



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

Análisis de la regulación génica mediada por el
complejo epigenético Trithorax en *Arabidopsis*
thaliana

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

BIÓLOGO

P R E S E N T A :

CARLOS EMILIANO CORTÉS
QUÍÑONES

TUTORA:

DRA. MA. DE LA PAZ SÁNCHEZ
JIMÉNEZ



FACULTAD DE CIENCIAS
UNAM

Cd. Mx. 2022



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

1. Datos del alumno

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

2. Datos del tutor

Grado
Nombre(s)
Apellido paterno
Apellido materno

3. Datos del sinodal 1

Grado
Nombre(s)
Apellido paterno
Apellido materno

4. Datos del sinodal 2

Grado
Nombre(s)
Apellido paterno
Apellido materno

5. Datos del sinodal 3

Grado
Nombre(s)
Apellido paterno
Apellido materno

6. Datos del sinodal 4

Grado
Nombre(s)
Apellido paterno
Apellido materno

7. Datos del trabajo escrito.

Título

Subtítulo

Número de páginas

Año

1. Datos del alumno

Cortés
Quiñones
Carlos Emiliano
7771926644
Universidad Nacional Autónoma de México
Facultad de Ciencias
Biología
417092609

2. Datos del tutor

Dra.
Ma. de la Paz
Sánchez
Jiménez

3. Datos del sinodal 1

Dra.
Alicia
Gamboa
de Buen

4. Datos del sinodal 2

Dra.
Rosalinda
Tapia
López

5. Datos del sinodal 3

Dr.
Ulises Yunuen
Rosas
López

6. Datos del sinodal 4

Dr.
Rodrigo
González
Barrios de la Parra

7. Datos del trabajo escrito

Análisis de la regulación génica mediada por el complejo epigenético Trithorax en *Arabidopsis thaliana*

60 p

2022

Agradecimientos académicos

Esta tesis se realizó bajo la dirección de la Dra. Ma de la Paz Sánchez Jiménez, en el Laboratorio de Genética Molecular, Epigenética, Desarrollo y Evolución de Plantas del Instituto de Ecología, de la Universidad Nacional Autónoma de México, bajo la coordinación académico-científica de las Dras. Elena Alvarez-Buylla Roces, Adriana Garay Arroyo y Berenice García Ponce de León.

Agradezco la coordinación administrativa y logística, así como el apoyo técnico de la Dra. Diana Belén Sánchez Rodríguez.

Además, quiero agradecer al Dr. Diego Arturo Ornelas Ayala por su apoyo en enseñarme a utilizar todas las herramientas y técnicas experimentales que utilicé en este trabajo.

Por último, quisiera agradecer por su tiempo a los miembros del jurado al revisar este trabajo.

El financiamiento para esta investigación provino de proyectos de investigación:

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

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

Agradecimientos personales

A Jesús, porque, aunque a veces haces muy complicado el acercarme a ti, agradezco infinitamente que llegaras a mi vida, saber que estás ahí siempre me da sentido en lo caótica que puede llegar a ser nuestra familia. Lo único que le pido a la vida es que nunca nos convirtamos en extraños. Y pues, además, en sentido estricto sí me ayudaste en esta tesis jaja.

A ustedes papás por estar siempre ahí para mí, por siempre apoyarme en todo. Espero estén orgullosos. A ti ma por creer siempre en mí y echarme porras sin importar lo que sea. A ti pa por siempre ser una fuente inagotable de inspiración. Gracias a ustedes, hoy, soy científico.

A ti Abue Chio porque tú llevas siempre un pedazo de mi corazón contigo y aunque siempre digas que eres la rasposa de la familia, es difícil pensar en alguien que me haya dado más cariño que tú. Gracias por siempre ser mi cómplice en todo.

A José, porque en gran medida, gracias a ti logré tener las fuerzas para acabar esta cosa. Hemos pasado tantas cosas juntos que me es imposible imaginar la vida sin ti. En ti siempre encuentro la luz que le falta a mi vida.

A Melanie, que ya para estas alturas, eres parte de mí y espero también ser parte de ti. Me da mucha felicidad que después de tantos años, te pueda poner en esta tesis. Las cosas y nosotros hemos cambiado tanto, pero aquí seguimos juntos.

A Ingrid, por siempre darme apoyo, risas, complicidad, confianza y hasta peleas. Estaré eternamente agradecido con el destino por juntarnos en tantas clases en aquel lejano inicio de la carrera y así haber conocido a una de las personas más importantes en mi vida.

A Paloma, que estuvo chingue y chingue que quería su párrafo de agradecimiento. Y la realidad es que no había necesidad de insistir porque tu amistad es de las cosas más preciadas que tengo. En serio agradezco eternamente haber conocido a alguien tan especial como tú y que siempre estemos en sintonía.

A Emiliano por aceptarme no solo en su casa, sino en su vida, por haber construido entre los dos ese caótico hogar en Andrés Molina, y por nuestra conexión astral que siempre existirá. Deseo que siempre podamos seguir teniendo muchas aventuras juntos.

A Rodrigo por ser el mejor regalo que la carrera me pudo dar desde el primer día. Te admiro muchísimo y agradezco poder haber visto cómo te convertías en el mejor biólogo-menos biólogo de todos.

