneas rectas rojas con flechas indican regulación positiva; las líneas amarillas con barras indican represión mediada por EMF1-PRC1/PRC2; las líneas rosas con barras indican represión mediada por AtBMI1-PRC1/PRC2 (la línea rosa discontinua indica una posible regulación negativa y la línea discontinua negra con flecha indica el movimiento de FT de las hojas al meristemo apical (SAM). Los genes reprimidos se indican en color azul claro y los genes activados con azul oscuro. Las proteínas y miRNAs están indicados en negro. *FUL* (*FRUITFULL*).

Aunque los papeles de estos miRNA han sido estudiados extensivamente, los mecanismos de su regulación todavía no se conocen, en particular aquellos relacionados con la disminución de *miR156* en función de la edad. Nuestros datos muestran que las plantas que carecen de la función de AtBMI1 presentan elevados niveles de *MIR156A/C* cuando los niveles de *miR156* tendrían que estar bajos, lo que indica que las proteínas AtBMI1 se requieren para la represión de *miR156*. Proponemos que los altos niveles de *miR156* en los mutantes *atbmi1a/b* contribuyen a reducir los niveles de *FT* en hojas y a reducir la habilidad de *FT* para inducir la floración en el SAM mediante la represión de la actividad de los SPL, dando lugar a una fase juvenil más extendida. Por el contrario, hemos visto que EMF1-PRC1 se requiere para mantener la represión de varios genes *SPL* y *MIR172* durante la fase juvenil, retrasando así la adquisición de la competencia para florecer (Fig. 26). En consecuencia, las plantas que carecen de la función de EMF1 presentan una sobre-expresión de *SPL3*, *SPL9* y *pri-MIR172* en una fase temprana del desarrollo, lo que puede desencadenar una sobre-expresión de *FT* independiente de *CO* y una adquisición precoz de la competencia para florecer.

4.2. El análisis de las redes de regulación de los distintos componentes PRC1 apunta tanto a funciones compartidas entre ellos como a específicas

Para investigar más a fondo el papel de los diferentes componentes de PRC1 y la posible interacción entre ellos a lo largo del desarrollo, integramos datos transcriptómicos con datos de localización genómica de las marcas H3K27me3 y de las diferentes proteínas relacionadas con PcG.

➤ Redundancia funcional entre las proteínas AtBMI1

La identificación de los tres parálogos de AtBMI1 en *Arabidopsis* planteó la pregunta de si presentaban una divergencia funcional (Sanchez-Pulido *et al.*, 2008). Hemos encontrado que AtBMI1A y B tienen principalmente funciones redundantes a lo largo del desarrollo, aunque un pequeño número de genes están específicamente regulados por AtBMI1B. Las diferencias de expresión génica entre los mutantes *atbmi1a* y *atbmi1b* se podría explicar por las funciones alternativas de la isoforma variante de AtBMI1B que no contiene el dominio RING finger en el amino terminal (Fig. 27). Por el contrario, AtBMI1C regula un subconjunto de los genes diana de AtBMI1A/B. El hecho de que la expresión ectópica de AtBMI1C en los mutantes dobles (Yang *et al.*, 2013) no pueda rescatar los defectos de *atbmi1a/b* en las partes aéreas de las plántulas, indica que se requieren factores específicos de tejido para la represión mediada por AtBMI1C. Por consiguiente, AtBMI1C actúa de forma redundante con AtBMI1A/B en la regulación de un número considerable de genes implicados en el desarrollo de la raíz. Las diferencias en la secuencia de proteína entre AtBMI1C y AtBMI1A/B (Bratzel *et al.*, 2010, 2012; Chen *et al.*, 2010) pueden haber restringido las posibilidades de AtBMI1C para interactuar con algunos factores y/o favorecer la interacción con otros. De forma similar, el componente de PRC2 MEA no puede compensar la pérdida de función de CLF y SWN a pesar de su expresión ectópica en los mutantes dobles *clf/swn* (Farrona *et al.*, 2011).

En cualquier caso, AtBMI1A, AtBMI1B y en parte AtBMI1C presentan una redundancia funcional, indicando la importancia de garantizar la función de AtBMI1 a lo largo del desarrollo.



Figura 27: Representación esquemática de las variantes de *splicing* de *AtBMI1B* (*At1g06770*) (izquierda) y la comparación de la secuencia de proteína predicha (derecha). Las cajas claras representan las regiones no traducidas, las cajas azules los exones y las líneas negras los intrones.

➤ Papel de la función AtBMI1 en el desarrollo de plantas

El análisis transcriptómico desveló que el 20% del transcriptoma estudiado fue desregulado en los mutantes *atbmi1a/b/c*, un porcentaje mucho más alto que en otros mutantes PcG, incluido *clf/swn* (Bouyer *et al.*, 2011; Kim *et al.*, 2012; H. Wang *et al.*, 2016), por lo tanto, destacando el papel central de AtBMI1 en la regulación génica. Para determinar la red de genes regulados por AtBMI1, nos centramos en los genes sobre-expresados en los mutantes *atbmi1* y marcados con H3K27me3 en plántulas WT de la misma edad, aunque estos genes pueden representar solo un subgrupo de los genes diana de AtBMI1 candidatos. Nuestro análisis apoya el requerimiento de la función de AtBMI1 para la represión del programa de maduración de semilla/dormición después de la germinación (Bratzel *et al.*, 2010; Chen *et al.*, 2010; Molitor *et al.*, 2014; H. Wang *et al.*, 2016); sin embargo, también desveló el papel crucial de estas proteínas en la promoción de la transición de una fase de desarrollo a la siguiente a lo largo del desarrollo (Fig. 28a). Después de la embriogénesis, las plantas experimentan la transición de la dormición de semilla a la germinación. Esta transición está regulada de forma antagonista por dos hormonas, ácido abscísico (ABA) y gibberelinas (GA); (Shu *et al.*, 2016). Durante la maduración de semilla, el ABA endógeno se acumula en la semilla, induciendo y manteniendo la dormición de la semilla. Por el contrario, antes del inicio de la germinación, los niveles de ABA endógeno en la semilla están reprimidos, mientras que el contenido de GA está aumentado. Entre los genes sobre-expresados en los mutantes *atbmi1a/b/c*, se encontraban los genes implicados en la inducción de ABA y/o en la reducción de la señalización por GA (e.g. *ABI3*, *ABI4*, *DOG1*, *PLT5*, *SOM*; Fig. 28a), indicando que la represión mediada por AtBMI1 de estos genes promueve esta transición de fase de desarrollo.

Después de la germinación, las plantas pasan por una fase de crecimiento vegetativo que se puede dividir en una fase juvenil y una fase vegetativa adulta. El *miR156* regula un conjunto de factores de transcripción SPL que ha sido demostrado que promueven la transición de la fase juvenil a la fase adulta (Wu y Poethig, 2006); por lo tanto, para permitir la transición de fase, los niveles de *miR156* tienen que bajar. Aunque nuestro análisis transcriptómico no pudo detectar los miRNA maduros, hemos demostrado mediante qRT-PCR que *pri-miR156* está sobre-expresado en los mutantes *atbmi1a/b* (Fig. 12); por consiguiente, varios SPL están reprimidos (e.g. *SPL2/3/4/8/12*; Fig. 28a), confirmando el papel crucial que juega AtBMI1 para que esta transición tenga lugar.

Finalmente, las plantas experimentan la transición del desarrollo vegetativo al reproductivo. Esta transición requiere la represión de varios represores de la floración como *FLC*, *MAF4/5* (Gu *et al.*, 2013) y *AGL15* (Fernandez *et al.*, 2014), que están sobre-expresados en los mutantes *atbmi1* dobles y triples (Fig. 28a). Por consiguiente, los genes de floración como *FT*, *SOC1* y *AGL24* estaban reprimidos en los mutantes *atbmi1*; por lo tanto, la actividad de AtBMI1 es necesaria también para pasar del desarrollo vegetativo al reproductivo.

Además, nuestros datos desvelan el papel clave de la actividad de AtBMI1 en controlar la especificación del nicho de células madre y la proliferación celular para un crecimiento y desarrollo adecuados del órgano, a través de la represión de varios reguladores centrales (e.g. los genes de las familias *PLT* y *WOX*; Fig. 28b), lo que está de acuerdo con la adquisición generalizada de la capacidad de proliferación de los mutantes *atbmi1a/b* severos y la alteración de la raíz, hoja y desarrollo de la flor observados en los diferentes mutantes *atbmi1* (Bratzel *et al.*, 2010; Yang *et al.*, 2013).

➤ Interacción de AtBMI1 con otros factores relacionados con PcG

La función de AtBMI1 ha sido relacionada con la función de las proteínas VAL1/2 en la regulación de varios genes de maduración de semilla (Yang *et al.*, 2013). Nuestros datos muestran que VAL1/2 y AtBMI1 actúan juntos en la regulación del programa de maduración/dormición de semilla; sin embargo, no parece que colaboran en la regulación de otros procesos de desarrollo. Hemos encontrado un enriquecimiento específico de los elementos de regulación en *cis*; *LEC2/ABI3/VP1* y *ABRE/G-box* solapando con *GT-box* en los promotores de los genes co-regulados por

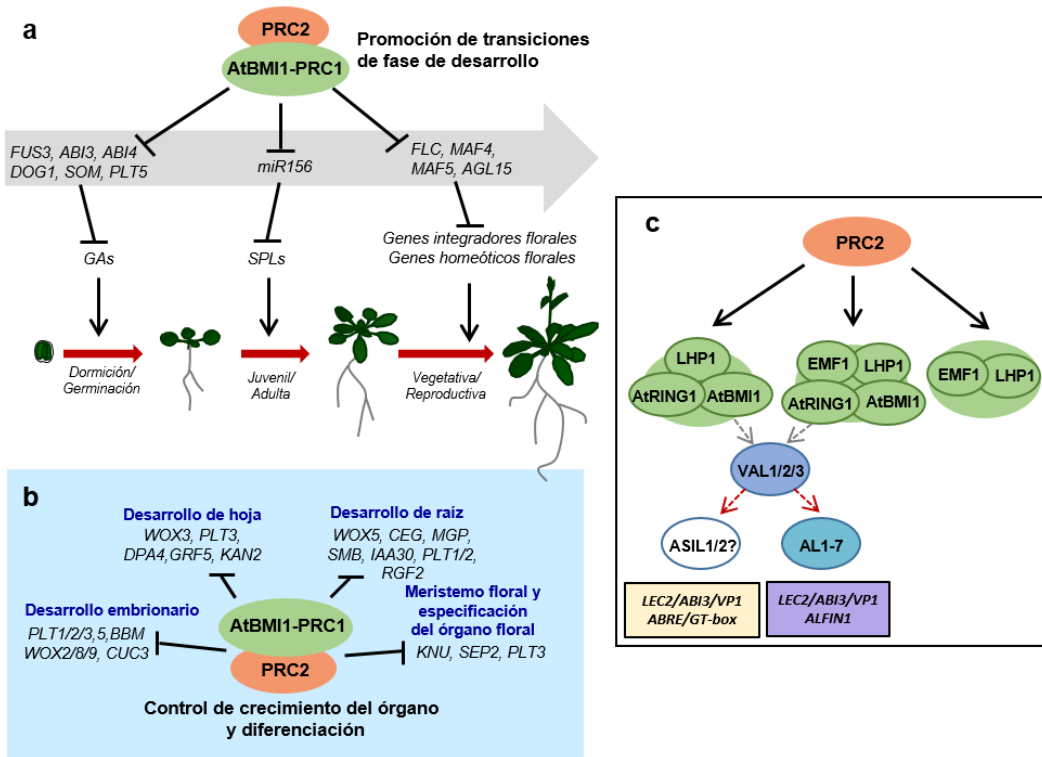


Figura 28: Papel de las proteínas AtBMI1 en la regulación del desarrollo en plantas. (a) Las proteínas AtBMI1 y PRC2 promueven las transiciones de fase de desarrollo mediante la represión de genes claves de regulación. (b) AtBMI1 y PRC2 se requieren para el control de la proliferación y diferenciación celular durante el crecimiento y el desarrollo del órgano mediante la represión de los genes reguladores centrales. (c) Las variantes de PRC1 con diferente composición y propiedades bioquímicas pueden colaborar con el PRC2 en la regulación de las transiciones de fase y en la regulación de diferentes procesos a lo largo del desarrollo de la planta. Las proteínas VAL y ASIL1/2 o AL1-7 pueden reclutar los complejos que contienen AtBMI1 a los promotores de los genes diana mediante la unión a una combinación de los elementos de regulación *in cis*.

las proteínas AtBMI1 y VAL1/2. Datos previos muestran que los elementos *LEC2/ABI3/VP1* y *ABRE BINDING FACTOR1* están enriquecidos en los promotores de los genes sobre-expresados en los mutantes *atbmi1a/b* débiles (H. Wang *et al.*, 2016). Los genes co-regulados por las proteínas ABI3/VP1-like y ABA contienen estos motivos en sus promotores (Suzuki, Ketterling y McCarty, 2005). Por lo que, ABI3 y ABI5 regulan la expresión génica sinérgicamente. Además, ABI3 interacciona físicamente con ABI5, por lo tanto, ABI3 también se recluta a los promotores de los genes diana mediante interacción proteína-proteína (Nakamura, Lynch y Finkelstein, 2001). Se podría proponer un mecanismo similar para la represión, en el que las proteínas VAL1/2 se unen a los elementos *LEC2/ABI3/VP1* y las proteínas ASIL1 a *GT-box*, compitiendo directamente con los activadores transcripcionales. La unión de VAL1/2 y probablemente las proteínas ASIL1 puede reclutar AtBMI1 y otras proteínas PcG para establecer las modificaciones de la cromatina que mantienen la represión génica. Aún queda por investigar si la represión

mediada por ASIL1 implica una interacción *in vivo* con VAL y/o proteínas PcG; sin embargo, de acuerdo con esto, se ha demostrado previamente que EMF1 interacciona con ASIL1 (llamada EIP7) en experimentos de doble híbrido en levaduras (Park *et al.*, 2011).

También hemos encontrado un enriquecimiento de los elementos *LEC2/ABI3/VP1*, pero no de *ABRE* o *GT-box* en los promotores de los genes exclusivamente *up_K27* en los mutantes *atbmi1a/b/c*, sugiriendo una implicación de los factores B3 en la regulación de estos genes también. Curiosamente, dos variantes de *splicing* alternativo de *VAL1* han sido identificadas mediante el análisis de secuenciación de RNA: una forma completa y una forma truncada que carece del dominio *homeodomain-like* de plantas similar a *VAL3*, que también carece del dominio *homeodomain-like* (Schneider *et al.*, 2016). Es posible que este grupo de genes sean diana de *VAL1* truncada y *VAL3*, explicando su no sobre-expresión en los mutantes *val1/2*.

Alternativamente, puesto que la superfamilia B3 engloba otras subfamilias, como las subfamilias de *AUXIN RESPONSE FACTORS*, *RELATED ABI3/VP1* y *REPRODUCTIVE MERISTEM* (Swaminathan, Peterson y Jack, 2008), algunos miembros no-caracterizados de estas podrían unirse al elemento *LEC2/ABI3/VP1* o a una variación de ese. En cualquier caso, los promotores de los genes independientes de *VAL1/2* también están enriquecidos en otros elementos de regulación en *cis*, como los motivos *ALFIN1* que son reconocidos en Arabidopsis por las proteínas AL. Puesto que las proteínas AL interaccionan con AtBMI1 (Molitor *et al.*, 2014), es posible que la combinación de los factores B3 y AL participe en la regulación de un subgrupo de estos genes.

La relación entre AtBMI1 y EMF1 es controversa. Por un lado, mutantes en ambos presentan fenotipos muy diferentes y diferentes grupos de genes diana de PRC2 desregulados (Kim, Zhu y Sung, 2010; Pu *et al.*, 2013; Yang *et al.*, 2013), lo que ha permitido proponer la existencia de variantes de PRC1 (Calonje *et al.*, 2008; Merini y Calonje, 2015); y por otro lado, también co-regulan un subgrupo de dianas (e.g. *ABI3*, *ABI4*, *FLC*) e interaccionan *in vitro*. Se ha demostrado que EMF1 co-purifica con los componentes de PRC2 (Liang *et al.*, 2015), cuestionando su exclusiva asociación con PRC1. Sin embargo, EMF1 co-localiza con solo el 45% de los genes marcados con H3K27me3 siendo su distribución en los genes diana más pequeña que la de las marcas H3K27me3 (Kim *et al.*, 2012). Otro componente de PRC1 propuesto es LHP1, que colocaliza con los sitios marcados con H3K27me3 (Turck *et al.*, 2007; Zhang *et al.*, 2007; Engelhorn *et al.*, 2012), pero también co-purifica con PRC2 (Derkacheva *et al.*, 2013; Liang *et al.*, 2015) e interacciona con las proteínas AtBMI1 y AtRING1 *in vitro* (Xu y Shen, 2008; Bratzel *et al.*, 2010). Sin embargo, ni EMF1 ni LHP1 parecen ser componentes de PRC2, puesto que se

requieren para establecer las marcas H3K27me3 en solo un subgrupo de las dianas de PRC2 (Kim *et al.*, 2012; H. Wang *et al.*, 2016).

Curiosamente, cuando comparamos los genes marcados con H3K27me3 que se sobre-expresan en *atbmi1a/b/c* con las dianas directas de EMF1 marcadas con H3K27me3 (K27_EMF1), encontramos que el 50% de los genes sobre-expresados en los mutantes *atbmi1* también son dianas de EMF1, sugiriendo que AtBMI1 y EMF1 podrían formar parte de un complejo y que ambas proteínas podrían afectar potencialmente a la expresión de estos genes. Lo mismo se puede decir del componente LHP1, puesto que ocupa el 93,9% de los genes up_K27 en WT. Sin embargo, el pequeño solapamiento entre los genes sobre-expresados en *atbmi1a/b/c* y *emf1-2* o *lhp1* sugiere un papel decisivo de la función de AtBMI1 en mantener la represión de los genes co-ocupados por las tres proteínas. Un estudio reciente ha desvelado que las dos proteínas homólogas SHORT LIFE (SHL) y EARLY BOLTING IN SHORT DAYS (EBS) interaccionan directamente con EMF1 mediante su dominio BAH formando dos complejos llamados BAH-EMF1c con funciones PRC1-like (Li *et al.*, 2018). Se ha visto que los dominios BAH se unen específicamente a las marcas H3K27me3 y se ha propuesto que los complejos BAH-EMF1 juegan un papel importante en la extensión de las marcas H3K27me3 mediadas por PRC2 y en su re-establecimiento después de la replicación, participando así en el mantenimiento de la represión de los genes diana (Li *et al.*, 2018). Además, se ha visto que SHL y EBS interaccionan con EMF1 y AtBMI1 (Li *et al.*, 2018). Por lo tanto, diferentes componentes podrían participar de forma simultánea en la represión de los genes mediada por PcG lo que explica el alto porcentaje de genes diana de EMF1 que no se activan en los mutantes *emf1-2*.

También encontramos genes que estaban exclusivamente sobre-expresados en *emf1-2* o *lhp1*, la mayoría de ellos están implicados en el desarrollo floral, estos genes no estaban sobre-expresados en los mutantes *atbmi1a/b/c*. Una posibilidad interesante podría ser que el mecanismo PcG dependiente de las actividades de EMF1, LHP1 y PRC2 estuviera implicado en la regulación específica del programa de desarrollo floral, lo que estaría de acuerdo con el hecho de que estas proteínas co-purifican con el PRC2 (Liang *et al.*, 2015).

En resumen, nuestros datos apuntan a redes funcionales de PRC1 en las que los genes pueden ser regulados por AtBMI1 y/o EMF1 junto con LHP1 y PRC2 y que se requieren proteínas adicionales para regular diferentes subgrupos de genes. Este es el caso de las proteínas VAL1/2 en el programa de desarrollo de semilla, que construye una red que aparentemente también incluye los factores de unión ABRE/GT-box (Fig. 28c). Además, parece que es muy probable que otros factores de transcripción con dominio B3 y las proteínas AL formen parte del circuito represivo de AtBMI1. Por el contrario, parece

que hay poco o no solapamiento en la regulación génica mediante AtBMI1 por un lado y por EMF1 y LHP1 por el otro, aunque existen factores que pueden interactuar físicamente y estar simultáneamente presentes en los genes diana.

CAPÍTULO II

Efecto de las marcas H2AK121ub y H3K27me3 en la organización de nucleosomas en Arabidopsis.

CAPÍTULO II: Efecto de las marcas H2AK121ub y H3K27me3 en la organización de nucleosomas en Arabidopsis

1. Introducción

Los nucleosomas no se distribuyen de forma aleatoria a lo largo de la cromatina, sino que se organizan en un patrón regular y controlado (Bai y Morozov, 2010; Sadeh y Allis, 2011; Struhl y Segal, 2013; Li *et al.*, 2014). La cromatina tiene el papel de organizar y compactar el DNA genómico dentro del núcleo, pero también representa una barrera física a cualquier proceso celular que requiera el acceso al DNA, como la transcripción génica. EL DNA enrollado en torno al nucleosoma es menos accesible que el *DNA linker*, por lo tanto, el posicionamiento de los nucleosomas puede facilitar o impedir la unión de los factores de transcripción al DNA, lo que afecta a las respuestas transcripcionales (Segal y Widom, 2009; Weber, Ramachandran y Henikoff, 2014; Lai y Pugh, 2017a).

Una de las cuestiones centrales de la biología es desvelar los factores y los mecanismos que determinan el posicionamiento de los nucleosomas y cómo afecta esto a la actividad génica. En los últimos años, gracias a la implementación de técnicas de secuenciación masiva, se están empezando a definir los principios generales que guían el posicionamiento de nucleosomas y las formas en que la organización nucleosomal afecta a los procesos relacionados con el DNA (Radman-Livaja y Rando, 2010).

1.1. Alineamiento de los nucleosomas en el genoma

Existen diferentes técnicas para estudiar la accesibilidad de la cromatina, aquellas que permiten identificar directamente las zonas accesibles de la cromatina y otras que evalúan la accesibilidad de forma indirecta mediante la localización de los nucleosomas.

La técnica de DNasa-seq se basa en la digestión de la cromatina mediante la endonucleasa DNasa I (Weintraub y Groudine, 1976). Las regiones abiertas del genoma son hipersensibles a la DNasa (DHS); sin embargo, dentro de las DHS, las regiones que están ocupadas por algún elemento de regulación en *trans*, se encuentran protegidas y por lo tanto no se digieren, lo que permite identificar los elementos que interaccionan con el DNA (Keene *et al.*, 1981). Con esta técnica se digiere toda la cromatina obteniéndose fragmentos de diferentes tamaños, de media de 300 pb, que se separan por electroforesis en gel. Se purifican los fragmentos pequeños entre 50 y 100 pb (< 147pb) para descartar los fragmentos ocupados por nucleosomas. Finalmente, estos fragmentos se secuencian obteniéndose las localizaciones a nivel genómico de las DHS (Song y Crawford, 2010; John *et al.*, 2013).

FAIRE-seq (Formaldehyde-Assisted Isolation of Regulatory Elements-seq) es uno de los métodos más sencillos para identificar las regiones del genoma con baja ocupación. Esta técnica se basa en la fragmentación de la cromatina fijada con formaldehído mediante sonicación y la posterior extracción del DNA con fenol-cloroformo. La fase acuosa contiene la fracción de la cromatina con mayor accesibilidad (con baja ocupación de nucleosomas), mientras que la fase fenólica contiene el DNA unido a las histonas por la fijación con formaldehído. Los fragmentos de la fase acuosa se purifican para su posterior secuenciación (Giresi *et al.*, 2007; Giresi y Lieb, 2009; Simon *et al.*, 2012, 2013).

ATAC-seq (Transposase Accessible Chromatin-seq) es una técnica más reciente basada en la capacidad de la transposasa hiperactiva Tn5 de integrarse dentro de las regiones abiertas de la cromatina y fragmentar el DNA incorporando unas secuencias (adaptadores). Los fragmentos de DNA se purifican y se amplifican en base a los adaptadores para generar la librería de secuenciación (Buenrostro *et al.*, 2013).

En esta tesis nos centramos en la técnica de MNasa-seq que ha sido la más usada durante los últimos años y se basa en un ensayo de digestión de la cromatina con la nucleasa micrococcal (MNasa) y la posterior secuenciación de los fragmentos obtenidos (Cui y Zhao, 2012; Rizzo y Sinha, 2014). El DNA ocupado por los nucleosomas es relativamente insensible a la MNasa, por lo tanto, la digestión de la cromatina con esta enzima nos permite estudiar el posicionamiento y la ocupación de los nucleosomas. La MNasa es una enzima que corta preferentemente el DNA internucleosomal, primero induce la rotura de una hebra y posteriormente la segunda, además, tiene una actividad exonucleasa que después de la rotura en una determinada posición, el DNA se digiere hasta alcanzar una posición protegida. Seguidamente, el DNA protegido se purifica y se analiza mediante secuenciación masiva para determinar la localización a nivel genómico de los nucleosomas (A. Weiner *et al.*, 2010; Chen *et al.*, 2013, 2014; Pajoro *et al.*, 2018). La MNasa tiene algunas limitaciones, por ejemplo, no corta de forma completamente aleatoria en el genoma, sino que tiene una cierta preferencia por algunas secuencias, además, se puede producir una sobre-digestión de las zonas más accesibles del genoma apareciendo a altas concentraciones de enzima como si fueran regiones libres de nucleosomas. En cualquier caso, las diferencias en el grado de digestión solo afectan a la ocupación de los nucleosomas, pero no alteran el posicionamiento y el espaciado entre los nucleosomas.

Para estudiar la distribución de los nucleosomas a nivel genómico, es importante definir los términos de posicionamiento, espaciado y ocupación. El posicionamiento de nucleosomas describe el lugar del DNA donde se localiza el nucleosoma. Cuando un nucleosoma se posiciona en la misma región genómica en toda una población de células

se dice que es un nucleosoma bien-posicionado. A diferencia del posicionamiento que describe la relación entre los nucleosomas y el DNA, el espaciado de nucleosomas se refiere a la distancia entre dos nucleosomas adyacentes y se determina por la longitud del *DNA linker*. Cuando los nucleosomas están espaciados de una forma uniforme (con la misma distancia entre nucleosomas adyacentes) se dice que forman una secuencia regular. Los conceptos de posicionamiento y espaciado son independientes, puesto que los nucleosomas pueden estar espaciados de una forma uniforme, pero mal-posicionados. Cuando los nucleosomas distribuidos en una secuencia regular también están bien-posicionados se dice que están en fase. Por último, la ocupación de nucleosomas indica para cada par de bases del genoma, la fracción de moléculas de DNA de la población que está ocupada por nucleosomas (Chereji, Ocampo y Clark, 2017).

1.2. La organización de los nucleosomas en las regiones codificantes del genoma

El alineamiento de los nucleosomas a nivel genómico nos permite explorar las propiedades de la organización de los nucleosomas alrededor de los genes. En los últimos años, gracias a aproximaciones a escala genómica de ensayos con MNasa, se han generado mapas de posicionamiento de nucleosomas en la mayoría de los organismos modelo, aunque los genomas más estudiados son los de levaduras y humanos (Hughes y Rando, 2014).

En los mapas obtenidos se pueden observar unas características generales de distribución de nucleosomas en regiones génicas, comunes en los diferentes organismos. La región del promotor inmediatamente aguas arriba del TSS presenta frecuentemente un menor grado de ocupación de nucleosomas por presentar una mayor accesibilidad, denominándose región libre de nucleosomas (del inglés *Nucleosome Depleted Region*, NDR) (Jiang y B. Franklin Pugh, 2009; Lai y Pugh, 2017b). La NDR está generalmente flanqueada por dos nucleosomas bien posicionados: los nucleosomas -1 y +1 localizados en los extremos 5' y 3' de esta región, respectivamente. Se observan también varios nucleosomas bien-posicionados aguas abajo del nucleosoma +1 aunque el nivel de posicionamiento va disminuyendo gradualmente desde el extremo 5' al 3' del gen. En general, todas estas características se aprecian bien en levaduras; sin embargo, en animales y plantas en algunos casos se complica la detección, al estar la mayoría de los nucleosomas pobremente posicionados (Valouev *et al.*, 2011; Zhang, Zhang y Jiang, 2015), probablemente por diferencias de expresión en poblaciones celulares o en distintos tejidos. En cualquier caso, estudios recientes realizados en una única célula en animales confirman estas características generales (Lai *et al.*, 2018).

A pesar de estas características generales, se han podido observar diferencias en el posicionamiento y ocupación de nucleosomas en base a distintos factores, por lo que,

en los últimos años, diferentes estudios se han centrado en investigar los determinantes del posicionamiento de los nucleosomas y cómo pueden afectar a la organización de la cromatina.

1.3. Determinantes del posicionamiento de los nucleosomas

Desde hace tiempo se sabe que la formación de los nucleosomas está termodinámicamente favorecida por algunas secuencias en comparación con otras, lo que sugiere que la conformación de la cromatina está en cierto modo programada por la secuencia del DNA (Drew y Travers, 1985). Sin embargo, también existen evidencias de que proteínas como los factores de transcripción, RNA polimerasa y remodeladores de la cromatina dependientes de ATP contribuyen de forma significativa en el posicionamiento de los nucleosomas. Por lo tanto, parece que tanto factores en *cis* (secuencia del DNA) como en *trans* (remodeladores, etc.) tienen influencia sobre la organización de los nucleosomas (Hughes y Rando, 2014).

1.3.1. La secuencia del DNA

La secuencia del DNA juega un papel crítico en el posicionamiento rotacional de la hélice de DNA, pero además es un factor determinante en la ocupación de los nucleosomas. En particular, los elementos *poly(dA:dT)* y *poly(dG:dC)* son inhibidores intrínsecos de la formación del nucleosoma; sin embargo, las regiones no homopoliméricas enriquecidas en G+C la favorecen (Struhl y Segal, 2013). Los análisis de alineamiento a nivel genómico de los nucleosomas han demostrado un mayor nivel de ocupación de nucleosomas en los exones que en los intrones, debido al alto contenido en secuencias G+C en los exones comparado con el alto porcentaje de enriquecimiento en secuencias A+T en los intrones (Nishida, 2012; Hughes y Rando, 2014). Por otro lado, los elementos en *cis* pueden afectar de forma indirecta al posicionamiento de los nucleosomas, guiando a factores en *trans*, como son los remodeladores de la cromatina capaces de desplazar o mover nucleosomas (Winger y Bowman, 2017), o los factores de transcripción que tras unirse a la secuencia de DNA pueden alterar la longitud de la NDR e influir en el posicionamiento de los nucleosomas flanqueantes. Por tanto, la secuencia del DNA por sí es importante para determinar el posicionamiento de los nucleosomas *in vivo*, pero son necesarios otros factores adicionales para regular la conformación de la cromatina (Chereji y Clark, 2018) (Partensky y Narlikar, 2009).

1.3.2. Transcripción

La organización de los nucleosomas afecta a la accesibilidad al DNA genómico, por lo que juega un papel importante en la regulación de la transcripción (Chen *et al.*, 2017).

Por otro lado, la unión y el paso de la maquinaria de transcripción a su vez afectan al posicionamiento de nucleosomas. Ensayos en distintos organismos han demostrado que la localización del complejo de pre-iniciación (PIC) de la transcripción influye en el posicionamiento del nucleosoma +1. Los componentes del PIC pueden ser necesarios para el reclutamiento de complejos remodeladores que anulan la preferencia de los nucleosomas por una secuencia determinada de DNA para posicionarlos en una localización precisa. Además, el espaciado de los nucleosomas que se encuentran aguas abajo del nucleosoma +1 depende de la maquinaria de elongación transcripcional (Struhl y Segal, 2013). Durante la transcripción, el paso de la RNA Polimerasa II (Pol II) causa un des-ensamblaje y re-ensamblaje de los nucleosomas. Esto está dirigido por las chaperonas de histonas que tienen una alta afinidad por los dímeros H2A-H2B o el tetrámero H3-H4 y que facilitan la disrupción de los nucleosomas o la deposición de las histonas. Estas chaperonas pueden asociarse a histonas específicas, modificaciones de histonas y factores de elongación. Por ejemplo, las chaperonas Nap1 (NUCLEOSOME ASSEMBLY PROTEIN 1) y FACT (FACILITATES CHROMATIN TRANSCRIPTION) facilitan el paso de la maquinaria de la transcripción a través de los nucleosomas mediante su contribución en la disrupción del dímero H2A-H2B (Petesch y Lis, 2012).

También se ha visto que el nivel de transcripción influye en la organización de los nucleosomas. Estudios a nivel genómico de los perfiles de distribución de los nucleosomas en levaduras, animales y plantas han desvelado que existe una alta correlación entre el patrón de distribución de los nucleosomas y los niveles de transcripción (Shivaswamy *et al.*, 2008; Assaf Weiner *et al.*, 2010; Yadon *et al.*, 2010; Sala *et al.*, 2011; Fu *et al.*, 2012; Gaffney *et al.*, 2012; Chen *et al.*, 2013; Li *et al.*, 2014; Zhang, Zhang y Jiang, 2015). Se ha demostrado que, en general, los genes con mayores niveles de expresión tienden a tener una menor densidad de nucleosomas en las regiones promotoras y un mejor posicionamiento de nucleosomas con un espaciado más regular a lo largo del cuerpo del gen comparados con los genes con niveles de expresión bajos (Schones *et al.*, 2008; Li *et al.*, 2014; Zhang, Zhang y Jiang, 2015).

1.3.3. Remodeladores de la cromatina dependientes de ATP

Los complejos remodeladores de la cromatina hidrolizan el ATP y convierten la energía generada en un movimiento mecánico que incluye el deslizamiento de los nucleosomas a lo largo del DNA, desensamblando los nucleosomas, intercambiando las histonas o reposicionando los nucleosomas (Bowman y McKnight, 2017; Hajheidari, Koncz y Bucher, 2019). Los remodeladores de la cromatina dependientes de ATP se agrupan en 4 subfamilias de acuerdo con sus dominios ATPasa conservados (INO80/SWR1, CHD, ISWI y SWI/SNF), aunque existen otros tipos de remodeladores como los pertenecientes a las

familias LSH, RAD54 y RAD5/16 (Knizewski, Ginalski y Jerzmanowski, 2008; Clapier y Cairns, 2009). Cada subfamilia, además del dominio catalítico ATPasa, contiene dominios únicos como el cromo- o bromodominio o el dominio PHD, lo que les permite actuar en una región específica de la cromatina (Cairns, 2009; Hargreaves y Crabtree, 2011; Narlikar, Sundaramoorthy y Owen-Hughes, 2013).

La función más importante *in vivo* de los remodeladores de la cromatina de las familias ISWI y CHD es organizar los nucleosomas en secuencias ordenadas a lo largo del gen y promover el ensamblaje de la cromatina después de la replicación (Corona y Tamkun, 2004; Petesch y Lis, 2012). En plantas, se ha visto que los remodeladores de la familia ISWI, CHROMATIN REMODELING 11 (CHR11) y CHR17 juegan un papel en el mantenimiento del espaciado de los nucleosomas, lo que influye tanto en la activación como en la represión transcripcional (Li *et al.*, 2014). Por otra parte, se ha visto que PICKLE (PKL), perteneciente a la familia CHD3, es capaz de movilizar los nucleosomas en un fragmento de DNA *in vitro* (Ho *et al.* 2013). PKL ha sido implicado tanto en represión, promoviendo la actividad de las proteínas PcG (Zhang *et al.*, 2012), como en activación de la transcripción, contrarrestando la represión mediada por las proteínas PcG (Aichinger *et al.*, 2009, 2011; Jing, Guo y Lin, 2019a). Este doble papel en represión y activación también se ha visto en factores de esta familia en animales (Murawska *et al.*, 2008; Fasulo *et al.*, 2012; Egan *et al.*, 2013).

Los remodeladores de la cromatina del tipo SWI/SNF en plantas como BRAHMA (BRM) y SPLAYED (SYD) tienen un papel activo en la regulación de la expresión génica y están implicados en numerosos procesos del desarrollo en plantas (Farrona *et al.*, 2004; Hurtado, Farrona y Reyes, 2006; Kwon *et al.*, 2006; Bezhani *et al.*, 2007; Tang *et al.*, 2008; Wu *et al.*, 2012). Estos factores han sido propuestos como activadores transcripcionales por su capacidad de desensamblar o desplazar los nucleosomas pero carecen de la función de ensamblaje de nucleosomas (Clapier y Cairns, 2009; Han *et al.*, 2015). Además, actúan en colaboración con las enzimas acetiltransferasas de histonas y de forma antagonista a las marcas de represión H3K27me3 (Ho *et al.*, 2011; Wu *et al.*, 2012; Li *et al.*, 2015). Sin embargo, se ha visto que los remodeladores del tipo SWI/SNF también pueden reprimir directamente la expresión génica (Ho *et al.*, 2009, 2011; Han *et al.*, 2015), mediante diferentes mecanismos que incluyen su participación en las interacciones entre enhancers y promotores (Heintzman *et al.*, 2007; Kim, Bresnick y Bultman, 2009), facilitando el reclutamiento de reguladores con función represora (Ho *et al.*, 2009), o estabilizando los nucleosomas en el TSS, lo que interfiere con la transcripción (Rafati *et al.*, 2011; Han *et al.*, 2012, 2015; Zhu *et al.*, 2013; Torres y Deal, 2019).

Los factores de la familia INO80 y *SWI/SNF related protein 1* (SWR1) tienen la capacidad de desensamblar los nucleosomas e intercambiar histonas (Phelan, Schnitzler y Kingston, 2000). Estos factores se agrupan en complejos multiproteicos. En levaduras, el complejo SWR1 reemplaza la histona H2A por la H2A.Z mientras que INO80 reemplaza H2A.Z por H2A (Mizuguchi *et al.*, 2004; Luk *et al.*, 2010; Papamichos-Chronakis *et al.*, 2011). En plantas, se ha demostrado que SWR1 tiene la misma actividad (Morrison y Shen, 2009; Jarillo y Piñeiro, 2015; Crevillén *et al.*, 2019). Este complejo está implicado en procesos de regulación génica, reparación del DNA y la recombinación del DNA. La presencia de H2A.Z en lugar de H2A en el octámero de histonas afecta a la estabilidad de los nucleosomas (Jin y Felsenfeld, 2007; Jin *et al.*, 2009). Todavía no se ha demostrado si el complejo INO80 de plantas intercambia la H2A.Z por H2A, lo que sí se ha visto es que está implicado en los procesos de recombinación homóloga y reparación del DNA (Fritsch *et al.*, 2004; Shaked, Avivi-Ragolsky y Levy, 2006).