A Diego por básicamente enseñarme todo lo que sé hacer en el laboratorio. Por darme la confianza desde el primer día y dejarme mezclarle las semillas y echarle a perder incontables PCRs jaja. Gracias por ayudarnos a todxs en el laboratorio.

A Marypaz a la que agradezco el siempre tenerme paciencia aun cuando los *deadlines* no son lo mío. Gracias (y no gracias) por no dejarme tirar la toalla en esta tesis en la que tantas dificultades me he (y nos hemos) enfrentado. Gracias por ser una tutora que siempre está al pendiente y se preocupa de sus alumnxs, la academia necesita más gente como tú.

A Ani porque sin ti este último pedazo de la tesis no hubiera salido y por hacer esas interminables horas en el microscopio y viendo fotos de raíces más llevaderas.

Y finalmente, de mí para mí: *No seas tan exigente contigo*

ÍNDICE

RESUMEN	1
INTRODUCCIÓN	2
1. La epigenética	2
1.1 Conceptualización del término	2
1.2 Lamarckismo y epigenética	3
1.3 ¿Una definición moderna?: la necesidad de claridad conceptual ----	4
2. Mecanismos epigenéticos	5
2.1 Metilación del DNA	5
2.2 Regulación mediada por RNAs no codificantes	6
2.3 Estructura tridimensional del genoma	7
2.4 Variantes de Histonas	8
2.5 Modificaciones postraduccionales de las histonas	8
2.5.1 Polycomb y Trithorax	11
2.5.1.1 El grupo Polycomb	12
2.5.1.2 El grupo Trithorax	13
HIPÓTESIS	16
OBJETIVOS	16
1. Objetivo general	16
2. Objetivos particulares	17
MATERIALES Y MÉTODOS	17
1. Análisis Bioinformáticos	17
1.1. Obtención de datos	17
1.2. Procesamiento de datos y análisis bioinformáticos.....	17
2. Estudios experimentales	18
2.1. Material biológico	18
2.2. Condiciones de crecimiento	18
2.3. Obtención de las plantas dobles y triples mutantes de <i>clf ult1</i> y <i>atx1</i>	
<i>clf ult1</i>	19

2.3.1 Genotipificación de las mutantes -----	19
2.4 Cinética de crecimiento y caracterización del fenotipo de diferenciación de columela -----	19
RESULTADOS -----	20
1. Análisis del número de genes diferencialmente expresados en las mutantes <i>atx1, sdg2, ult1</i> y <i>clf</i> -----	20
2. Genes regulados de forma conjunta e independiente entre proteínas del grupo Trithorax -----	23
3. Relación de los genes regulados de forma conjunta e independiente entre proteínas del grupo Trithorax y CLF como representante de Polycomb -----	25
3.1 Interacciones con <i>ult1</i> -----	25
3.2 Interacciones entre CLF y las HKMTs de Trithorax -----	26
4. Procesos biológicos asociados a los genes desregulados en las mutantes de <i>atx1, sdg2, ult1</i> y <i>clf</i> -----	28
5. Procesos biológicos asociados a los genes desregulados en la mutante de <i>ult1</i> -----	30
6. Generación de la doble y triple mutante de <i>clf ult1</i> y <i>atx1 clf ult1</i> como herramientas para describir la interacción entre TrxG y PcG -----	31
6.1 Efectos de la mutación de CLF, ULT1 y ATX1 en el fenotipo: usando a la raíz como modelo -----	34
DISCUSIÓN -----	38
1. La desregulación de genes en la mutante <i>sdg2</i> corresponde a su fenotipo --	38
2. La proporción entre los subconjuntos de genes sub- y sobrexpresados en las mutantes no corresponden a su actividad -----	39
2.1 Polycomb y Trithorax pueden regular genes de forma independiente el uno del otro -----	40
3. Trithorax presenta interacciones intracomplejo distintas a las esperadas --	41
3.1 ULT1: un caso especial dentro de Trithorax -----	41
3.2 SDG2 y ATX1 no absorben sus blancos de regulación -----	42
4. Los defectos en la vía de las auxinas presentes en <i>ult1</i> pueden deberse a problemas en la síntesis de la hormona -----	42

5. ULT1 tiene el fenotipo más prevalente respecto a la diferenciación de las células de columela	43
CONCLUSIONES	44
ANEXO	45
REFERENCIAS	46