1.3.4. Modificaciones post-traduccionales de histonas (PTM)

Las modificaciones post-traduccionales (PTM) de las histonas (acetilación, metilación, fosforilación, monoubiquitinación, biotinilación, glicosilación, etc) pueden regular la composición y la estructura de los nucleosomas en los genes. Estas modificaciones pueden alterar la accesibilidad a la cromatina mediante dos mecanismos: 1) alterando las interacciones histona-histona o histona-DNA, o 2) actuando como sitios de reclutamiento específicos de factores remodeladores de la cromatina, otros factores con actividades modificadoras de histonas, coactivadores, o corepresores transcripcionales (Musselman *et al.*, 2012; Allis y Jenuwein, 2016).

Existen PTM correlacionadas positiva o negativamente con la actividad transcripcional de los genes; sin embargo, se conoce poco sobre el efecto de muchas de estas modificaciones en la organización de la cromatina (Petesch y Lis, 2012). La modificación de histonas más estudiada que afecta directamente a la estructura de la cromatina es la acetilación de los residuos de lisina de las histonas. De forma general, la acetilación neutraliza la carga positiva de las histonas debilitando la interacción electrostática con el DNA cargado negativamente, lo que resulta en un aumento de la accesibilidad del DNA nucleosomal a los factores de transcripción y la RNA PolIII. Mediante ensayos *in vitro*, se ha demostrado que la acetilación del dímero H2A-H2B debilita las interacciones presentes a 40 pb alrededor del sitio de la marca y la acetilación de H3-H4 reduce la formación de estructuras de cromatina de orden superior y reduce la cantidad de DNA unido a los nucleosomas (Brower-Toland *et al.*, 2005; Petesch y Lis, 2012).

A diferencia de la acetilación o fosforilación, la metilación de histonas no produce alteraciones en la carga. Esta es una de las razones por las cuales las metilaciones pueden ser tanto marcas de activación como de represión dependiendo del sitio y el grado de la metilación. Por ejemplo, H3K4me1-2-3 se encuentra típicamente en sitios de transcripción activos, mientras que H3K9me2-3 y H3K27me2-3 son marcas de represión (Zhang *et al.*, 2009; Köhler y Hennig, 2010). Sin embargo, estas PTM pueden actuar como señales de reclutamiento de otros factores que alteran la conformación del nucleosoma o de la cromatina. Estos factores contienen dominios conservados evolutivamente que reconocen PTM de histonas específicas; tales como los bromodominios que son específicos para las acetil-lisinas y los cromodominios para las metil-lisinas (Mujtaba, Zeng y Zhou, 2007). Algunas modificaciones tienen efecto distinto dependiendo en qué histona se incorporen, como por ejemplo la monoubiquitinación. Las histonas H2B y H2A pueden ser monoubiquitinadas en su extremo C-terminal. La monoubiquitinación de la H2B favorece la creación de una conformación abierta y accesible mediante un mecanismo independiente y sinérgico con la acetilación de la histona H4 (Fierz *et al.*, 2011), mientras que la monoubiquitinación de la H2A está implicada en la represión transcripcional (Flanagan *et al.*, 2005; Scheuermann, Gutiérrez y Müller, 2012) aunque no se conoce el papel exacto que juega.

En general, se conoce mejor el efecto de las modificaciones de histonas que participan en el establecimiento de una conformación de cromatina abierta para facilitar el proceso de transcripción que el de las implicadas en el mantenimiento de una conformación de cromatina refractaria a la transcripción.

2. Objetivo

Nos propusimos determinar el papel exacto que tienen las marcas incorporadas por la maquinaria PcG en la organización de los nucleosomas en *Arabidopsis*. Como se ha discutido en el capítulo anterior de esta tesis, los complejos PRC1 y PRC2 son responsables de la incorporación de la monoubiquitinación de H2A y de la trimetilación de H3, respectivamente, que actúan como marcas represoras de la transcripción (Calonje, 2014). Sin embargo, a pesar de los avances logrados en el conocimiento de la represión PcG, aún no se conoce el mecanismo mediante el cual estas modificaciones crean una cromatina inaccesible a la maquinaria de transcripción. Ensayos *in vitro* en vertebrados, *Drosophila* y *Arabidopsis* indican que algunos de los componentes PcG son capaces de compactar la cromatina inhibiendo el efecto de factores remodeladores y la transcripción. En concreto, en *Drosophila* el componente del PRC1 Psc es capaz por sí solo de tener este efecto *in vitro* (King, Francis y Kingston, 2002; Francis, Kingston y Woodcock, 2004), en vertebrados un homólogo de la subunidad Pc de *Drosophila*, M33, también tiene esta capacidad (Grau *et al.*, 2011) y en *Arabidopsis*, EMF1 parece ser el equivalente funcional (Calonje *et al.*,

2008; Beh, Colwell y Francis, 2012). Por otro lado, se ha visto que a nivel genómico las marcas H3K27me3 están asociadas con zonas de baja accesibilidad tanto en animales (Bell *et al.*, 2010; Kelly *et al.*, 2012; Beck *et al.*, 2014; Deaton *et al.*, 2016) como en *Arabidopsis* (Shu *et al.*, 2012). Aparte de esto, excepto por un trabajo reciente en vertebrados que propone que el PRC1 conlleva a un aumento de la ocupación de nucleosomas y a una disminución del espacio entre nucleosomas en los promotores de los genes diana PcG (King *et al.*, 2018), se desconoce el papel exacto que juegan las marcas H2AK121ub y H3K27me3 en modificar el posicionamiento y/u ocupación de los nucleosomas para crear una conformación inaccesible de la cromatina, especialmente en plantas.

Resultados previos del laboratorio sobre la distribución de las marcas H2AK121ub y H3K27me3 a nivel genómico (Zhou *et al.*, 2017), mostraron que las dos marcas se localizan generalmente en regiones génicas, siendo los picos de H3K27me3 más extensos que los picos de H2AK121ub. El 80% de los picos de H2AK121ub solapan con la región inmediatamente aguas abajo del TSS, mientras que los picos H3K27me3 solapan con el promotor y el cuerpo del gen. Además, a pesar de que las dos marcas co-localizan en un gran número de genes, existen también conjuntos de genes exclusivamente marcados con H2AK121ub o H3K27me3 (Zhou *et al.*, 2017), lo que sugiere papeles diferentes en la represión transcripcional. De acuerdo con esto, la mayoría de los genes marcados con H2AK121ub/H3K27me3 y con H3K27me3 presentan bajos niveles de expresión o no se expresan, mientras que un porcentaje alto de los genes marcados solo con H2AK121ub presentan actividad transcripcional, aunque su nivel de expresión aumenta en los mutantes de pérdida de función de PRC1 (Zhou *et al.*, 2017). Para determinar el efecto de cada una de estas marcas hemos analizado la distribución de los nucleosomas a nivel genómico en los distintos conjuntos de genes diana PcG y el impacto que tiene la pérdida de función de AtBMI1 en el posicionamiento y la ocupación de los nucleosomas.

3. Resultados

3.1. Obtención de muestras para MNase-seq

Para investigar un posible efecto de las marcas establecidas por los complejos PRC1 y PRC2 de *Arabidopsis* en la organización de los nucleosomas, se llevó a cabo un ensayo de digestión de la cromatina de plantas WT-Col0 recogidas a los 7 días después de la germinación con la enzima nucleasa micrococcal (MNasa) y la posterior secuenciación de los fragmentos obtenidos (MNase-seq). La MNasa es una endo- y exonucleasa que corta preferentemente el DNA que conecta dos nucleosomas, mientras que el DNA nucleosómico está protegido frente a la digestión, al menos a concentraciones bajas de enzima. Esta técnica ha permitido mapear el posicionamiento y ocupación de nucleosomas en el genoma de múltiples especies (Mavrigh *et al.*, 2008; Valouev *et al.*, 2008; Teif *et al.*, 2012; Zhang, Zhang y Jiang, 2015; Voong *et al.*, 2016; Chen *et al.*, 2017; Zhang *et al.*, 2018), incluida *Arabidopsis* (Li *et al.*, 2014; Sacharowski *et al.*, 2015; Zhang, Zhang y Jiang, 2015).

La secuenciación se realizó sobre la fracción de DNA mononucleosomal (FM) purificada que representaba el 70-80% del DNA digerido. Esta fracción contenía fragmentos de un tamaño de alrededor de 150 pares de bases (pb) correspondiente a la longitud de DNA requerido para enrollar un nucleosoma; sin embargo, se podía observar la presencia de otras bandas minoritarias correspondientes a las fracciones di- y trinucleosomales (Fig. 29) lo que indicaba que las muestras no fueron sometidas a una digestión excesiva.

Estudios previos de los perfiles de densidad media de nucleosomas utilizando datos de MNase-seq en distintos organismos han identificado un patrón de organización de nucleosomas común alrededor de los sitios de inicio de la transcripción (TSS) (Hughes y Rando, 2014). En general, justo aguas arriba del TSS se localiza una región libre de nucleosomas (NDR), flanqueada por dos nucleosomas bien posicionados, el nucleosoma +1, que es el primer nucleosoma dentro de la región transcrita, y el nucleosoma -1, que se encuentra al otro lado de la NDR. Por otro lado, los genes con actividad transcripcional suelen presentar una serie de nucleosomas bien posicionados aguas abajo del TSS (+1, +2, +3...). Esto es debido a que los nucleosomas se sitúan y espacian de forma similar a partir del TSS. Sin embargo, este buen posicionamiento de nucleosomas no se observa en los genes reprimidos ya que, a pesar de que los nucleosomas puedan presentar un espaciamiento regular, no se posicionan de la misma forma (Baldi, 2019).

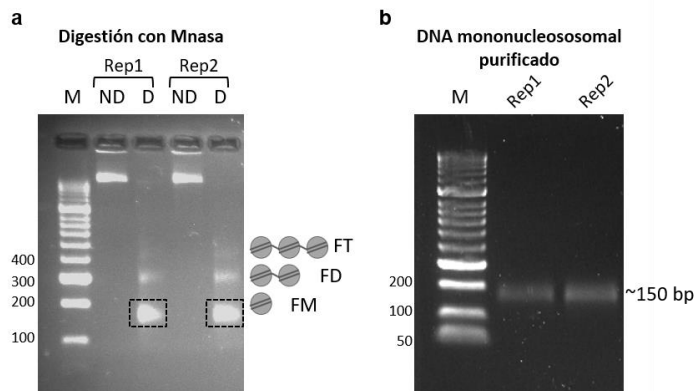


Figura 29: Digestión con MNase y purificación de fracción mononucleosomal (FM). (a) DNA no digerido (ND) y digerido (D) de las réplicas 1 y 2 de WT-Col0. Fracción Dinucleosomal (FD); Fracción Trinucleosomal (FT). (b) FM purificada del DNA de WT digerido de las réplicas 1 y 2. (M) marcador de peso molecular de fragmentos de DNA.

Para analizar el perfil de densidad de nucleosomas en la región del TSS en nuestras muestras, en primer lugar, seleccionamos todos los genes codificantes de proteínas detectados en un análisis de RNA-seq, previamente realizado en el grupo, en plántulas WT de la misma edad (Zhou *et al.*, 2017). Los genes fueron agrupados según su nivel de expresión (<1, 1-10 y >10 fragmentos por kilobase de exón por millón de lecturas mapeadas (FPKM) (Fig. 30a). Encontramos que, de acuerdo con datos previamente publicados en *Arabidopsis* (Li *et al.*, 2014; Zhang, Zhang y Jiang, 2015; Pass *et al.*, 2017), todos los genes codificantes de proteínas, independientemente de su nivel de expresión, presentaban una menor densidad de nucleosomas en la región aguas arriba del TSS que en la región aguas abajo del TSS. Esta menor densidad, aunque más pronunciada en la NDR, se extendía a lo largo de toda la zona promotora (750 pb aguas-arriba del TSS) (Fig. 30a). Encontramos que los genes con mayor nivel de expresión presentaban una NDR con una menor ocupación de nucleosomas que los genes que se expresaban menos; sin embargo, ninguno de estos grupos presentaba un buen posicionamiento de nucleosomas en toda la zona aguas arriba del TSS. Por otro lado, se observó que los genes con niveles de expresión intermedio y alto (1-10 y >10 FPKM, respectivamente) presentaban una serie de nucleosomas bien posicionados en la región aguas abajo del TSS, mientras que en el caso de los genes que no se expresan o se expresan poco (<1 FPKM), solo se podía distinguir con claridad el posicionamiento del nucleosoma +1. Asimismo, como ha sido descrito anteriormente (Li *et al.*, 2014), encontramos que la posición del nucleosoma +1 se correlacionaba con los niveles de expresión de los genes, de tal forma que cuanto más altos eran los niveles de expresión génica, más lejos se posicionaba el nucleosoma +1 del TSS. La distribución regular de los nucleosomas en la región aguas abajo del TSS y el desplazamiento del nucleosoma +1 en los genes transcripcionalmente activos han sido

relacionados con un posible papel de la RNA polimerasa II en determinar la conformación de la cromatina (Radman-Livaja y Rando, 2010). Por tanto, el análisis general de nuestras muestras estaba en concordancia con resultados anteriores.

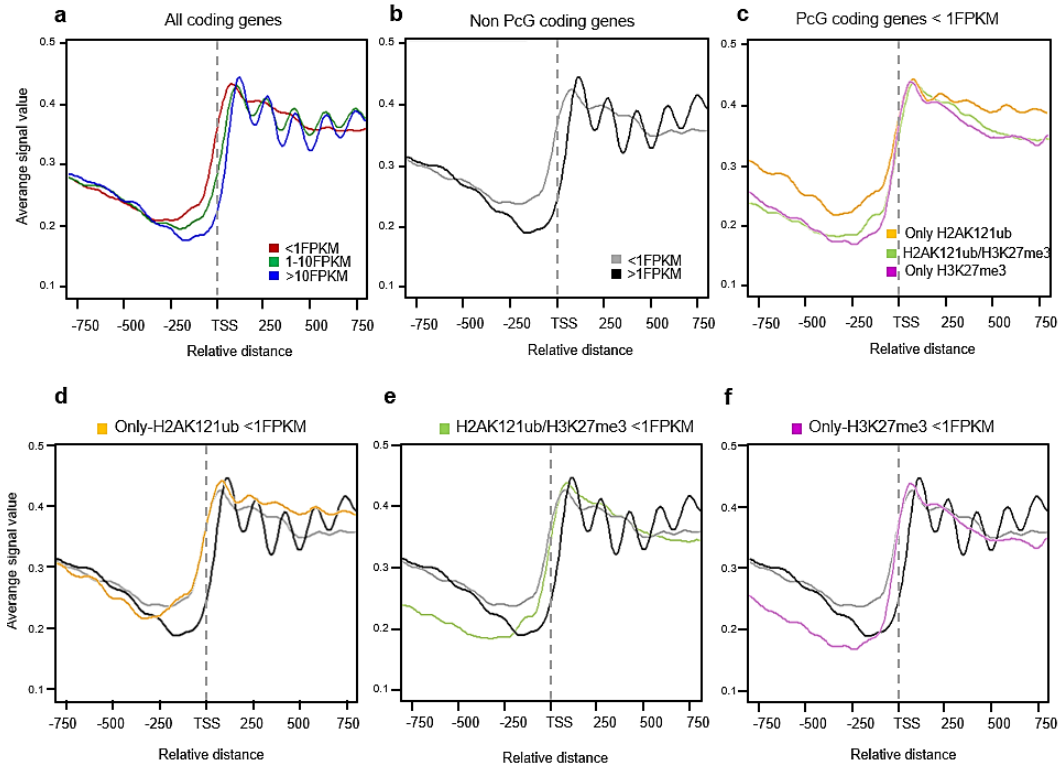


Figura 30: Comparación de los perfiles de distribución media de nucleosomas alrededor del sitio de inicio de transcripción (TSS) en diferentes conjuntos de genes en la fracción mononucleosomal (FM). (a) Perfiles de los genes codificantes de proteínas agrupados según su nivel de expresión en WT (< 1 FPKM (rojo), 1-10 FPKM (verde), > 1 FPKM (azul)). (b) Perfiles de los genes non-PcG activos (> 1 FPKM, negro) y reprimidos (< 1 FPKM, gris). (c) Comparación de los perfiles de los genes only-H2AK121ub, H2AK121ub/H3K27me3 y only-H3K27me3 (< 1 FPKM). (d) Perfiles de los genes only-H2AK121ub (< 1 FPKM) comparados con los genes non-PcG activos y reprimidos. (e) Perfiles de los genes H2AK121ub/H3K27me3 (< 1 FPKM) comparados con los genes non-PcG activos y reprimidos. (f) Perfiles de los genes only-H3K27me3 (< 1 FPKM) comparados con los genes non-PcG activos y reprimidos.

3.2. Los distintos grupos de genes diana PcG presentan diferencias en la organización de nucleosomas

Pasamos a analizar la organización de los nucleosomas en los genes diana de la maquinaria PcG. Ensayos anteriores del grupo de inmunoprecipitación de cromatina seguida de secuenciación (ChiP-seq) en plantas WT de 7 DAG, utilizando anticuerpos frente a las marcas H2AK121ub y H3K27me3 (Zhou *et al.*, 2017), pusieron de manifiesto la presencia de tres subgrupos diferentes de genes diana PcG en base a la presencia de

una de estas marcas (bien H2AK121ub o H3K27me3) o de las dos. Los niveles de expresión de los genes encontrados en los diferentes subgrupos fueron, en general, consistentes con el papel represivo de estas modificaciones. De acuerdo con esto, la mayoría de estos genes se encuentran sobre-expresados en los mutantes PRC1 o PRC2 (Bouyer *et al.*, 2011; Lafos *et al.*, 2011; Merini *et al.*, 2017; Zhou *et al.*, 2017). Sin embargo, se encontró que un porcentaje de genes en cada subgrupo presentaban actividad transcripcional (>5 FPKM). Este porcentaje resultó ser alto en los genes marcados solo con H2AK121ub -only-H2AK121ub- (alrededor de 50-60%), menor en el subgrupo de genes con ambas marcas -H2AK121ub/H3K27me3 (alrededor de 25%), y aún más bajo en los genes marcados solo con H3K27me3 -only-H3K27me3 (alrededor del 10%) (Zhou *et al.*, 2017). Por tanto, para investigar un posible papel de estas modificaciones en la organización de los nucleosomas, pero evitando el posible efecto que pueda causar la maquinaria de la transcripción en ello, se analizó la distribución de los nucleosomas en los genes codificantes de los tres subgrupos de dianas PcG con niveles de expresión entre 0 y 1 FPKM (<1 FPKM) de forma independiente. Elegimos este corte porque, de acuerdo con los análisis preliminares de la distribución de nucleosomas en función de la actividad, los genes englobados en este rango de expresión presentan el perfil característico de los genes reprimidos (Fig. 30a). Analizamos también el perfil de densidad media de nucleosomas en los genes codificantes que no son dianas PcG (non_PcG) reprimidos (<1 FPKM) y activados (>1 FPKM) (Fig. 30b) para compararlos con aquellos de los distintos grupos de dianas PcG.

Los resultados mostraron diferencias destacables en la distribución de nucleosomas entre los tres subgrupos de genes, tanto en la región aguas abajo como aguas arriba del TSS (Fig. 30c). Por un lado, encontramos que en los genes only-H2AK121ub la ocupación de nucleosomas en la zona aguas abajo del TSS era más parecida a la de los genes non-PcG activados que a la de los reprimidos y se podía distinguir el posicionamiento de los nucleosomas del +1 al +5, aunque de forma más pobre que en el caso de los genes non-PcG activados (Fig. 30d). Por el contrario, en los genes only-H3K27me3 la ocupación de nucleosomas en esta zona era similar a la de los genes non-PcG reprimidos y no se distinguía un posicionamiento claro de nucleosomas, a excepción del nucleosoma +1 (Fig. 30f). Los genes H2AK121ub/H3K27me3 en la zona aguas abajo del TSS presentaban una ocupación y un posicionamiento de nucleosomas similar a la de los genes only-H3K27me3 (Fig. 30e). En cualquier caso, los tres subgrupos de genes, aunque con algunas diferencias importantes, como se verá más adelante, mostraron un posicionamiento del nucleosoma +1 más próximo al TSS consistente con su estado transcripcional reprimido.

En cuanto a la región aguas arriba del TSS, en la que se englobaría la región promotora, encontramos que en los genes only-H2AK121ub la densidad nucleosomal era similar a la mostrada por los genes non-PcG reprimidos (Fig. 30d); sin embargo, en los genes únicamente marcados con H3K27me3 la densidad era significativamente inferior que a la de los genes non-PcG reprimidos, especialmente en la región distal del promotor (de -250 pb a -750 pb) donde era incluso menor que la mostrada por los genes non-PcG activados (Fig. 30f). En los genes H2AK121ub/H3K27me3, encontramos una situación intermedia (Fig. 30e). Aunque estos resultados sugerían un efecto claro de las marcas H3K27me3 sobre la ocupación de los nucleosomas en la región promotora, nos resultaron sorprendentes ya que ha sido descrito en animales que los promotores de los genes diana PcG presentan una sensibilidad reducida a la digestión con nucleasas en comparación con los promotores de los genes que no lo son (Bell *et al.*, 2010; Kelly *et al.*, 2012; Deaton *et al.*, 2016), por lo que cabría esperar una densidad de nucleosomas, si no mayor, al menos igual a la de los genes non-PcG reprimidos.

Respecto al posicionamiento de los nucleosomas en esta región, encontramos que solo en los genes only-H2AK121ub y H2AK121ub/H3K27me3 se podía intuir el posicionamiento del nucleosoma -1, aunque con una baja ocupación (Fig. 30d, e) y los subsiguientes nucleosomas (-2, -3...) solo se distinguían en los genes only-H2AK121ub. Por otro lado, en los genes non-PcG activos, aunque no se distinguía claramente el nucleosoma -1, se detectaba un cierto posicionamiento de los nucleosomas -2, -3... (Fig. 30b).

3.3. La comparación de perfiles de nucleosomas obtenidos en distintas condiciones de ensayo proporciona información complementaria

Se han descrito diferencias importantes en cuanto a la ocupación de los nucleosomas dependiendo de si las muestras son digeridas en condiciones de alta o baja concentración de MNasa (Mieczkowski *et al.*, 2016); de tal forma que, algunas regiones presentan mayor ocupación a bajas concentraciones y otras a altas concentraciones. Las regiones más accesibles del genoma se liberan a bajas concentraciones de enzima, detectándose así la ocupación de los nucleosomas “frágiles”; sin embargo, en condiciones de altas concentraciones, la digestión de estas regiones da lugar a fragmentos de tamaños menores al correspondiente a un nucleosoma, por lo que son excluidas del análisis en el caso de purificar la fracción mononucleosomal (~ 150 bp). Por otro lado, las regiones de baja accesibilidad pueden sufrir una digestión parcial generándose fragmentos de mayor tamaño, por lo que tampoco serían detectadas al llevarse a cabo la purificación de esta fracción. Por estas razones, se ha demostrado que el análisis de muestras tratadas en distintas condiciones puede proporcionar una información biológicamente importante que no se obtiene usando una única condición (Mieczkowski *et al.*, 2016).

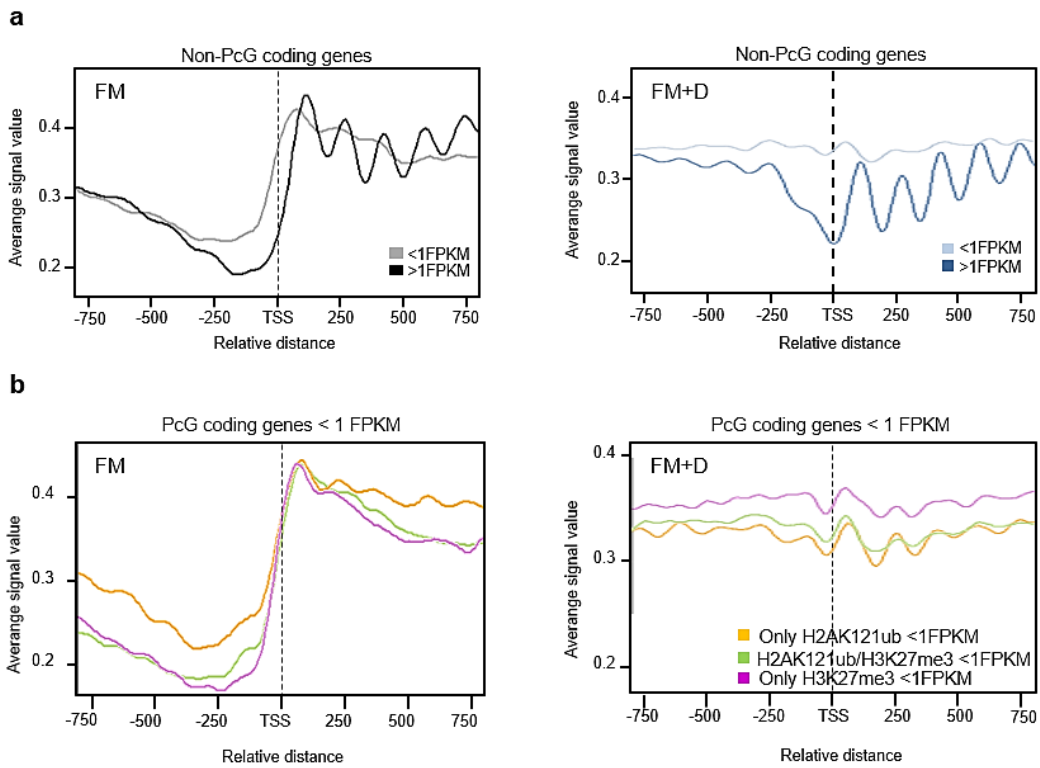


Figura 31: Las diferentes condiciones de ensayo de Mnase-seq proporcionan diferentes resultados en los perfiles de ocupación de nucleosomas en Arabidopsis. Perfiles de densidad media de nucleosomas alrededor del TSS de los genes en los diferentes grupos de genes. **(a)** Perfiles de los genes non-PcG agrupados según su nivel de expresión en la fracción mononucleosomal (FM), (panel izquierdo; genes activos (negro) y reprimidos (gris)) y en la fracción mono y dinucleosomal (FM+D), (panel derecho; genes activos (azul) y reprimidos (azul claro)). **(b)** Comparación de los perfiles de los genes only-H2AK121ub, H2AK121ub/H3K27me3 y only-H3K27me3 (< 1 FPKM) en la fracción mononucleosomal (FM, panel izquierdo y en la fracción mono y dinucleosomal (FM+D, panel derecho). En el eje X se representa la distancia relativa al TSS y en el eje Y se representa el valor medio de la señal. Las líneas discontinuas indican la posición del TSS.

Bajo nuestras condiciones de digestión obtuvimos una fracción mayoritaria de mononucleosomas (FM) que fue purificada para su secuenciación; sin embargo, se podía apreciar también la presencia de bandas de tamaño superior correspondientes a las fracciones di- y trinucleosomales (Fig. 29), por lo que pensamos que podríamos estar perdiendo información relevante para entender el papel de las marcas PcG. Para comprobar esta posibilidad, se analizaron datos de Mnase-seq previamente publicados en plántulas WT de edad similar. Estos datos fueron obtenidos a partir de muestras en las que no se había llevado a cabo la purificación de la fracción mononucleosomal sino que la secuenciación se había llevado a cabo de una fracción de fragmentos de tamaño ≤ 400

pb, incluyendo por tanto también la fracción dinucleosomal (FM+D), y posiblemente otros tamaños intermedios (Torres y Deal, 2019).

En primer lugar, se compararon los perfiles de densidad media de nucleosomas en los genes non-PcG reprimidos y activados en las dos muestras (FM y FM+D, Fig. 31a). En ambos análisis, en la región aguas abajo del TSS se observaba el patrón típico de distribución de nucleosomas con un nivel alto de ocupación en los dos grupos de genes y un buen posicionamiento de nucleosomas en los genes activados que no se observaba en los genes reprimidos. Sin embargo, se encontraron diferencias significativas en cuanto a la densidad de nucleosomas entre FM y la FM+D en general, particularmente en la región aguas arriba del TSS. A pesar de que en ambas muestras la densidad de nucleosomas era mayor en los genes reprimidos que en los activados, en la FM+D se podían apreciar detalles que no se distinguían en la FM. De tal forma que, por ejemplo, la NDR, aunque en las dos muestras había claras diferencias entre los genes non-PcG activados y reprimidos, en la FM+D esta región aparecía mejor definida y claramente flanqueada por los nucleosomas +1 y -1. Además, en la FM+D se observaba un posicionamiento más claro de los nucleosomas aguas arriba del nucleosoma -1, tanto en los genes non-PcG reprimidos como activados. Por tanto, pudimos comprobar que el análisis de los datos de la FM+D, en efecto, proporcionaba una información complementaria a la obtenida con la FM, por lo que procedimos a analizar los perfiles de densidad media de nucleosomas de los genes PcG en estas muestras.

3.4. El análisis comparativo de los perfiles de nucleosomas obtenidos en distintas condiciones desvela importantes diferencias entre los grupos de los genes PcG

Seguidamente, se obtuvieron los perfiles de densidad media de nucleosomas de los tres subconjuntos de dianas PcG con expresión <1 FPKM procedentes de la FM+D y se compararon con los presentados por los genes non-PcG reprimidos y activados (Fig. 31b, panel derecho, Fig. 32a, b, c, paneles inferiores). En general, los tres grupos de dianas PcG mostraron un perfil más similar al de los genes non-PcG reprimidos que al de los activados, lo que concuerda con su nivel de expresión. Sin embargo, se apreciaban diferencias interesantes tanto en la región aguas arriba del TSS como en la región aguas abajo. Estas diferencias en unos casos coincidían con los resultados de la FM (Fig. 31b, panel izquierdo, Fig. 32a, b, c, paneles superiores) y en otros proporcionaban información adicional.

3.5. H2AK121ub afecta al posicionamiento de nucleosomas +1 mientras que H3K27me3 altera la organización de nucleosomas en la región aguas abajo del +1

En primer lugar, comparamos el perfil de nucleosomas en la región aguas abajo del TSS en las dos muestras. En términos generales, encontramos que en la FM+D los genes *only-H2AK121ub*, al igual que en la FM (Fig. 32a), presentaban un posicionamiento del nucleosoma +1 más próximo al TSS que el de los *non-PcG* activados, una densidad de nucleosomas similar a la de los genes *non-PcG* activados y un mejor posicionamiento de los nucleosomas del +1 a +5 en comparación con los genes *non-PcG* reprimidos (Fig. 32a, panel inferior). Sin embargo, el perfil de nucleosomas de los genes *only-H3K27me3* difería claramente del obtenido con los datos de la FM (Fig. 32c). A pesar de que también mostraba un posicionamiento pobre de nucleosomas, la densidad de nucleosomas en toda esta zona era considerablemente mayor a la de los genes *non-PcG* reprimidos (Fig. 32c, panel inferior), cosa que no se observaba en el análisis de la FM. Estos resultados podrían indicar que los genes marcados con H3K27me3 presentan una menor accesibilidad a la digestión por MNasa, lo que encajaría con datos anteriores (Bell *et al.*, 2010; Kelly *et al.*, 2012; Shu *et al.*, 2012; Deaton *et al.*, 2016) y explicaría por qué en el perfil obtenido de la FM se encuentra una densidad de mononucleosomas menor. Curiosamente, los genes *H2AK121ub/H3K27me3* presentaban características intermedias entre los *only-H2AK121ub* y los *only-H3K27me3* en cuanto a la ocupación y el posicionamiento de nucleosomas en esta zona (Fig. 32b, panel inferior).

Al analizar tanto la FM como la FM+D nos llamó la atención que el posicionamiento del nucleosoma +1 en las dianas *PcG*, aunque era similar al de los genes *non-PcG* reprimidos, parecía mostrar una cierta variabilidad. Para investigar esto, calculamos el posicionamiento más frecuente de este nucleosoma en los distintos grupos de genes tanto en los perfiles de la FM como en los de la FM+D (Fig. 32d). Para ello, determinamos para cada gen de los distintos subgrupos la posición en que se detectaba el valor máximo de señal dentro del intervalo en que, de acuerdo con los perfiles de densidad de nucleosomas, se delimita el nucleosoma +1. Encontramos que el posicionamiento del nucleosoma +1 en los genes *only-H2AK121ub* y *H2AK121ub/H3K27me3* estaba desplazado de forma significativa unas 8 pb aguas abajo en comparación con el de los genes *non-PcG* reprimidos, tanto en la FM como en la FM+D. Estos resultados respaldan que las marcas H2AK121ub juegan un papel en determinar un posicionamiento concreto del nucleosoma +1. Por el contrario, la posición del nucleosoma +1 en los genes *only-H3K27me3* era similar a la de los genes *non-PcG* reprimidos en la FM. Esto también ocurría en la FM+D cuando considerábamos como el centro del intervalo predefinido para el nucleosoma +1 el valor de señal de ocupación promedio máximo (Fig. 32d, e); sin embargo, la forma del perfil promedio del nucleosoma +1 en la FM+D era

asimétrica, por tanto, al considerar el intervalo completo (Fig. 32e), encontramos diferencias significativas en comparación con los genes non-PcG reprimidos (Fig. 32d, panel derecho). Esta distribución asimétrica del nucleosoma +1 en los genes only-H3K27me3 sugiere una accesibilidad reducida a la MNasa en esta región.

A continuación, pasamos a analizar el posicionamiento y espaciado de los cuatro primeros nucleosomas aguas abajo del TSS. Para ello, nos centramos en los datos procedentes de la FM+D, al ser donde se podían apreciar diferencias más claras (Fig. 33a). Cuando un nucleosoma está bien posicionado en la misma región en los distintos genes y/o en las distintas células se refleja en una mayor altura del pico de localización del nucleosoma (Fig. 33b). Por otro lado, la regularidad de la secuencia de nucleosomas en un grupo de genes se puede intuir en función de la relación entre la anchura de los picos (Fig. 33b). Estas características son independientes, así por ejemplo la secuencia de nucleosomas en un grupo de genes puede estar espaciada regularmente, pero presentar un posicionamiento pobre (Baldi, 2019).

Encontramos que los cuatro primeros nucleosomas aguas abajo del TSS presentaban una altura considerablemente mayor en los genes non-PcG activados que en los otros grupos de genes. En cuanto a los genes only-H3K27me3, H2AK121ub/H3K27me3 y non-PcG reprimidos, encontramos que la altura de los picos en las distintas posiciones era similar entre ellos; sin embargo, en los genes only-H2AK121ub era comparativamente más alta, particularmente en las posiciones de los nucleosomas +2 y +3, lo que sugiere un mejor posicionamiento de los nucleosomas en only-H2AK121ub que en el resto de los genes reprimidos (Fig. 33c). Por otro lado, encontramos que los distintos picos en los genes non-PcG activos mostraban anchos similares. Este también era el caso en los genes non-PcG reprimidos, aunque la anchura de los picos era mayor que en los genes activos (Fig. 33d). En contraste, el ancho de los picos de los cuatro nucleosomas en los genes PcG era más irregular, especialmente en los genes marcados con H3K27me3 (Fig. 33d). Estos resultados, en conjunto, indican que los nucleosomas aguas abajo del TSS de los genes non-PcG se disponen en secuencias regulares con un buen posicionamiento de nucleosomas en el caso de los activados y con un posicionamiento pobre en el de los reprimidos (Fig. 33e). Sin embargo, en los genes diana PcG estos nucleosomas se organizan en secuencias irregulares, presentando además un posicionamiento pobre de nucleosomas, especialmente en los marcados con H3K27me3 (Fig. 33e), lo que sugiere que estas marcas tienen un efecto en el posicionamiento de los nucleosomas en el cuerpo del gen.

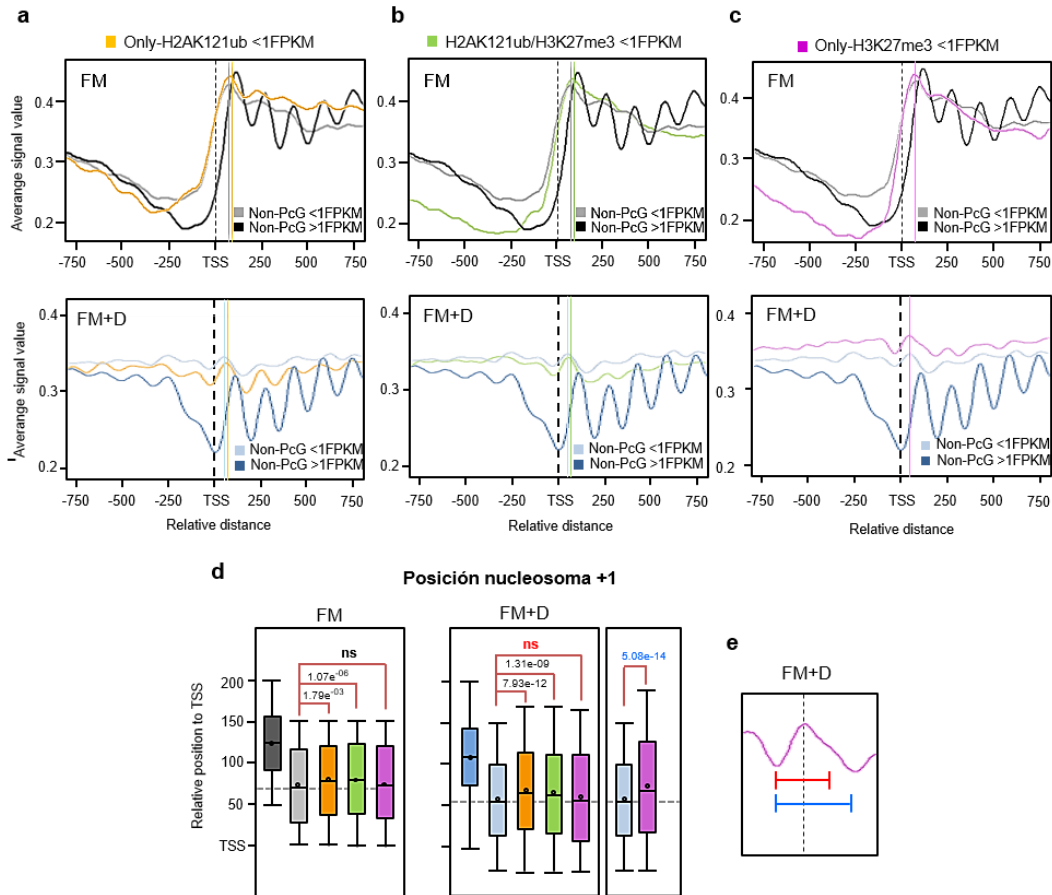


Figura 32: Perfiles de densidad media de nucleosomas de los diferentes genes PcG reprimidos en FM y FM+D. (a) Perfiles de los genes only-H2AK121ub comparados con los non-PcG activos y reprimidos en FM (panel superior) y FD+M (panel inferior). (b) Perfiles de los genes H2AK121ub/H3K27me3 comparados con los non-PcG activos y reprimidos en FM (panel superior) y FD+M (panel inferior). (c) Perfiles de los genes H3K27me3 comparados con los non-PcG activos y reprimidos en FM (panel superior) y FD+M (panel inferior). (d) Diagrama de cajas y bigotes representando las diferencias en la posición del nucleosomas +1 entre los diferentes grupos de genes. En cada caja, se representa la mediana (segmento dentro del rectángulo), la media (punto dentro del rectángulo), cuartiles superior e inferior (cajas) y los valores máximo y mínimo (bigotes). Se indican los valores de p-valor de las diferencias significativas determinados mediante el test de Wilcoxon. Cada grupo de genes se representa en un color diferente. Las líneas verticales indican el valor máximo del posicionamiento del nucleosoma +1. La línea negra discontinua indica la posición del TSS. (e) Se consideraron dos intervalos diferentes para determinar la posición del nucleosoma +1 en los genes only-H3K27me3; unos considerando el valor máximo de la señal como el centro del intervalo (rojo) y otro incluyendo los extremos distribuidos asimétricamente (azul).

3.6. Los genes marcados con H3K27me3 muestran una accesibilidad reducida a la digestión por MNasa en la región del promotor

Seguidamente, analizamos la región aguas arriba del TSS en los distintos grupos de genes. Como se ha mencionado anteriormente, los perfiles de los genes non-PcG derivados de los datos de la FM+D mostraron importantes diferencias en comparación con los obtenidos de la FM en esta región. Los genes only-H2AK121ub presentaban un perfil de densidad similar al de los non-PcG reprimidos en esta zona tanto en la FM como en la FM+D (Fig. 32a; Fig. 33a); sin embargo, los genes marcados con H3K27me3 (tanto los H2AK121ub/H3K27me3 como los only-H3K27me3) mostraron una densidad de nucleosomas considerablemente menor que la de los genes non-PcG reprimidos en la FM (Fig. 32b, c, paneles superiores) en comparación con la FM+D (Fig. 32b, c, paneles inferiores). En el caso de los genes H2AK121ub/H3K27me3 en la FM+D, esta zona presentaba una densidad similar a la de los genes non-PcG reprimidos y en el caso de los only-H3K27me3 la densidad era incluso mayor (Fig. 32b, c, paneles inferiores), lo que sugiere que los genes marcados con H3K27me3 presentan una menor accesibilidad a la digestión por MNasa. Esto hace que se detecte una menor proporción de mononucleosomas en la FM en esta zona que en los non-PcG reprimidos. Estos resultados encajan con el hecho de que las marcas H2AK121ub se localizan mayoritariamente en la zona inmediatamente después del TSS, por lo que no debería tener efecto en la región promotora, mientras que las H3K27me3 ocupan el cuerpo del gen pero también se extienden a la región promotora (Zhou *et al.*, 2017), por lo que podrían tener un papel en el establecimiento de una conformación cerrada de la cromatina en esta región, como ha sido propuesto anteriormente (Bell *et al.*, 2010; Kelly *et al.*, 2012; Shu *et al.*, 2012; Deaton *et al.*, 2016).

A diferencia de FM, en FM + D la posición del nucleosoma -1 podría identificarse en los perfiles de los distintos subconjuntos de genes (Fig. 33a); sin embargo, en los genes non-PcG activos, la ocupación en esta posición estaba considerablemente reducida en comparación con el resto de los genes, lo que respalda la idea de que el nucleosoma -1 es desplazado por la maquinaria de transcripción (Jiang y B Franklin Pugh, 2009). Curiosamente, encontramos que la posición del nucleosoma -1 era similar en los diferentes subconjuntos de genes reprimidos (Fig. 33f). Determinamos también la ocupación en la NDR, ya que se ha demostrado que esta región es fundamental para la unión de factores de transcripción y de la maquinaria de transcripción (Zhang y Reese, 2007; Liu *et al.*, 2015). La ocupación más baja en esta región se encontró en genes non-PcG activos, seguidos de los genes only-H2AK121ub, H2AK121ub/H3K27me3, non-PcG reprimidos y only-H3K27me3 (Fig. 33g). El hecho de que la posición del nucleosoma -1 sea similar en todos los genes reprimidos, indica que las diferencias en la ocupación de NDR dependen del grado de invasión de nucleosomas +1 en NDR. El distinto grado de invasión por parte del nucleosoma +1 debe obstaculizar en distinta medida la accesibilidad de secuencias de DNA reguladoras funcionalmente importantes.

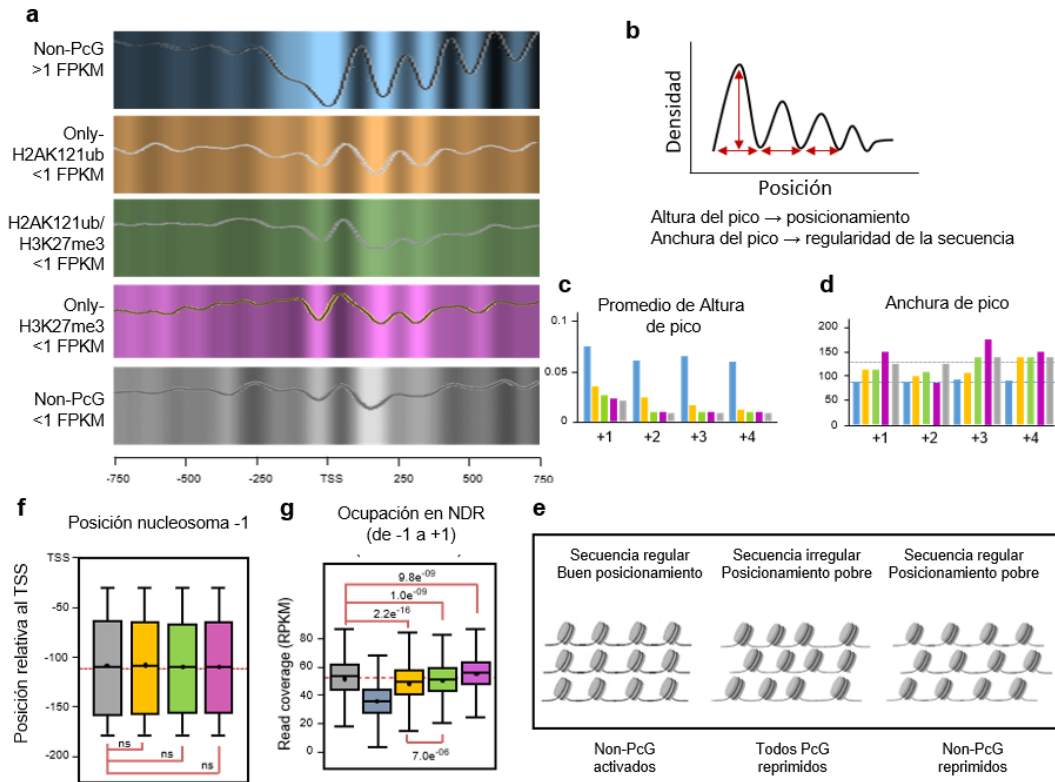


Figura 33: Los genes PcG presentan secuencias irregulares y un posicionamiento pobre de los nucleosomas. (a) Mapa de calor representando la densidad media de nucleosomas en los genes PcG y non-PcG alrededor del TSS en FM+D. (b) Parámetros para determinar el posicionamiento y la regularidad de la secuencia de nucleosomas. Las flechas rojas indican la altura y la anchura del pico. (c) La altura relativa del pico de los cuatro primeros nucleosomas aguas abajo del TSS en los diferentes grupos de genes. (d) La anchura relativa del pico de los cuatro primeros nucleosomas aguas abajo del TSS en los diferentes grupos de genes. (e) Esquema mostrando la organización de los cuatro primeros nucleosomas en los diferentes grupos de genes. (f) Diagrama de cajas y bigotes representando las diferencias en la posición del nucleosomas -1 entre los diferentes grupos de genes. En cada caja, se representa la mediana (segmento dentro del rectángulo), la media (punto dentro del rectángulo), cuartiles superior e inferior (cajas) y los valores máximo y mínimo (bigotes). Se indican los valores de p-valor de las diferencias significativas determinados mediante el test de Wilcoxon. Cada grupo de genes se representa en un color diferente. Las líneas verticales indican el valor máximo del posicionamiento del nucleosoma -1. La línea negra discontinua indica la posición del TSS. (g) Diagramas de cajas y bigotes representando las diferencias en la ocupación en la NDR. En cada caja, se representa la mediana (segmento dentro del rectángulo), la media (punto dentro del rectángulo), cuartiles superior e inferior (cajas) y los valores máximo y mínimo (bigotes). Se indican los valores de p-valor de las diferencias significativas determinados mediante el test de Wilcoxon.

3.7. Los perfiles de nucleosomas de los genes only-H2AK121ub o H2AK121ub/H3K27me3 activos desvelan una heterogeneidad transcripcional en los genes de estos conjuntos

Como se ha mencionado anteriormente, resultados de ChIP-seq y RNA-seq del laboratorio en plántulas WT enteras a 7 DAG mostraron que un porcentaje de los genes only-H2AK121ub y H2AK121ub/H3K27me3 presentaban actividad transcripcional > 5 FPKM (50-60% y 25%, respectivamente, (Zhou *et al.*, 2017)). La alta proporción de genes only-H2AK121ub con actividad transcripcional llevó a proponer que o bien el efecto de represión de esta modificación puede ser superado por la maquinaria transcripcional gracias a la presencia de modificaciones adicionales implicadas en su activación, o que estas marcas juegan un papel en la activación (Zhou *et al.*, 2017). Sin embargo, al comparar los genes codificantes de proteínas con niveles de expresión >1 FPKM en estos dos grupos de dianas PcG con los genes sobre-expresados en el mutante *atbmi1a/b/c* (FC >2 , p-value $< 0,05$ (Gómez-Zambrano, Merini y Calonje, 2019)), encontramos que había un alto número de genes solapantes (Fig. 34a, b), lo que apoya que la incorporación de H2AK121ub en estos genes tenga un efecto de represión.

Seguidamente, pasamos a comparar el perfil de densidad de nucleosomas de los genes only-H2AK121ub y H2AK121ub/H3K27me3 con niveles de expresión >1 FPKM con el de los genes non-PcG activos (>1 FPKM) y reprimidos (<1 FPKM) (Fig. 34c, d, paneles superiores). Encontramos que, en ambos casos, especialmente en el de los only-H2AK121ub, el perfil se aproximaba al de los non-PcG activados; sin embargo, la ocupación en la NDR, la altura de los picos y el posicionamiento del nucleosoma +1 no llegaban a ser iguales a los de los genes non-PcG activados (Fig. 34c, d, paneles superiores, Fig. 34e, panel izquierdo). Para comprobar que esto no era debido a que hubiese una mayor proporción de genes con mayor actividad en los non-PcG, seleccionamos en ambos casos los genes con un intervalo de expresión entre 1 y 10 FPKM, pero los resultados fueron similares (Fig. 34c, d, paneles inferiores, Fig. 34e, panel derecho).

Es importante tener en cuenta que los análisis de ChIP-seq, de RNA-seq y de MNasa-seq al haberse realizado en plántulas enteras proporcionan datos promedio del análisis de distintos tejidos, por tanto, nos preguntamos si estos resultados podían ser debidos a que hubiera una mezcla de una población de células con genes marcados/reprimidos y otras con genes no-marcados/activados, apareciendo en conjunto como marcados y expresados. Esto explicaría la presencia de marcas PcG en genes con actividad transcripcional y el hecho de que los genes PcG y non-PcG con el mismo nivel de expresión presenten perfiles de nucleosomas distintos. Para evaluar esta posibilidad, tomamos una muestra aleatoria de genes non-PcG activados y de genes only-H2AK121ub o H2AK121ub/H3K27me3 reprimidos y los mezclamos en distintas proporciones para obtener el perfil de densidad promedio. Los perfiles mostrados por el conjunto de genes

con proporción 1:2 de genes only-H2AK121ub <1 FPKM y non-PcG >1 FPKM y con 9:1 de H2AK121ub/H3K27me3 <1 FPKM y non-PcG >1 FPKM se asemejaban a los mostrados por los genes con actividad transcripcional de 1-10 FPKM only-H2AK121ub y H2AK121ub/H3K27me3, respectivamente (Fig. 35a, b), lo que apoya esta posibilidad.

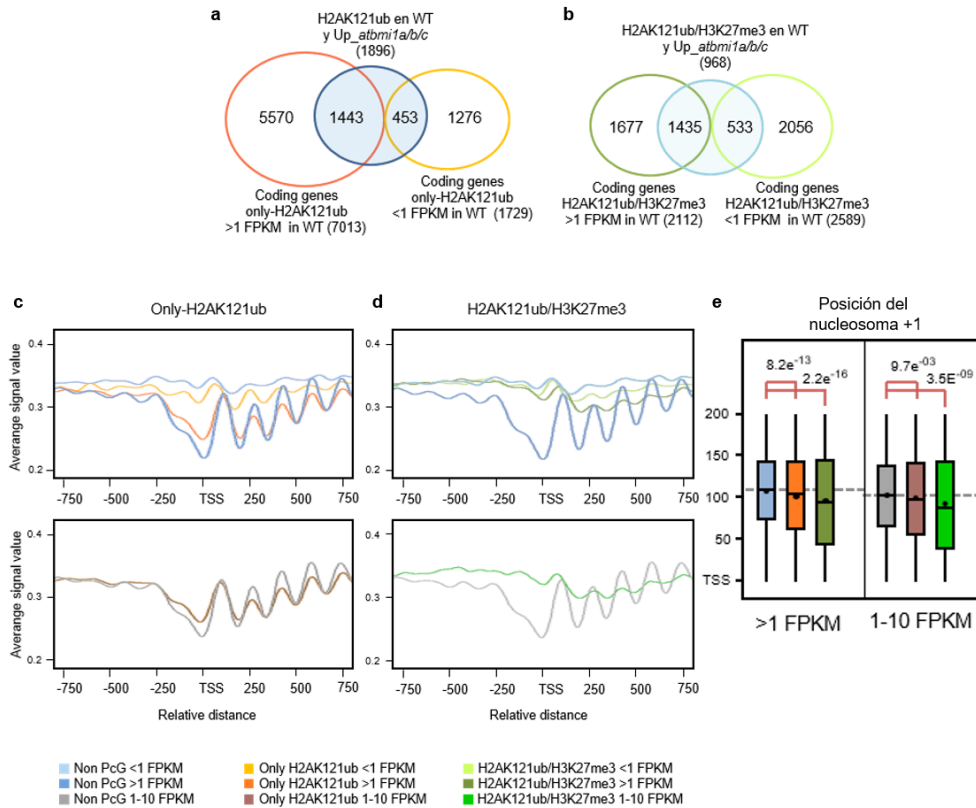


Figura 34: Perfiles de densidad media de nucleosomas de los diferentes genes PcG activados en FM+D. (a) Diagrama de Venn representando el solapamiento entre los genes only-H2AK121ub con niveles de expresión >1 o <1 FPKM en WT y los genes sobre-expresados en el mutante *atbmi1a/b/c* y marcados con H2AK121ub en WT. **(b)** Diagrama de Venn representando el solapamiento entre los genes H2AK121ub/H3K27me3 con niveles de expresión >1 o <1 FPKM en WT y los genes sobre-expresados en el mutante *atbmi1a/b/c* y marcados con H2AK121ub/H3K27me3 en WT. **(c)** Perfiles de los genes only-H2AK121ub con niveles de expresión >1 FPKM comparados con los genes non-PcG activos y reprimidos con niveles de expresión >1 FPKM y <1 FPKM, respectivamente (panel superior); Comparación de los perfiles de los genes only-H2AK121ub y non-PcG activos con niveles de expresión 1-10 FPKM (panel inferior). **(d)** Perfiles de los genes H2AK121ub/H3K27me3 con niveles de expresión >1 FPKM comparados con los genes non-PcG activos y reprimidos con niveles de expresión >1 FPKM y <1 FPKM, respectivamente (panel superior); Comparación de los perfiles de los genes H2AK121ub/H3K27me3 y non-PcG activos con niveles de expresión 1-10 FPKM (panel inferior). **(e)** Diagramas de cajas y bigotes representando las diferencias en el posicionamiento del nucleosoma +1 entre los diferentes grupos de genes con niveles de expresión >1 FPKM o 1-10 FPKM. En cada caja, se representa la mediana (segmento dentro del rectángulo), la media (punto dentro del rectángulo), cuartiles superior e inferior (cajas) y los valores máximo y mínimo (bigotes). Se indican los valores de p-valor de las diferencias significativas determinados mediante el test de Wilcoxon.

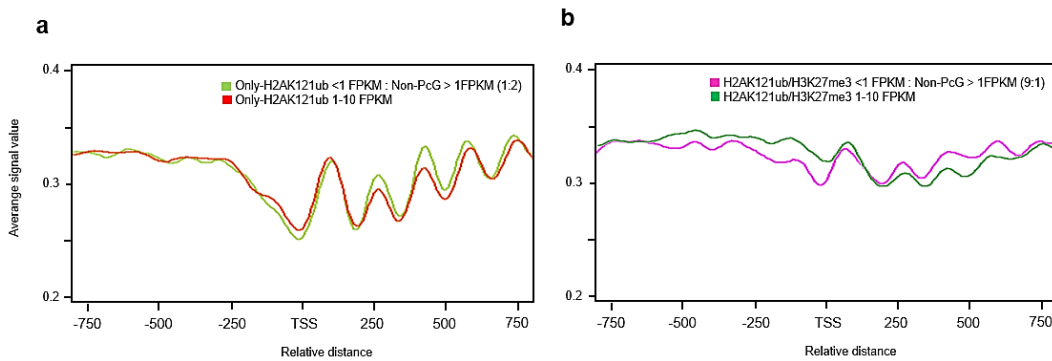


Figura 35. Perfiles de densidad media de nucleosomas de los grupos de genes con niveles de expresión mixtos. (a) Perfiles de la mezcla de genes only-H2AK121ub reprimidos y non-PcG activos (1:2) comparados a los genes only-H2AK121ub activos 1-10 FPKM. **(b)** Perfiles de la mezcla de genes H2AK121ub/H3K27me3 reprimidos y non-PcG activos (9:1) comparados a los genes H2AK121ub/H3K27me3 activos 1-10 FPKM.

El hecho de que haya una proporción mayor de genes activados en los only-H2AK121ub que en H2AK121ub/H3K27me3 encaja con el tipo de genes presentes en estos dos grupos. Los only-H2AK121ub están enriquecidos en términos de GO relacionados con metabolismo, procesos celulares y respuesta a estímulos, por lo que se espera una regulación más dinámica de estos genes en las distintas células y, por tanto, una mayor heterogeneidad de actividad transcripcional. Sin embargo, los H2AK121ub/H3K27me3 están enriquecidos en términos de GO relacionados con el desarrollo y diferenciación de distintos tejidos, por lo que se espera que, si se expresan, lo hagan en un número reducido de células, lo que explicaría el poco impacto que tienen los genes activos en el perfil de densidad de nucleosomas en este grupo de genes. Esto apoyaría un papel de las marcas H2AK121ub en mediar una represión inicial y más flexible y un papel de las marcas H3K27me3 en el mantenimiento de una represión estable.

3.8. Los perfiles de densidad de nucleosomas en *atbmi1a/b/c* apoyan el papel de las marcas H2AK121ub en determinar el posicionamiento del nucleosoma +1

Para investigar más a fondo el papel de las marcas PcG en la organización de los nucleosomas, llevamos a cabo la digestión con MNasa de la cromatina del mutante *atbmi1a/b/c* a los 7 DAG y la subsiguiente purificación de la FM para su secuenciación, como en el caso de las plántulas WT. Estos mutantes experimentan una pérdida sustancial, aunque no total de las marcas H2AK121ub en los genes only-H2AK121ub y H2AK121ub/H3K27me3. Además, los genes H2AK121ub/H3K27me3 experimentan una pérdida de H3K27me3 que se correlaciona con la pérdida de H2AK121ub, lo que indica que el establecimiento de H2AK121ub es necesario para la incorporación de H3K27me3 en estos genes. Sin embargo, los niveles de H3K27me3 en los genes únicamente marcados

con esta modificación no están alterados (Zhou *et al.*, 2017), lo que sugiere que la incorporación de H3K27me3 en estos genes se lleva a cabo a través de un mecanismo diferente (Merini y Calonje, 2015; Zhou *et al.*, 2017).

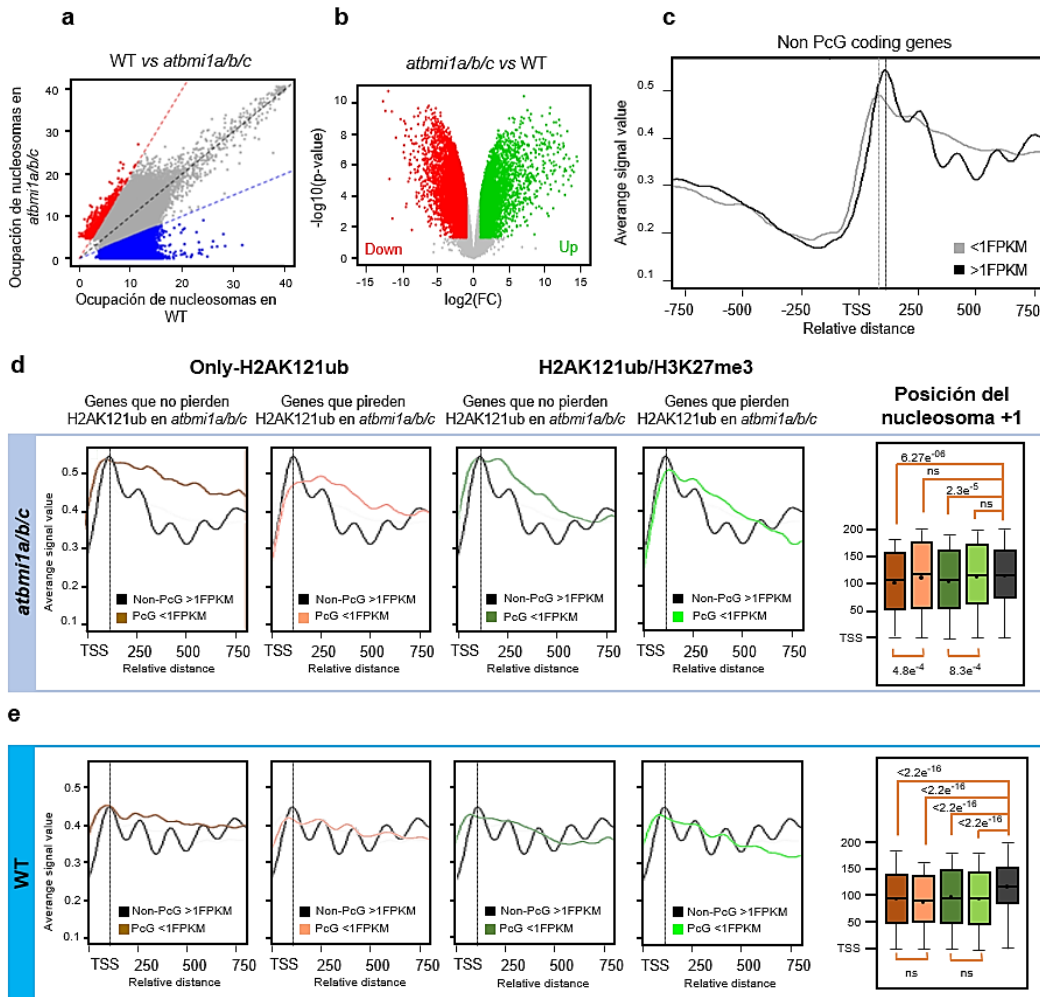


Figura 36: La organización de los nucleosomas está alterada en el mutante *atbmi1a/b/c*. (a) Diagrama de dispersión (Scatterplot) representando la comparación de los nucleosomas detectados en WT y *atbmi1a/b/c*. (b) Diagrama de dispersión (Volcano plot) representando los genes reprimidos (rojo) y sobre-expresados (verde) en *atbmi1a/b/c* comparados con WT. (c) Perfiles de densidad media de nucleosomas en los genes non-PcG en el mutante *atbmi1a/b/c*. (d) Comparación de los perfiles de diferentes grupos de genes only-H2AK121ub o H2AK121ub/H3K27me3 con los genes non-PcG activos en el mutante *atbmi1a/b/c*. Diagrama de cajas y bigotes representando las diferencias en el posicionamiento del nucleosoma +1 entre los diferentes grupos de genes. En cada caja, se representa la mediana (segmento dentro del rectángulo), la media (punto dentro del rectángulo), cuartiles superior e inferior (cajas) y los valores máximo y mínimo (bigotes). Se indican los valores de p-valor de las diferencias significativas determinados mediante el test de Wilcoxon. (e) Igual que en (d) pero en WT.

Al comparar los nucleosomas detectados en el mutante con los detectados en WT en las mismas condiciones de digestión, encontramos que un gran número de nucleosomas en WT no se detectaban o se detectaban con menor intensidad en el mutante y que había otros nucleosomas (aunque un número menor) en los que ocurría lo contrario (Fig. 36a). La pérdida de función de la actividad AtBMI1 da lugar a la desregulación de un número muy grande de genes, tanto al alza como a la baja (Fig. 36b). Entre los genes sobre-expresados se encuentran los genes embrionarios, lo que hace que estos mutantes no logren salir de la fase embrionaria. Esta alteración generalizada del transcriptoma, así como la permanencia de las células en un estadio embrionario deben impactar considerablemente en la organización de los nucleosomas, como sugieren las diferencias encontradas en la detección de nucleosomas con las plántulas WT. Por otro lado, parte de estas diferencias podrían estar incrementadas como consecuencia del proceso de generación de la librería de secuenciación, ya que al haber una proporción menor de mononucleosomas, los que sí están serán amplificados con mayor eficiencia. De acuerdo con esto, al analizar el perfil de los genes non-PcG activados y reprimidos en *atbmi1a/b/c* (Fig. 36c) encontramos que, aunque estos genes seguían presentando las características generales de los genes activos y reprimidos, había diferencias importantes en cuanto a la ocupación de nucleosomas, tanto en la región aguas arriba del TSS como en la región aguas abajo, con los perfiles en WT (Fig. 30b). Los cambios observados en los genes que no son dianas directas de la regulación PcG ponen de manifiesto el impacto generalizado, directo y/o indirecto, en la organización de los nucleosomas en el mutante *atbmi1a/b/c*. Esto nos impidió llevar a cabo el análisis comparativo directo de los perfiles de nucleosomas obtenidos en plántulas *atbmi1a/b/c* y WT.

En cualquier caso, como en el mutante *atbmi1a/b/c* los genes non-PcG activos y reprimidos mostraban un posicionamiento de nucleosomas consistente con sus niveles de expresión (Fig. 36c), para determinar el efecto de la pérdida de H2AK121ub comparamos el perfil de nucleosomas en distintos subconjuntos de genes en *atbmi1a/b/c*. Centramos el análisis en las dianas directas de regulación PcG, es decir, los genes marcados en WT con H2AK121ub o H2AK121ub/H3K27me3. Dividimos estos genes en dos subgrupos: los que sufrían una pérdida considerable de marcas H2AK121ub (niveles $\leq 60\%$ de los niveles en WT); y los que no experimentaban una pérdida sustancial de las mismas (niveles $>80\%$ de los niveles en WT). Para evitar el efecto de la maquinaria de transcripción en la organización de los nucleosomas en estos genes, restringimos el análisis a los genes que estaban reprimidos en WT ($< 1\text{FPKM}$) y que no experimentaban un cambio de expresión en *atbmi1a/b/c* (Fig. 37). Muchos de los genes diana de PcG, a pesar de perder las marcas en *atbmi1a/b/c*, no se activan transcripcionalmente al no encontrarse probablemente el factor requerido para ello (Merini *et al.*, 2017).

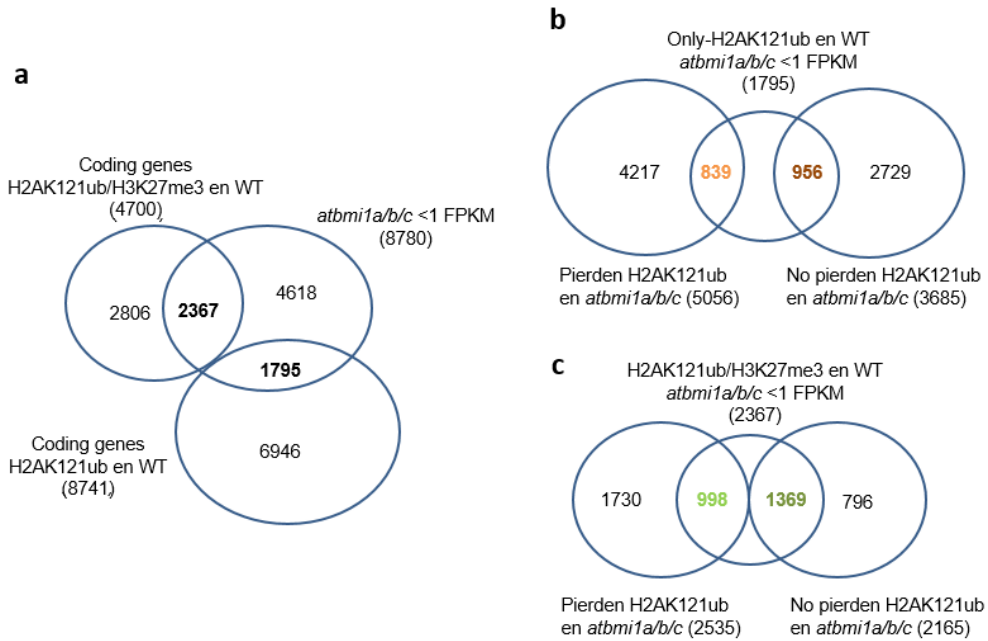


Figura 37: Selección de diferentes conjuntos de genes en *atbmi1a/b/c* para la comparación de los perfiles de nucleosomas. (a) Diagramas de Venn representando el solapamiento entre los genes codificantes con niveles de expresión <1 FPKM y los genes marcados con H2AK121ub/H3K27me3 o H2AK121ub en WT. **(b)** Diagramas de Venn representando el solapamiento entre los genes only-H2AK121ub en WT con niveles de expresión <1 FPKM en *atbmi1a/b/c* y los genes only-H2AK121ub en WT que pierden o no pierden la marca H2AK121ub en *atbmi1a/b/c*. **(c)** Diagramas de Venn representando el solapamiento entre los genes H2AK121ub/H3K27me3 en WT con niveles de expresión <1 FPKM en *atbmi1a/b/c* y los genes H2AK121ub/H3K27me3 en WT que pierden o no pierden la marca H2AK121ub en *atbmi1a/b/c*.

La comparación del perfil de nucleosomas de los genes que perdían y no perdían marca en *atbmi1a/b/c* (Fig. 36d, paneles izquierdos) indicaba un efecto de H2AK121ub en el posicionamiento del nucleosoma +1. Tanto en el caso de los genes only-H2AK121ub como en los H2AK121ub/H3K27me3, los genes que no perdían las marcas H2K121ub presentaban un posicionamiento del nucleosoma +1 desplazado aguas arriba comparado con el de los genes non-PcG activados, similar a lo que ocurre en el caso de los perfiles en WT (Fig. 36e), mientras que los genes que sí perdían las marcas, el posicionamiento del nucleosoma +1 era similar al de los genes activados (Fig. 36d, último panel). Estos resultados, una vez más, apoyan el papel de H2AK121ub en determinar el posicionamiento del nucleosoma +1.

4. Discusión

Nuestros datos apoyan un papel de las marcas H2AK121ub y H3K27me3 en la organización de los nucleosomas para favorecer el establecimiento de una conformación de cromatina refractaria a la transcripción.

La técnica de MNasa-seq ha sido extensivamente utilizada para generar mapas de nucleosomas en muchos organismos y tipos celulares (Mavrich *et al.*, 2008; Valouev *et al.*, 2008; Teif *et al.*, 2012; Zhang, Zhang y Jiang, 2015; Voong *et al.*, 2016; Chen *et al.*, 2017; Zhang *et al.*, 2018). Sin embargo, uno de los problemas de esta aproximación en muestras que contienen distintos tipos celulares es que los perfiles que se obtienen están basados en la densidad media en el conjunto de células, de tal forma que los nucleosomas que no ocupan la misma posición en los distintos genes no aparecen como picos definidos. Esto se acentúa además si hay heterogeneidad en los niveles de expresión de los genes de célula a célula. Otro de los problemas de esta técnica es la variabilidad de los resultados dependiendo de las diferencias en el grado de digestión de las muestras, aunque esto afecta más al nivel de ocupación de los nucleosomas que al posicionamiento y espaciado. En cualquier caso, ensayos recientes de MNasa-seq realizados en una única célula confirman muchas de las conclusiones obtenidas en mezclas complejas de células (Lai *et al.*, 2018; Baldi 2019). Para minimizar estas limitaciones, llevamos a cabo el análisis en muestras tratadas en distintas condiciones y seleccionamos grupos de genes basándonos en el nivel de expresión y en el tipo de marcas que presentan. Esto, por un lado, nos ha permitido obtener una información que no se hubiera obtenido usando una sola condición, y por otro establecer diferencias claras entre genes activados y reprimidos, y entre las distintas dianas de regulación PcG.

Excepto por un trabajo recientemente publicado en animales en el que se ha analizado el efecto de los complejos PRC1 y PRC2 en la organización de nucleosomas mediante ATAC-seq en las islas GC de los promotores de las dianas PcG (King *et al.*, 2018), poco se sabe del papel de los complejos PcG en la organización de los nucleosomas. En este trabajo, se propone que el PRC1 juega un papel en incrementar la ocupación de nucleosomas y disminuir el espaciado entre ellos en los promotores diana. Aparte de esto, se ha visto que las proteínas PcG son capaces de mediar *in vitro* la inhibición de la actividad de complejos remodeladores de nucleosomas y la compactación de la cromatina (Francis, Kingston y Woodcock, 2004; Trojer *et al.*, 2007, 2011; Grau *et al.*, 2011), lo que encaja con el hecho de que los promotores regulados por los complejos PcG presentan una menor accesibilidad a la digestión por nucleasas en comparación con los que no son dianas PcG (Bell *et al.*, 2010; Calabrese *et al.*, 2012; Kelly *et al.*, 2012; Beck *et al.*, 2014; Deaton *et al.*, 2016).

En plantas, la distribución de las marcas H2AK121ub y H3K27me3 difiere considerablemente de la encontrada en animales (Zhou *et al.*, 2017). H2AK121ub se localizan preferentemente en la zona inmediatamente después del TSS, mientras que las marcas H3K27me3 se localizan a lo largo del cuerpo del gen expandiéndose también a la región promotora (Turck *et al.*, 2007; Zhang y Reese, 2007; Bouyer *et al.*, 2011; Zhou *et al.*, 2017). La distinta distribución de estas marcas en animales y plantas podría indicar un efecto distinto en la organización de nucleosomas. En cualquier caso, hay trabajos que muestran que las proteínas PcG en plantas también son capaces de inhibir la actividad de complejos remodeladores de nucleosomas y de compactar la cromatina *in vitro* (Calonje *et al.*, 2008; Beh, Colwell y Francis, 2012; Calonje, 2014). Asimismo, se ha visto en *Arabidopsis* que los niveles de H3K27me3 están positivamente correlacionados con la inaccesibilidad del DNA a la digestión por endonucleasas (Shu *et al.*, 2012).

En conjunto, nuestros resultados indican que los nucleosomas aguas abajo del TSS en los genes diana PcG se organizan en secuencias irregulares presentando además un posicionamiento pobre de nucleosomas, aunque como se discutirá a continuación, se aprecian diferencias importantes dependiendo de la presencia de una u otra marca. Esta organización difiere de la encontrada en los genes non-PcG que se disponen en secuencias más regulares con un buen posicionamiento de nucleosomas en el caso de los activados y un posicionamiento pobre en el de los reprimidos.

Estos datos apuntan a un papel de las marcas H2AK121ub en determinar el posicionamiento del nucleosoma +1. Este nucleosoma en plántulas WT se encuentra desplazado unos 8 pares de bases aguas abajo respecto al de los genes reprimidos no marcados. Este posicionamiento característico se pierde en el mutante *atbmi1a/b/c* en los genes que experimentan una pérdida sustancial de H2AK121ub. El hecho de que en este mutante el posicionamiento del nucleosoma +1 en estos genes sea similar al de los genes activados a pesar de que no presenten una activación transcripcional, indica que es un efecto independiente de la maquinaria de transcripción. Esto sugiere una posible interacción de estas marcas con una actividad remodeladora de nucleosomas. El reconocimiento de esta modificación por un complejo remodelador podría hacer que el nucleosoma +1 se desplace a una posición que bloquee o desfavorezca la transcripción. En relación con esto, se ha visto que la monoubiquitinación de H2A (H2AK119ub) en animales es reconocida por un componente del complejo remodelador de cromatina RSF (REMODELING AND SPACING FACTOR) llamado REMODELING AND SPACING FACTOR 1 (RSF1) (Zhang *et al.*, 2017). Este complejo tiene dos subunidades, un homólogo de Imitation Switch (ISWI) SNF2H, y RSF1 (Loyola *et al.*, 2003). La unión de RSF1 a H2AK119ub parece ser necesaria para la represión de los genes diana de H2AK119ub y para el mantenimiento de un patrón de nucleosomas (Zhang *et al.*, 2017). Por otro lado, se ha

visto que la monoubiquitinación de la H2A interfiere con el proceso de elongación por parte de la RNA Polimerasa II (Stock *et al.*, 2007) y con el reclutamiento de la chaperona FACT (Zhou *et al.*, 2008). FACT es necesaria para el proceso de elongación, pero también para superar la barrera nucleosomal en la transición entre la fase de iniciación y elongación (Jimeno-Gonzalez *et al.*, 2006), por lo que podría ser que el complejo remodelador reconozca los nucleosomas marcados con H2Aub y modifique su posicionamiento de tal forma que interfiera con el reclutamiento de FACT. En *Arabidopsis*, hasta ahora no se ha identificado ningún factor remodelador que reconozca las marcas de H2AK121ub; sin embargo, existen distintos complejos remodeladores capaces de desplazar los nucleosomas tanto de la familia de ISWI como de SWI/SNF y Chromodomain Helicase DNA-Binding (CHD) (Han *et al.*, 2015; Ojolo *et al.*, 2018) que podrían llevar a cabo esta función.