ABREVIATURAS Y CONTRACCIONES ÚTILES

5mC:	5-metil Citosina
ATX1	ARABIDOPSIS TRITHORAX1
CLF	CURLY LEAF
Col-0	Ecotipo Columbia
CSC	<i>Columella Stem Cells</i> /Células iniciales de Columela
DCC	<i>Differentiated Columella Cells</i> /Células de Columela diferenciada
DNA	Ácido Desoxirribonucleico
DNMT	DNA metiltransferasa
<i>Fold Change</i>	Tasa de cambio respecto a una condición control
H2AK121ub	Monoubiquitinación de la lisina 121 de la Histona H2A
H3K27me3	Trimetilación de la lisina 27 de la Histona H3
H3K36me3	Trimetilación de la lisina 36 de la Histona H3
H3K4me3	Trimetilación de la lisina 4 de la Histona H3
H3K9me3	Trimetilación de la lisina 9 de la Histona H3
HKMT	Histona lisina metiltransferasa
lncRNA	<i>long non-coding RNA</i> / RNA largo no codificante
pb	Pares de bases
PcG	Polycomb
PCR	<i>Polymerase Chain Reaction</i> /Reacción en Cadena de la Polimerasa
QC	<i>Quiescent center</i> /Centro quiescente
RNA	Ácido Ribonucleico
SCN	<i>Stem Cell Niche</i> /Nicho de células troncales
SDG2	SET DOMAIN GROUP2
sRNA	RNA pequeño
T-DNA	<i>Transfer DNA</i> /DNA de transferencia, DNA insercional
TrxG	Trithorax
ULT1	ULTRAPETALA1
Wt	<i>Wild type</i> /Plantas silvestres

RESUMEN

En plantas, la habilidad de tener respuestas plásticas ante los cambios ambientales resulta indispensable, y muchas de estas respuestas están mediadas por mecanismos epigenéticos. Dentro de los mecanismos epigenéticos, las modificaciones postraduccionales de las histonas parecen tener una particular importancia y en particular existen dos complejos proteicos de naturaleza antagónica que son importantes para mantener el estado de expresión o represión de los genes: el complejo Trithorax (TrxG) y el complejo Polycomb (PcG). En lo general las proteínas del grupo PcG ponen las marcas de H3K27me3 y H2AK119ub; ambas asociadas a la represión. Por su parte, las proteínas del complejo TrxG catalizan las marcas de la H3K4me3 y H3K36me3, canónicamente asociadas a activación transcripcional. En la planta modelo *Arabidopsis thaliana*, dentro del complejo Trithorax, se encuentra la proteína ULTRAPETALA 1 (ULT1) una proteína que se ha visto que tiene un papel importante en la regulación del meristemo apical. Fue descrita como un miembro de TrxG debido a su antagonismo con CLF (una de las 3 histonas metiltransferasas de PcG en *Arabidopsis*), además en el meristemo aéreo, ULT1 interactúa físicamente con ATX1, una de las varias enzimas de TrxG. A pesar de lo anterior, parece ser que este patrón no es conservado en toda la planta, debido a que, en la raíz, esta proteína sigue jugando un papel importante en el mantenimiento del nicho de células troncales, pero no lo hace de forma conjunta con ATX1. Cabe destacar que ATX1 no es la única enzima encargada de la H3K4me3 y particularmente SDG2 parece ser de gran importancia para la deposición de esta marca ya que la mutante de pérdida de función genera una disminución de alrededor del 34% de la H3K4me3 total. Dentro de los fenotipos presentados por la mutante *sdg2*, resulta interesante que la disrupción de la señalización de auxinas y la diferenciación prematura de células iniciales columela son compartidos con la mutante de ULT1, lo que sugiere la existencia de una regulación conjunta de estas dos proteínas particularmente en la raíz. Adicionalmente, se ha reportado que ULT1 puede tener actividad en conjunto a miembros de PcG, lo cual da pie a cuestionamientos sobre la posibilidad de la existencia de interacciones que contra intuitivamente salen del modelo antagónico clásico de PcG vs TrxG, pero que abren un nuevo paradigma en donde proteínas canónicamente descritas como represoras o activadoras pueden tener funciones duales que dependan del contexto. Debido a todo lo anterior, con el propósito de aportar más conocimiento al estudio de la regulación epigenética en plantas, este proyecto buscó analizar el comportamiento de la expresión génica en *Arabidopsis* mediada por diferentes proteínas del grupo TrxG y su comparación con proteínas del grupo PcG. Esto principalmente a través de un enfoque bioinformático que permitió de forma exploratoria a través de datos in silico observar comportamientos individuales y conjuntos de la expresión génica en las mutantes de los factores epigenéticos ULT1, SDG2, ATX1 y CLF y de esta forma brindar información sobre la relación existente entre cada uno.

INTRODUCCIÓN

1. La epigenética

1.1 Conceptualización del término

Desde su nacimiento, la epigenética ha sido un concepto complejo que ha cambiado significativamente a lo largo de la historia. Inicialmente fue definido (o mejor dicho planteado) por Conrad Waddington como un área de estudio de la biología que se encargaba de analizar las relaciones causales entre los genes, sus productos y el fenotipo que ultimadamente presenta el organismo (Jablonka & Lamb, 2006). Lo anterior hacía referencia a un tema atractivo pero un tanto general y esto provocó que el área de estudio tuviera límites un tanto difusos, generando una disciplina naciente con múltiples enfoques que intentaban explicar el fenómeno del “desacople” entre el fenotipo y genotipo (Holliday, 2006; Noble, 2015). Es decir, la epigenética en su sentido más amplio, busca encontrar una explicación a aquellas situaciones en las que la variación genética no conlleva a una variación en el fenotipo y a su vez, aquellas en las que la variación fenotípica no está relacionada con un cambio en la secuencia de los genes (Jablonka & Lamb, 2006).