La formación de secuencias regulares de nucleosomas puede entenderse como la interacción de dos tipos de actividades de remodelación de cromatina. Una actividad que posiciona el nucleosoma +1 mediante la lectura directa de la secuencia de DNA por parte del remodelador o en cooperación con otros factores que actúan en *trans*, y una segunda actividad de espaciado que genera secuencias de nucleosomas regulares. La combinación de posicionamiento y espaciado conduce a la aparición de una secuencia de nucleosomas regular a partir del nucleosoma +1. Estas secuencias regulares pueden estar o no bien posicionadas (Baldi, 2019), lo que coincide con lo que encontramos en el caso de los genes non-PcG activados y reprimidos, que presentan una secuencia regular de nucleosomas bien y mal posicionados, respectivamente. Sin embargo, en los genes diana PcG se observa un posicionamiento irregular de los nucleosomas en la región aguas abajo del TSS, esto está especialmente acentuado en los genes marcados con H3K27me3. Se ha propuesto que el espaciamiento de nucleosomas determina la formación de estructuras de orden superior en la cromatina (Routh, Sandin y Rhodes, 2008; Grigoryev, 2012), lo que encajaría con un papel de estas marcas en el mantenimiento de una conformación de la cromatina cerrada. Esta hipótesis es consistente con la localización a nivel genómico de estas marcas y la accesibilidad al DNA localmente reducida asociada con las regiones marcadas con H3K27me3 en *Arabidopsis* (Shu *et al.*, 2012).

Por otro lado, encontramos también que las marcas PcG afectan de forma diferente a la organización de nucleosomas en la región aguas arriba del TSS. Mientras que los genes only-H2AK121ub presentan una densidad de nucleosomas similar a la de los non-PcG reprimidos en esta región, los genes marcados con H3K27me3 presentan una ocupación mayor y una menor accesibilidad a la digestión por MNasa, lo que sugiere que las marcas H3K27me3 juegan un papel en el establecimiento de una cromatina menos

accesible también en la región promotora, lo que concuerda con el hecho de que estas marcas se extienden aguas arriba del TSS y las H2AK121ub no (Zhou *et al.*, 2017).

Hasta la fecha, no se ha descrito ningún complejo remodelador de nucleosomas en animales o plantas que reconozca las marcas H3K27me₃; sin embargo, existen numerosos ejemplos en *Arabidopsis* de factores remodeladores que actúan de forma antagonista a estas marcas (Han *et al.*, 2015; Ojolo *et al.*, 2018). Por ejemplo, la pérdida de actividad de las ATPasas BRAHMA (BRM) o SPLAYED (SYD) conlleva a la ganancia de marcas H3K27me₃ en diversos genes (Li *et al.*, 2015). Además, se ha visto que tanto BRM como SYD se necesitan para activar la expresión de genes marcados con H3K27me₃ (Wu *et al.*, 2012). Se ha propuesto que el reclutamiento de BRM a los genes diana tiene lugar a través de la interacción con la desmetilasa de H3K27 RELATIVE OF EARLY FLOWERING 6 (REF6) (Li *et al.*, 2016), lo que conlleva a antagonizar la represión mediada por PRC2 (Xu *et al.*, 2016). Estos datos sugieren que la organización de la cromatina inducida por las marcas H3K27me₃ es refractaria a la actividad de complejos remodeladores de nucleosomas implicados en la activación transcripcional. En concordancia con esto, se ha propuesto que la histona H1.2 reconoce las marcas H3K27me₃ en animales (Kim *et al.*, 2015). La histona de unión H1 ha sido relacionada con una compactación de cromatina en estructuras de orden superior y con el silenciamiento de los genes, por tanto, un mecanismo similar puede explicar la reducida accesibilidad de los genes marcados con H3K27me₃.

Por otro lado, se ha comprobado que el factor remodelador de cromatina de tipo CHD3 PICKLE (PKL) es capaz de desplazar nucleosomas *in vitro* (Ho *et al.*, 2013). PKL se encuentra asociado a genes enriquecidos en H3K27me₃ (Zhang *et al.*, 2012) y diversos trabajos apoyan que PKL tiene un papel en promover la represión transcripcional de los genes marcados con H3K27me₃ (Perruc, Kinoshita y Lopez-Molina, 2007; H. Zhang *et al.*, 2008; Ho *et al.*, 2013; Xu *et al.*, 2016). Además, se ha visto que los niveles de H3K27me₃ se reducen considerablemente a nivel genómico en los mutantes *pkl* (Carter *et al.*, 2018) aunque solo la expresión de algunos genes marcados con H3K27me₃ parece estar afectada en estos mutantes. Sin embargo, PKL también ha sido detectado en genes altamente expresados como *ACTIN 7* o *UBIQUITIN 10* (Aichinger *et al.*, 2009, 2011; Zhang *et al.*, 2012; Jing, Guo y Lin, 2019a), y se ha visto que antagoniza la represión PcG en algunos genes (Jing, Guo y Lin, 2019a), por lo que puede ser que actúe de forma diferente de acuerdo a distintos factores, como se ha visto que ocurre con miembros de la misma familia en otros organismos (Ho *et al.*, 2013). Una posibilidad interesante sería que PKL, cuando actúa en combinación con las marcas H3K27me₃, contribuya a mantener un espaciado de nucleosomas favorable para lograr un mayor grado de compactación.

Nuestros datos también apoyan que la incorporación de H2AK121ub debe estar sometida a una regulación dinámica, lo que permite una represión más flexible de los genes diana. Esto encaja con el hecho de que los genes only-H2AK121ub están relacionados con procesos metabólicos y de respuesta a estímulos, cuya expresión no es extraño que esté regulada de forma diferente en las distintas células. Por el contrario, la regulación establecida por la combinación de marcas H2AK121ub y H3K27me3 es mucho más estable y dirigida a reprimir programas génicos en todos aquellos tejidos en los que no sean requeridos. Esto está de acuerdo con la baja proporción de genes con perfil de activación encontrada entre los genes H2AK121ub/H3K27me3, y con que estos genes correspondan a genes específicos de tejido. El hecho de que para la represión de los genes H2AK121ub/H3K27me3 se requiera primero la incorporación de H2AK121ub sugiere que estas marcas se necesitan para mediar una represión inicial y que la incorporación de las marcas H3K27me3 es necesaria para mantener una represión estable. Esto está en concordancia con el tipo de efecto que tiene cada una de las marcas en la organización de nucleosomas y el perfil con características mezcladas en la organización de nucleosomas que presentan los genes que llevan las dos marcas.

Por otro lado, encontramos que los genes only-H3K27me3 muestran una mayor ocupación de nucleosomas que los genes H2AK121ub/H3K27me3. Cabe destacar que la incorporación de las marcas en estos genes es independiente de H2AK121ub (Yang *et al.*, 2013; Merini y Calonje, 2015; Zhou *et al.*, 2017). Contrariamente a lo que pasa con los genes que portan las dos modificaciones, prácticamente la totalidad de los genes only-H3K27me3 están reprimidos en plántulas de 7 DAG, lo que sugiere que la represión de estos genes es estable y generalizada. Entre los genes only-H3K27me3 hay muchos genes homeóticos implicados en el desarrollo de órganos florales. Estos genes deben permanecer apagados en todos los tejidos hasta el momento en el que se activan en el meristemo floral, por lo que no pasan de un estado activo a uno reprimido, como ocurre con la mayoría de los genes regulados por la combinación de H2AK121ub y H3K27me3 (e.j. los genes de maduración de semilla, *FLC*, *MIR156*). Aunque no se conoce el mecanismo de reclutamiento de las marcas H3K27me3 a estos genes, parece que, en su represión, juega un papel importante EMF1 ya que en el mutante *emf1* su expresión se activa de forma ectópica (Sung *et al.*, 1992; Yang, Chen y Sung, 1995; Aubert *et al.*, 2001; Calonje *et al.*, 2008; Kim, Zhu y Sung, 2010; Kim *et al.*, 2012). Curiosamente, EMF1 tiene la capacidad de compactar cromatina *in vitro* (Calonje *et al.*, 2008; Beh, Colwell y Francis, 2012; Calonje, 2014), por tanto, el efecto combinado de EMF1 y H3K27me3 podría ser la causa de la menor accesibilidad a la digestión por MNasa de estos genes.

CAPÍTULO III

Mecanismo de reclutamiento de la actividad Polycomb en plantas

CAPÍTULO III: Mecanismo de reclutamiento de la actividad Polycomb en plantas.

1. Introducción

La incorporación de modificaciones post-traduccionales de histonas produce una alteración de la organización estructural de la cromatina y por tanto juega un papel importante en la regulación de la expresión génica. De acuerdo con esto, distintos tipos de modificaciones de histonas se correlacionan con regiones transcripcionalmente activas y otras con regiones reprimidas (Bell et al., 2011).

Las proteínas PcG mediante la incorporación de modificaciones en dos histonas, la H3 y la H2A, mantienen la represión de los genes que no se requieren en una etapa específica del desarrollo. Dado que los complejos PcG por sí no tienen una capacidad inherente para reconocer regiones o secuencias específicas en el genoma, una cuestión clave que todavía queda por resolver es determinar cómo la maquinaria PcG se dirige a sus distintas dianas en las diferentes fases del desarrollo. En *Drosophila*, que es donde el reclutamiento de los complejos PcG ha sido más extensivamente estudiado, se ha visto que las regiones enriquecidas en marcas PcG están asociadas a unas secuencias llamadas Elementos de Respuesta Polycomb o PRE (del inglés *Polycomb responsive Elements*) (Müller y Kassis, 2006). Se ha demostrado que existen distintas proteínas de unión a DNA que reconocen motivos específicos dentro de las secuencias PRE y que actúan de forma cooperativa en el reclutamiento de los complejos PcG (Brown y Kassis, 2013). Los PREs de *Drosophila* son regiones que van desde cientos de pares de bases a algunos kilopares de bases (Ringrose y Paro, 2007; Buzas, 2017). Los sitios de unión de los factores Pleiohomeotic (Pho) y Pho-like son elementos frecuentes de los PREs de *Drosophila*, pero también contienen sitios de unión de otras proteínas como los factores GAGA, Pipsqueak y Zeste (Okulski et al., 2011). A pesar de esto, estudios a nivel genómico indican que los sitios de unión de estos factores solapan solo parcialmente con las diana de PRC1 y PRC2, por lo que no se sabe hasta qué punto contribuyen al reclutamiento (Simon y Kingston, 2009; Deng et al., 2013).

En animales, se han identificado algunas secuencias que actúan como PRE, por ejemplo, se ha demostrado que la región promotora de 3 kb del gen *MafB* de ratón y una región de 1,8 kb situada entre los genes *HOXD11* y *HOXD12* llamada D11.12 en humanos, son necesarias para una represión dependiente de PcG en sistemas de expresión utilizando genes reporteros (Sing et al., 2009; Woo et al., 2010). Ambos elementos contienen sitios de unión del factor de transcripción YY1 (el homólogo de Pho en animales), lo que sugiere que existe una cierta conservación de los mecanismos

de reclutamiento PcG entre animales y *Drosophila*. Sin embargo, en animales existen otros mecanismos de reclutamiento de las proteínas PcG a los promotores de los genes diana. Se ha visto que secuencias enriquecidas en dinucleótidos CpG (islas CpG) no metilados solapan con las regiones marcadas con H3K27me3 y que hay una alta correlación entre las dianas de PRC2 y las islas CpG (Ku *et al.*, 2008; Mendenhall *et al.*, 2010). Además, se ha demostrado que el factor Polycomb-Like 3 (PCL3) favorece la unión del complejo PRC2 a las islas CpG (Hunkapiller *et al.*, 2012). Las proteínas PCL reconocen específicamente las islas CpG no-metiladas mediante sus dominios EH_{WH} y reclutan al PRC2 a estas regiones (Li *et al.*, 2017). Más aún, las proteínas Ten-eleven translocation 1 (Tet1) que median la desmetilación de las islas CpG facilitan el reclutamiento del PRC2 (Wu *et al.*, 2011). Por otro lado, la variante del complejo PRC1 en animales PRC1.1 tiene como componente la desmetilasa de la histona H3K36 KDM2B que contiene un dominio CXXG el cual tiene una alta afinidad por las islas CpG no-metiladas (Lagarou *et al.*, 2008; Tavares *et al.*, 2012; Blackledge *et al.*, 2014; Gil y O’Loghlen, 2014). Las proteínas KDM2B reclutan al complejo PRC1.1 a las islas CpG no-metiladas para que medie la ubiquitinación de H2AK119 de forma independiente de PRC2 (Farcas *et al.*, 2012; He, Huang y Xu, 2013), por lo que parece que las islas CpG participan tanto en el reclutamiento de PRC2 como de PRC1. Además, estos resultados sugieren que el estado de metilación de las islas CpG debe modular la ocupación de los complejos PcG.

Por otro lado, varios RNA de cadena larga no codificantes (lncRNA) han sido relacionados con la función PRC2 en animales (Jeffrey A Simon y Kingston, 2013; Mozgova, Köhler y Hennig, 2015). Por ejemplo, se ha demostrado que el lncRNA Xist interacciona directamente con la subunidad EZH2 del PRC2 a través de la región RepA/Xist (región de repeticiones A) para iniciar la inactivación del cromosoma X (Zhao *et al.*, 2008). Por otro lado, estudios recientes muestran que el silenciamiento del cromosoma X mediado por Xist se inicia por la incorporación de las marcas H2AK119ub por la variante del complejo PCGF3/5-PRC1. Esto resulta en el reclutamiento de otras variantes de PRC1 y PRC2, lo que permite la deposición de las marcas H3K27me3 en el cromosoma (Almeida *et al.*, 2017). Parece que es la región XN de Xist la que está implicada en el reclutamiento de los complejos PRC1; sin embargo, todavía no se conoce el mecanismo (Almeida *et al.*, 2017). Otro lncRNA relacionado con PcG es el RNA intergénico *HOTAIR* que interacciona con el PRC2 a través de su dominio 5' y lo recluta a los genes *HOXD*. *HOTAIR* se requiere para la trimetilación de H3K27 y el silenciamiento de los genes *HOXD* (Rinn *et al.*, 2007; Tsai *et al.*, 2010; Deng *et al.*, 2013).

En *Arabidopsis*, diversos estudios han relacionado determinadas secuencias de DNA con el reclutamiento PcG. Los genes diana de PRC2 se encuentran enriquecidos en los motivos *GAGA* y los motivos *TELOBOX* (Xiao *et al.*, 2017; Zhou *et al.*, 2018). Los motivos *GAGA* son reconocidos por los factores de transcripción BARLEY B RECOMBINANT/BASIC PENTACYSSTEINE (BBR/BPC) a través de su dominio BPC (Hecker *et al.*, 2015). Se ha propuesto que el factor de transcripción BPC6 se une a los motivos *GAGA* y recluta al PRC2 mediante su interacción con LHP1. De acuerdo con esto, se ha encontrado un enriquecimiento de los motivos *GAGA* en la región promotora de los genes diana de LHP1, lo que coincide también con las regiones marcadas con H3K27me3 (Hecker *et al.*, 2015; Molitor *et al.*, 2016). Además, los factores de transcripción BPC1 y ARABIDOPSIS ZINC FINGER1 (AZF1), reclutan al PRC2 a través de su interacción con FIE. BPC1 se une a los motivos *GAGA*, y AZF1 parece tener afinidad por los motivos *TELOBOX* (Deng *et al.*, 2013; Xiao *et al.*, 2017), ambos factores colaborarían en el reclutamiento de PRC2 (Xiao *et al.*, 2017). Por otro lado, un estudio reciente propone que los factores Telomere-Repeat-Binding Factors (TRB), que también reconocen los motivos *TELOBOX*, reclutan al PRC2 mediante su interacción con los componentes CLF y SWN. Los TRB parecen ser esenciales para la deposición de las marcas H3K27me3 en un grupo de genes diana PcG (Zhou *et al.*, 2018).

Además, se han identificado en *Arabidopsis* secuencias que actúan de forma similar a los PREs (PRE-like) para el reclutamiento de PRC2 a genes específicos (Berger *et al.*, 2011; Liu *et al.*, 2011; Deng *et al.*, 2013). El primer PRE-like caracterizado, fue el elemento RLE (Repressive LEC2 Element) de 50 pb localizado en la región promotora del regulador del desarrollo de semilla *LEAFY COTYLEDON2 (LEC2)*. RLE juega un papel importante en el reclutamiento del PRC2 y el establecimiento de las marcas H3K27me3 para la represión de *LEC2* en los tejidos vegetativos (Berger *et al.*, 2011). Se ha visto que la incorporación del elemento RLE en un promotor que no es diana PcG es suficiente para promover la incorporación de las marcas H3K27me3 y regular la actividad del gen (Berger *et al.*, 2011). El promotor de *LEC2*, además, contiene un motivo *GAGA* localizado cerca del elemento RLE, que debe contribuir al reclutamiento de PRC2 (Berger *et al.*, 2011).

Se ha visto también que la unión de algunos factores de transcripción a sus sitios de reconocimiento conduce a la incorporación de H3K27me3. Por ejemplo, el complejo heterodimérico ASYMMETRIC LEAVES (AS) se une a los promotores de los genes *BREVIPEDICELLUS (BP)* y *KNAT2* e interacciona físicamente con el PRC2 reclutándolo para mantener la represión de estos genes durante la diferenciación de los órganos laterales. (Lodha, Marco y Timmermans, 2013). También los genes *WUSCHEL (WUS)* y *KNUCKLES (KNU)* contienen sitios de unión de AG (motivos

CC(A/T)₆GG o *CArG*) que son necesarios para el establecimiento y la propagación de las marcas H3K27me3 mediadas por PRC2 (Sun *et al.*, 2009, 2014; Liu *et al.*, 2011). AG reprime *WUS* en el meristemo floral (Liu *et al.*, 2011). La mutación de los dos motivos *CArG* en la región situada a 1 kb aguas abajo del sitio de inicio de transcripción de *WUS* resulta en la activación de la expresión de *WUS*. Por otro lado, la mutación de AG resulta en una disminución de las marcas H3K27me3 y la pérdida de la unión de LHP1, por lo que los motivos *CArG* actúan como elementos reclutadores de PRC2 en la represión de *WUS* (Liu *et al.* 2011). AG también reprime *WUS* de forma indirecta mediante la activación de *KNU* que a su vez reprime *WUS* (Sun *et al.*, 2009). Se ha caracterizado una secuencia de 153 pb en la región promotora de *KNU* (*KNU*-PRE) suficiente para el reclutamiento de las proteínas PcG FIE y EMF2, y para la incorporación de las marcas H3K27me3. Esta secuencia contiene también tres sitios de unión de AG, por lo que se ha propuesto que AG activa *KNU* en este caso actuando de forma antagonista a las proteínas PcG (Sun *et al.*, 2014). Además, se ha encontrado una implicación directa de los factores de transcripción VAL1/2/3 en el reclutamiento de la maquinaria PcG, particularmente en el de PRC1 a los genes de maduración de semilla (Yang *et al.*, 2013; Merini *et al.*, 2017). VAL1/2/3 poseen un dominio B3 de unión a DNA específico de plantas que reconoce los elementos *Sph/Ry* (5'-CATGCA-3'/5'-TGCATG-3') (Suzuki, 1997; Ho *et al.*, 2009; Yuan *et al.*, 2016; N. Chen *et al.*, 2018). El análisis estructural del dominio B3 de VAL1 indica que se une al motivo *RY* mediante la formación de puentes de hidrógeno con las seis bases del motivo y que la sustitución de una única base impide la formación del complejo (Sasnauskas, Kauneckaitė y Siksnyš, 2018). Las proteínas VAL1 reconocen y se unen también a dos motivos *RY* en la región de nucleación del gen *FLC*. VAL1 interacciona con AtBMI1A y también con SAP18 que es un componente del complejo Apoptosis and Splicing-Associated Proteín (ASAP) que a su vez interacciona con la desacetilasa de histonas HISTONE DEACETYLASE 19 (HDA19). Esto permite el reclutamiento del PHD-PRC2 para que incorpore H3K27me3 (Qüesta *et al.*, 2016; Yuan *et al.*, 2016). Otro gen regulado por las proteínas VAL es *AGAMOUS-LIKE15* (*AGL15*). El promotor de *AGL15* contiene elementos *RY* que son reconocidos por las proteínas VAL que reclutan el PRC2 mediante la interacción directa con el componente MULTICOPY SUPPRESSOR OF IRA1 (*MSI1*) para el establecimiento de H3K27me3 (N. Chen *et al.*, 2018). Recientemente, también se ha implicado a VAL1 en el reclutamiento de LHP1 y MSI1 al gen *FT* (Jing, Guo y Lin, 2019a).

Las proteínas VAL, además del dominio B3 de unión a DNA, contienen tres otros dominios: el dominio PHD-like (PHD-L) y el dominio CW, ambos implicados en el reconocimiento de modificaciones de histonas, y el dominio de represión transcripcional EAR (Suzuki y McCarty, 2008; Jing, Guo y Lin, 2019b). El dominio CW de VAL1 se ha visto que se une directamente a H3K4me3 (Hoppmann *et al.*, 2011) y que interacciona con

HDA19 (Zhou *et al.*, 2013). Por otro lado, el dominio PHD-L parece que está implicado en promover la incorporación de H3K27me3 (Veerappan *et al.*, 2012, 2014).

En cuanto al motivo EAR, varios estudios han puesto de manifiesto que participa en la represión transcripcional (Kagale y Rozwadowski, 2011; Causier *et al.*, 2012; Ma *et al.*, 2017; Wu y Citovsky, 2017; Choi, Seo y Cho, 2018). Las proteínas que contienen un motivo EAR pueden llevar a cabo la represión de dos modos distintos: 1) actuando como factores de transcripción (TF) al contener también dominios de unión a DNA específicos o 2) actuando como reguladores transcripcionales (TRs) participando en la represión mediante una interacción indirecta con los genes diana. El motivo EAR interacciona con co-supresores como SAP18 y TOPLESS (TPL) que a su vez reclutan a desacetilasas de histonas (HDAC) (Kagale y Rozwadowski, 2011). Por ejemplo, los miembros de la familia Ethylene Response Factor (ERF) AtERF3 y AtERF4 se unen a AtSAP18 y reclutan AtHDA19 para formar un complejo HDAC para reprimir la expresión génica (Song y Galbraith, 2006). Un ejemplo de las proteínas EAR que actúan como reguladores transcripcionales son las proteínas JASMONATE ZIM domain (JAZ), que reprimen diferentes activadores transcripcionales claves en las respuestas a jasmonato, como MYC2, MYC3, MYC4 y MYC5 (Chini *et al.*, 2007; Kagale, Links y Rozwadowski, 2010; Goossens *et al.*, 2015; Thireault *et al.*, 2015; Song *et al.*, 2017; Yang *et al.*, 2018).

Otro factor que interacciona con las proteínas PcG es la desmetilasa de H3K4me3 PKDM7B/JMJ14. Se ha demostrado que MJ14 interacciona *in vivo* con los componentes EMF1 y AtBMI1 de PRC1 para la represión de *FT* (Wang *et al.*, 2014). También, se ha visto que modificaciones post-traduccionales de la cromatina pueden participar en el reclutamiento del PRC2. Se ha propuesto que los componentes LHP1, AtRING1 y AtBMI1 se reclutan a los genes de desarrollo de semilla mediante la interacción con las proteínas ALFIN-LIKE 1-7 (ALs) formando un complejo PRC1-like para la represión de estos genes. Las proteínas ALs poseen un dominio PHD que reconoce las marcas de activación H3K4me3 y el dominio PAL que interacciona con el dominio RAWUL de AtBMI1B y AtRING1A. El reclutamiento de las proteínas PcG mediado por las marcas H3K4me3 ha sido propuesto como un mecanismo para cambiar de un estado de activación transcripcional asociado a H3K4me3 a un estado reprimido mediado por las marcas H3K27me3 (Molitor *et al.*, 2014).

Aparte de todos estos factores, estudios sobre el mecanismo del silenciamiento de *FLC* inducido por vernalización han relacionado los lncRNAs con el reclutamiento de las proteínas PcG a la región de nucleación de *FLC*. La represión de *FLC* se mantiene de forma dependiente de PRC2 después de la vernalización (Gendall *et al.*, 2001). Se han identificado dos lncRNA que se transcriben en el mismo sentido que *FLC*, *COLDAIR* que

se origina del primer intrón y *COLDWRAP* que se origina en el promotor. Estos lncRNAs parece que interactúan físicamente con el PRC2 y lo reclutan al *locus FLC* durante la vernalización (Kim y Sung, 2017; Kim, Xi y Sung, 2017). También se ha demostrado que la inducción de la transcripción del lncRNA antisentido *COOLAIR* está relacionada con la represión de *FLC* independiente de PcG durante la primera fase de la vernalización (Swiezewski *et al.*, 2009) y que cuando aumenta la exposición al frío, *COOLAIR* se requiere para la coordinación del aumento de las marcas H3K27me3 y de la disminución de las marcas H3K36me3 en la región de nucleación de *FLC* (Yang, Howard y Dean, 2014).

Colectivamente, estos datos sugieren que existen diferentes factores que participan en el reclutamiento de las proteínas PcG en plantas, algunos parecen actuar de forma general, como son los factores que se unen a los motivos *GAGA* y *TELOBOX* o los que se unen a modificaciones de histonas, y otros de forma específica participando en el reclutamiento selectivo de las proteínas PcG a determinados genes en determinados momentos, como parece ser el caso de factores de transcripción o lncRNA.

2. Objetivo

Nuestro objetivo es entender cómo se establece la represión PcG y de qué depende. Nos planteamos determinar si se necesita la acción de un factor reclutador para que se inicie la represión o si la unión de un único componente PRC1 o PRC2 es capaz de recapitular la maquinaria de represión, además de la implicación de los distintos componentes en el mecanismo. Para ello, por un lado nos propusimos: 1) llevar a cabo la identificación a nivel genómico de posibles motivos reguladores en *cis* que se requieren para el reclutamiento de los complejos PcG a sus genes diana, y 2) desarrollar un sistema que nos permita reclutar de forma independiente bien a los factores que reconozcan las secuencias *cis* identificadas, o a los diferentes componentes PcG a un promotor sintético que esté libre de la complejidad y secuencias reguladoras presentes en los promotores de los genes diana PcG. Este promotor fusionado a un gen reportero nos permitirá determinar el impacto que tienen las distintas proteínas en la incorporación de las marcas y en la represión transcripcional.

3. Resultados

3.1. Las zonas de decaimiento de las marcas H2AK121ub y H3K27me3 se localizan mayoritariamente en regiones promotoras

Mediante análisis de ChIP-seq, se llevó a cabo en el laboratorio la localización a nivel genómico de las marcas PcG en WT de 7 DAG (Zhou *et al.*, 2017). Este análisis desveló que las marcas H2AK121ub se encuentran distribuidas a lo largo de todo el genoma, co-localizando en un gran número de genes con las marcas H3K27me3. Sin embargo, también se encontró que hay un conjunto de genes únicamente marcados con H2AK121ub o con H3K27me3. El análisis comparativo de estas marcas en distintos mutantes PcG puso de manifiesto que la incorporación de H2AK121ub es necesaria para que el PRC2 medie la trimetilación de H3K27 en aquellos genes que presentan las dos marcas (Zhou *et al.*, 2017). De acuerdo con esto, los niveles de ambas marcas estaban considerablemente disminuidos en los mutantes *atbmi1a/b/c* (Fig. 38a). Sin embargo, el grado de disminución de las marcas no era igual en todos los genes, encontrándose genes en los que las marcas desaparecían casi en su totalidad y otros que presentaban una disminución localizada de ellas (Fig. 38b).

Dado que varios estudios han relacionado a distintos factores de transcripción con el reclutamiento de los complejos PcG (Deng *et al.*, 2013; Xiao *et al.*, 2017; Zhou *et al.* 2018; Merini *et al.* 2017), decidimos investigar a nivel genómico los elementos reguladores en *cis* que se encuentran en las regiones reguladoras de los genes que presentaban las dos marcas. Para ello, nos centramos en los genes marcados con las dos marcas en WT y que mostraban pérdida de ambas marcas en los mutantes *atbmi1a/b/c*. Como el estudio utilizando todas las regiones que experimentan pérdida de marcas en el mutante puede ser poco informativo al ser marcas que se extienden a lo largo de los genes, decidimos centrarnos en aquellas regiones en las que se producía un decaimiento localizado de las dos marcas. Estas regiones más sensibles a la pérdida de la función PRC1 podrían ser buenas candidatas a contener elementos reguladores reconocidos por los factores implicados en el reclutamiento de los complejos PcG (*hot spots*) (Fig. 38b). Aunque también cabe la posibilidad de que estas regiones contengan secuencias reguladoras reconocidas por factores antagonistas cuya accesibilidad quede reducida por el efecto de las marcas.

Para identificar estas regiones a nivel genómico, en primer lugar, seleccionamos las regiones donde co-localizan las dos marcas H2AK121ub y H3K27me3 en WT. Para ello, después de alinear las lecturas cortas utilizando la herramienta *Bowtie* (Langmead *et al.*, 2009), identificamos la localización de los picos de cada una de las marcas H2AK121ub y

H3K27me3 mediante la llamada de picos utilizando la herramienta *software MACS2* (Y. Zhang *et al.*, 2008). A continuación, intersectamos los picos de H2AK121ub y H3K27me3 utilizando la función *intersectBed* de la herramienta *BedTools*, obteniendo así las regiones donde coinciden las dos marcas.

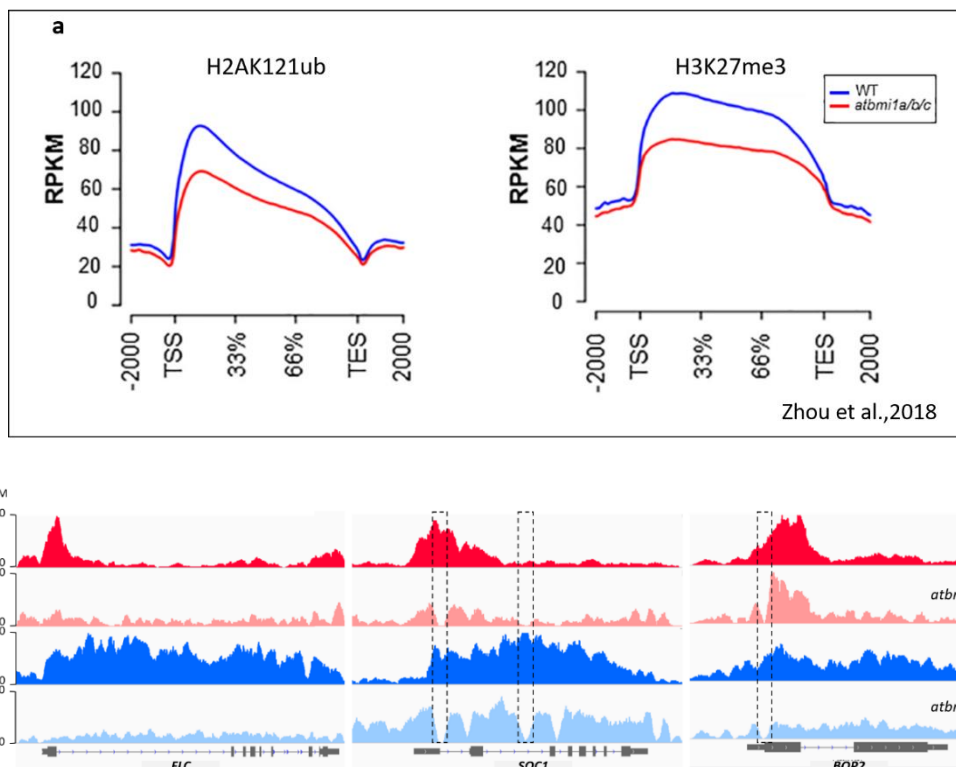


Figura 38: Las marcas PcG disminuyen de forma global en el mutante *atbmi1a/b/c*. (a) Metagenes plot de la distribución de las marcas H2AK121ub en todos los genes marcados (panel izquierdo). Metagenes plot de la distribución de las marcas H3K27me3 en todos los genes marcados (panel derecho). (b) Visualización IGV de la ocupación de las marcas H2AK121ub y H3K27me3 en genes seleccionados: *FLC* un gen que pierde las marcas de forma completa en el mutante *atbmi1 a/b/c*, *SOC1* y *BOP2* genes que pierden las marcas en zonas localizadas. El nombre y la estructura del gen se representan debajo de cada panel. Las líneas discontinuas muestran las zonas localizadas de decaimiento de las marcas H2AK121ub y H3K27me3.

Para identificar las regiones que presentaban una pérdida de marcas de forma localizada, lo primero que hicimos fue definir estas regiones (regiones de decaimiento) como aquellas con niveles altos de marcas H2AK121ub y H3K27me3 en el WT que presentan una región de pérdida drástica flanqueada por dos regiones sin pérdida o con una pérdida menor en el mutante *atbmi1a/b/c* (Fig. 39). En primer lugar, alineamos las lecturas cortas obtenidas en los dos genotipos WT y *atbmi1a/b/c* al genoma de referencia utilizando la herramienta *Bowtie* y a continuación calculamos el número de lecturas por cada base (o *coverage*) y lo normalizamos por el factor de escala, lecturas por kilobase por millón de lecturas mapeadas o RPKM (del inglés *Reads Per Kilobase per Million*

mapped reads) utilizando la herramienta *bamCoverage* (<https://deeptools.readthedocs.io/en/develop/content/tools/bamCoverage.html>).

Para identificar las regiones de decaimiento de las marcas a nivel genómico impusimos ciertas condiciones. Por un lado, los niveles de H2AK121ub y H3K27me3 en el mutante (medidos en función del número de lecturas (RPKM)) debían de ser iguales o inferiores a 100 RPKM, ya que si utilizábamos un umbral de corte inferior (20, 40, 60 o 80 RPKM) se perdían muchas zonas de decaimiento parcial, es decir, regiones con pérdida sustancial pero no drástica. El mutante, además, en estas regiones debía presentar unos niveles al menos 4 veces inferiores que el WT ($FC \leq 4$). Por último, las regiones de pérdida de marcas en el mutante debían estar flanqueadas (50 pb a un lado y a otro del centro del decaimiento) por regiones con niveles al menos 5 veces superiores ($FC \geq 5$) a los presentados por el centro de la zona de decaimiento (Fig. 39a). Esta última condición era crucial para discriminar las zonas de decaimiento localizado de las de pérdida generalizada.

Nuestro análisis detectó 3243 regiones del genoma con un decaimiento significativo de las dos marcas PcG de acuerdo con nuestros criterios en el mutante *atbmi1a/b/c*. La representación del perfil de las marcas H2AK121ub y H3K27me3 en estas regiones en un intervalo de 200 pb alrededor del centro de la región (Fig. 39b) mostraba un claro decaimiento de las marcas en el mutante *atbmi1a/b/c* en comparación con el WT.

A continuación, investigamos la localización de las regiones identificadas en el genoma (Fig. 39c) utilizando el paquete de R ChipSeeker (Yu, Wang y He, 2015), considerando como región promotora, aquella secuencia de DNA de 1000 pb que se encuentra aguas abajo del TSS. Las regiones intergénicas son aquellas que se encuentran entre el sitio de terminación de la transcripción (TES) de un gen y el promotor del gen siguiente. La región del cuerpo del gen incluye todos los intrones y los exones excepto el primer exón. Encontramos que las regiones de decaimiento estaban enriquecidas mayoritariamente en las regiones promotoras y las regiones intergénicas en comparación con otras regiones anotadas del genoma de *Arabidopsis* (Fig. 39c). Puesto que las regiones intergénicas son más amplias y gran parte todavía no se conoce su papel en la regulación génica, centramos el estudio en aquellas regiones detectadas en los promotores.

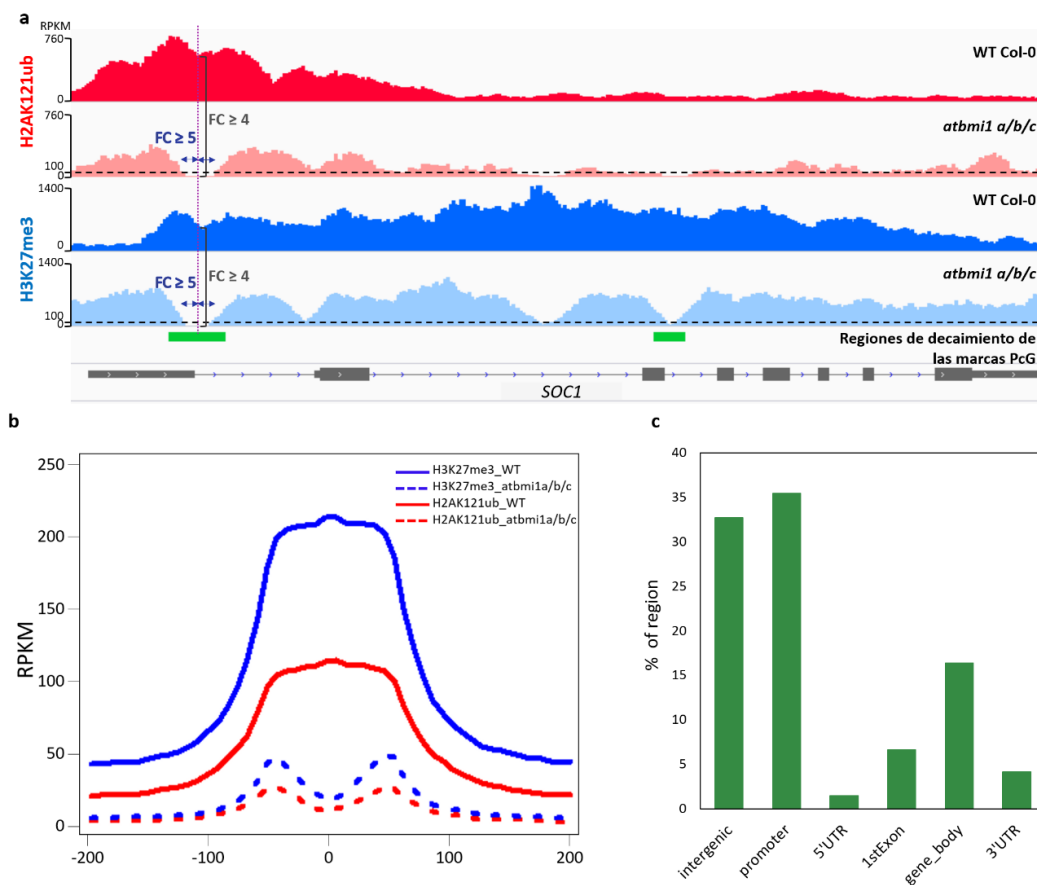


Figura 39: (a) Visualización IGV de la ocupación de las marcas H2AK121ub y H3K27me3 en WT y en el mutante *atbmi1a/b/c* en el gen *SOC1*. Se indican las condiciones para detectar las regiones de decaimiento de las marcas: La línea continua negra indica el umbral de nivel de marca en el mutante *atbmi1a/b/c* (100 RPKM). El nivel de marca en WT es superior a *atbmi1a/b/c* con un $FC \geq 4$. El nivel de marca en las regiones 50 pb alrededor del centro de la región es superior a aquel en el centro de la región con un $FC \geq 5$. Las cajas verdes muestran las regiones de decaimiento de las marcas detectadas. **(b)** Nivel medio de las marcas H2K121ub y H3K27me3 en WT y *atbmi1a/b/c* en las regiones de decaimiento de las marcas H2AK121ub y H3K27me3. El eje x representa la distancia al centro de la región y el eje y representa el nivel de la señal por número de lecturas normalizadas por RPKM (Reads per Kilobase per Million mapped reads). **(c)** Porcentaje de regiones con decaimiento localizado de las marcas H2AK121ub y H3K27me3 en las regiones génicas anotadas e intergénicas en el genoma de *A. thaliana*.

3.2. Identificación de los motivos 6-mer enriquecidos en las regiones del promotor que presentan un decaimiento de las marcas PcG

A continuación, pasamos a investigar las posibles secuencias reguladoras en *cis* contenidas en las regiones de decaimiento detectadas en los promotores. Para ello, realizamos un análisis computacional *de novo* del enriquecimiento de los motivos de DNA de 6 pb (6-mer) presentes. Este análisis consistió en calcular para cada motivo posible de 6-mer (4096 motivos) su frecuencia de aparición en las regiones de decaimiento en comparación con su frecuencia en estas regiones cuando fueron distribuidas de forma aleatoria a lo largo de todas las regiones promotoras del genoma (se han generado $10E^6$ regiones aleatorias). Los motivos enriquecidos de forma significativa fueron seleccionados en base a un p-valor $< 0,001$ (el p-valor se calculó sumando el número de veces que la frecuencia del motivo en las regiones aleatorias era mayor que la frecuencia en las regiones de decaimiento, dividido por el número de las generaciones aleatorias ($10E^6$)) y un enriquecimiento mayor de 1,2 (se entiende por enriquecimiento es el número de veces de más que aparece el motivo en las regiones de decaimiento, comparado con el número de veces que aparece en las regiones aleatorias).

Identificamos 115 motivos enriquecidos de forma significativa en las regiones de decaimiento de las marcas H2AK121ub y H3K27me3 en los promotores (Fig. 40a). Seguidamente llevamos a cabo un análisis utilizando la función *TOMTOM* de la herramienta *software MEME-Suite* (Bailey *et al.*, 2009) para determinar si estos motivos estaban asociados a secuencias reconocidas por factores de transcripción. Pudimos relacionar estos motivos con el sitio de unión de 23 familias de factores de transcripción en *Arabidopsis* (Fig. 40b). Uno de los motivos más enriquecidos en estas regiones era el motivo *RY* (o elementos *LEC2/ABI3/VP1*) "*catgca*" que es el motivo de unión específico de las proteínas VAL1/2/3. Se ha demostrado previamente que estas proteínas intervienen en el reclutamiento del complejo PRC1 (Yang *et al.*, 2013; Merini *et al.*, 2017), lo que apoyaba nuestro método de búsqueda de posibles factores reclutadores.

Nos planteamos por tanto qué podían tener en común todos estos factores de transcripción para poder mediar el reclutamiento de los complejos PcG. En términos generales a nivel de secuencia encontramos que eran todos muy dispares, sin embargo, en todas estas familias de factores de transcripción encontramos algunos miembros que al igual que las proteínas VAL contenían uno o varios motivos de represión EAR (Kagale, Links y Rozwadowski, 2010) (Fig. 40b). De hecho, algunos de estos factores han sido ya relacionados con el reclutamiento de complejos PcG, como por ejemplo AS2. (Lodha, Marco y Timmermans, 2013), FLC (Kim y Sung, 2017; Kim, Xi y Sung, 2017), SVP (Cui y Benfey, 2009; Liu *et al.*, 2009; Li y Luan, 2011) o SUPERMAN (SUP) (Xu *et al.*, 2018). Por tanto, aunque se ha propuesto que el motivo EAR interacciona con los co-supresores

SAP18 y/o TOPLESS (TPL) y que estos por su parte reclutan a desacetilasas de histonas (HDAC) cuya actividad conlleva a la represión transcripcional (Song *et al.*, 2005; Song y Galbraith, 2006; Hill, Wang y Perry, 2008; Causier *et al.*, 2012), podría ser posible que el motivo EAR también mediara el reclutamiento de los complejos PcG de forma directa o indirecta.

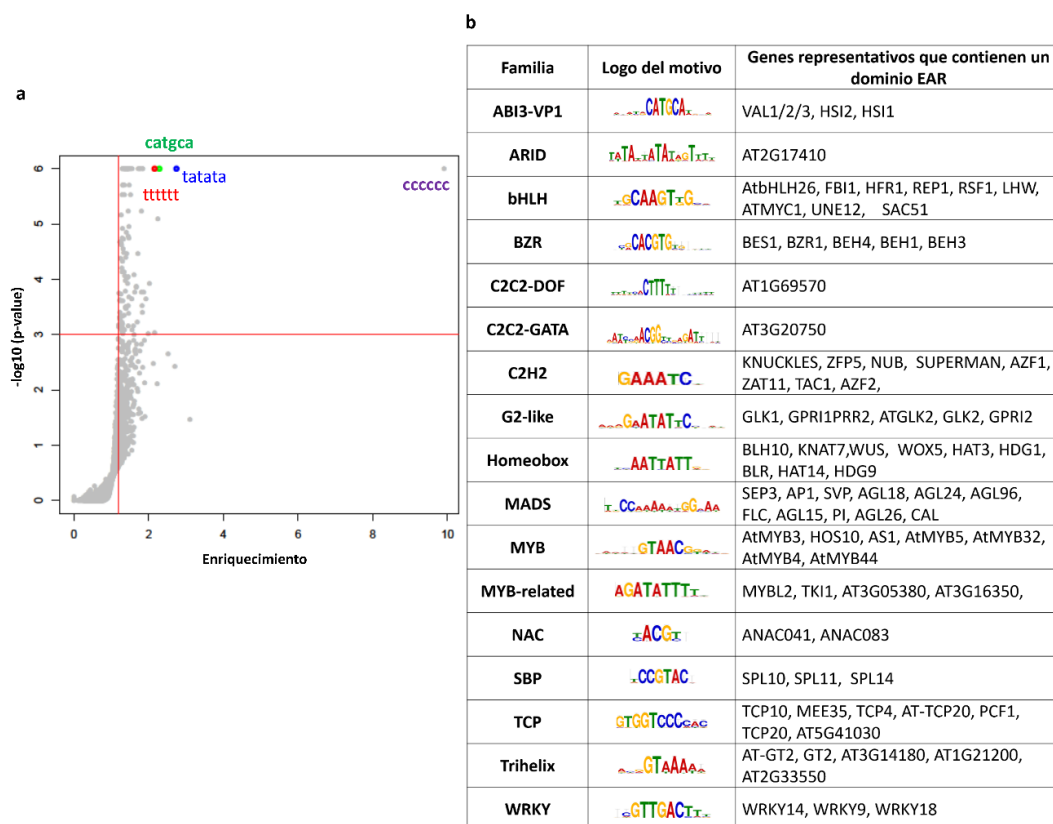


Figura 40: Enriquecimiento de los motivos 6-mer en las regiones de decaimiento de las marcas PcG. (a) Representación de los motivos identificados en base a su enriquecimiento y el p-valor. El eje x representa el enriquecimiento y el eje y representa $-\log_{10}(\text{p-valor})$. Los motivos que están enriquecidos de forma significativa son aquellos que tienen un enriquecimiento mayor de 1,4 y un p-valor $< 0,001$ ($-\log_{10}(\text{p-valor}) = 3$). Las líneas rojas delimitan los umbrales del enriquecimiento = 1,2 y $-\log_{10}(\text{p-valor}) = 3$. **(b)** Motivos que aparecen de forma significativa en las regiones de decaimiento de las marcas PcG y las familias de factores de transcripción a los que se asocian. Se indican algunos ejemplos de los miembros de estas familias que contienen un motivo EAR en su secuencia de acuerdo con los datos del estudio (Kagale, Links y Rozwadowski, 2010).

3.3. Recapitulación sintética de los componentes de la maquinaria PcG

Para investigar a fondo el papel de los diferentes factores implicados en la represión mediada por las proteínas PcG y entender mejor la secuencia de eventos moleculares, hemos desarrollado un sistema para poder reclutar diferentes componentes a un promotor sintético y estudiar su impacto en el establecimiento de las marcas H2AK121ub y H3K27me3 y en la represión transcripcional. Este sistema consiste en una construcción que permite la expresión en *Arabidopsis* de la proteína de interés fusionada al dominio de unión de DNA de la proteína bacteriana LexA (LexABD) y otra, introducida en la misma planta, en la que la expresión del gen reportero *GUS* está controlada por el promotor *CaMV35s* al que se le ha introducido la secuencia de reconocimiento de la proteína LexA (LexA operador; LexAO). Este sistema permite determinar el efecto de la proteína de interés sobre el gen *GUS* tras unirse al LexAO con la ventaja de que puede ser usado en un fondo WT al no existir el problema de que se una la proteína endógena. Además, al ser un promotor constitutivo, tampoco hay posibilidad de que haya interferencia por la unión de otros factores que actúen como represores al ser un promotor constitutivo. Por otro lado, el hecho de que sean construcciones estables en *Arabidopsis* permite que todos los componentes de la maquinaria PcG estén disponibles y puedan incorporarse en el sistema en el caso que se den las condiciones apropiadas para ello.

En primer lugar, utilizamos este sistema para estudiar el efecto que tienen los componentes PRC1, AtBMI1A y AtRING1B, los cuales están implicados en la incorporación de H2AK121ub (Bratzel *et al.*, 2010; Chen *et al.*, 2010; Yang *et al.*, 2013), y VAL1, que ha sido propuesto como reclutador tanto de PRC1 como de PRC2 (Suzuki y McCarty, 2008; Guerriero *et al.*, 2009; Bratzel *et al.*, 2010; Yang *et al.*, 2013), y cuyo motivo de unión hemos encontrado enriquecido en las regiones de decaimiento de marcas en los mutantes *atbmi1a/b/c*. Para ello, hemos generado plantas transgénicas *pLexAO-35S:GUS/pG10-90:LexABD-AtBMI1A* (*LexAO/BD-AtBMI1A*), *pLexAO-35S:GUS/pG10-90:LexABD-AtRING1B* (*LexAO/BD-AtRING1B*) y *pLexAO-35S:GUS/pG10-90:LexABD-VAL1* (*LexAO/BD-VAL1*) en fondo WT. En primer lugar se generaron plantas transgénicas en las que se introdujeron las construcciones *pG10-90:LexABD-AtBMI1A*, *pG10-90:LexABD-AtRING1B* y *pG10-90:LexABD-VAL1* y se seleccionaron, mediante análisis de Western blot sobre extractos totales de proteínas usando el anticuerpo anti-LexABD, las líneas que expresaban las proteínas de fusión (Fig. 41). Estas líneas se cruzaron con la misma línea *pLexAO-35S:GUS* (*LexAO:GUS*) tras comprobar por análisis de segregación que contenía una única inserción del transgen. El hecho de usar la misma línea *LexAO:GUS* es importante para evitar posibles variaciones por inserción del transgen en localizaciones genómicas de características distintas.

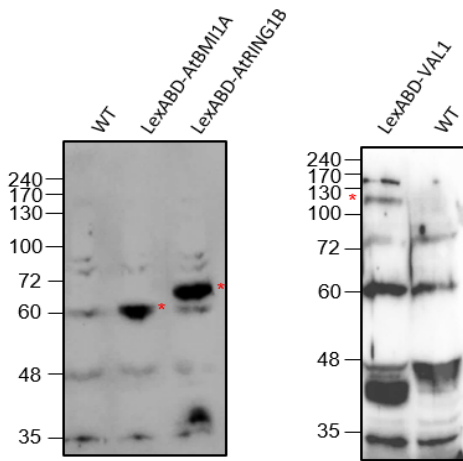


Figura 41: Comprobación de las líneas transgénicas LexABD-AtBMI1A, LexABD-AtRING1B y LexABD-VAL1. Ensayo de western blot utilizando el anticuerpo anti-LexABD. Se han obtenido los tamaños de las proteínas de fusión predichos: LexABD-AtBMI1A (60 KDa), LexABD-AtRING1A (65 KDa) y LexABD-VAL1 (100 KDa).

A continuación, determinamos por ChIP-qPCR usando el anticuerpo anti-LexABD si las proteínas de fusión eran capaces de unirse al *LexAO* (Fig. 42a). Para ello, tras la inmunoprecipitación del DNA analizamos por qPCR un posible enriquecimiento en distintas regiones contenidas tanto en el promotor como en la secuencia codificante de *GUS*. Como control negativo de la unión utilizamos la línea *LexAO:GUS* que no contiene ninguna proteína de fusión. Este ensayo mostró un enriquecimiento especialmente significativo de las tres proteínas en la región correspondiente al *LexAO*.

Una vez verificado que las proteínas de fusión se unían al promotor sintético, evaluamos el efecto que tenían en la expresión del gen reportero *GUS*. Para ello, cuantificamos la actividad *GUS* en 15-20 plántulas T2 de 7 DAG de edad de las distintas líneas (Fig. 42b). Se observó una disminución de la actividad *GUS* en todos los casos respecto a la línea *LexAO:GUS* control, aunque en las plantas *LexAO/BD-VAL1* la disminución era considerablemente mayor. Mientras que las plantas *LexAO/BD-AtBMI1A* y *LexAO/BD-AtRING1B* mostraron alrededor de un 60-80% de la actividad *GUS* detectada en las plantas control, las plantas *LexAO/BD-VAL1* mostraron una actividad en torno a un 40%, indicando que la unión de VAL1 es capaz de mediar de forma mucho más efectiva la represión del gen.

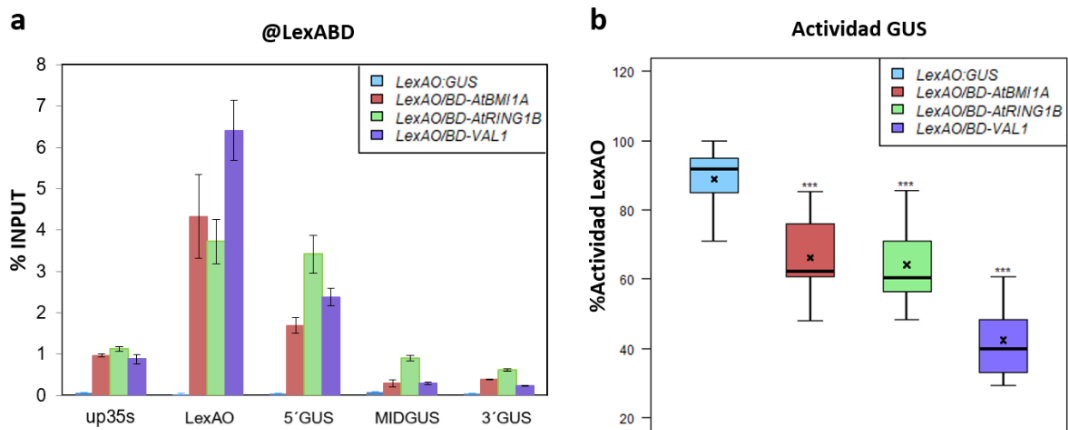


Figura 42: Unión de las proteínas de fusión LexABD-PcG a LexAO y su efecto en la expresión del reportero GUS (a) Análisis de CHIP-qPCR utilizando el anticuerpo anti-LexABD, de las líneas *LexAO:GUS*, *LexAO/BD-AtBMI1A*, *LexAO/BD-AtRING1B* y *LexAO/BD-VAL1*. **(b)** Diagrama de cajas y bigotes que muestra el porcentaje de la actividad de GUS en las diferentes líneas en relación al control negativo *LexAO:GUS*, que no contiene proteína de fusión. Los valores fueron normalizados al mayor valor de actividad obtenido en el control negativo que se considera como 100% de actividad. Se utilizaron 15 plantas de cada línea transgénica. En cada caja, se representa la mediana (segmento dentro del rectángulo), la media (cruz dentro del rectángulo), cuartiles superior e inferior (cajas) y los valores máximo y mínimo (bigotes). Las diferencias significativas se determinaron mediante el test de Wilcoxon y se indican con 3 asteriscos (p-valor < 0,001).

Seguidamente, estudiamos el impacto que tiene la unión de estos componentes en la deposición y/o eliminación de las modificaciones de histonas. Para ello, analizamos los niveles de las marcas represoras H2AK121ub y H3K27me3 y la marca de activación H3ac a lo largo de la construcción *pLexAO-35S:GUS* en las distintas plantas transgénicas (Fig. 43). Los resultados mostraron que la unión de la proteína VAL1 al *LexAO* conduce a la incorporación de niveles altos de H2AK121ub y H3K27me3 y a una considerable disminución de los niveles de H3ac en comparación con los niveles detectados en las plantas control (Fig. 43). Sin embargo, el reclutamiento de las proteínas AtBMI1A y AtRING1B conllevaba a un ligero aumento de H2AK121ub y H3K27me3 y a una ligera disminución de los niveles de H3ac, lo que concordaba con los niveles de actividad GUS detectados en las distintas plantas.

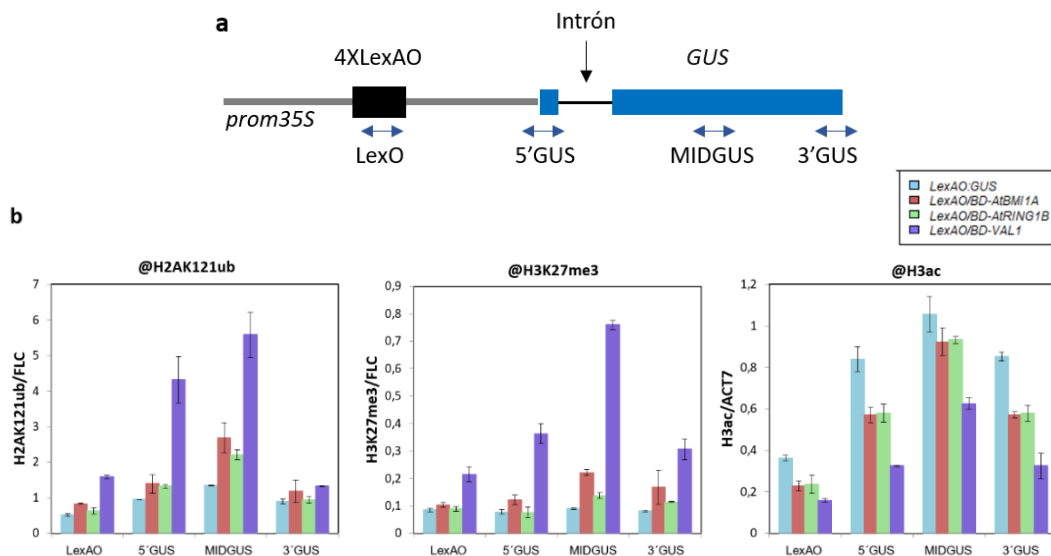


Figura 43: Análisis ChIP-qPCR de los niveles de H2AK121ub, H3K27me3 y H3ac en las líneas *LexAO:GUS*, *LexAO/BD-AtBMI1A*, *LexAO/BD-AtRING1B* y *LexAO/BD-VAL1* (a) Estructura del locus reportero *pLexAO-35S:GUS* y la localización de las regiones amplificadas por ChIP-qPCR. (b) Niveles de H2AK121ub, H3K27me3 y acetilación de H3 (H3ac) a lo largo del locus reportero. En el caso de H2AK121ub y H3K27me3, los resultados fueron normalizados al control positivo *FLC*, y en el caso de H3ac, los resultados fueron normalizados a *ACT7*. Las barras de error indican la desviación estándar entre dos réplicas biológicas. Las diferencias con *pLexAO-35S:GUS* fueron obtenidas mediante la prueba *t* de Student.

Varios estudios han relacionado las secuencias *GAGA* y *TELOBOX* con la represión mediada por la maquinaria PcG en *Arabidopsis*, en concreto, se ha propuesto que los factores que reconocen estos motivos participan en el reclutamiento de PRC2 y son necesarios para la incorporación de H3K27me3 (Deng *et al.*, 2013; Hecker *et al.*, 2015; Molitor *et al.*, 2016; Xiao *et al.*, 2017; Zhou *et al.*, 2018). Para investigar el efecto que pueden ejercer estas secuencias en la incorporación de las marcas mediadas por los complejos PcG, incluimos en el promotor sintético, tanto aguas arriba del *LexAO* como aguas abajo una secuencia de 110 pb de la región reguladora del gen *ABI3* que contiene un motivo *GAGA* y dos motivos *TELOBOX* (Fig. 44a). Esta secuencia está incluida en un fragmento de DNA de 600 pb que ha sido previamente relacionado con el reclutamiento del complejo PRC2 (Xiao *et al.*, 2017). Esta secuencia podría tener una actividad PRE, por lo que, para simplificar, la llamaremos de aquí en adelante PRE. La inserción del fragmento se llevó a cabo en ambas posiciones para evaluar si su posicionamiento jugaba un papel importante en la represión, de tal forma que las dos nuevas plantas transgénicas en fondo WT contenían las construcciones *pPRE-LexAO-35S:GUS* (*PRE-LexAO:GUS*) y *pLexAO-PRE-35S:GUS* (*LexAO-PRE:GUS*), respectivamente. Evaluamos si el simple hecho de incluir estas secuencias en el promotor sintético tenía algún efecto en la incorporación de H3K27me3 o en la expresión de *GUS* (Fig. 44b y c). No encontramos diferencias significativas en cuanto al nivel de H3K27me3 entre las tres

líneas transgénicas *LexAO:GUS*, *PRE-LexAO:GUS* y *LexAO-PRE:GUS* (Fig. 44b), y las variaciones de la actividad GUS no fueron grandes (Fig. 44c), por lo que parece que la incorporación de la secuencia PRE aguas arriba o aguas abajo del LexAO por sí sola no conlleva a cambios significativos.

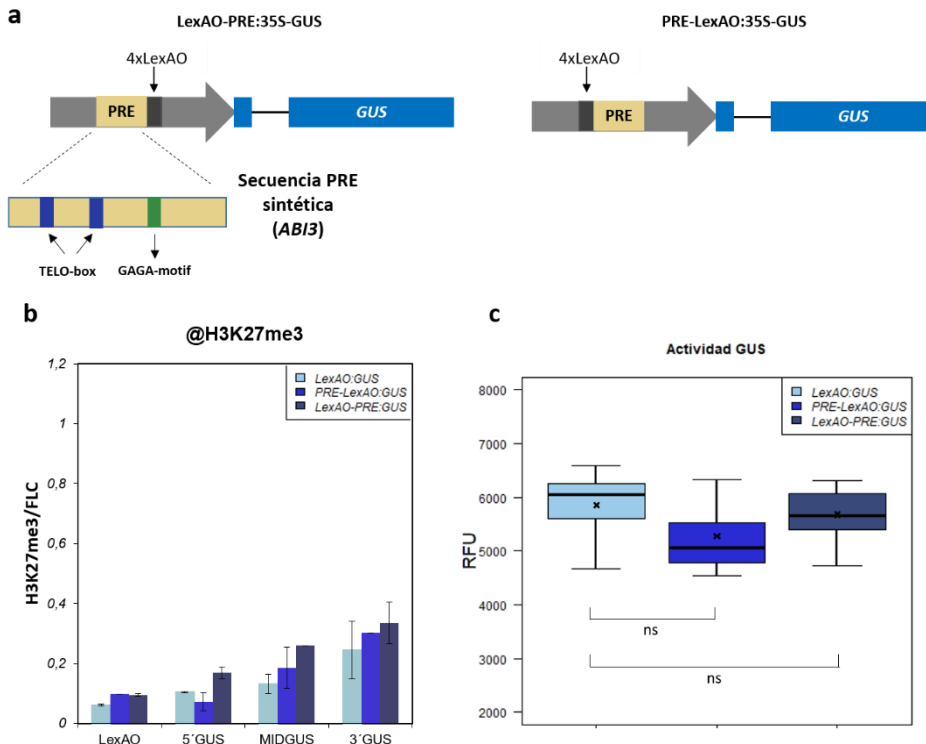


Figura 44: Los niveles de H3K27me3 y la expresión del reportero GUS no se afectan por la presencia de la secuencia PRE. (a) Esquema representativo de la construcción del promotor sintético conteniendo la secuencia PRE *LexAO-PRE:GUS* y *PRE-LexAO:GUS*. (b) Niveles de H3K27me3 a lo largo de *pLexAO-35S:GUS*. Los resultados fueron normalizados al control positivo *FLC*. Las barras de error indican la desviación estándar entre dos réplicas biológicas. Las diferencias con *pLexAO-35S:GUS* fueron obtenidas mediante la prueba *t de Student*. (c) Diagrama de cajas y bigotes que muestra la actividad de GUS. Se utilizaron 15 plantas T1 en cada caso. En cada caja, se representa la mediana (segmento dentro del rectángulo), la media (cruz dentro del rectángulo), cuartiles superior e inferior (cajas) y los valores máximo y mínimo (bigotes). Las diferencias significativas se determinaron mediante el test de Wilcoxon. ns: no hay diferencias significativas

Para estudiar la influencia de la secuencia PRE en la incorporación de las marcas H3K27me3 o en la expresión de GUS, cruzamos las nuevas líneas con las plantas *BD-AtBMI1A*, *BD-AtRING1B* y *BD-VAL1* y llevamos a cabo los correspondientes análisis en las transgénicas dobles. Encontramos que la incorporación del PRE tanto en una posición como en otra, llevaba a un aumento de los niveles de H3K27me3 significativo tras la unión de las tres proteínas de fusión al LexAO (Fig. 45a, b); sin embargo, no se observaron

grandes cambios en cuanto a la actividad GUS, lo que sugiere que el aumento de los niveles de H3K27me3 tras la unión de LexABD-AtBMI1A y LexABD-AtRING1B no conlleva a priori a una represión más efectiva de GUS al no haber una disminución de la actividad (Fig. 45c).

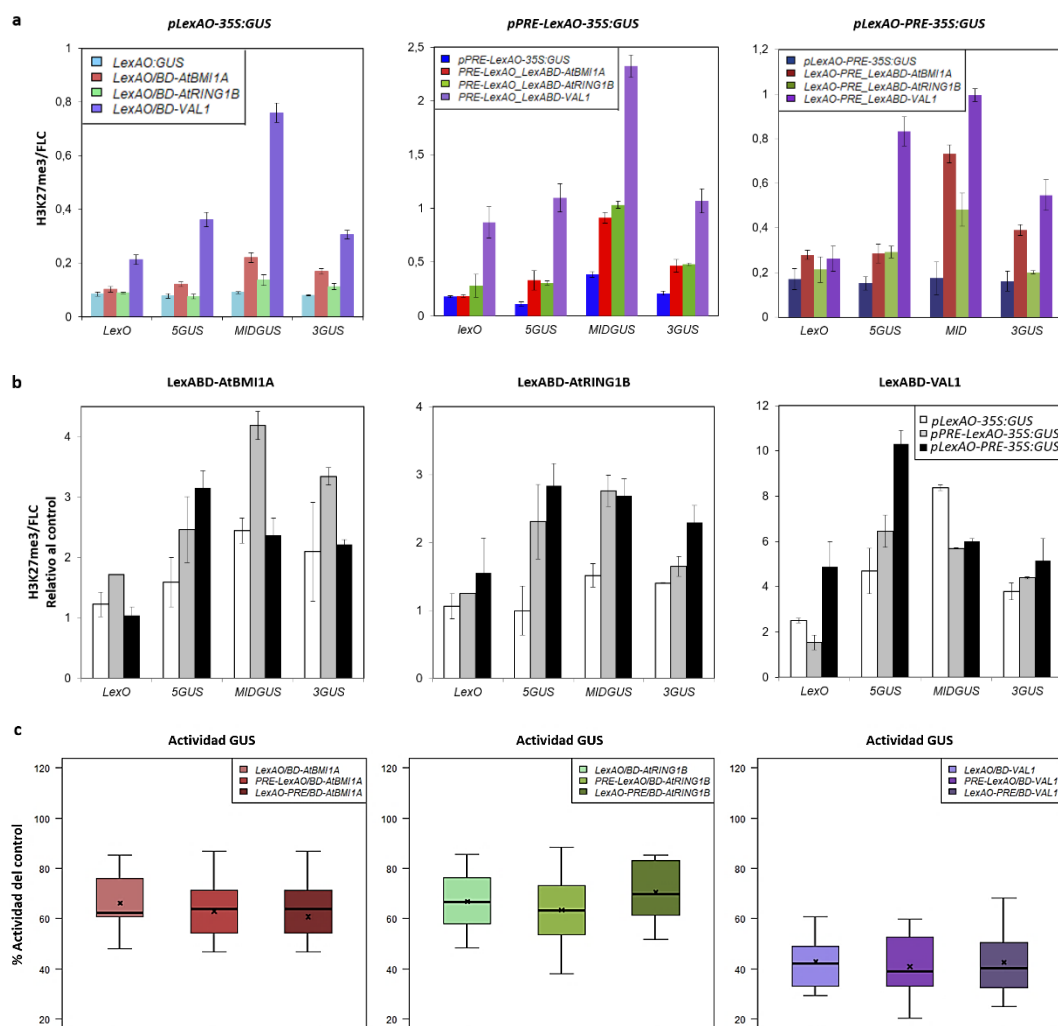


Figura 45: Efecto de las proteínas LexABD-PcG en los niveles de H3K27me3 y expresión de GUS en presencia de la secuencia PRE. (a) Niveles de H3K27me3 a lo largo del *locus* reportero en las líneas que expresan las proteínas de fusión LexABD-AtBMI1A, AtRING1B o VAL1 con los distintos promotores. Los resultados fueron normalizados al control positivo *FLC*. **(b)** Comparación de los niveles de la marca H3K27me3 referidos a los valores obtenidos en los controles sin proteína de fusión. **(c)** Diagrama de cajas y bigotes que muestra la actividad GUS en las diferentes líneas. Se utilizaron 15 plantas de cada línea transgénica. En cada caja, se representa la mediana (segmento dentro del rectángulo), la media (cruz dentro del rectángulo), cuartiles superior e inferior (cajas) y los valores máximo y mínimo (bigotes). Las diferencias significativas se determinaron mediante el test de Wilcoxon.

Encontramos que cuando la secuencia PRE se localiza aguas abajo del sitio de unión *LexAO* los niveles de H3K27me3 eran más altos en la región 5'GUS que en la región MIDGUS, lo que puede ser debido al aumento de la longitud de la región entre *LexAO* y el TSS en comparación con las construcciones *LexAO:GUS* y *PRE-LexAO:GUS* (Fig. 44a).

A continuación, investigamos el posible aumento de la incorporación de las marcas H2AK121ub en las líneas *LexAO-PRE/BD-AtBMI1A* y *LexAO-PRE/BD-AtRING1B* debido a la presencia de la secuencia PRE (Fig. 46). No observamos un aumento significativo de estas marcas en ninguno de los casos en las plantas en presencia o ausencia del PRE. Es cierto que observamos un cierto aumento en las regiones *LexAO* y 5'GUS tras la unión de *AtRING1B* en presencia del PRE, pero este efecto probablemente era debido al incremento de la longitud de la región entre *LexAO* y el TSS, en cualquier caso, a pesar del incremento, los niveles se mantenían bajos. Por lo tanto, la secuencia PRE no parece tener un efecto importante en la incorporación de las marcas H2AK121ub.

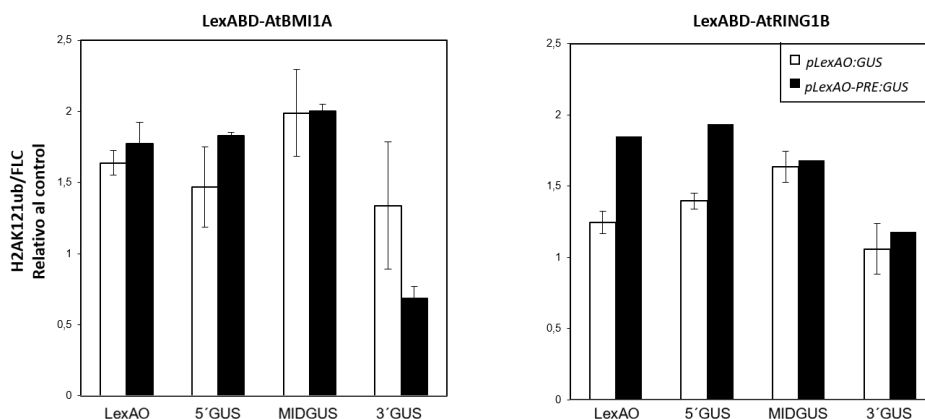


Figura 46: Efecto de las proteínas LexABD-AtBMI1 y AtRING1B en los niveles de H2AK121ub en presencia de la secuencia PRE. Los valores son referidos a los valores obtenidos en los controles *LexAO:GUS* y *LexAO-PRE:GUS*.

3.4. La expresión del dominio EAR fusionado a LexABD conduce a la eliminación de la acetilación de H3

EL dominio EAR ha sido identificado como un dominio activo de represión en varios factores de transcripción (Kagale y Rozwadowski, 2011). Puesto que la represión mediada por PcG es dependiente de VAL1, nos preguntamos si el motivo EAR es el responsable de esta represión o alternativamente existen otros dominios implicados. Para ello, hemos generado las líneas transgénicas *LexAO/BD-EAR* en las que el dominio EAR esta fusionado al dominio LexABD. Para hacer que la proteína de fusión se dirija al núcleo,

incluimos una secuencia de localización nuclear (NLS) en la construcción (Fig. 47a). La expresión de la proteína de fusión fue verificada mediante análisis de western blot en extracto total de proteínas de las plantas transgénicas (Fig. 47b). Mediante análisis de Chip-qPCR utilizando el anticuerpo anti-LexABD, comprobamos también la unión de las proteínas de fusión LexABD-EAR al operador LexAO (Fig. 47c). El nivel de unión de LexABD-EAR aparentemente era más bajo al obtenido con otras proteínas, probablemente a consecuencia del tamaño de la proteína de fusión, sin embargo, al analizar la actividad GUS encontramos una disminución significativa comparada con el control *LexAO:GUS* (Fig. 47d), lo que indicaba que el EAR era funcional. Sin embargo, el grado de represión fue inferior al obtenido tras la unión de VAL1, lo que indica que VAL1 tiene otros motivos que deben participar en el reclutamiento de los factores necesarios para la represión.

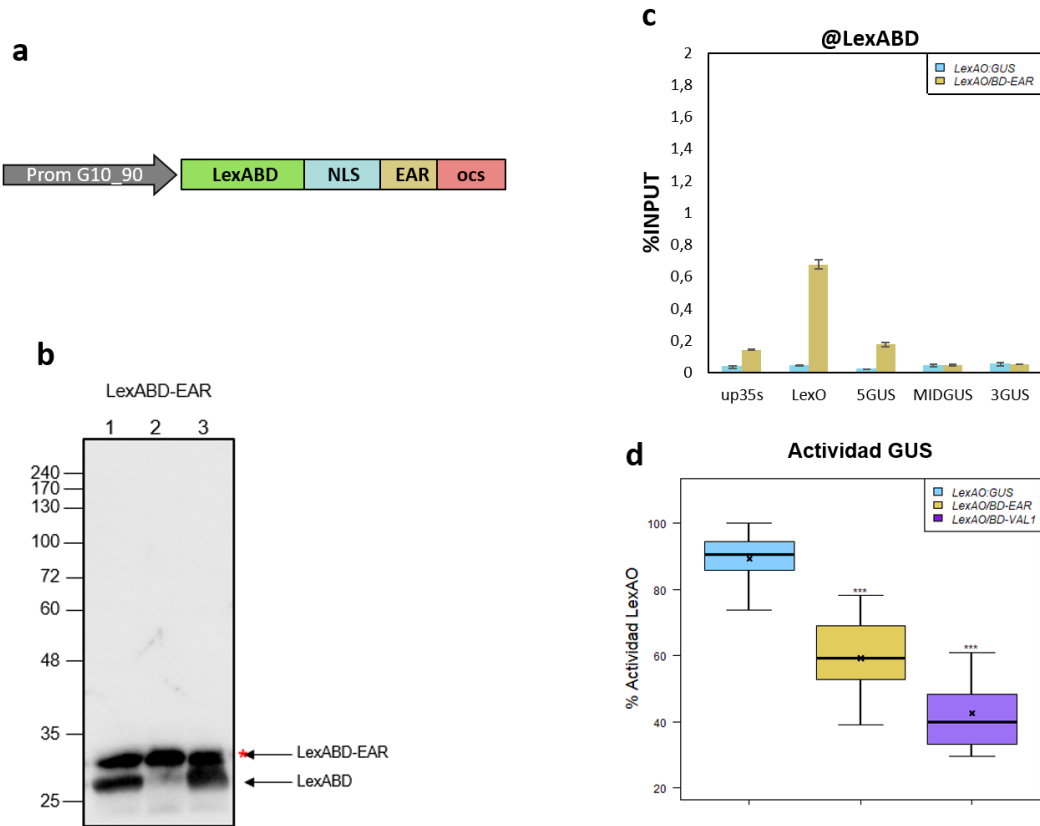


Figura 47: Efecto del motivo EAR en la represión mediada por PcG. (a) Esquema representativo de la construcción *pG10-90:LexABD-EAR*. (b) Western-blot utilizando el anticuerpo anti-LexABD de extracto total de proteínas de las plantas *LexAO/BD-EAR*. (c) Western-blot utilizando el anticuerpo anti-LexABD de extracto total de proteínas de las plantas *LexAO:GUS*, *LexAO/BD-EAR* y *LexAO/BD-VAL1*. Se utilizaron 15 plantas de cada línea transgénica. En cada caja, se representa la mediana (segmento dentro del rectángulo), la media (cruz dentro del rectángulo), cuartiles superior e inferior (cajas) y los valores máximo y mínimo (bigotes). Las diferencias significativas con el control se determinaron mediante el test de Wilcoxon y se representan con asteriscos.