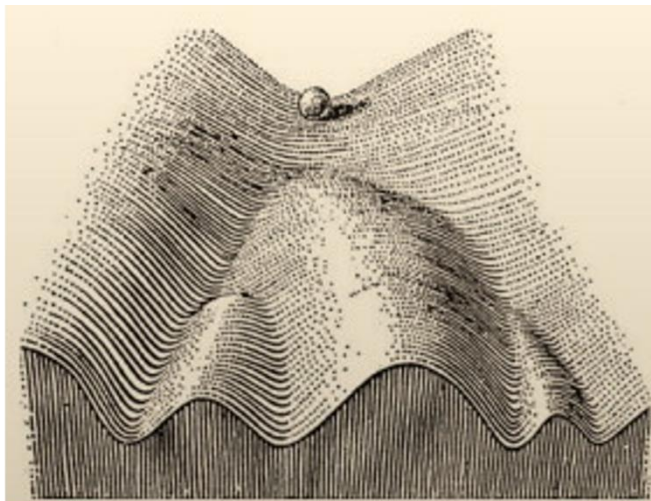


Figura 1. Paisaje epigenético según Waddington. Una metáfora para describir el camino que seguirá el embrión (la pelota) en su proceso de desarrollo embrionario a través de distintos canales, resultado de la interacción de las redes génicas y el ambiente (Jablonka & Lamb, 2006).

Inicialmente, la epigenética se concentró casi exclusivamente en los procesos del desarrollo embrionario. De hecho, esta idea de Waddington del cómo interactúan las redes génicas para dar origen al fenotipo (explicado en su famoso diagrama del paisaje epigenético (Figura 1), fue exclusivamente en un contexto del paso del cigoto al organismo; tanto así, que mucho del debate inicial en torno al concepto no fue sobre su validez sino sobre si era o no un sinónimo de la biología del desarrollo (Holliday, 2006; Jablonka & Lamb, 2006). Poco a poco el concepto fue adquiriendo

características más definidas y en una época en que la síntesis moderna de la evolución recién había sido planteada, la reciente reconciliación de la genética con la evolución, contrastaba mucho con las afirmaciones *waddingtonianas* de un proceso mucho menos directo entre el gen y el fenotipo (Haig, 2004). Lo anterior, sumado a que la epigenética no provenía más que de observaciones superficiales y carecía de algún mecanismo, generó escepticismo entre la comunidad científica. Sin embargo, personajes importantes como Huxley, derivado de sus observaciones en células cancerígenas utilizó el término epigenética para referirse al proceso por el cual un genotipo constante da origen a células diferenciadas y tejidos, y su perturbación (tal vez en la regulación de la expresión génica)

podían ser la causa del cáncer (Haig, 2004; Holliday, 2006). Afirmación que, a pesar de no haber sido interpretada así de momento, ya puede desprenderse del estudio del desarrollo embrionario y extrapolarse a otros contextos de los organismos.

Probablemente uno de los cambios más importantes en la conceptualización de la epigenética, fue el realizado por Nanney a finales de los 50s. Él hizo una separación importante entre dos niveles de control en las células, un componente genético estático y uno secundario “auxiliar” que determina la expresión particular de ciertos genes en ciertas células particulares; es decir, un nivel de regulación de la expresión. Relacionando así este mecanismo auxiliar al concepto de la epigenética, haciéndolo uno que se refiriera a un proceso global de los organismos (ya no exclusivo al desarrollo embrionario), pero a su vez especificándolo como un mecanismo regulatorio concreto sobre la expresión de los genes (Deichmann, 2016a; Haig, 2004).

A pesar de que Nanney únicamente consideraba la herencia de estos patrones auxiliares de forma citoplásmica (y no nuclear) y otras varias diferencias sutiles, es posible considerar su definición, como el origen de las subsecuentes formas “modernas” de ver a la epigenética (Nicoglou & Merlin, 2017). Los avances en las técnicas de biología molecular a mediados de los 90s y con ellos la identificación de estos mecanismos de regulación, llevaron a la noción moderna de la epigenética en su sentido bioquímico y molecular, la cual es de gran relevancia para este trabajo y se discutirá más adelante.

1.2 Lamarckismo y epigenética

Uno de los puntos más interesantes que fueron abordados por este concepto inicial de epigenética, y que siguen siendo de gran importancia para todas las definiciones modernas, es la influencia del ambiente en el fenotipo y la heredabilidad de este. Esto fue de los aspectos más controversiales ya que Waddington, en contraste a otros científicos como Weismann sí pudo observar experimentalmente una adquisición de caracteres influenciados por el ambiente y su subsecuente herencia en generaciones futuras (Noble, 2015). Lo anterior, no solo chocaba con el paradigma de herencia que se tenía, sino que también contrastaba con un debate mucho más antiguo aparentemente zanjado: la teoría de la evolución. Erróneamente pensando a la epigenética como un resurgir del lamarckismo.

Lamarck había propuesto en 1802 su teoría de la evolución, el núcleo conceptual de la misma se puede resumir en dos postulados, primero, que la evolución es un proceso direccional y que los organismos a lo largo de su vida adquieren o pierden características dependiendo del grado en que usen (o no) dichas partes de sus cuerpos y segundo, que estas características que adquirió o perdió el organismo son directamente heredadas a su progenie. El debate de porqué estas ideas no se sostienen al ponerlas a prueba en el mundo real no es de mucha relevancia para los efectos de este trabajo, sin embargo, lo importante es que la teoría de la evolución de Darwin llegó para plantear una idea de cómo ocurren los procesos evolutivos que refleja mucho más lo que ocurre realmente y que además es incompatible con lo dicho por Lamarck.