A nivel de incorporación de las marcas epigenéticas, observamos que tras la unión de EAR disminuyen los niveles de la acetilación de H3 y aumentan los niveles de H3K27me3 de forma significativa. Por el contrario, los niveles de las marcas H2AK121ub no varían mucho respecto al control *LexAO:GUS* (Fig. 48).

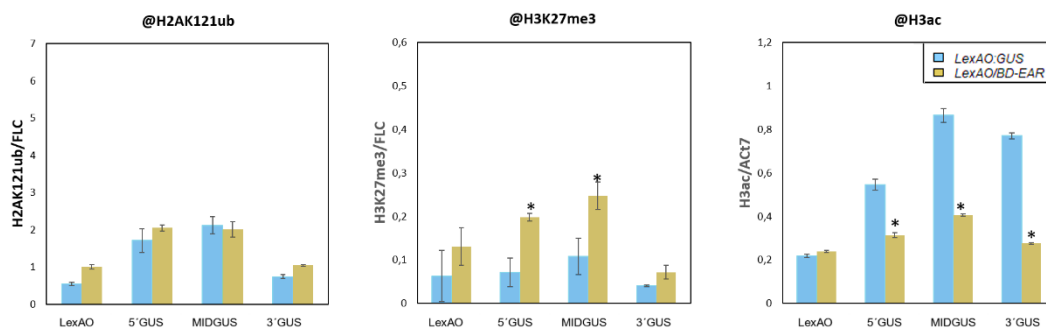


Figura 48: Efecto del motivo EAR en la incorporación de las marcas H2AK121ub, H3K27me3 y acetilación de H3 (H3ac). En el caso de H2AK121ub y H3K27me3, los resultados fueron normalizados al control positivo *FLC*, y en el caso de H3ac, los resultados fueron normalizados a *ACT7*. Las barras de error indican la desviación estándar entre dos réplicas biológicas. Las diferencias con *LexAO:GUS* fueron obtenidas mediante la *prueba t de Student*. Las diferencias significativas con respecto al control se representan con asterisco.

4. Discusión

La represión de un gran número de genes en *Arabidopsis* requiere la actividad conjunta de PRC1 y PRC2; sin embargo, todavía no está claro si la unión y/o actividad de un complejo lleva a la unión del otro. Tampoco está claro cómo se reclutan estos complejos y si la incorporación de las marcas H2AK121ub y H3K27me3 son las únicas responsables de la represión. Muchos trabajos apoyan la participación de factores de transcripción en el reclutamiento de PRC2. Estudios bioquímicos han puesto de manifiesto la pérdida de marcas H3K27me3 en determinados genes en mutantes de pérdida de función de factores como AS2 (Lodha, Marco y Timmermans, 2013), WUS (Sun *et al.*, 2009, 2014; Liu *et al.*, 2011) o las proteínas VAL1/2/3 (Veerappan *et al.*, 2012, 2014; Qüesta *et al.*, 2016; Yuan *et al.*, 2016; N. Chen *et al.*, 2018). Estas últimas también interactúan con los componentes PRC1 (Yang *et al.*, 2013; Qüesta *et al.*, 2016) y son necesarias *in vivo* para la incorporación de H2AK121ub (Yang *et al.*, 2013). Además, la pérdida de función de distintos factores que reconocen las secuencias *GAGA* y *TELOBOX*, las cuales están enriquecidas en los promotores de los genes marcados con H3K27me3, también se ha visto que afecta a los niveles de H3K27me3 (Deng *et al.*, 2013; Hecker *et al.*, 2015; Molitor *et al.*, 2016; Xiao *et al.*, 2017; Zhou *et al.*, 2018).

Para profundizar en el estudio de la implicación de factores de transcripción en el reclutamiento de los complejos PcG, analizamos los posibles sitios de unión de factores de transcripción presentes en los promotores de los genes diana PcG. Para ello nos centramos en las regiones de decaimiento localizado de las marcas H2AK121ub y H3K27me3 en el mutante *atbmi1a/b/c*. Con este análisis se identificaron los sitios de unión propuestos para 23 familias de factores de transcripción. Entre los componentes de estas familias se encuentran muchos de los factores mencionados previamente relacionados con la pérdida de marcas PcG, entre ellos, el sitio de unión de las proteínas VAL, el motivo *RY*. Sin embargo, estos resultados apoyan de forma muy indirecta una posible implicación de estos factores en la represión PcG, ya que cabría también la posibilidad de que estas secuencias no estuviesen relacionadas con el establecimiento de la represión, sino que podría ser que la incorporación de las marcas PcG restringiesen el acceso de los factores que las reconocen.

Para aclarar tanto el papel de los distintos componentes PcG como la posible implicación de factores de transcripción en el reclutamiento de PRC1 y PRC2, desarrollamos un sistema para reclutar de forma específica e independiente a un promotor sintético esta batería de factores. A pesar de que durante esta tesis doctoral se generaron otras construcciones necesarias para responder a estas preguntas (como aquellas para expresar distintos componentes PRC2 y varios factores de transcripción),

por cuestiones de tiempo no se ha podido analizar el efecto de todos estos factores. Por tanto, al no poder disponer de toda la información, algunas de las conclusiones podrían variar una vez llevado a cabo el análisis en su totalidad.

Nuestros datos apoyan que la unión de la proteína VAL1 actúa como plataforma o “*scaffold*” de reclutamiento de los factores necesarios para que se dé una represión eficaz. Esto queda reflejado en los niveles de actividad de GUS, en la disminución de las marcas de activación H3ac y en el aumento de las marcas de represión H2AK121ub y H3K27me3. La disminución de las marcas H3ac está de acuerdo con la capacidad de las proteínas VAL para interactuar con desacetilasas de histonas, entre ellas, HDA9 (Zeng *et al.*, 2019), HDA6 (Chhun *et al.*, 2016) y HDA19 (Zhou *et al.*, 2013). Por otro lado, el aumento de las marcas H2AK121ub confirma que las proteínas VAL pueden interactuar y reclutar a los componentes de PRC1 responsables de la incorporación de estas marcas (Yang *et al.*, 2013), y la incorporación de H3K27me3 confirma los resultados que relacionaban la actividad VAL con PRC2 (Veerappan *et al.*, 2012, 2014; Qüesta *et al.*, 2016; Yuan *et al.*, 2016; N. Chen *et al.*, 2018; Zeng *et al.*, 2019).

Por otro lado, encontramos que el reclutamiento directo de AtBMI1A y AtRING1B al promotor sintético no tiene un impacto alto en la actividad del gen reportero. Esto puede indicar que bien estos factores por sí mismos no pueden reclutar a todos los componentes necesarios para la represión al no darse la secuencia apropiada de acontecimientos, o bien que la incorporación del LexABD afecte a la actividad de estas proteínas. Esto último podría ser el caso de LexABD-AtRING1B; sin embargo, en el caso de LexABD-AtBMI1A el hecho de que se observen ciertos niveles de H2AK121ub tras su unión al promotor sugiere que su actividad no está afectada y que lo más probable es que no se dé la secuencia apropiada para la represión. Esto está apoyado por el hecho de que la unión de AtBMI1A no afecte significativamente a los niveles de acetilación de la histona H3 y que se encuentren niveles apreciables pero muy bajos de H3K27me3, lo que sugiere que las marcas de activación deben de ser eliminadas para que se dé la incorporación apropiada de las marcas PcG.

Se ha visto que en los mutantes *atbmi1a/b/c* los niveles de H3K27me3 están considerablemente reducidos en los genes que en WT presentan las dos marcas (Zhou *et al.*, 2017). Por tanto, los altos niveles de H3K27me3 en el *locus* reportero tras la unión de la proteína VAL1 podría ser una consecuencia indirecta del reclutamiento de PRC1 y/o de la incorporación de marcas H2AK121ub. El hecho de que se detecte algo de H3K27me3 tras la unión de AtBMI1 apoya una inter-relación entre PRC1 y PRC2, pero no descarta que VAL tenga un efecto directo en el reclutamiento de PRC2. Ambos mecanismos pueden actuar de forma sinérgica.

Por otro lado, diversos estudios han mostrado que la pérdida de función de los factores que reconocen las secuencias *GAGA* y *TELOBOX* tiene un efecto en los niveles de H3K27me3 en una serie de genes diana PcG (Deng *et al.*, 2013; Hecker *et al.*, 2015; Molitor *et al.*, 2016; Xiao *et al.*, 2017; Zhou *et al.*, 2018). De acuerdo con esto, la incorporación de un fragmento de DNA con estas secuencias en nuestro promotor sintético ejercía un efecto positivo en la incorporación de H3K27me3 tras la unión tanto de VAL1 como de AtBMI1A y AtRING1B; sin embargo, no tenía efecto si ninguna de estas proteínas era reclutada con anterioridad. Esto indica que los factores BPC1, AZF1 o TRB no son suficientes para reclutar al PRC2, pero sí colaboran en el establecimiento de H3K27me3. Curiosamente, la inserción del PRE no parece tener un efecto en la incorporación de las marcas H2AK121ub puesto que, en general, no se observaba un cambio de los niveles de estas marcas en el *locus* reportero.

Por último, de acuerdo con nuestro análisis de los promotores de los genes diana PcG y con otros datos experimentales, existen otros factores de transcripción que podrían funcionar de forma similar a las proteínas VAL. Curiosamente, a pesar de que la diferencia a nivel de secuencia entre todos estos factores es muy grande, encontramos que, como es el caso de VAL, un gran número de ellos contienen al menos un dominio de represión EAR, lo que sugiere una posible implicación directa o indirecta de este motivo en la represión mediada por PcG. Para investigar esto, llevamos a cabo el reclutamiento de una proteína quimérica que contenía el LexABD y el dominio EAR. Nuestros resultados sugieren una posible contribución de EAR en la represión PcG. Encontramos que esta proteína quimérica era capaz de reprimir en cierta medida la expresión del gen *GUS*. También observamos una reducción de los niveles de H3ac, lo que está de acuerdo con el hecho de que este dominio puede ser reconocido por los co-represores TPL y SAP18 que a su vez reclutan a desacetilasas de histonas (Song *et al.*, 2005; Song y Galbraith, 2006; Hill, Wang y Perry, 2008; Causier *et al.*, 2012). Por otro lado, aunque este dominio no parece tener un efecto en la incorporación de las marcas de represión H2AK121ub, sí que se observaba un aumento de H3K27me3, lo que abre la posibilidad de que exista una conexión entre este dominio y la represión PcG. En cualquier caso, estos datos tienen que confirmarse y ser apoyados por los experimentos con construcciones que contienen otros factores y componentes PcG.

CONCLUSIONES

CONCLUSIONES

1. Nuestros datos apuntan a la existencia de variantes de PRC1 que actúan en combinación con el PRC2 para controlar las distintas transiciones de fase a lo largo del desarrollo de *Arabidopsis* y regular el desarrollo y crecimiento de los órganos. Las distintas variantes PRC1 se asocian con factores específicos para regular la expresión de subgrupos de genes diferentes, de tal forma que la acción combinada de distintas variantes permite que los distintos procesos de desarrollo ocurran en el momento apropiado.
2. Las marcas H2AK121ub juegan un papel en el ajuste de la posición del nucleosoma +1, mientras que las marcas H3K27me3 alteran el espaciamiento de los nucleosomas y reducen la accesibilidad de la cromatina en el promotor y el cuerpo del gen. Estas dos marcas tienen un efecto aditivo cuando se localizan en los mismos genes. La represión mediada solo por las marcas H2AK121ub es más flexible y regulada dinámicamente, mientras que H3K27me3 es necesaria para mantener una represión estable de los genes diana.
3. Las proteínas VAL actúan como plataforma o *scaffold* de reclutamiento no solo de PRC1 y PRC2 para la incorporación de las marcas de represión H2AK121ub y H3K27me3, sino que también de otros factores que eliminan las marcas de activación preexistentes en los genes diana. El reclutamiento de todos estos factores se necesita para que tenga lugar la represión transcripcional.

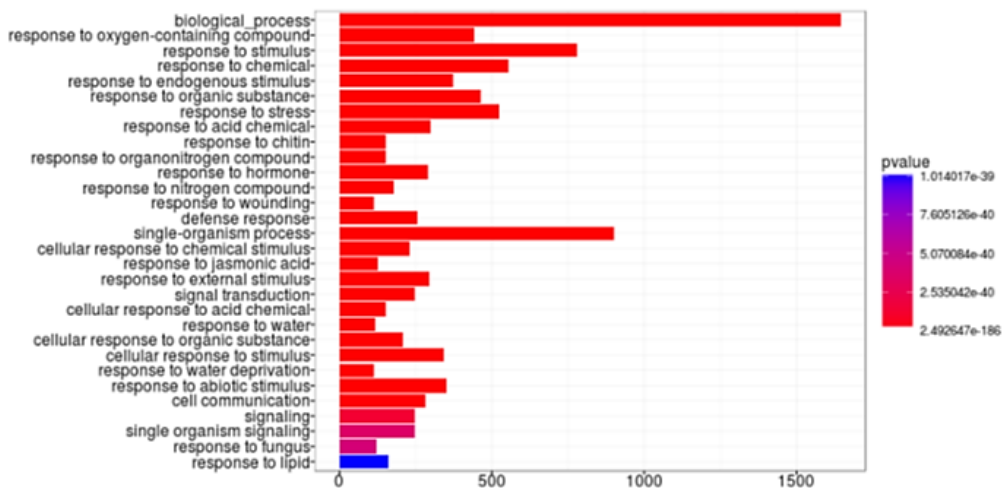
ANEXOS

Anexo1

Análisis de enriquecimiento en términos de ontología génica (GO) de los genes sobre-expresados y reprimidos en los mutantes *atbmi1a/b*. Distribución de los términos GO enriquecidos dentro de las diferentes categorías de “procesos biológicos” como vienen definidas por TAIR. Los p-valores están indicados.

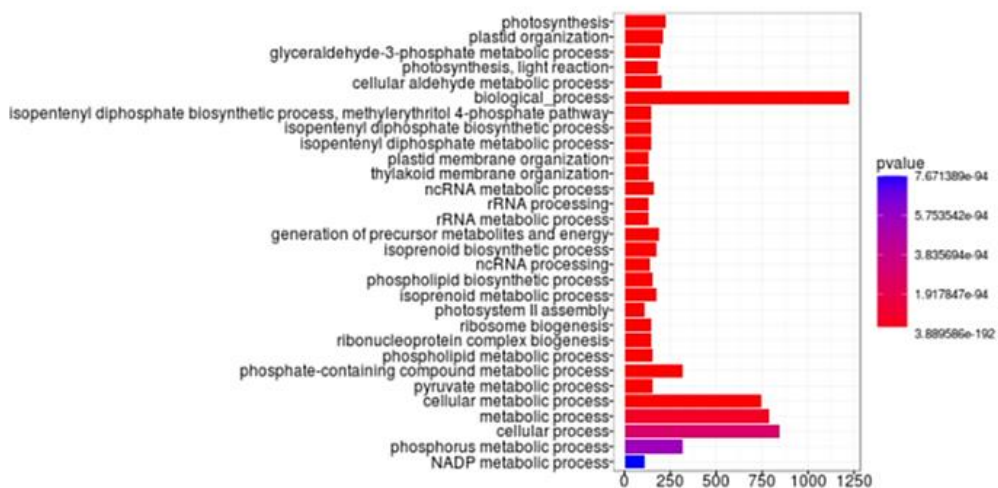
a

Enriquecimiento GO de los genes sobre-expresados en *atbmi1a/b*



b

Enriquecimiento GO de los genes reprimidos en *atbmi1a/b*

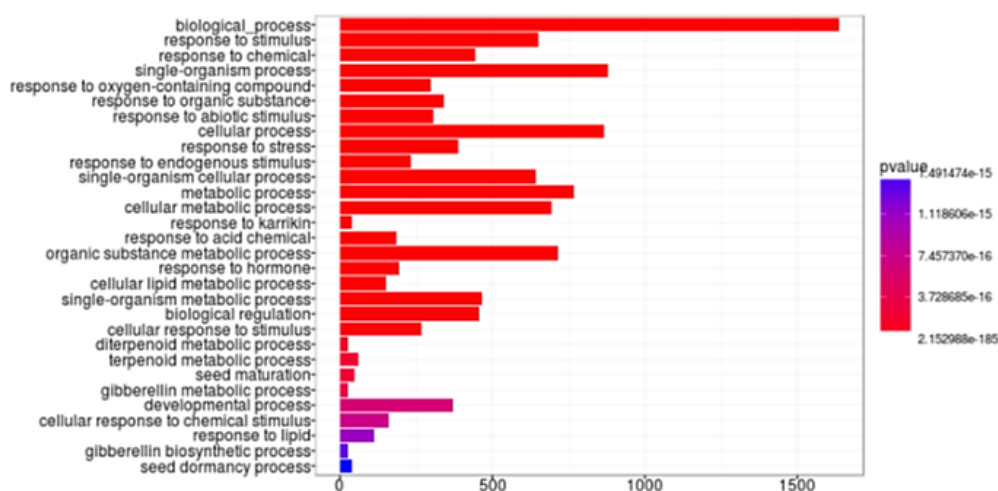


Anexo2

Análisis de enriquecimiento en términos de ontología génica (GO) de los genes sobre-expresados y reprimidos en los mutantes *atbmi1a/b/c*. Distribución de los términos GO enriquecidos dentro de las diferentes categorías de "procesos biológicos" como vienen definidas por TAIR. Los p-valores están indicados.

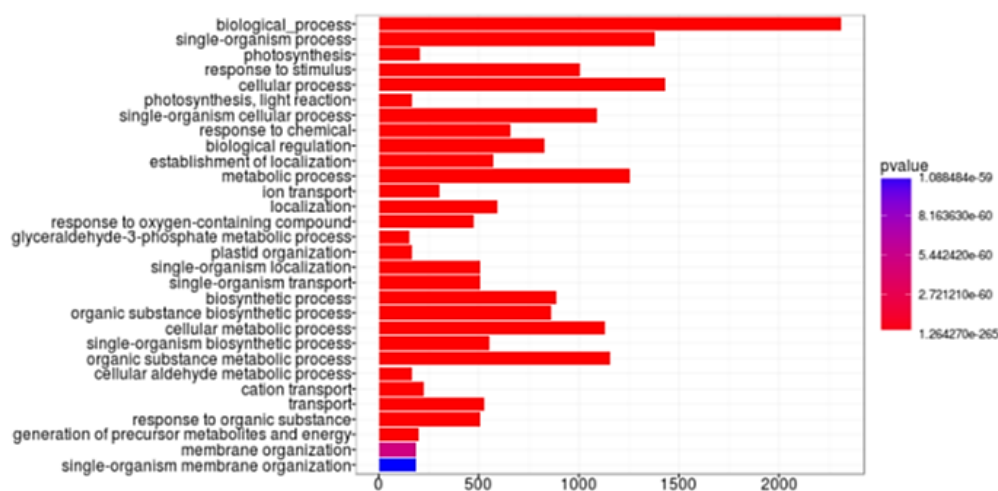
a

Enriquecimiento GO de los genes sobre-expresados en *atbmi1a/b/c*



b

Enriquecimiento GO de los genes reprimidos en *atbmi1a/b/c*



Anexo3: Lista de oligonucleótidos

Lista de oligonucleótidos empleados para reacciones de CHIP-qPCR

Oligonucleótido	Secuencia 5'-3'
MAF4 ChIP Fw	CCCGGTAGATTTGTTGAGAAAC
MAF4 ChIP Rev	CACTTGAAATTAACCAAGGAATGC
MAF5ChIP Fw	CAAGTCATCTTAACCTTTGTCTTGCT
MAF5ChIP Rev	GGCACTCGTTTCCACTAGATT
MIR156A ChIP Fw	CTCTCAAATCTCAAGTTCATTGCC
MIR156A ChIP Rev	GGCTCTTGTCGCTTTCTTTATC
MIR156C ChIP Fw	TCTCCGGTTTTGCTTGTTAAC
MIR156C ChIP Rev	AGAAGATTGGAAAGGAGGCAG
FUS3 ChIP Fw	ACTTTTGCTACACTTGTTCAACATG
FUS3 ChIP Rev	CGCAACAAGATCTAATGCCACT
WUS ChIP Fw	CAAACCTCTCTTTCTCTTGTTCTCTC
WUS ChIP Rev	GGCTCCATGTGTGTTTGATTGAC
FLC ChIP Fw	TCTGGTTATCGATTGCGATTCT
FLC ChIP Rev	CGTGCATATACAAATCCAAGAGAAC
up35S Fw	CCCAGGCTTTACACTTTATGCT
up35S Rev	GGTCATAGCTGTTTCTGTGTG
LexO Fw	GGGCTGCAGGTCGAGGCTAAAAACT
LexO Rev	GATCCTAGAGTCGAGCATATTACA
5'GUS Fw	CGCAAGACCCTTCCTCTAT
5'GUS Rev	AAAGGGTCCTAACCAAGAAA
MIDGUS Fw	GGAACCACTGAACACGTATCT
MIDGUS Rev	ACTTGCCGTCGTTGACTT
3'GUS Fw	CCGATCGTTCAAACATTTGGC
3'GUS Rev	CCCATCTCATAAATAACGTCATGC
FLC-10 Fw	CGTACTTATCGCCGGAGGAG
FLC-10 Rev	CATCGAGAAAGCTCGTCAGC
ACT7 Fw	GCGATGTTTGAGTTTCAATAAACGCTGC
ACT7 Rev	CTCACCTTCACCATTCCAGTTCCA

Lista de oligonucleótidos empleados para reacciones de q-RT-PCR

Oligonucleótido	Secuencia 5'-3'
FLC Fw	GCCACCTTAAATCGGCGGTTG
FLC Rev	CACAAAGTCTCTTGCCAAAGAGAGAG
MAF1 Fw	CGGCTGAGTTTTACCTTAAACTCAAAGCC
MAF1 Rev	GAGGAAGATAAAAGGTTTGAGATTACACAGC
MAF2 Fw	GGCTGAGCTTTCACCTTAAACTTACAGC
MAF2 Rev	CCACATTGGCGCGAGGAAGATAAAAAGG
MAF3 Fw	GGCTGAGCTTTCACCTTAAACTTACAGC
MAF3 Rev	GCTTCGTTTTGTTTTACCTTTATTTCCACATTGGG
MAF4 Fw	GCTACGGAAAAGTCATCCAAGGAGATGC
MAF4 Rev	CGAAAGTAAATACTATATCATCCTGTCTCCGAAGG
MAF5 Fw	CCACCAATCATCAACGGCTGATTTTTTCATCATCC
MAF5 Rev	CCGTATGCAGGGGGAGAAGAGG
FT Fw	CGAACGGTGATGATGCCTATAGTAG
FT Rev	CACTCTCATTTTCCTCCCCCTCTC
CO Fw	CCAATGGACAGAGAAGCCAGG
CO Rev	GCATCGTGTGAACCCTTGC
Pri-MIR156A Fw	CTTCGTTCTCTATGTCTCAATCTCTC
Pri-MIR156A Rev	TGATTAAAGGCTAAAGGTCTCCTC
Pri-MIR156C Fw	GTGATAATGAGTGATGACTGATG
Pri-MIR156C Rev	GAAAACGTGACCGGGACCGAATCG
FUS3 Fw	TCATGGTCTGCAGCTAGGTGACTT
FUS3 Rev	CGTCTACTTCTTCTTCCGATGC
LEC1 Fw	TGGAGCTCCCTTCTCTCACTATCA
LEC1 Rev	CTGCTGGACCACGATACCATTGTT
pri-miR172b Fw	CGGATTAGGGCGTTAATTACAATG
pri-miR172b Rev	GGTCTCTGGACGA ACTATTCTGTA
SLP3 Fw	CTTAGCTGGACACAACGAGAGAAGGC
SLP3 Rev	GAGAAACAGACAGAGACACAGAGGA
SLP9 Fw	CAAGGTTCA GTTGGTGGAGGA
SLP9 Rev	TGAAGAAGCTCGCCATGTATTG
WOX9-RT-Fw	ACTGTCGGAGGGTTTGAAGGTATC
WOX9-RT-Rev	AGTGGTAGCGTAACAAATCTGAGTCT
WOX2-RT-Fw	GCTTACTTCAATCGCCTCCTCCACAA
WOX2-RT-Rev	GTCCGTTTCTCGTAGCCACCACTTG
SMB-RT-Fw	ACGAATATCGCTTGGACGATAG
SMB-RT-Rev	GCTCTTGTTCTTGGTGAAATCC
ACT2 Fw	CACTTGCACCAAGCAGCATGAAGA
ACT2 Rev	AATGGAACCACCGATCCAGACACT

Lista de oligonucleótidos empleados para clonaje molecular

Oligonucleótido	Secuencia 5'-3'
PromG10-90-LexABD EcoRI FW	aaaGAATTCATAGTTTAAACTGAAGGCGGGA
PromG10-90-LexABD SacI Rev	aaaGAGCTCCGGTTCACCGGCAGCCACA
PromG10-90-LexABD KpnI Rev	aaaGGTACCCGGTTCACCGGCAGCCACA
AtBMI1A SacI Fw	aaaGAGCTCATGGAAGGAGACATGGTG
AtBMI1A SmaI Rev	aaaCCCGGGTTAGTTGTTGCATTGAGGGA
AtRING1B SacI Fw	aaaGAGCTCATGCCTTCCTTGAAGAGCT
AtRING1B SmaI Rev	aaaCCCGGGCTACGCGATTTGCTTTCTCC
VAL1 (SacI) Fw	aaaGAGCTCATGTTTGAAGTCAAAATGGGGTCA
VAL1 (SmaI) Rev	aaaCCCGGGTCAGCTTGAAACTCTCGGCTCTCAC
EAR domain (SacI) Fw	aaaGAGCTCAGGTTCAAGACGCTTATGATGAGGA
NLS (SmaI) Rev	aaaCCCGGGTTATTCTCGCTGTGTTCACTTCC

BIBIOGRAFÍA

BIBLIOGRAFÍA

Aichinger *et al.* (2009) «CHD3 Proteins and Polycomb Group Proteins Antagonistically Determine Cell Identity in Arabidopsis», *PLoS Genetics*. Editado por G. P. Copenhaver, 5(8), p. e1000605. doi: 10.1371/journal.pgen.1000605.

Aichinger *et al.* (2011) «The CHD3 chromatin remodeler PICKLE and polycomb group proteins antagonistically regulate meristem activity in the Arabidopsis root», *Plant Cell*, 23(3), pp. 1047-1060. doi: 10.1105/tpc.111.083352.

Allis y Jenuwein (2016) «The molecular hallmarks of epigenetic control», *Nature Reviews Genetics*. Nature Publishing Group, 17(8), pp. 487-500. doi: 10.1038/nrg.2016.59.

Almeida *et al.* (2017) «PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation», *Science (New York, N.Y.)*. Europe PMC Funders, 356(6342), p. 1081. doi: 10.1126/SCIENCE.AAL2512.

Aubert *et al.* (2001) «EMF1, a novel protein involved in the control of shoot architecture and flowering in Arabidopsis», *Plant Cell*, 13(8), pp. 1865-1875. doi: 10.1105/tpc.13.8.1865.

Bai y Morozov (2010) «Gene regulation by nucleosome positioning», *Trends in Genetics*. Elsevier Current Trends, 26(11), pp. 476-483. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0168952510001678?via%3Dihub> (Accedido: 29 de julio de 2019).

Bailey *et al.* (2009) «MEME Suite: Tools for motif discovery and searching», *Nucleic Acids Research*, 37(SUPPL. 2). doi: 10.1093/nar/gkp335.

Baldi (2019) «Nucleosome positioning and spacing: from genome-wide maps to single arrays.», *Essays in biochemistry*. Portland Press Limited, 63(1), pp. 5-14. doi: 10.1042/EBC20180058.

Bastow *et al.* (2004) «Vernalization requires epigenetic silencing of FLC by histone methylation.», *Nature*, 427(6970), pp. 164-7. doi: 10.1038/nature02269.

Baumbusch (2001) «The Arabidopsis thaliana genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes», *Nucleic Acids Research*. Oxford University Press (OUP), 29(21), pp. 4319-4333. doi: 10.1093/nar/29.21.4319.

Bechtold y Pelletier (1998) «In Planta Agrobacterium Mediated Transformation of Adult Arabidopsis thaliana Plants by Vacuum Infiltration», en *Arabidopsis Protocols*. Totowa, NJ: Humana Press, pp. 259-266. doi: 10.1385/0-89603-391-0:259.

Beck *et al.* (2014) «CpG island-mediated global gene regulatory modes in mouse embryonic stem cells», *Nature Communications*, 5(1), p. 5490. doi: 10.1038/ncomms6490.

Beh, Colwell y Francis (2012) «A core subunit of Polycomb repressive complex 1 is broadly conserved in function but not primary sequence», *Proceedings of the National Academy of Sciences of the United States of America*, 109(18). doi: 10.1073/pnas.1118678109.

Bell *et al.* (2010) «Accessibility of the *Drosophila* genome discriminates PcG repression, H4K16 acetylation and replication timing.», *Nature structural & molecular biology*, 17(7), pp. 894-900. doi: 10.1038/nsmb.1825.

Berger *et al.* (2011) «Transcriptional Regulation of *Arabidopsis* *LEAFY COTYLEDON2* Involves *RLE*, a *cis*-Element That Regulates Trimethylation of Histone H3 at Lysine-27», *The Plant Cell*, 23(11), pp. 4065-4078. doi: 10.1105/tpc.111.087866.

Bezhanian *et al.* (2007) «Unique, shared, and redundant roles for the *Arabidopsis* SWI/SNF chromatin remodeling ATPases *BRAHMA* and *SPLAYED*.», *The Plant cell*, 19(2), pp. 403-16. doi: 10.1105/tpc.106.048272.

Blackledge *et al.* (2014) «Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation.», *Cell*, 157(6), pp. 1445-59. doi: 10.1016/j.cell.2014.05.004.

Blazquez (2007) «Quantitative GUS Activity Assay in Intact Plant Tissue», *Cold Spring Harbor Protocols*, 2007(2), p. pdb.prot4688-pdb.prot4688. doi: 10.1101/pdb.prot4688.

Bouyer *et al.* (2011) «Polycomb repressive complex 2 controls the embryo-to-seedling phase transition.», *PLoS genetics*. Public Library of Science, 7(3), p. e1002014. doi: 10.1371/journal.pgen.1002014.

Bowman y McKnight (2017) «Sequence-specific targeting of chromatin remodelers organizes precisely positioned nucleosomes throughout the genome», *BioEssays*, 39(1), p. e201600183. doi: 10.1002/bies.201600183.

Bratzel *et al.* (2010) «Keeping cell identity in *Arabidopsis* requires PRC1 RING-finger homologs that catalyze H2A monoubiquitination.», *Current biology : CB*. Elsevier, 20(20), pp. 1853-9. doi: 10.1016/j.cub.2010.09.046.

Bratzel *et al.* (2012) «Regulation of the new *Arabidopsis* imprinted gene *AtBMI1C* requires the interplay of different epigenetic mechanisms.», *Molecular plant*. Elsevier, 5(1), pp. 260-9. doi: 10.1093/mp/ssr078.

Brower-Toland *et al.* (2005) «Specific Contributions of Histone Tails and their Acetylation to the Mechanical Stability of Nucleosomes», *Journal of Molecular Biology*. Academic Press, 346(1), pp. 135-146. doi: 10.1016/J.JMB.2004.11.056.

Brown y Kassis (2013) «Architectural and Functional Diversity of Polycomb Group Response Elements in *Drosophila*», *Genetics*, 195(2), pp. 407-419. doi: 10.1534/genetics.113.153247.

Buenrostro *et al.* (2013) «Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position», *Nature Methods*, 10(12), pp. 1213-1218. doi: 10.1038/nmeth.2688.

Butenko y Ohad (2011) «Polycomb-group mediated epigenetic mechanisms through plant evolution», *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1809(8), pp. 395-406. doi: 10.1016/j.bbagr.2011.05.013.

Buzas (2017) «Capturing Environmental Plant Memories in DNA, with a Little Help from Chromatin», *Plant and Cell Physiology*. Narnia, 58(8), pp. 1302-1312. doi: 10.1093/pcp/pcx092.

Cairns (2009) «The logic of chromatin architecture and remodelling at promoters», *Nature*, 461(7261), pp. 193-198. doi: 10.1038/nature08450.

Calabrese *et al.* (2012) «Site-specific silencing of regulatory elements as a mechanism of X inactivation», *Cell*, 151(5), pp. 951-963. doi: 10.1016/j.cell.2012.10.037.

Calonje *et al.* (2008) «EMBRYONIC FLOWER1 Participates in Polycomb Group-Mediated AG Gene Silencing in Arabidopsis», *THE PLANT CELL ONLINE*, 20(2), pp. 277-291. doi: 10.1105/tpc.106.049957.

Calonje (2014) «PRC1 Marks the Difference in Plant PcG Repression», *Molecular Plant. Cell Press*, 7(3), pp. 459-471. doi: 10.1093/MP/SST150.

Cao *et al.* (2002) «Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing», *Science*, 298(5595), pp. 1039-1043. doi: 10.1126/science.1076997.

Carles *et al.* (2002) «Regulation of Arabidopsis thaliana Em genes: role of ABI5», *The Plant Journal*. John Wiley & Sons, Ltd (10.1111), 30(3), pp. 373-383. doi: 10.1046/j.1365-313X.2002.01295.x.

Carter *et al.* (2018) «The Chromatin Remodelers PKL and PIE1 Act in an Epigenetic Pathway That Determines H3K27me3 Homeostasis in Arabidopsis.», *The Plant cell*, 30(6), pp. 1337-1352. doi: 10.1105/tpc.17.00867.

Causier *et al.* (2012) «The TOPLESS interactome: A framework for gene repression in Arabidopsis», *Plant Physiology*, 158(1), pp. 423-438. doi: 10.1104/pp.111.186999.

Chanvivattana *et al.* (2004) «Interaction of Polycomb-group proteins controlling flowering in Arabidopsis.», *Development (Cambridge, England)*. The Company of Biologists Ltd, 131(21), pp. 5263-76. doi: 10.1242/dev.01400.

Chanvivattana (2004) «Interaction of Polycomb-group proteins controlling flowering in Arabidopsis», *Development*, 131(21), pp. 5263-5276. doi: 10.1242/dev.01400.

Chen *et al.* (1997) «EMF genes regulate Arabidopsis inflorescence development.», *The Plant Cell*, 9(11), pp. 2011-2024. doi: 10.1105/tpc.9.11.2011.

Chen *et al.* (2010) «The Arabidopsis PRC1-like ring-finger proteins are necessary for repression of embryonic traits during vegetative growth», *Cell Research*. Nature Publishing Group, 20(12), pp. 1332-1344. doi: 10.1038/cr.2010.151.

Chen *et al.* (2013) «DANPOS: Dynamic analysis of nucleosome position and occupancy by sequencing», *Genome Research*, 23(2), pp. 341-351. doi: 10.1101/gr.142067.112.

Chen *et al.* (2014) «Improved nucleosome-positioning algorithm iNPS for accurate nucleosome positioning from sequencing data», *Nature Communications*. Nature Publishing Group, 5. doi: 10.1038/ncomms5909.

Chen *et al.* (2015) «Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes.», *Nature genetics*. NIH Public Access, 47(10), pp. 1149-57. doi: 10.1038/ng.3385.

Chen *et al.* (2017) «Genome-wide Nucleosome Occupancy and Organization Modulates the Plasticity of Gene Transcriptional Status in Maize». doi: 10.1016/j.molp.2017.05.001.

Dong Hong Chen *et al.* (2018) «Chromatin-based regulation of plant root development», *Frontiers in Plant Science*. Frontiers Media S.A. doi: 10.3389/fpls.2018.01509.

Naichong Chen *et al.* (2018) «HSI2/VAL1 Silences AGL15 to Regulate the Developmental Transition from Seed Maturation to Vegetative Growth in Arabidopsis.», *The Plant cell*, 30(3), pp. 600-619. doi: 10.1105/tpc.17.00655.

Chereji y Clark (2018) «Major Determinants of Nucleosome Positioning», *Biophysical Journal*. Cell Press, 114(10), pp. 2279-2289. doi: 10.1016/J.BPJ.2018.03.015.

Chereji, Ocampo y Clark (2017) «MNase-Sensitive Complexes in Yeast: Nucleosomes and Non-histone Barriers.», *Molecular cell*. NIH Public Access, 65(3), p. 565-577.e3. doi: 10.1016/j.molcel.2016.12.009.

Chhun *et al.* (2016) «HSI2 Repressor Recruits MED13 and HDA6 to Down-Regulate Seed Maturation Gene Expression Directly During Arabidopsis Early Seedling Growth», *Plant and Cell Physiology*, 57(8), pp. 1689-1706. doi: 10.1093/pcp/pcw095.

Chini *et al.* (2007) «The JAZ family of repressors is the missing link in jasmonate signalling», *Nature*. Nature Publishing Group, 448(7154), pp. 666-671. doi: 10.1038/nature06006.

Choi *et al.* (2000) «ABFs, a family of ABA-responsive element binding factors.», *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 275(3), pp. 1723-30. doi: 10.1074/JBC.275.3.1723.

Choi, Seo y Cho (2018) «Two TPL-binding motifs of ARF2 are involved in repression of auxin responses», *Frontiers in Plant Science*. Frontiers Media S.A., 9. doi: 10.3389/fpls.2018.00372.

Clapier y Cairns (2009) «The Biology of Chromatin Remodeling Complexes», *Annual Review of Biochemistry*, 78(1), pp. 273-304. doi: 10.1146/annurev.biochem.77.062706.153223.

Clarke (2009) «Cetyltrimethyl Ammonium Bromide (CTAB) DNA Miniprep for Plant DNA Isolation», *Cold Spring Harbor Protocols*, 2009(3), p. pdb.prot5177-pdb.prot5177. doi: 10.1101/pdb.prot5177.

Comet y Helin (2014) «Revolution in the Polycomb hierarchy», *Nature Structural & Molecular Biology*, 21(7), pp. 573-575. doi: 10.1038/nsmb.2848.

Conesa *et al.* (2016) «A survey of best practices for RNA-seq data analysis», *Genome Biology*. BioMed Central Ltd. doi: 10.1186/s13059-016-0881-8.

Corona y Tamkun (2004) «Multiple roles for ISWI in transcription, chromosome organization and DNA replication», *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*, 1677(1-3), pp. 113-119. doi: 10.1016/j.bbaexp.2003.09.018.

Crevillén *et al.* (2019) «Arabidopsis YAF9 histone readers modulate flowering time through NuA4-complex-dependent H4 and H2A.Z histone acetylation at *FLC* chromatin», *New Phytologist*. John Wiley & Sons, Ltd (10.1111), 222(4), pp. 1893-1908. doi: 10.1111/nph.15737.