Esta teoría de Darwin, para empezar, plantea que el proceso evolutivo es de cambio solamente (es decir, no hay una dirección a la perfección como planteaba Lamarck) y requiere de azar como fuente de la variación biológica (que después será fijada o eliminada por las distintas “fuerzas evolutivas”) (Darwin, 1859). Esto último elimina hasta cierto punto la posibilidad de que, de forma directa, el ambiente fuera un agente (hasta cierto punto “inteligente”) de cambio en los organismos, y además que el organismo en una sola generación transmitiera este cambio adquirido a su descendencia (ya que la evolución ocurre en poblaciones, en varias generaciones, no en individuos), por lo tanto, es lógico que las observaciones de Waddington saltaran a la vista.

Sin embargo, con un análisis lo suficientemente cauteloso, es fácil darse cuenta por que la epigenética y el lamarckismo no son tan parecidos como puede parecer a simple vista. Waddington no estaba observando como el ambiente modelaba al organismo, lo que él observaba era como se expresaba la plasticidad en respuesta a estímulos externos. Es decir, siguiendo su modelo del paisaje epigenético (Figura 1), el ambiente sí canalizaba al organismo a cierto fenotipo, y esta canalización se estabilizaba y era pasada a su descendencia, pero la clave está en pensar que dicho paisaje epigenético y las “opciones plásticas” posibles forman parte de un proceso (evolutivo) intrínseco del organismo, y no son modeladas por el ambiente (Deichmann, 2016b; Noble, 2015). El ambiente no es un agente creador, únicamente influencia al programa epigenético para la canalización del fenotipo resultante, ocupando las reglas propias del organismo/célula en cuestión.

Además de lo anterior, cabe mencionar que la herencia de los programas epigenéticos no solo debe pensarse como una herencia transgeneracional (que sí ocurre y sobre todo en plantas al provenir sus gametos de células somáticas), sino como una herencia mitótica, que muchas veces es más relevante para los organismos, pero que cambios en estos programas usualmente no son heredados a generaciones futuras.

1.3 ¿Una definición moderna?: la necesidad de claridad conceptual

El mencionar el proceso histórico de la evolución del concepto de epigenética, es relevante por la simple razón de que hoy en día, la epigenética representa un campo de frontera en la investigación biológica, pero ¿Cuál de todas *las epigenéticas*? Pocos son los debates actuales entre si existen o no estos procesos de regulación que van más allá de las secuencias del DNA y que rompen con el determinismo que la mayoría conoce entre el genotipo y el fenotipo; más bien son debates mucho más específicos sobre como son los mecanismos en sí. Se sabe que los procesos dependen de interacciones complejas y redes de regulación que no están aisladas del ambiente y es impresionante como hoy en día la inclusividad de lo que estudia la epigenética, sigue siendo bastante amplia.

Por ejemplo, en el ámbito de la biología molecular y particularmente en la embriología, hoy en día varios embriólogos siguen manejando la definición de epigenética como “el conjunto de interacciones entre células y productos celulares que llevan a la morfogénesis y diferenciación” (Haig, 2004). Procesos netamente físicos como las fuerzas de mecánicas durante la formación del rostro mamífero o las fuerzas de tensión y tracción celular que dan origen a los patrones de colores de las serpientes, son considerados

epigenética, y de hecho existe hasta la propuesta de la existencia de una rama conocida como la epigenética cultural, muy alejado del ámbito molecular (Caballero et al., 2012; Jablonka, 2016).

Es muy válida y enriquecedora la existencia de tantas áreas bien definidas de la epigenética, y el pluralismo del concepto hace que, no solo sea un área interesante, sino que constantemente se esté renovando lo que sabemos sobre todo aquello que va “más allá” de lo codificado en el genoma. Sin embargo, es por ello que no debe tomarse a la ligera él como planteamos nuestra definición de trabajo, y en este caso, el concepto más útil es la epigenética como el conjunto de mecanismos moleculares y bioquímicos.

Una definición concreta, bajo esta perspectiva molecular, sería describir a la epigenética como el estudio de los cambios heredables mitótica o meióticamente en el control de la expresión génica independientes a la secuencia de DNA (Holliday, 2006; Jablonka & Lamb, 2006; Knight, 2015). Por lo anterior, es esta acepción de epigenética la que será usada para los fines de este proyecto. Dentro de este marco, el estudio de esta regulación molecular incluye áreas como la memoria celular, el desarrollo embrionario, la respuesta a estímulos, la diferenciación celular, la regulación de la respuesta a procesos cíclicos, entre otros (Jablonka, 2009).

2. Mecanismos epigenéticos

En plantas, la habilidad de tener respuestas plásticas ante los cambios ambientales resulta indispensable ante su incapacidad para desplazarse (Sanchez et al., 2015), y muchas de estas respuestas están mediadas por mecanismos epigenéticos. Tanto para plantas, como para el resto de los eucariontes, se puede hablar de cinco ejes principales en los que se agrupan dichos mecanismos: 1) Metilación del DNA, 2) Regulación mediada por RNAs (ej. RNAs largos no codificantes), 3) Modificación de la estructura tridimensional del genoma en el núcleo, 4) Incorporación de variantes de histonas y 5) Modificaciones postraduccionales de las histonas (Figura 2) (Grant-Downton & Dickinson, 2005; Jones & Sung, 2014). Todos juegan un papel importante en el desarrollo de los organismos vegetales, sin embargo, el estado de conocimiento actual dista de tener claro los procesos concretos que controlan y los mecanismos por los cuales llevan a cabo su función. A continuación, se describirán brevemente cada uno de los procesos, haciendo énfasis en las modificaciones postraduccionales de histonas, al ser el mecanismo que se abordará en este proyecto.

2.1 Metilación del DNA

Desde 1948 se conocía la existencia de modificaciones químicas a las bases nitrogenadas del DNA, entre ellas la metilación de citosinas y no fue sino hasta mediados de los 70s que se empezó a describir su papel en la regulación de la expresión genética (Allis & Jenuwein, 2016). En plantas la mayor parte de las citosinas metiladas (5-metilcitosinas, 5mC), se encuentran en un contexto simétrico de CG o CHG (donde H es cualquier nucleótido que no sea C) y en mucho menor medida en contextos asimétricos de tipo CHH (Grant-Downton & Dickinson, 2005). En su mayoría, la metilación del DNA se

encuentra asociada a represión transcripcional tanto para elementos transponibles y secuencias repetidas así como en algunos genes, teniendo una particular relevancia silenciando a los primeros dos, ya que la pérdida de esta marca, en muchas ocasiones tiene un impacto negativo en la integridad del genoma (Grant-Downton & Dickinson, 2005). Adicionalmente a lo mencionado, otra función destacable de la metilación del DNA en secuencias génicas tiene que ver con el fenómeno de la impronta. Donde, la represión mediada por esta marca dirige la expresión preferencial de un alelo sobre el otro dependiendo de su origen parental y en *Arabidopsis* es un fenómeno principalmente detectado en el endospermo durante el desarrollo de la semilla (Satyaki & Gehring, 2017).

Mecanísticamente, son enzimas denominadas DNA metiltransferasas (DNMTs) las encargadas en modificar a las citosinas agregando el grupo metilo, la 5mC es capaz de atraer a proteínas que reconocen la marca y a su vez reclutan otro tipo de proteínas modificadoras de la cromatina que tienen el impacto final en la represión del estado transcripcional; estas enzimas junto a desmetilasas del DNA y la maquinaria de RNA de interferencia mantienen la homeostasis y dinámica de la marca (Figura 2A) (Zhang et al., 2010). Se pueden dar dos tipos distintos de metilación, la metilación de mantenimiento, que perpetúa los patrones de metilación en los ciclos de replicación del DNA y la metilación *de novo* que establece los patrones de metilación en citosinas nunca antes modificadas (Zhang et al., 2010).

En *Arabidopsis thaliana*, la enzima DNA-METHYLTRANSFERASE 1 (MET1), se encarga de mantener la metilación en contexto CG, mientras que las enzimas CHROMOMETHYLASE 3 y 2 (CMT3/2) se encargan de la metilación en contexto CHG. La metilación *de novo* en contexto CHH es mantenida por DOMAINS REARRANGED METHYL- TRANSFERASE 1 y 2 (DRM1/2) a través de la vía de metilación dirigida por RNAs (*RNA directed DNA Methylation*, RdDM) y en algunas regiones es mantenida por CMT2 (Zhang et al., 2018). Entre las proteínas encargadas de reconocer la marca y tener un impacto en el estado de la cromatina, se encuentran VARIANT IN METHYLATION 1 (VIM1) con dominio de unión a la 5mC, DECREASE IN DNA METHYLATION 1 (DDM1) con actividad de remodeladora de la cromatina, HISTONE DEACETYLASE 6 (HDA6) y KRYPTONITE (KYP) que a través de modificar histonas favorecen la represión transcripcional (Zhang et al., 2010). Finalmente, existen dos formas de borrar la metilación del DNA en plantas: una pasiva que simplemente depende de los ciclos de replicación asociados a baja actividad de las DNA metiltransferasas, que provocan dilución de la marca, y una desmetilación activa mediada por las enzimas, DEMETER (DME) y REPRESSOR OF SILENCING 1 (ROS1) (Choi et al., 2002; Gong et al., 2002).

2.2 Regulación mediada por RNAs no codificantes

Los RNAs no codificantes (regiones transcritas a RNA pero que no se traducen a proteína), clasificados por su tamaño en largos y pequeños (lncRNAs y sRNAs) participan en diversos mecanismos de regulación de la expresión fungiendo como andamiaje, puente o acarreo de otras proteínas (Costa, 2010). En plantas se ha identificado su papel mediador en la metilación del DNA o en procesos tan importantes como la regulación de la floración a

través del proceso de vernalización (Grant-Downton & Dickinson, 2005; Jones & Sung, 2014).