Cui y Benfey (2009) «Interplay between SCARECROW, GA and LIKE HETEROCHROMATIN PROTEIN 1 in ground tissue patterning in the Arabidopsis root», *The*

Plant Journal, 58(6), pp. 1016-1027. doi: 10.1111/j.1365-313X.2009.03839.x.

Cui y Zhao (2012) «Genome-Wide Approaches to Determining Nucleosome Occupancy in Metazoans Using MNase-Seq», en, pp. 413-419. doi: 10.1007/978-1-61779-477-3_24.

Cutler *et al.* (2010) «Abscisic Acid: Emergence of a Core Signaling Network», *Annual Review of Plant Biology*. Annual Reviews , 61(1), pp. 651-679. doi: 10.1146/annurev-arplant-042809-112122.

Czermin *et al.* (2002) «Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites», *Cell*. Cell Press, 111(2), pp. 185-196. doi: 10.1016/S0092-8674(02)00975-3.

Davuluri *et al.* (2003) «AGRIS: Arabidopsis Gene Regulatory Information Server, an information resource of Arabidopsis cis -regulatory elements and transcription factors», *BMC Bioinformatics*. BioMed Central, 4(1), p. 25. doi: 10.1186/1471-2105-4-25.

Deaton *et al.* (2016) «Enhancer regions show high histone H3.3 turnover that changes during differentiation», *eLife*, 5. doi: 10.7554/eLife.15316.

Delmas *et al.* (2013) «ABI3 controls embryo degreening through Mendel's I locus», *Proceedings of the National Academy of Sciences of the United States of America*, 110(40), pp. E3888-3894. doi: 10.1073/pnas.1308114110.

Deng *et al.* (2013) «Arabidopsis Polycomb Repressive Complex 2 binding sites contain putative GAGA factor binding motifs within coding regions of genes», *BMC Genomics*, 14(1), p. 593. doi: 10.1186/1471-2164-14-593.

Derkacheva *et al.* (2013) «Arabidopsis MSI1 connects LHP1 to PRC2 complexes», *The EMBO Journal*, 32(14), pp. 2073-2085. doi: 10.1038/emboj.2013.145.

Derkacheva y Hennig (2014) «Variations on a theme: Polycomb group proteins in plants», *Journal of Experimental Botany*. Narnia, 65(10), pp. 2769-2784. doi: 10.1093/jxb/ert410.

Drew y Travers (1985) «DNA bending and its relation to nucleosome positioning», *Journal of Molecular Biology*, 186(4), pp. 773-790. doi: 10.1016/0022-2836(85)90396-1.

Du *et al.* (2010) «agriGO: a GO analysis toolkit for the agricultural community», *Nucleic Acids Research*, 38(Web Server issue), pp. W64-70. doi: 10.1093/nar/gkq310.

Egan *et al.* (2013) «CHD5 is required for neurogenesis and has a dual role in facilitating gene expression and polycomb gene repression», *Developmental Cell*, 26(3), pp. 223-236. doi: 10.1016/j.devcel.2013.07.008.

Eggleston (2005) «Unraveling chromatin organization», *Nature Structural & Molecular Biology*. Nature Publishing Group, 12(1), pp. 6-6. doi: 10.1038/nsmb0105-6.

Eichten, Schmitz y Springer (2014) «Epigenetics: Beyond chromatin modifications and complex genetic regulation», *Plant Physiology*. American Society of Plant Biologists, 165(3), pp. 933-947. doi: 10.1104/pp.113.234211.

Engelhorn *et al.* (2012) «Development-related PcG target in the apex 4 controls leaf margin architecture in Arabidopsis thaliana.», *Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 139(14), pp. 2566-75. doi: 10.1242/dev.078618.

Farcas *et al.* (2012) «KDM2B links the polycomb repressive complex 1 (PRC1) to recognition of CpG islands», *eLife*, 2012(1). doi: 10.7554/eLife.00205.

Farrona *et al.* (2004) «The Arabidopsis thaliana SNF2 homolog AtBRM controls shoot development and flowering», *Development*, 131(20), pp. 4965-4975. doi: 10.1242/dev.01363.

Farrona *et al.* (2011) «Tissue-Specific Expression of *FLOWERING LOCUS T* in *Arabidopsis* Is Maintained Independently of Polycomb Group Protein Repression», *The Plant Cell*, 23(9), pp. 3204-3214. doi: 10.1105/tpc.111.087809.

Fasulo *et al.* (2012) «The Drosophila Mi-2 Chromatin-Remodeling Factor Regulates Higher-Order Chromatin Structure and Cohesin Dynamics In Vivo», *PLoS Genetics*, 8(8). doi: 10.1371/journal.pgen.1002878.

Feng y Lu (2017) «LHP1 Could Act as an Activator and a Repressor of Transcription in Plants.», *Frontiers in plant science*. Frontiers Media SA, 8, p. 2041. doi: 10.3389/fpls.2017.02041.

Feng y Shen (2014) «Dynamic regulation and function of histone monoubiquitination in plants», *Frontiers in Plant Science*. Frontiers Research Foundation. doi: 10.3389/fpls.2014.00083.

Fernandez *et al.* (2014) «The MADS-Domain Factors AGAMOUS-LIKE15 and AGAMOUS-LIKE18, along with SHORT VEGETATIVE PHASE and AGAMOUS-LIKE24, Are Necessary to Block Floral Gene Expression during the Vegetative Phase.», *Plant physiology*. American Society of Plant Biologists, 165(4), pp. 1591-1603. doi: 10.1104/pp.114.242990.

Fierz *et al.* (2011) «Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction», *Nature Chemical Biology*. Nature Publishing Group, 7(2), pp. 113-119. doi: 10.1038/nchembio.501.

Flanagan *et al.* (2005) «Double chromodomains cooperate to recognize the methylated histone H3 tail», *Nature*. Nature Publishing Group, 438(7071), pp. 1181-1185. doi: 10.1038/nature04290.

Flores y Orozco (2011) «nucleR: a package for non-parametric nucleosome positioning.», *Bioinformatics (Oxford, England)*, 27(15), pp. 2149-50. doi: 10.1093/bioinformatics/btr345.

Förderer, Zhou y Turck (2016) «The age of multiplexity: recruitment and interactions of Polycomb complexes in plants», *Current Opinion in Plant Biology*. Elsevier Current Trends, 29, pp. 169-178. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/26826786> (Accedido: 28 de febrero de 2019).

Francis *et al.* (2001) «Reconstitution of a functional core polycomb repressive complex.», *Molecular cell*, 8(3), pp. 545-56. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/11583617> (Accedido: 14 de abril de 2019).

Francis, Kingston y Woodcock (2004) «Chromatin Compaction by a Polycomb Group Protein Complex», *Science*. American Association for the Advancement of Science, 306(5701), pp. 1574-1577. Disponible en: <http://science.sciencemag.org/content/306/5701/1574?ijkey=40ab1878047adc64973f8>

63481d42fb1af90fa55&keytype=tf_ipsecsha (Accedido: 28 de febrero de 2019).

Fritsch *et al.* (2004) «The INO80 protein controls homologous recombination in *Arabidopsis thaliana*», *Molecular Cell*, 16(3), pp. 479-485. doi: 10.1016/j.molcel.2004.09.034.

Fu *et al.* (2012) «DiNuP: a systematic approach to identify regions of differential nucleosome positioning», *Bioinformatics*. Narnia, 28(15), pp. 1965-1971. doi: 10.1093/bioinformatics/bts329.

Fyodorov *et al.* (2018) «Emerging roles of linker histones in regulating chromatin structure and function», *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 19(3), pp. 192-206. doi: 10.1038/nrm.2017.94.

Gaffney *et al.* (2012) «Controls of Nucleosome Positioning in the Human Genome», *PLoS Genetics*. Editado por E. Segal, 8(11), p. e1003036. doi: 10.1371/journal.pgen.1003036.

Gangappa *et al.* (2013) «Z-box binding transcription factors (ZBFs): a new class of transcription factors in *Arabidopsis* seedling development.», *Molecular plant*. Elsevier, 6(6), pp. 1758-68. doi: 10.1093/mp/sst140.

Gao *et al.* (2009) «Repression of Seed Maturation Genes by a Trihelix Transcriptional Repressor in *Arabidopsis* Seedlings», *The Plant Cell*, 21(1), pp. 54-71. doi: 10.1105/tpc.108.061309.

Gao *et al.* (2012) «PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes», *Molecular Cell*, 45(3), pp. 344-356. doi: 10.1016/j.molcel.2012.01.002.

Gaudin *et al.* (2001) «Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in *Arabidopsis*», *Development (Cambridge, England)*, 128(23), pp. 4847-4858.

Gendall *et al.* (2001) «The VERNALIZATION 2 gene mediates the epigenetic regulation of vernalization in *Arabidopsis*.», *Cell*. Elsevier, 107(4), pp. 525-35. doi: 10.1016/S0092-8674(01)00573-6.

Gentry y Hennig (2014) «Remodelling chromatin to shape development of plants», *Experimental Cell Research*. Academic Press, 321(1), pp. 40-46. doi: 10.1016/J.YEXCR.2013.11.010.

Gil y O'Loughlen (2014) «PRC1 complex diversity: Where is it taking us?», *Trends in Cell Biology*. Elsevier Ltd, pp. 632-641. doi: 10.1016/j.tcb.2014.06.005.

Giresi *et al.* (2007) «FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin», *Genome Research*, 17(6), pp. 877-885. doi: 10.1101/gr.5533506.

Giresi y Lieb (2009) «Isolation of active regulatory elements from eukaryotic chromatin using FAIRE (Formaldehyde Assisted Isolation of Regulatory Elements)», *Methods*, 48(3), pp. 233-239. doi: 10.1016/j.ymeth.2009.03.003.

Goff, Trapnell y Kelley (2014) *CummeRbund: Visualization and Exploration of Cufflinks High-throughput Sequencing Data*. Disponible en: <https://bioconductor.statistik.tu->

dortmund.de/packages/3.1/bioc/vignettes/cummeRbund/inst/doc/cummeRbund-manual.pdf (Accedido: 3 de abril de 2019).

Gómez-Zambrano, Merini y Calonje (2019) «The repressive role of Arabidopsis H2A.Z in transcriptional regulation depends on AtBMI1 activity», *Nature Communications*. Nature Publishing Group, 10(1). doi: 10.1038/s41467-019-10773-1.

Goodrich *et al.* (1997) «A Polycomb-group gene regulates homeotic gene expression in Arabidopsis», *Nature*. Nature Publishing Group, 386(6620), pp. 44-51. doi: 10.1038/386044a0.

Goodstein *et al.* (2012) «Phytozome: a comparative platform for green plant genomics», *Nucleic Acids Research*, 40(Database issue), pp. D1178-1186. doi: 10.1093/nar/gkr944.

Goossens *et al.* (2015) «Change of a conserved amino acid in the MYC2 and MYC3 transcription factors leads to release of JAZ repression and increased activity», *New Phytologist*, 206(4), pp. 1229-1237. doi: 10.1111/nph.13398.

van der Graaff, Laux y Rensing (2009) «The WUS homeobox-containing (WOX) protein family», *Genome Biology*. BioMed Central, 10(12), p. 248. doi: 10.1186/gb-2009-10-12-248.

Grau *et al.* (2011) «Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge.», *Genes & development*, 25(20), pp. 2210-21. doi: 10.1101/gad.17288211.

Greb *et al.* (2007) «The PHD finger protein VRN5 functions in the epigenetic silencing of Arabidopsis FLC.», *Current biology: CB*, 17(1), pp. 73-8. doi: 10.1016/j.cub.2006.11.052.

Grigoryev (2012) «Nucleosome spacing and chromatin higher-order folding», *Nucleus*. Taylor & Francis, 3(6), p. 493. doi: 10.4161/NUCL.22168.

Grossniklaus *et al.* (1998) «Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis.», *Science (New York, N.Y.)*, 280(5362), pp. 446-50. doi: 10.1126/science.280.5362.446.

Gu *et al.* (2013) «Arabidopsis FLC clade members form flowering-repressor complexes coordinating responses to endogenous and environmental cues», *Nature Communications*. Nature Publishing Group, 4(1), p. 1947. doi: 10.1038/ncomms2947.

Guerriero *et al.* (2009) «The RY/Sph element mediates transcriptional repression of maturation genes from late maturation to early seedling growth», *New Phytologist*, 184(3), pp. 552-565. doi: 10.1111/j.1469-8137.2009.02977.x.

Hajheidari, Koncz y Bucher (2019) «Chromatin Evolution-Key Innovations Underpinning Morphological Complexity», *Frontiers in Plant Science*, 10, p. 454. doi: 10.3389/fpls.2019.00454.

Han *et al.* (2012) «The SWI2/SNF2 chromatin remodeling ATPase BRAHMA represses abscisic acid responses in the absence of the stress stimulus in Arabidopsis.», *The Plant cell*, 24(12), pp. 4892-906. doi: 10.1105/tpc.112.105114.

Han *et al.* (2015) «Roles and activities of chromatin remodeling ATPases in plants». doi: 10.1111/tpj.12877.

Hands, Rabiger y Koltunow (2016) «Mechanisms of endosperm initiation», *Plant Reproduction*. Springer Berlin Heidelberg, 29(3), pp. 215-225. doi: 10.1007/s00497-016-0290-x.

Hargreaves y Crabtree (2011) «ATP-dependent chromatin remodeling: genetics, genomics and mechanisms.», *Cell research*, 21(3), pp. 396-420. doi: 10.1038/cr.2011.32.

He, Huang y Xu (2013) «Mechanisms guiding Polycomb activities during gene silencing in *Arabidopsis thaliana*», *Frontiers in Plant Science*. Frontiers Research Foundation. doi: 10.3389/fpls.2013.00454.

Hecker *et al.* (2015) «The *Arabidopsis* GAGA-Binding Factor BASIC PENTACYSTEINE6 Recruits the POLYCOMB-REPRESSIVE COMPLEX1 Component LIKE HETEROCHROMATIN PROTEIN1 to GAGA DNA Motifs», *Plant Physiology*, 168(3), pp. 1013-1024. doi: 10.1104/pp.15.00409.

Hehenberger *et al.* (2012) «Endosperm cellularization defines an important developmental transition for embryo development.», *Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 139(11), pp. 2031-9. doi: 10.1242/dev.077057.

Heintzman *et al.* (2007) «Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome», *Nature Genetics*, 39(3), pp. 311-318. doi: 10.1038/ng1966.

Heinz *et al.* (2010) «Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities», *Molecular Cell*. Elsevier, 38(4), pp. 576-589. doi: 10.1016/J.MOLCEL.2010.05.004.

Hennig y Derkacheva (2009) «Diversity of Polycomb group complexes in plants: same rules, different players?», *Trends in Genetics*. Elsevier Current Trends, 25(9), pp. 414-423. doi: 10.1016/J.TIG.2009.07.002.

Hill, Wang y Perry (2008) «A transcriptional repression motif in the MADS factor AGL15 is involved in recruitment of histone deacetylase complex components», *The Plant Journal*, 53(1), pp. 172-185. doi: 10.1111/j.1365-313X.2007.03336.x.

Ho *et al.* (2009) «An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency», *Proceedings of the National Academy of Sciences*, 106(13), pp. 5181-5186. doi: 10.1073/pnas.0812889106.

Ho *et al.* (2011) «esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function.», *Nature cell biology*, 13(8), pp. 903-13. doi: 10.1038/ncb2285.

Ho *et al.* (2013) «PICKLE is a CHD subfamily II ATP-dependent chromatin remodeling factor», *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. Elsevier, 1829(2), pp. 199-210. doi: 10.1016/J.BBAGRM.2012.10.011.

Holec y Berger (2012) «Polycomb group complexes mediate developmental transitions in plants.», *Plant physiology*. American Society of Plant Biologists, 158(1), pp. 35-43. doi: 10.1104/pp.111.186445.

Hong *et al.* (2009) «Temporal and spatial expression patterns of nine *Arabidopsis* genes encoding Jumonji C-domain proteins», *Molecules and Cells*, 27(4), pp. 481-490. doi:

10.1007/s10059-009-0054-7.

Hoppmann *et al.* (2011) «The CW domain, a new histone recognition module in chromatin proteins», *The EMBO Journal*, 30(10), pp. 1939-1952. doi: 10.1038/emboj.2011.108.

Horstman *et al.* (2014) «AINTEGUMENTA-LIKE proteins: hubs in a plethora of networks», *Trends in Plant Science*, 19(3), pp. 146-157. doi: 10.1016/j.tplants.2013.10.010.

Huang *et al.* (2019) «Evolution and conservation of polycomb repressive complex 1 core components and putative associated factors in the green lineage», *BMC Genomics*. BioMed Central Ltd., 20(1). doi: 10.1186/s12864-019-5905-9.

Hughes y Rando (2014) «Mechanisms Underlying Nucleosome Positioning In Vivo», *Annual Review of Biophysics*. Annual Reviews , 43(1), pp. 41-63. doi: 10.1146/annurev-biophys-051013-023114.

Huijser y Schmid (2011) «The control of developmental phase transitions in plants», *Development*, 138(19), pp. 4117-4129. doi: 10.1242/dev.063511.

Hunkapiller *et al.* (2012) «Polycomb-Like 3 Promotes Polycomb Repressive Complex 2 Binding to CpG Islands and Embryonic Stem Cell Self-Renewal», *PLoS Genetics*. Editado por G. Cavalli. Public Library of Science, 8(3), p. e1002576. doi: 10.1371/journal.pgen.1002576.

Hurtado, Farrona y Reyes (2006) «The putative SWI/SNF complex subunit BRAHMA activates flower homeotic genes in Arabidopsis thaliana.», *Plant molecular biology*, 62(1-2), pp. 291-304. doi: 10.1007/s11103-006-9021-2.

Hwan Lee, Joon Kim y Ahn (2012) «Role of *SEPALLATA3* (*SEP3*) as a downstream gene of *miR156-SPL3-FT* circuitry in ambient temperature-responsive flowering», *Plant Signaling & Behavior*, 7(9), pp. 1151-1154. doi: 10.4161/psb.21366.

Ingouff, Haseloff y Berger (2005) «Polycomb group genes control developmental timing of endosperm», *The Plant Journal*, 42(5), pp. 663-674. doi: 10.1111/j.1365-313X.2005.02404.x.

Jarillo y Piñeiro (2015) «H2A.Z mediates different aspects of chromatin function and modulates flowering responses in Arabidopsis.», *The Plant journal : for cell and molecular biology*, 83(1), pp. 96-109. doi: 10.1111/tpj.12873.

Jiang *et al.* (2008) «Repression of FLOWERING LOCUS C and FLOWERING LOCUS T by the Arabidopsis Polycomb Repressive Complex 2 Components», *PLoS ONE*. Editado por B. P. Dilkes, 3(10), p. e3404. doi: 10.1371/journal.pone.0003404.

Jiang y B. Franklin Pugh (2009) «A compiled and systematic reference map of nucleosome positions across the Saccharomyces cerevisiae genome», *Genome Biology*, 10(10), p. R109. doi: 10.1186/gb-2009-10-10-r109.

Jiang y B. Franklin Pugh (2009) «Nucleosome positioning and gene regulation: advances through genomics», *Nature Reviews Genetics*, 10(3), pp. 161-172. doi: 10.1038/nrg2522.

Jimeno-Gonzalez *et al.* (2006) «A Gene-Specific Requirement for FACT during Transcription Is Related to the Chromatin Organization of the Transcribed Region»,

Molecular and Cellular Biology. American Society for Microbiology, 26(23), pp. 8710-8721. doi: 10.1128/mcb.01129-06.

Jin *et al.* (2009) «H3.3/H2A.Z double variant-containing nucleosomes mark “nucleosome-free regions” of active promoters and other regulatory regions.», *Nature genetics*, 41(8), pp. 941-5. doi: 10.1038/ng.409.

Jin y Felsenfeld (2007) «Nucleosome stability mediated by histone variants H3.3 and H2A.Z», *Genes & Development*, 21(12), pp. 1519-1529. doi: 10.1101/gad.1547707.

Jing, Guo y Lin (2019a) « The Chromatin-Remodeling Factor PICKLE Antagonizes Polycomb Repression of FT to Promote Flowering », *Plant Physiology*. American Society of Plant Biologists (ASPB), 181(2), pp. 656-668. doi: 10.1104/pp.19.00596.

Jing, Guo y Lin (2019b) «The B3-domain Transcription Factor VAL1 Regulates the Floral Transition by Repressing FLOWERING LOCUS T», *Plant Physiology*, p. 00642.2019. doi: 10.1104/pp.19.00642.

John *et al.* (2013) «Genome-Scale Mapping of DNase I Hypersensitivity», en *Current Protocols in Molecular Biology*. Hoboken, NJ, USA: John Wiley & Sons, Inc., p. 21.27.1-21.27.20. doi: 10.1002/0471142727.mb2127s103.

Kagale, Links y Rozwadowski (2010) «Genome-wide analysis of ethylene-responsive element binding factor-associated amphiphilic repression motif-containing transcriptional regulators in arabidopsis», *Plant Physiology*, 152(3), pp. 1109-1134. doi: 10.1104/pp.109.151704.

Kagale y Rozwadowski (2011) «EAR motif-mediated transcriptional repression in plants», *Epigenetics*, 6(2), pp. 141-146. doi: 10.4161/epi.6.2.13627.

Kasajima *et al.* (2004) «A protocol for rapid DNA extraction from *Arabidopsis thaliana* for PCR analysis», *Plant Molecular Biology Reporter*. Springer-Verlag, 22(1), pp. 49-52. doi: 10.1007/BF02773348.

Keene *et al.* (1981) «DNase I hypersensitive sites in *Drosophila* chromatin occur at the 5' ends of regions of transcription.», *Proceedings of the National Academy of Sciences*, 78(1), pp. 143-146. doi: 10.1073/pnas.78.1.143.

Kelly *et al.* (2012) «Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules», *Genome Research*, 22(12), pp. 2497-2506. doi: 10.1101/gr.143008.112.

Kim *et al.* (2012) «EMF1 and PRC2 cooperate to repress key regulators of *Arabidopsis* development», *PLoS genetics*, 8(3), p. e1002512. doi: 10.1371/journal.pgen.1002512.

Kim *et al.* (2015) «Linker histone H1.2 establishes chromatin compaction and gene silencing through recognition of H3K27me3», *Scientific Reports*. Nature Publishing Group, 5. doi: 10.1038/srep16714.

Kim, Bresnick y Bultman (2009) «BRG1 directly regulates nucleosome structure and chromatin looping of the alpha globin locus to activate transcription.», *Nucleic acids research*, 37(18), pp. 6019-27. doi: 10.1093/nar/gkp677.

Kim y Sung (2010) «Role of *VIN3-LIKE 2* in facultative photoperiodic flowering response in *Arabidopsis*», *Plant Signaling & Behavior*, 5(12), pp. 1672-1673. doi:

10.4161/psb.5.12.14035.

Kim y Sung (2013) «Coordination of the Vernalization Response through a VIN3 and FLC Gene Family Regulatory Network in Arabidopsis», *The Plant Cell*, 25(2), pp. 454-469. doi: 10.1105/tpc.112.104760.

Kim y Sung (2017) «Vernalization-Triggered Intragenic Chromatin Loop Formation by Long Noncoding RNAs», *Developmental Cell*. Cell Press, 40(3), p. 302-312.e4. doi: 10.1016/J.DEVCEL.2016.12.021.

Kim, Xi y Sung (2017) «Modular function of long noncoding RNA, COLDAIR, in the vernalization response.», *PLoS genetics*. Public Library of Science, 13(7), p. e1006939. doi: 10.1371/journal.pgen.1006939.

Kim, Zhu y Sung (2010) «Epigenetic regulation of gene programs by EMF1 and EMF2 in Arabidopsis», *Plant Physiology*, 152(2), pp. 516-528. doi: 10.1104/pp.109.143495.

King *et al.* (2018) «Polycomb repressive complex 1 shapes the nucleosome landscape but not accessibility at target genes.», *Genome research*. Cold Spring Harbor Laboratory Press, 28(10), pp. 1494-1507. doi: 10.1101/gr.237180.118.

King, Francis y Kingston (2002) «Native and Recombinant Polycomb Group Complexes Establish a Selective Block to Template Accessibility To Repress Transcription In Vitro», *Molecular and Cellular Biology*, 22(22), pp. 7919-7928. doi: 10.1128/MCB.22.22.7919-7928.2002.

Kinoshita *et al.* (2001) «Polycomb repression of flowering during early plant development», *Proceedings of the National Academy of Sciences*, 98(24), pp. 14156-14161. doi: 10.1073/pnas.241507798.

Knizewski, Ginalski y Jerzmanowski (2008) «Snf2 proteins in plants: gene silencing and beyond», *Trends in Plant Science*, 13(10), pp. 557-565. doi: 10.1016/j.tplants.2008.08.004.

Kohler (2003) «Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development», *The EMBO Journal*, 22(18), pp. 4804-4814. doi: 10.1093/emboj/cdg444.

Köhler *et al.* (2003) «Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development.», *The EMBO journal*. EMBO Press, 22(18), pp. 4804-14. doi: 10.1093/emboj/cdg444.

Köhler y Hennig (2010) «Regulation of cell identity by plant Polycomb and trithorax group proteins.», *Current opinion in genetics & development*, 20(5), pp. 541-7. doi: 10.1016/j.gde.2010.04.015.

Kotake *et al.* (2003) «Arabidopsis TERMINAL FLOWER 2 Gene Encodes a Heterochromatin Protein 1 Homolog and Represses both FLOWERING LOCUS T to Regulate Flowering Time and Several Floral Homeotic Genes», *Plant and Cell Physiology*, 44(6), pp. 555-564. doi: 10.1093/pcp/pcg091.

Kradolfer, Hennig y Köhler (2013) «Increased Maternal Genome Dosage Bypasses the Requirement of the FIS Polycomb Repressive Complex 2 in Arabidopsis Seed Development», *PLoS Genetics*. Editado por J. F. Gutierrez-Marcos. Public Library of Science, 9(1), p. e1003163. doi: 10.1371/journal.pgen.1003163.

Ku *et al.* (2008) «Genomewide Analysis of PRC1 and PRC2 Occupancy Identifies Two Classes of Bivalent Domains», *PLoS Genetics*. Editado por B. van Steensel, 4(10), p. e1000242. doi: 10.1371/journal.pgen.1000242.

Kuzmichev (2002) «Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein», *Genes & Development*, 16(22), pp. 2893-2905. doi: 10.1101/gad.1035902.

Kuzmichev *et al.* (2005) «Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation», *Proceedings of the National Academy of Sciences*, 102(6), pp. 1859-1864. doi: 10.1073/pnas.0409875102.

Kwon *et al.* (2006) «A role for chromatin remodeling in regulation of CUC gene expression in the Arabidopsis cotyledon boundary.», *Development (Cambridge, England)*, 133(16), pp. 3223-30. doi: 10.1242/dev.02508.

LAEMMLI (1970) «Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4», *Nature*. Nature Publishing Group, 227(5259), pp. 680-685. doi: 10.1038/227680a0.

Lafos *et al.* (2011) «Dynamic Regulation of H3K27 Trimethylation during Arabidopsis Differentiation», *PLoS Genetics*. Editado por T. Kakutani, 7(4), p. e1002040. doi: 10.1371/journal.pgen.1002040.

Lagarou *et al.* (2008) «dKDM2 couples histone H2A ubiquitylation to histone H3 demethylation during Polycomb group silencing», *Genes & Development*, 22(20), pp. 2799-2810. doi: 10.1101/gad.484208.

Lai *et al.* (2018) «Principles of nucleosome organization revealed by single-cell micrococcal nuclease sequencing», *Nature*, 562(7726), pp. 281-285. doi: 10.1038/s41586-018-0567-3.

Lai y Pugh (2017a) «Genome-wide uniformity of human “open” pre-initiation complexes.», *Genome research*. Cold Spring Harbor Laboratory Press, 27(1), pp. 15-26. doi: 10.1101/gr.210955.116.

Lai y Pugh (2017b) «Understanding nucleosome dynamics and their links to gene expression and DNA replication.», *Nature reviews. Molecular cell biology*. NIH Public Access, 18(9), pp. 548-562. doi: 10.1038/nrm.2017.47.

Lamesch *et al.* (2012) «The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools», *Nucleic Acids Research*, 40(D1), pp. D1202-D1210. doi: 10.1093/nar/gkr1090.

Langmead *et al.* (2009) «Ultrafast and memory-efficient alignment of short DNA sequences to the human genome», *Genome Biology*. BioMed Central, 10(3), p. R25. doi: 10.1186/gb-2009-10-3-r25.

Larsson, Landberg y Meeks-Wagner (1998) «The TERMINAL FLOWER2 (TFL2) gene controls the reproductive transition and meristem identity in Arabidopsis thaliana.», *Genetics*, 149(2), pp. 597-605. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/9611176> (Accedido: 16 de octubre de 2019).

Lee *et al.* (2009) «Arabidopsis ING and Alfin1-like protein families localize to the nucleus and bind to H3K4me3/2 via plant homeodomain fingers», *The Plant Journal*. John

Wiley & Sons, Ltd (10.1111), 58(3), pp. 511-524. doi: 10.1111/j.1365-313X.2009.03795.x.

Levine *et al.* (2002) «The Core of the Polycomb Repressive Complex Is Compositionally and Functionally Conserved in Flies and Humans», *Molecular and Cellular Biology*, 22(17), pp. 6070-6078. doi: 10.1128/MCB.22.17.6070-6078.2002.

Lewis (1978) «A Gene Complex Controlling Segmentation in Drosophila», en *Genes, Development and Cancer*. Boston, MA: Springer US, pp. 205-217. doi: 10.1007/978-1-4419-8981-9_13.

Li *et al.* (2009) «The Sequence Alignment/Map format and SAMtools», *Bioinformatics*, 25(16), pp. 2078-2079. doi: 10.1093/bioinformatics/btp352.

Li *et al.* (2011) «Overexpression of AtBMI1C, a Polycomb Group Protein Gene, Accelerates Flowering in Arabidopsis», *PLoS ONE*. Editado por I. Baxter, 6(6), p. e21364. doi: 10.1371/journal.pone.0021364.

Li *et al.* (2014) «ISWI proteins participate in the genome-wide nucleosome distribution in Arabidopsis», *The Plant Journal*. John Wiley & Sons, Ltd (10.1111), 78(4), pp. 706-714. doi: 10.1111/tpj.12499.

Li *et al.* (2015) «The Arabidopsis SWI2/SNF2 chromatin Remodeler BRAHMA regulates polycomb function during vegetative development and directly activates the flowering repressor gene SVP.», *PLoS genetics*. Editado por D. Schubert, 11(1), p. e1004944. doi: 10.1371/journal.pgen.1004944.

Li *et al.* (2016) «Concerted genomic targeting of H3K27 demethylase REF6 and chromatin-remodeling ATPase BRM in Arabidopsis.», *Nature genetics*. NIH Public Access, 48(6), pp. 687-93. doi: 10.1038/ng.3555.

Li *et al.* (2017) «Polycomb-like proteins link the PRC2 complex to CpG islands», *Nature*. Nature Publishing Group, 549(7671), pp. 287-291. doi: 10.1038/nature23881.

Li *et al.* (2018) «Polycomb-mediated gene silencing by the BAH-EMF1 complex in plants», *Nature Genetics*. Nature Publishing Group, 50(9), pp. 1254-1261. doi: 10.1038/s41588-018-0190-0.

Li y Luan (2011) «The cyclophilin AtCYP71 Interacts with CAF-1 and LHP1 and functions in multiple chromatin remodeling processes», *Molecular Plant*. Oxford University Press, 4(4), pp. 748-758. doi: 10.1093/mp/ssr036.

Liang *et al.* (2015) «Kicking against the PRCs – A Domesticated Transposase Antagonises Silencing Mediated by Polycomb Group Proteins and Is an Accessory Component of Polycomb Repressive Complex 2», *PLoS Genetics*. Editado por C. Feschotte. Public Library of Science, 11(12), p. e1005660. doi: 10.1371/journal.pgen.1005660.

Liu *et al.* (2009) «Regulation of Floral Patterning by Flowering Time Genes», *Developmental Cell*, 16(5), pp. 711-722. doi: 10.1016/j.devcel.2009.03.011.

Liu *et al.* (2011) «AGAMOUS terminates floral stem cell maintenance in Arabidopsis by directly repressing WUSCHEL through recruitment of Polycomb Group proteins.», *The Plant cell*. American Society of Plant Biologists, 23(10), pp. 3654-70. doi: 10.1105/tpc.111.091538.

Liu *et al.* (2015) «Determinants of nucleosome positioning and their influence on plant gene expression», *Genome Research*. Cold Spring Harbor Laboratory Press, 25(8),

pp. 1182-1195. doi: 10.1101/gr.188680.114.

Lodha, Marco y Timmermans (2013) «The ASYMMETRIC LEAVES complex maintains repression of KNOX homeobox genes via direct recruitment of Polycomb-repressive complex2.», *Genes & development*. Cold Spring Harbor Laboratory Press, 27(6), pp. 596-601. doi: 10.1101/gad.211425.112.

Loyola *et al.* (2003) «Functional Analysis of the Subunits of the Chromatin Assembly Factor RSF», *Molecular and Cellular Biology*. American Society for Microbiology, 23(19), pp. 6759-6768. doi: 10.1128/mcb.23.19.6759-6768.2003.

Lu *et al.* (2008) «Comparative Analysis of JmjC Domain-containing Proteins Reveals the Potential Histone Demethylases in *Arabidopsis* and Rice», *Journal of Integrative Plant Biology*, 50(7), pp. 886-896. doi: 10.1111/j.1744-7909.2008.00692.x.

De Lucia *et al.* (2008) «A PHD-Polycomb Repressive Complex 2 triggers the epigenetic silencing of FLC during vernalization», *Proceedings of the National Academy of Sciences*, 105(44), pp. 16831-16836. doi: 10.1073/pnas.0808687105.

Luk *et al.* (2010) «Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome.», *Cell*, 143(5), pp. 725-36. doi: 10.1016/j.cell.2010.10.019.

Ma *et al.* (2017) «A D53 repression motif induces oligomerization of TOPLESS corepressors and promotes assembly of a corepressor-nucleosome complex», *Science Advances*. American Association for the Advancement of Science, 3(6). doi: 10.1126/sciadv.1601217.

Mavrich *et al.* (2008) «Nucleosome organization in the *Drosophila* genome», *Nature*, 453(7193), pp. 358-362. doi: 10.1038/nature06929.

Mendenhall *et al.* (2010) «GC-rich sequence elements recruit PRC2 in mammalian ES cells», *PLoS Genetics*, 6(12), pp. 1-10. doi: 10.1371/journal.pgen.1001244.

Merini *et al.* (2017) «The *Arabidopsis* Polycomb Repressive Complex 1 (PRC1) Components AtBMI1A, B, and C Impact Gene Networks throughout All Stages of Plant Development», *Plant Physiology*. American Society of Plant Biologists, 173(1), pp. 627-641. doi: 10.1104/pp.16.01259.

Merini y Calonje (2015) «PRC1 is taking the lead in PcG repression», *Plant Journal*, 83(1). doi: 10.1111/tpj.12818.

Mieczkowski *et al.* (2016) «MNase titration reveals differences between nucleosome occupancy and chromatin accessibility», *Nature Communications*. Nature Publishing Group, 7(1), p. 11485. doi: 10.1038/ncomms11485.

Mizuguchi *et al.* (2004) «ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex.», *Science (New York, N.Y.)*, 303(5656), pp. 343-8. doi: 10.1126/science.1090701.

Molitor *et al.* (2014) «*Arabidopsis* AL PHD-PRC1 Complexes Promote Seed Germination through H3K4me3-to-H3K27me3 Chromatin State Switch in Repression of Seed Developmental Genes», *PLoS Genetics*. Editado por J. Goodrich. Public Library of Science, 10(1), p. e1004091. doi: 10.1371/journal.pgen.1004091.

Molitor *et al.* (2016) «The *Arabidopsis* hnRNP-Q Protein LIF2 and the PRC1 Subunit

LHP1 Function in Concert to Regulate the Transcription of Stress-Responsive Genes.», *The Plant Cell*. American Society of Plant Biologists, 28(9), pp. 2197-2211. doi: 10.1105/tpc.16.00244.

Moon *et al.* (2003) «EMF Genes Maintain Vegetative Development by Repressing the Flower Program in Arabidopsis», *The Plant Cell*, 15(3), pp. 681-693. doi: 10.1105/tpc.007831.

Morao, Bouyer y Roudier (2016) «Emerging concepts in chromatin-level regulation of plant cell differentiation: timing, counting, sensing and maintaining», *Current Opinion in Plant Biology*. Elsevier Ltd, pp. 27-34. doi: 10.1016/j.pbi.2016.07.010.

Morrison y Shen (2009) «Chromatin remodelling beyond transcription: the INO80 and SWR1 complexes», *Nature Reviews Molecular Cell Biology*, 10(6), pp. 373-384. doi: 10.1038/nrm2693.

Mortazavi *et al.* (2008) «Mapping and quantifying mammalian transcriptomes by RNA-Seq», *Nature Methods*, 5(7), pp. 621-628. doi: 10.1038/nmeth.1226.

Mozgova, Köhler y Hennig (2015) «Keeping the gate closed: functions of the polycomb repressive complex PRC2 in development», *The Plant Journal*. John Wiley & Sons, Ltd (10.1111), 83(1), pp. 121-132. doi: 10.1111/tpj.12828.

Mujtaba, Zeng y Zhou (2007) «Structure and acetyl-lysine recognition of the bromodomain», *Oncogene*. Nature Publishing Group, 26(37), pp. 5521-5527. doi: 10.1038/sj.onc.1210618.

Müller *et al.* (2002) «Histone methyltransferase activity of a Drosophila Polycomb group repressor complex.», *Cell*. Elsevier, 111(2), pp. 197-208. doi: 10.1016/S0092-8674(02)00976-5.

Müller y Kassis (2006) «Polycomb response elements and targeting of Polycomb group proteins in Drosophila», *Current Opinion in Genetics & Development*, 16(5), pp. 476-484. doi: 10.1016/j.gde.2006.08.005.

Murashige y Skoog (1962) «A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures», *Physiologia Plantarum*. John Wiley & Sons, Ltd (10.1111), 15(3), pp. 473-497. doi: 10.1111/j.1399-3054.1962.tb08052.x.

Murawska *et al.* (2008) «dCHD3, a novel ATP-dependent chromatin remodeler associated with sites of active transcription.», *Molecular and cellular biology*, 28(8), pp. 2745-57. doi: 10.1128/MCB.01839-07.