Los sRNAs, clasificados de esta forma por tener menos de 200 nucleótidos, pueden a su vez ser subdivididos en *micro RNAs* (miRNAs), *small interfering RNA* (siRNAs) y *Piwi-interacting RNAs* (piRNAs) (Heo et al., 2013). El papel de los sRNAs en el silenciamiento mediado por metilación del DNA vía RdDM, es un buen ejemplo de un proceso que se tiene bien caracterizado. Uno de los mecanismos descritos es que estos RNAs pequeños son generados a partir de precursores transcritos por las polimerasas RNA polimerasa IV (RNAP IV) y RDR2, que posteriormente son procesados por DICER-LIKE 3 (DCL3) (Heo et al., 2013; Jones & Sung, 2014; M. Zhang et al., 2010). Una vez procesado, el sRNA se une a proteínas de la familia ARGONAUTE (AGO) permitiendo la unión subsecuente a transcritos de la RNA polimerasa V (RNAP V), funcionando así como andamios para la formación de complejos de silenciamiento (RNA-induced Silencing Complexes, RISC) con metilasas del DNA *de novo* como DRM2 (Figura 2B) (Grant-Downton & Dickinson, 2005; Jones & Sung, 2014).

En cuanto a los RNAs largos, tienen papeles mucho más heterogéneos y menos caracterizados, sin embargo, uno de los procesos biológicos de *Arabidopsis* en los que participan es en la floración inducida por vernalización. La vernalización describe como una temporada de frío prolongada promueve la floración, a través de la represión del gen *FLOWERING LOCUS C (FLC)* (Sung & Amasino, 2004). Este silenciamiento está mediado por proteínas del grupo Polycomb (que se describirán más adelante) pero se ha identificado a tres tipos distintos de lncRNAs que regulan esta represión: *COOLAIR* transcrito en antisentido a *FLC*, *COLDWRAP* transcrito del promotor y *COLDAIR* transcrito del primer intrón de *FLC* (Heo et al., 2013; Kim & Sung, 2017). Detalles del mecanismo de acción de *COOLAIR* siguen siendo poco claros, sin embargo se hipotetiza su función a través del reclutamiento de desmetilasas de la H3K36me3 (marca de activación) y de algún modo favorecer la deposición de la marca de la H3K27me3 (marca de represión); respecto a *COLDAIR*, se ha identificado su unión a enzimas del complejo Polycomb y de esta forma favorecer la deposición de la H3K27me3 en *FLC* (Figura 2B) (Berry & Dean, 2015; Heo et al., 2013). Recientemente se ha reportado que *COLDAIR* y *COLDWRAP* forman una asa de cromatina requerida para mantener la represión estable de *FLC* (Kim & Sung, 2017).

2.3 Estructura tridimensional del genoma

Poco se sabe sobre el impacto que tiene la estructura tridimensional de la cromatina en la expresión génica en plantas, sin embargo en poáceas se ha observado configuraciones cromosómicas específicas y aunque en *Arabidopsis thaliana* los compartimentos de jerarquía mayor parecen no existir, es poco clara la relevancia funcional de las interacciones que sí están presentes (Pontvianne & Grob, 2020).

Se hipotetiza que la distribución relativamente uniforme de las marcas epigenéticas en *Arabidopsis* (es decir, no hay zonas grandes enriquecidas en una sola marca en particular) hace que la delimitación de compartimentos topológicamente asociados sea

complicada y por lo tanto, se ha sugerido que estos sean virtualmente inexistentes (Pontvianne & Grob, 2020). Sin embargo, se han descrito, asas de cromatina con un impacto en la represión (Ariel et al., 2020; Veluchamy et al., 2016), dominios asociados a proteínas nucleares que facilitan la formación de heterocromatina (Poulet et al., 2017) e inclusive impactos globales como la activación de transposones en respuesta a cambios en la estructura tridimensional del genoma (Sun et al., 2020) (Figura 2C). Esto hace evidente el hecho de que la dinámica de la estructura y posición del genoma en el núcleo, así como su regulación fina, tengan un papel biológico importante en *A. thaliana* pero se necesitan más estudios al respecto.

2.4 Variantes de Histonas

Como se evidenció con los mecanismos anteriores, la cromatina juega un papel central en la regulación epigenética, en su concepto más general, es la asociación de proteínas y en ocasiones RNAs con el DNA dentro del núcleo eucarionte. Su función más básica es proteger al DNA evitando daños a su integridad por rupturas o desorganización espacial y ayudar a la formación de estructuras que faciliten la segregación del material genético durante la división celular. Ésta, en general puede crear estructuras físicas a distintos niveles, desde el nucleosoma (que es su unidad básica), hasta el cromosoma metafásico, que es la forma más compactada de un genoma y es, justo esta compactación de la cromatina la que está relacionada (en parte) con la expresión de los genes y sobre la cuál actúan muchos de los mecanismos epigenéticos (Klemm et al., 2019).

El nucleosoma consiste en alrededor de 150 pb de DNA enrolladas en un octámero de histonas. Cada nucleosoma está conformado por 4 histonas distintas: H3, H4, H2A y H2B, dos de cada una para formar el octámero y adicionalmente existe la histona *linker* H1 fuera del nucleosoma que permite la formación de las estructuras de orden superior que se mencionaron (Karp, 2013). Dos mecanismos tienen como blanco a las histonas: la incorporación de las variantes de histonas y las modificaciones postraduccionales de las mismas (Allis & Jenuwein, 2016).