Musselman *et al.* (2012) «Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1», *Nature Structural & Molecular Biology*, 19(12), pp. 1266-1272. doi: 10.1038/nsmb.2435.

Mylne *et al.* (2006) «LHP1, the Arabidopsis homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC», *Proceedings of the National Academy of Sciences*, 103(13), pp. 5012-5017. doi: 10.1073/pnas.0507427103.

Nakamura, Lynch y Finkelstein (2001) «Physical interactions between ABA response loci of Arabidopsis», *The Plant Journal*. John Wiley & Sons, Ltd (10.1111), 26(6), pp. 627-635. doi: 10.1046/j.1365-313x.2001.01069.x.

Narlikar, Sundaramoorthy y Owen-Hughes (2013) «Mechanisms and Functions of

ATP-Dependent Chromatin-Remodeling Enzymes», *Cell*, 154(3), pp. 490-503. doi: 10.1016/j.cell.2013.07.011.

Nekrasov *et al.* (2007) «Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes», *The EMBO Journal*. EMBO Press, 26(18), pp. 4078-4088. doi: 10.1038/sj.emboj.7601837.

Ng *et al.* (2007) «Plant SET domain-containing proteins: Structure, function and regulation», *Biochimica et Biophysica Acta - Gene Structure and Expression*, pp. 316-329. doi: 10.1016/j.bbaexp.2007.04.003.

Nishida (2012) «Nucleosome Positioning.», *ISRN molecular biology*. Hindawi Limited, 2012, p. 245706. doi: 10.5402/2012/245706.

Ohad *et al.* (1999) «Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization.», *The Plant cell*, 11(3), pp. 407-16. doi: 10.1105/tpc.11.3.407.

Ojolo *et al.* (2018) «Regulation of plant growth and development: a review from a chromatin remodeling perspective», *Frontiers in Plant Science*. Frontiers Media S.A. doi: 10.3389/fpls.2018.01232.

Okulski *et al.* (2011) «Quantitative analysis of polycomb response elements (PREs) at identical genomic locations distinguishes contributions of PRE sequence and genomic environment», *Epigenetics and Chromatin*, 4(1). doi: 10.1186/1756-8935-4-4.

Pajoro *et al.* (2018) «Profiling Nucleosome Occupancy by MNase-seq: Experimental Protocol and Computational Analysis», en *Methods in molecular biology (Clifton, N.J.)*, pp. 167-181. doi: 10.1007/978-1-4939-7318-7_11.

Pandey (2002) «Analysis of histone acetyltransferase and histone deacetylase families of *Arabidopsis thaliana* suggests functional diversification of chromatin modification among multicellular eukaryotes», *Nucleic Acids Research*, 30(23), pp. 5036-5055. doi: 10.1093/nar/gkf660.

Papamichos-Chronakis *et al.* (2011) «Global Regulation of H2A.Z Localization by the INO80 Chromatin-Remodeling Enzyme Is Essential for Genome Integrity», *Cell*, 144(2), pp. 200-213. doi: 10.1016/j.cell.2010.12.021.

Park *et al.* (2011) «EMF1 Interacts with EIP1, EIP6 or EIP9 Involved in the Regulation of Flowering Time in *Arabidopsis*», *Plant and Cell Physiology*. Narnia, 52(8), pp. 1376-1388. doi: 10.1093/pcp/pcr084.

Partensky y Narlikar (2009) «Chromatin Remodelers Act Globally, Sequence Positions Nucleosomes Locally», *Journal of Molecular Biology*, 391(1), pp. 12-25. doi: 10.1016/j.jmb.2009.04.085.

Pass *et al.* (2017) «Genome-wide chromatin mapping with size resolution reveals a dynamic sub-nucleosomal landscape in *Arabidopsis*», *PLOS Genetics*. Editado por O. Mittelsten Scheid. Public Library of Science, 13(9), p. e1006988. doi: 10.1371/journal.pgen.1006988.

Peng *et al.* (2018) «Structural Analysis of the *Arabidopsis* AL2-PAL and PRC1 Complex Provides Mechanistic Insight into Active-to-Repressive Chromatin State Switch», *Journal of Molecular Biology*. Academic Press, 430(21), pp. 4245-4259. doi:

10.1016/j.jmb.2018.08.021.

Perruc, Kinoshita y Lopez-Molina (2007) «The role of chromatin-remodeling factor PKL in balancing osmotic stress responses during Arabidopsis seed germination.», *The Plant journal : for cell and molecular biology*, 52(5), pp. 927-36. doi: 10.1111/j.1365-313X.2007.03288.x.

Peterson *et al.* (2004) «Requirement for sex comb on midleg protein interactions in Drosophila polycomb group repression.», *Genetics*. Genetics Society of America, 167(3), pp. 1225-39. doi: 10.1534/genetics.104.027474.

Petesch y Lis (2012) «Overcoming the nucleosome barrier during transcript elongation», *Trends in Genetics*. Elsevier Current Trends, 28(6), pp. 285-294. doi: 10.1016/J.TIG.2012.02.005.

Phelan, Schnitzler y Kingston (2000) «Octamer Transfer and Creation of Stably Remodeled Nucleosomes by Human SWI-SNF and Its Isolated ATPases», *Molecular and Cellular Biology*, 20(17), pp. 6380-6389. doi: 10.1128/MCB.20.17.6380-6389.2000.

Pien y Grossniklaus (2007) «Polycomb group and trithorax group proteins in Arabidopsis», *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*, 1769(5-6), pp. 375-382. doi: 10.1016/j.bbaexp.2007.01.010.

Pu *et al.* (2013) «EMBRYONIC FLOWER1 and ULTRAPETALA1 Act Antagonistically on Arabidopsis Development and Stress Response.», *Plant physiology*. American Society of Plant Biologists, 162(2), pp. 812-30. doi: 10.1104/pp.112.213223.

Qüesta *et al.* (2016) «Arabidopsis transcriptional repressor VAL1 triggers Polycomb silencing at FLC during vernalization.», *Science (New York, N.Y.)*. American Association for the Advancement of Science, 353(6298), pp. 485-8. doi: 10.1126/science.aaf7354.

Radman-Livaja y Rando (2010) «Nucleosome positioning: How is it established, and why does it matter?», *Developmental Biology*. Academic Press, 339(2), pp. 258-266. doi: 10.1016/J.YDBIO.2009.06.012.

Rafati *et al.* (2011) «Repressive LTR Nucleosome Positioning by the BAF Complex Is Required for HIV Latency», *PLoS Biology*. Editado por M. Emerman, 9(11), p. e1001206. doi: 10.1371/journal.pbio.1001206.

Reinhart *et al.* (2002) «MicroRNAs in plants», *Genes & Development*, 16(13), pp. 1616-1626. doi: 10.1101/gad.1004402.

Ringrose y Paro (2007) «Polycomb/Trithorax response elements and epigenetic memory of cell identity.», *Development (Cambridge, England)*. The Company of Biologists Ltd, 134(2), pp. 223-32. doi: 10.1242/dev.02723.

Rinn *et al.* (2007) «Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs», *Cell*, 129(7), pp. 1311-1323. doi: 10.1016/j.cell.2007.05.022.

Rizzo y Sinha (2014) «Analyzing the Global Chromatin Structure of Keratinocytes by MNase-Seq», en pp. 49-59. doi: 10.1007/7651_2014_77.

Rodrigues *et al.* (2010) «Polycomb group gene function in sexual and asexual seed development in angiosperms», *Sexual Plant Reproduction*, 23(2), pp. 123-133. doi: 10.1007/s00497-009-0131-2.

Rosa y Shaw (2013) «Insights into chromatin structure and dynamics in plants», *Biology*. MDPI AG, pp. 1378-1410. doi: 10.3390/biology2041378.

Roudier *et al.* (2011) «Integrative epigenomic mapping defines four main chromatin states in Arabidopsis», *The EMBO Journal*, 30(10), pp. 1928-1938. doi: 10.1038/emboj.2011.103.

Routh, Sandin y Rhodes (2008) «Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure», *Proceedings of the National Academy of Sciences of the United States of America*, 105(26), pp. 8872-8877. doi: 10.1073/pnas.0802336105.

Sacharowski *et al.* (2015) «SWP73 Subunits of Arabidopsis SWI/SNF Chromatin Remodeling Complexes Play Distinct Roles in Leaf and Flower Development», *The Plant Cell*, 27(7), pp. 1889-1906. doi: 10.1105/tpc.15.00233.

Sadeh y Allis (2011) «Genome-wide “Re”-Modeling of Nucleosome Positions», *Cell*. Cell Press, 147(2), pp. 263-266. doi: 10.1016/J.CELL.2011.09.042.

Sala *et al.* (2011) «Genome-wide characterization of chromatin binding and nucleosome spacing activity of the nucleosome remodelling ATPase ISWI», *The EMBO Journal*, 30(9), pp. 1766-1777. doi: 10.1038/emboj.2011.98.

Sanchez-Pulido *et al.* (2008) «RAWUL: A new ubiquitin-like domain in PRC1 Ring finger proteins that unveils putative plant and worm PRC1 orthologs», *BMC Genomics*. BioMed Central, 9(1), p. 308. doi: 10.1186/1471-2164-9-308.

Sandelin (2004) «JASPAR: an open-access database for eukaryotic transcription factor binding profiles», *Nucleic Acids Research*. Oxford University Press (OUP), 32(90001), p. 91D-94. doi: 10.1093/nar/gkh012.

Santos-Mendoza *et al.* (2008) «Deciphering gene regulatory networks that control seed development and maturation in Arabidopsis», *The Plant Journal*. John Wiley & Sons, Ltd (10.1111), 54(4), pp. 608-620. doi: 10.1111/j.1365-313X.2008.03461.x.

Sarma *et al.* (2008) «Ezh2 Requires PHF1 To Efficiently Catalyze H3 Lysine 27 Trimethylation In Vivo», *Molecular and Cellular Biology*, 28(8), pp. 2718-2731. doi: 10.1128/MCB.02017-07.

Sasnauskas, Kauneckaitė y Siksnyš (2018) «Structural basis of DNA target recognition by the B3 domain of Arabidopsis epigenome reader VAL1», *Nucleic Acids Research*, 46(8), pp. 4316-4324. doi: 10.1093/nar/gky256.

Schaumberg *et al.* (2015) «Quantitative characterization of genetic parts and circuits for plant synthetic biology», *Nature Methods*. Nature Publishing Group, 13(1), pp. 94-100. doi: 10.1038/nmeth.3659.

Scheuermann *et al.* (2010) «Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB», *Nature*, 465(7295), pp. 243-247. doi: 10.1038/nature08966.

Scheuermann, Gutiérrez y Müller (2012) «Histone H2A monoubiquitination and Polycomb repression», *Fly*, 6(3), pp. 162-168. doi: 10.4161/fly.20986.

Schmidt *et al.* (2013) «The Polycomb group protein MEDEA and the DNA methyltransferase MET1 interact to repress autonomous endosperm development in Arabidopsis», *The Plant Journal*, 73(5), pp. 776-787. doi: 10.1111/tpj.12070.

Schneider *et al.* (2016) «Potential targets of VIVIPAROUS1/ABI3-LIKE1 (VAL1) repression in developing *Arabidopsis thaliana* embryos», *The Plant Journal*, 85(2), pp. 305-319. doi: 10.1111/tpj.13106.

Schones *et al.* (2008) «Dynamic Regulation of Nucleosome Positioning in the Human Genome», *Cell*, 132(5), pp. 887-898. doi: 10.1016/j.cell.2008.02.022.

Schonrock *et al.* (2006) «Polycomb-group proteins repress the floral activator AGL19 in the FLC-independent vernalization pathway», *Genes & Development*, 20(12), pp. 1667-1678. doi: 10.1101/gad.377206.

Schoorlemmer *et al.* (1997) «Ring1A is a transcriptional repressor that interacts with the Polycomb-M33 protein and is expressed at rhombomere boundaries in the mouse hindbrain», *The EMBO Journal*, 16(19), pp. 5930-5942. doi: 10.1093/emboj/16.19.5930.

Schubert (2019) «Evolution of polycomb-group function in the green lineage [version 1; referees: 2 approved]», *F1000Research*. F1000 Research Ltd. doi: 10.12688/f1000research.16986.1.

Schwartz y Pirrotta (2013) «A new world of Polycombs: unexpected partnerships and emerging functions», *Nature Reviews Genetics*. Nature Publishing Group, 14(12), pp. 853-864. doi: 10.1038/nrg3603.

Segal y Widom (2009) «What controls nucleosome positions?», *Trends in Genetics*. Elsevier Current Trends, 25(8), pp. 335-343. doi: 10.1016/J.TIG.2009.06.002.

Sequeira-Mendes *et al.* (2014) «The functional topography of the Arabidopsis genome is organized in a reduced number of linear motifs of chromatin states», *Plant Cell*. American Society of Plant Biologists, 26(6), pp. 2351-2366. doi: 10.1105/tpc.114.124578.

Servet, Conde E Silva y Zhou (2010) «Histone acetyltransferase AtGCN5/HAG1 is a versatile regulator of developmental and inducible gene expression in Arabidopsis», *Molecular Plant*. Oxford University Press, 3(4), pp. 670-677. doi: 10.1093/mp/ssq018.

Shaked, Avivi-Ragolsky y Levy (2006) «Involvement of the Arabidopsis SWI2/SNF2 chromatin remodeling gene family in DNA damage response and recombination», *Genetics*, 173(2), pp. 985-994. doi: 10.1534/genetics.105.051664.

Shao *et al.* (1999) «Stabilization of chromatin structure by PRC1, a Polycomb complex.», *Cell*. Elsevier, 98(1), pp. 37-46. doi: 10.1016/S0092-8674(00)80604-2.

Shen *et al.* (2014) «The putative PRC1 RING-finger protein AtRING1A regulates flowering through repressing MADS AFFECTING FLOWERING genes in Arabidopsis», *Development*, 141(6), pp. 1303-1312. doi: 10.1242/dev.104513.

Shivaswamy *et al.* (2008) «Dynamic Remodeling of Individual Nucleosomes Across a Eukaryotic Genome in Response to Transcriptional Perturbation», *PLoS Biology*. Editado por O. J. Rando, 6(3), p. e65. doi: 10.1371/journal.pbio.0060065.

Shu *et al.* (2012) «Distinct modes of DNA accessibility in plant chromatin», *Nature Communications*. Nature Publishing Group, 3(1), p. 1281. doi: 10.1038/ncomms2259.

Shu *et al.* (2016) «Two Faces of One Seed: Hormonal Regulation of Dormancy and Germination.», *Molecular plant*. Elsevier, 9(1), pp. 34-45. doi: 10.1016/j.molp.2015.08.010.

Simon *et al.* (2012) «Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA», *Nature Protocols*, 7(2), pp. 256-267. doi: 10.1038/nprot.2011.444.

Simon *et al.* (2013) «A Detailed Protocol for Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE)», en *Current Protocols in Molecular Biology*. Hoboken, NJ, USA: John Wiley & Sons, Inc., p. Unit21.26. doi: 10.1002/0471142727.mb2126s102.

Simon y Kingston (2009) «Mechanisms of Polycomb gene silencing: knowns and unknowns», *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 10(10), pp. 697-708. doi: 10.1038/nrm2763.

Jeffrey A. Simon y Kingston (2013) «Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put.», *Molecular cell*. Elsevier, 49(5), pp. 808-24. doi: 10.1016/j.molcel.2013.02.013.

Jeffrey A. Simon y Kingston (2013) «Occupying Chromatin: Polycomb Mechanisms for Getting to Genomic Targets, Stopping Transcriptional Traffic, and Staying Put», *Molecular Cell*, pp. 808-824. doi: 10.1016/j.molcel.2013.02.013.

Jeffrey A.A. Simon y Kingston (2013) «Occupying Chromatin: Polycomb Mechanisms for Getting to Genomic Targets, Stopping Transcriptional Traffic, and Staying Put», *Molecular Cell*. Cell Press, 49(5), pp. 808-824. doi: 10.1016/J.MOLCEL.2013.02.013.

Sing *et al.* (2009) «A Vertebrate Polycomb Response Element Governs Segmentation of the Posterior Hindbrain», *Cell*, 138(5), pp. 885-897. doi: 10.1016/j.cell.2009.08.020.

Song *et al.* (2005) «Role of an Arabidopsis AP2/EREBP-type transcriptional repressor in abscisic acid and drought stress responses», *Plant Cell*, 17(8), pp. 2384-2396. doi: 10.1105/tpc.105.033043.

Song *et al.* (2017) «MYC5 is Involved in Jasmonate-Regulated Plant Growth, Leaf Senescence and Defense Responses», *Plant and Cell Physiology*, 58(10), pp. 1752-1763. doi: 10.1093/pcp/pcx112.

Song y Crawford (2010) «DNase-seq: A High-Resolution Technique for Mapping Active Gene Regulatory Elements across the Genome from Mammalian Cells», *Cold Spring Harbor Protocols*, 2010(2), p. pdb.prot5384-pdb.prot5384. doi: 10.1101/pdb.prot5384.

Song y Galbraith (2006) «AtSAP18, an orthologue of human SAP18, is involved in the regulation of salt stress and mediates transcriptional repression in Arabidopsis», *Plant Molecular Biology*, 60(2), pp. 241-257. doi: 10.1007/s11103-005-3880-9.

Sowpati, Ramamoorthy y Mishra (2015) «Expansion of the polycomb system and evolution of complexity», *Mechanisms of Development*. Elsevier Ireland Ltd, 138, pp. 97-112. doi: 10.1016/j.mod.2015.07.013.

Springer *et al.* (2003) «Comparative analysis of SET domain proteins in maize and Arabidopsis reveals multiple duplications preceding the divergence of monocots and dicots», *Plant Physiology*, 132(2), pp. 907-925. doi: 10.1104/pp.102.013722.

Steffens (2004) «AthaMap: an online resource for in silico transcription factor binding sites in the Arabidopsis thaliana genome», *Nucleic Acids Research*. Oxford University Press (OUP), 32(90001), p. 368D-372. doi: 10.1093/nar/gkh017.

Stock *et al.* (2007) «Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells», *Nature Cell Biology*, 9(12), pp. 1428-1435. doi: 10.1038/ncb1663.

Strahl y Allis (2000) «The language of covalent histone modifications», *Nature*, pp. 41-45. doi: 10.1038/47412.

Struhl y Segal (2013) «Determinants of nucleosome positioning», *Nature Structural & Molecular Biology*. Nature Publishing Group, 20(3), pp. 267-273. doi: 10.1038/nsmb.2506.

Sun *et al.* (2009) «A timing mechanism for stem cell maintenance and differentiation in the Arabidopsis floral meristem», *Genes and Development*, 23(15), pp. 1791-1804. doi: 10.1101/gad.1800409.

Sun *et al.* (2014) «Timing mechanism dependent on cell division is invoked by Polycomb eviction in plant stem cells», *Science*. American Association for the Advancement of Science, 343(6170). doi: 10.1126/science.1248559.

Sung *et al.* (1992) «EMF, an Arabidopsis Gene Required for Vegetative Shoot Development», *Science*, 258(5088), pp. 1645-1647. doi: 10.1126/science.258.5088.1645.

Sung y Amasino (2004) «Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3», *Nature*, 427(6970), pp. 159-164. doi: 10.1038/nature02195.

Sung, Schmitz y Amasino (2006) «A PHD finger protein involved in both the vernalization and photoperiod pathways in Arabidopsis», *Genes & Development*, 20(23), pp. 3244-3248. doi: 10.1101/gad.1493306.

Supek *et al.* (2011) «REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms», *PLoS ONE*. Editado por C. Gibas. Public Library of Science, 6(7), p. e21800. doi: 10.1371/journal.pone.0021800.

Suzuki (1997) «The Conserved B3 Domain of VIVIPAROUS1 Has a Cooperative DNA Binding Activity», *THE PLANT CELL ONLINE*, 9(5), pp. 799-807. doi: 10.1105/tpc.9.5.799.

Suzuki, Ketterling y McCarty (2005) «Quantitative statistical analysis of cis-regulatory sequences in ABA/VP1- and CBF/DREB1-regulated genes of Arabidopsis.», *Plant physiology*. American Society of Plant Biologists, 139(1), pp. 437-47. doi: 10.1104/pp.104.058412.

Suzuki y McCarty (2008) «Functional symmetry of the B3 network controlling seed development», *Current Opinion in Plant Biology*, 11(5), pp. 548-553. doi: 10.1016/j.pbi.2008.06.015.

Suzuki, Wang y McCarty (2007) «Repression of the LEAFY COTYLEDON 1/B3 regulatory network in plant embryo development by VP1/ABSCISIC ACID INSENSITIVE 3-LIKE B3 genes», *Plant Physiology*, 143(2), pp. 902-911. doi: 10.1104/pp.106.092320.

Swaminathan, Peterson y Jack (2008) «The plant B3 superfamily.», *Trends in plant science*. Elsevier, 13(12), pp. 647-55. doi: 10.1016/j.tplants.2008.09.006.

Swiezewski *et al.* (2009) «Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target», *Nature*. Nature Publishing Group, 462(7274), pp. 799-802. doi: 10.1038/nature08618.

Swygert y Peterson (2014) «Chromatin dynamics: Interplay between remodeling

enzymes and histone modifications», *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. Elsevier, pp. 728-736. doi: 10.1016/j.bbagr.2014.02.013.

Talbert y Henikoff (2014) «Environmental responses mediated by histone variants», *Trends in Cell Biology*. Elsevier Ltd, pp. 642-650. doi: 10.1016/j.tcb.2014.07.006.

Tang *et al.* (2008) «The Arabidopsis BRAHMA chromatin-remodeling ATPase is involved in repression of seed maturation genes in leaves.», *Plant physiology*, 147(3), pp. 1143-57. doi: 10.1104/pp.108.121996.

Tavares *et al.* (2012) «RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3», *Cell*, 148(4), pp. 664-678. doi: 10.1016/j.cell.2011.12.029.

Teif *et al.* (2012) «Genome-wide nucleosome positioning during embryonic stem cell development», *Nature Structural & Molecular Biology*. Nature Publishing Group, 19(11), pp. 1185-1192. doi: 10.1038/nsmb.2419.

Thireault *et al.* (2015) «Repression of jasmonate signaling by a non-TIFY JAZ protein in Arabidopsis», *The Plant Journal*, 82(4), pp. 669-679. doi: 10.1111/tpj.12841.

Thoma, Koller y Klug (1979) «Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin.», *The Journal of cell biology*. Rockefeller University Press, 83(2 Pt 1), pp. 403-27. doi: 10.1083/jcb.83.2.403.

Tonkin *et al.* (2002) «Identification and characterisation of novel mammalian homologues of Drosophila polyhomeotic permits new insights into relationships between members of the polyhomeotic family», *Human Genetics*, 111(4-5), pp. 435-442. doi: 10.1007/s00439-002-0814-3.

Torres y Deal (2019) «The histone variant H2A.Z and chromatin remodeler BRAHMA act coordinately and antagonistically to regulate transcription and nucleosome dynamics in Arabidopsis», *The Plant Journal*. John Wiley & Sons, Ltd (10.1111), 99(1), p. tpj.14281. doi: 10.1111/tpj.14281.

Trapnell *et al.* (2012) «Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks», *Nature Protocols*. Nature Publishing Group, 7(3), pp. 562-578. doi: 10.1038/nprot.2012.016.

Trojer *et al.* (2007) «L3MBTL1, a Histone-Methylation-Dependent Chromatin Lock», *Cell*, 129(5), pp. 915-928. doi: 10.1016/j.cell.2007.03.048.

Trojer *et al.* (2011) «L3MBTL2 Protein Acts in Concert with PcG Protein-Mediated Monoubiquitination of H2A to Establish a Repressive Chromatin Structure», *Molecular Cell*, 42(4), pp. 438-450. doi: 10.1016/j.molcel.2011.04.004.

Tsai *et al.* (2010) «Long noncoding RNA as modular scaffold of histone modification complexes.», *Science (New York, N.Y.)*. NIH Public Access, 329(5992), pp. 689-93. doi: 10.1126/science.1192002.

Turck *et al.* (2007) «Arabidopsis TFL2/LHP1 Specifically Associates with Genes Marked by Trimethylation of Histone H3 Lysine 27», *PLoS Genetics*. Public Library of Science, 3(6), p. e86. doi: 10.1371/journal.pgen.0030086.

Valouev *et al.* (2008) «A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning.», *Genome research*. Cold Spring

Harbor Laboratory Press, 18(7), pp. 1051-63. doi: 10.1101/gr.076463.108.

Valouev *et al.* (2011) «Determinants of nucleosome organization in primary human cells», *Nature*, 474(7352), pp. 516-520. doi: 10.1038/nature10002.

Veerappan *et al.* (2012) «A novel HSI2 mutation in Arabidopsis affects the PHD-like domain and leads to derepression of seed-specific gene expression», *Planta*, 236(1), pp. 1-17. doi: 10.1007/s00425-012-1630-1.

Veerappan *et al.* (2014) «HSI2/VAL1 PHD-like domain promotes H3K27 trimethylation to repress the expression of seed maturation genes and complex transgenes in Arabidopsis seedlings», *BMC Plant Biology*. BioMed Central Ltd., 14(1). doi: 10.1186/s12870-014-0293-4.

Veluchamy *et al.* (2016) «LHP1 Regulates H3K27me3 Spreading and Shapes the Three-Dimensional Conformation of the Arabidopsis Genome», *PLOS ONE*. Editado por M. Bendahmane, 11(7), p. e0158936. doi: 10.1371/journal.pone.0158936.

Vicente-Carbajosa y Carbonero (2005) «Seed maturation: developing an intrusive phase to accomplish a quiescent state», *The International Journal of Developmental Biology*. UPV/EHU Press, 49(5-6), pp. 645-651. doi: 10.1387/ijdb.052046jc.

Voong *et al.* (2016) «Insights into Nucleosome Organization in Mouse Embryonic Stem Cells through Chemical Mapping», *Cell*. Cell Press, 167(6), p. 1555-1570.e15. doi: 10.1016/J.CELL.2016.10.049.

Wang *et al.* (2004) «Hierarchical Recruitment of Polycomb Group Silencing Complexes», *Molecular Cell*, 14(5), pp. 637-646. doi: 10.1016/j.molcel.2004.05.009.

Wang *et al.* (2006) «Partially redundant functions of two SET-domain polycomb-group proteins in controlling initiation of seed development in Arabidopsis.», *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 103(35), pp. 13244-9. doi: 10.1073/pnas.0605551103.

Wang *et al.* (2014) «Photoperiodic control of the floral transition through a distinct polycomb repressive complex.», *Developmental cell*. Elsevier, 28(6), pp. 727-36. doi: 10.1016/j.devcel.2014.01.029.

Hua Wang *et al.* (2016) «Arabidopsis Flower and Embryo Developmental Genes are Repressed in Seedlings by Different Combinations of Polycomb Group Proteins in Association with Distinct Sets of Cis-regulatory Elements», *PLOS Genetics*. Editado por V. Colot. Public Library of Science, 12(1), p. e1005771. doi: 10.1371/journal.pgen.1005771.

Jingjing Wang *et al.* (2016) «The roles of cross-talk epigenetic patterns in *Arabidopsis thaliana*», *Briefings in Functional Genomics*. Narnia, 15(4), pp. 278-287. doi: 10.1093/bfpg/elv025.

Wang, Czech y Weigel (2009) «miR156-Regulated SPL Transcription Factors Define an Endogenous Flowering Pathway in Arabidopsis thaliana», *Cell*, 138(4), pp. 738-749. doi: 10.1016/j.cell.2009.06.014.

Wang y Perry (2013) «Identification of Direct Targets of FUSCA3, a Key Regulator of Arabidopsis Seed Development», *Plant Physiology*, 161(3), pp. 1251-1264. doi: 10.1104/pp.112.212282.

Weber, Ramachandran y Henikoff (2014) «Nucleosomes Are Context-Specific,

H2A.Z-Modulated Barriers to RNA Polymerase», *Molecular Cell*, 53(5), pp. 819-830. doi: 10.1016/j.molcel.2014.02.014.

Wei *et al.* (2015) «The Alfin-like homeodomain finger protein AL5 suppresses multiple negative factors to confer abiotic stress tolerance in Arabidopsis», *The Plant Journal*. John Wiley & Sons, Ltd (10.1111), 81(6), pp. 871-883. doi: 10.1111/tpj.12773.

A. Weiner *et al.* (2010) «High-resolution nucleosome mapping reveals transcription-dependent promoter packaging», *Genome Research*, 20(1), pp. 90-100. doi: 10.1101/gr.098509.109.

Assaf Weiner *et al.* (2010) «High-resolution nucleosome mapping reveals transcription-dependent promoter packaging», *Genome Research*, 20(1), pp. 90-100. doi: 10.1101/gr.098509.109.

Weintraub y Groudine (1976) «Chromosomal subunits in active genes have an altered conformation», *Science*, 193(4256), pp. 848-856. doi: 10.1126/science.948749.

Winger y Bowman (2017) «The Sequence of Nucleosomal DNA Modulates Sliding by the Chd1 Chromatin Remodeler», *Journal of Molecular Biology*, 429(6), pp. 808-822. doi: 10.1016/j.jmb.2017.02.002.

Wolff *et al.* (2011) «High-Resolution Analysis of Parent-of-Origin Allelic Expression in the Arabidopsis Endosperm», *PLoS Genetics*. Editado por T. Kakutani. Public Library of Science, 7(6), p. e1002126. doi: 10.1371/journal.pgen.1002126.

Woo *et al.* (2010) «A Region of the Human HOXD Cluster that Confers Polycomb-Group Responsiveness», *Cell*, 140(1), pp. 99-110. doi: 10.1016/j.cell.2009.12.022.

Wood *et al.* (2006) «The Arabidopsis thaliana vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3», *Proceedings of the National Academy of Sciences*, 103(39), pp. 14631-14636. doi: 10.1073/pnas.0606385103.

Woodcock y Ghosh (2010) «Chromatin higher-order structure and dynamics.», *Cold Spring Harbor perspectives in biology*. Cold Spring Harbor Laboratory Press, 2(5), p. a000596. doi: 10.1101/cshperspect.a000596.

Wu *et al.* (2009) «The Sequential Action of miR156 and miR172 Regulates Developmental Timing in Arabidopsis», *Cell*, 138(4), pp. 750-759. doi: 10.1016/j.cell.2009.06.031.

Wu *et al.* (2011) «Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells», *Nature*. Nature Publishing Group, 473(7347), pp. 389-393. doi: 10.1038/nature09934.

Wu *et al.* (2012) «SWI2/SNF2 chromatin remodeling ATPases overcome polycomb repression and control floral organ identity with the LEAFY and SEPALLATA3 transcription factors.», *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 109(9), pp. 3576-81. doi: 10.1073/pnas.1113409109.

Wu y Citovsky (2017) «Adaptor proteins GIR1 and GIR2. II. Interaction with the co-repressor TOPLESS and promotion of histone deacetylation of target chromatin», *Biochemical and Biophysical Research Communications*. Elsevier B.V., 488(4), pp. 609-613.

doi: 10.1016/j.bbrc.2017.05.085.

Wu y Poethig (2006) «Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3.», *Development (Cambridge, England)*. The Company of Biologists Ltd, 133(18), pp. 3539-47. doi: 10.1242/dev.02521.

Xiao *et al.* (2017) «Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in *Arabidopsis*», *Nature Genetics*. Nature Publishing Group, 49(10), pp. 1546-1552. doi: 10.1038/ng.3937.

Xiao y Wagner (2015) «Polycomb repression in the regulation of growth and development in *Arabidopsis*», *Current Opinion in Plant Biology*. Elsevier Current Trends, 23, pp. 15-24. doi: 10.1016/J.PBI.2014.10.003.

Xu *et al.* (2016) «Epigenetic regulation of vegetative phase change in *Arabidopsis*», *Plant Cell*. American Society of Plant Biologists, 28(1), pp. 28-41. doi: 10.1105/tpc.15.00854.

Xu *et al.* (2018) «SUPERMAN regulates floral whorl boundaries through control of auxin biosynthesis », *The EMBO Journal*. EMBO, 37(11). doi: 10.15252/emboj.201797499.

Xu y Shen (2008) «Polycomb silencing of KNOX genes confines shoot stem cell niches in *Arabidopsis*.», *Current biology: CB*. Elsevier, 18(24), pp. 1966-71. doi: 10.1016/j.cub.2008.11.019.

Yadon *et al.* (2010) «Chromatin remodeling around nucleosome-free regions leads to repression of noncoding RNA transcription.», *Molecular and cellular biology*. American Society for Microbiology (ASM), 30(21), pp. 5110-22. doi: 10.1128/MCB.00602-10.

Yang *et al.* (2013) «VAL- and AtBMI1-mediated H2Aub initiate the switch from embryonic to postgerminative growth in *Arabidopsis*.», *Current biology: CB*. Elsevier, 23(14), pp. 1324-9. doi: 10.1016/j.cub.2013.05.050.

Yang *et al.* (2018) «PlantEAR: Functional Analysis Platform for Plant EAR Motif-Containing Proteins», *Frontiers in Genetics*. Frontiers Media SA, 9. doi: 10.3389/fgene.2018.00590.

Yang, Chen y Sung (1995) «Genetic Regulation of Shoot Development in *Arabidopsis*: Role of the EMF Genes», *Developmental Biology*. Academic Press, 169(2), pp. 421-435. doi: 10.1006/DBIO.1995.1158.

Yang, Howard y Dean (2014) «Antagonistic Roles for H3K36me3 and H3K27me3 in the Cold-Induced Epigenetic Switch at *Arabidopsis* FLC», *Current Biology*. Cell Press, 24(15), pp. 1793-1797. doi: 10.1016/J.CUB.2014.06.047.

Yoshida *et al.* (2001) «EMBRYONIC FLOWER2, a novel polycomb group protein homolog, mediates shoot development and flowering in *Arabidopsis*.», *The Plant cell*. American Society of Plant Biologists, 13(11), pp. 2471-81. doi: 10.1105/tpc.010227.

Yu *et al.* (2012) «clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters», *OMICS: A Journal of Integrative Biology*. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA , 16(5), pp. 284-287. doi: 10.1089/omi.2011.0118.

Yu, Wang y He (2015) «ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization», *Bioinformatics*, 31(14), pp. 2382-2383. doi:

10.1093/bioinformatics/btv145.

Yuan *et al.* (2016) «A cis cold memory element and a trans epigenome reader mediate Polycomb silencing of FLC by vernalization in Arabidopsis», *Nature Genetics*. Nature Publishing Group, 48(12), pp. 1527-1534. doi: 10.1038/ng.3712.

Zeng *et al.* (2019) «HISTONE DEACETYLASE 9 functions with Polycomb silencing to repress FLOWERING LOCUS C expression.», *Plant physiology*. doi: 10.1104/pp.19.00793.

Zhang *et al.* (2006) «Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in Arabidopsis», *Cell*, 126(6), pp. 1189-1201. doi: 10.1016/j.cell.2006.08.003.

Zhang *et al.* (2007) «The Arabidopsis LHP1 protein colocalizes with histone H3 Lys27 trimethylation», *Nature Structural & Molecular Biology*. Nature Publishing Group, 14(9), pp. 869-871. doi: 10.1038/nsmb1283.

Yong Zhang *et al.* (2008) «Model-based analysis of ChIP-Seq (MACS)», *Genome Biology*, 9(9). doi: 10.1186/gb-2008-9-9-r137.

Heng Zhang *et al.* (2008) «The CHD3 remodeler PICKLE promotes trimethylation of histone H3 lysine 27.», *The Journal of biological chemistry*, 283(33), pp. 22637-48. doi: 10.1074/jbc.M802129200.

Zhang *et al.* (2009) «Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in Arabidopsis thaliana», *Genome Biology*, 10(6). doi: 10.1186/gb-2009-10-6-r62.

Zhang *et al.* (2012) «The CHD3 Remodeler PICKLE Associates with Genes Enriched for Trimethylation of Histone H3 Lysine 27», *Plant Physiology*. American Society of Plant Biologists, 159(1), pp. 418-432. doi: 10.1104/PP.112.194878.

Zhang *et al.* (2017) «Role of remodeling and spacing factor 1 in histone H2A ubiquitination-mediated gene silencing», *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 114(38), pp. E7949-E7958. doi: 10.1073/pnas.1711158114.

Zhang *et al.* (2018) «Rice nucleosome patterns undergo remodeling coincident with stress-induced gene expression», *BMC Genomics*. BioMed Central Ltd., 19(1). doi: 10.1186/s12864-017-4397-8.

Zhang, Lang y Zhu (2018) «Dynamics and function of DNA methylation in plants», *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, pp. 489-506. doi: 10.1038/s41580-018-0016-z.

Zhang y Reese (2007) «Exposing the core promoter is sufficient to activate transcription and alter coactivator requirement at RNR3», *Proceedings of the National Academy of Sciences of the United States of America*, 104(21), pp. 8833-8838. doi: 10.1073/pnas.0701666104.

Zhang, Zhang y Jiang (2015) «Genome-Wide Nucleosome Occupancy and Positioning and Their Impact on Gene Expression and Evolution in Plants.», *Plant physiology*. American Society of Plant Biologists, 168(4), pp. 1406-16. doi: 10.1104/pp.15.00125.

Zhao *et al.* (2008) «Polycomb Proteins Targeted by a Short Repeat RNA to the

Mouse X Chromosome», *Science*, 322(5902), pp. 750-756. doi: 10.1126/science.1163045.

Zhou *et al.* (2008) «Histone H2A Monoubiquitination Represses Transcription by Inhibiting RNA Polymerase II Transcriptional Elongation», *Molecular Cell*, 29(1), pp. 69-80. doi: 10.1016/j.molcel.2007.11.002.

Zhou *et al.* (2013) «HISTONE DEACETYLASE19 Interacts with HSL1 and Participates in the Repression of Seed Maturation Genes in *Arabidopsis* Seedlings», *The Plant Cell*, 25(1), pp. 134-148. doi: 10.1105/tpc.112.096313.

Zhou *et al.* (2017) «H2A monoubiquitination in *Arabidopsis thaliana* is generally independent of LHP1 and PRC2 activity», *Genome Biology*, 18(1), p. 69. doi: 10.1186/s13059-017-1197-z.

Zhou *et al.* (2018) «Telobox motifs recruit CLF/SWN–PRC2 for H3K27me3 deposition via TRB factors in *Arabidopsis*», *Nature Genetics*. Nature Publishing Group, 50(5), pp. 638-644. doi: 10.1038/s41588-018-0109-9.

Zhu *et al.* (2013) «A SWI/SNF chromatin-remodeling complex acts in noncoding RNA-mediated transcriptional silencing.», *Molecular cell*, 49(2), pp. 298-309. doi: 10.1016/j.molcel.2012.11.011.