Respecto a las variantes de histonas, estas solamente difieren de las canónicas por unos cuantos aminoácidos, sin embargo, esto es suficiente para tener un impacto en el estado de la cromatina y en el fenotipo, tanto a nivel molecular (facilitando la deposición de ciertas marcas epigenéticas), como a nivel macroscópico. Por ejemplo, regulando la tasa de proliferación celular en el eje de la raíz (Probst et al., 2020). Como ejemplos particulares, la histona H2A.X, una variante de la histona H2A, marca sitios de daño al DNA (Probst et al., 2020) y la H2A.Z, otra de las variantes de esta misma histona, se ha identificado con un papel en la regulación transcripcional (en parte al evitar la metilación del DNA) y en la formación de fronteras de heterocromatina (Deal & Henikoff, 2011; Grant-Downton & Dickinson, 2005). También se tiene de ejemplo a la histona H3.3 (variante de la histona H3.1, que es la canónica) que parece favorecer la deposición de marcas de represión en sus dominios N terminales, particularmente facilita la deposición de la mono y di metilación de la lisina 27 (Deal & Henikoff, 2011; Wollmann et al., 2017) (Figura 2D).

2.5 Modificaciones postraduccionales de las histonas

Todas las histonas del nucleosoma (H3, H4, H2A y H2B) poseen dominios N terminales (*colas*) que protruyen fuera de su estructura principal y que son susceptibles a ser modificadas postraduccionalmente, siendo así el material para que ocurra este mecanismo de regulación epigenética. En plantas, esto parece tener una particular importancia en el mantenimiento del nicho de células troncales en los meristemos aéreo y de raíz, así como de otros procesos de desarrollo de órganos, vernalización, el control de la dormancia en la semilla, entre otros (Sanchez et al., 2015). Haciendo su estudio de vital interés si se busca comprender la biología de estos organismos.

Al haber cuatro diferentes tipos de histona formando el nucleosoma, cada una con sus propios residuos susceptibles a modificaciones, se generan varias posibles marcas y combinaciones de estas, que provocan a su vez distintos efectos en la expresión génica. Estos dominios en la región N-terminal de las histonas, son ricos en argininas, lisinas y serinas, residuos que son susceptibles a ser modificados al adicionarles grupos metilo, acetilaciones, ubiquitinaciones, fosforilaciones, entre otras (Figura 2E) (Allis & Jenuwein, 2016; Bannister & Kouzarides, 2011). Por sí mismas, salvo la acetilación (que tiene un efecto electrostático) las marcas no tienen un impacto directo en la cromatina y en la transcripción génica, por lo que dependen del reclutamiento de otros complejos proteicos para tener un efecto en el estado cromatínico. Teniendo así, como en los mecanismos anteriores, proteínas escritoras (enzimas que colocan la marca), lectoras (proteínas que reconocen la marca) y borradoras (enzimas que quitan estas marcas), siendo cada una igual de importante para la correcta regulación génica (Bannister & Kouzarides, 2011).

Dentro de todas las modificaciones posibles, la metilación en residuos de lisina y especialmente de las histonas H3 y H4 es la más entendida y estudiada. Cada lisina puede ser modificada en tres *niveles*, teniendo así residuos mono-, di- o trimetilados. Particularmente en *Arabidopsis* se tienen ejemplos de modificaciones postraduccionales (marcas) identificadas como la di y trimetilación de la lisina 36 de la histona 3 (H3K36me₂ y H3K36me₃) o la di y trimetilación de la lisina 4 de la histona 3 (H3K4me₂ y H3K4me₃) (Du et al., 2013; Lauria & Rossi, 2011). Mientras que las marcas de represión, que se han identificado son la dimetilación de la lisina 9 en la histona 3 (H3K9me₂) y la trimetilación de la lisina 27 de la histona 3 (H3K27me₃) (Lauria & Rossi, 2011; Pfluger & Wagner, 2007).

Para la metilación de lisinas, las proteínas escritoras mantienen la característica común de poseer un dominio catalítico conservado de tipo SET (Su(var)3-9, Enhancer-of-zeste and Trithorax) encargado de modificar las lisinas, y colectivamente se les conoce como Histonas metiltransferasas (HKMT) (Berr et al., 2011). En cuanto a las proteínas que reconocen las metilaciones, existe una diversidad mayor en cuanto a los dominios que presentan, pero destacan los dominios del tipo Tudor, CW, WD40, PHD, PWWP y Cromodominios. Estas proteínas a su vez tienen el efecto último en la expresión al poder fungir como reclutadoras de proteínas remodeladoras de la cromatina, otras escritoras de marcas distintas, factores de transcripción, proteínas de reparación de daño, etc. (Yun et al., 2011). Finalmente, en cuanto a las enzimas borradoras de la marca, se tienen reportadas varias proteínas con actividad de desmetilasas pertenecientes a dos familias dependiendo de

su dominio catalítico: las LSD y las JMJC (Chen et al., 2011; Marije Kooistra & Helin, 2012).

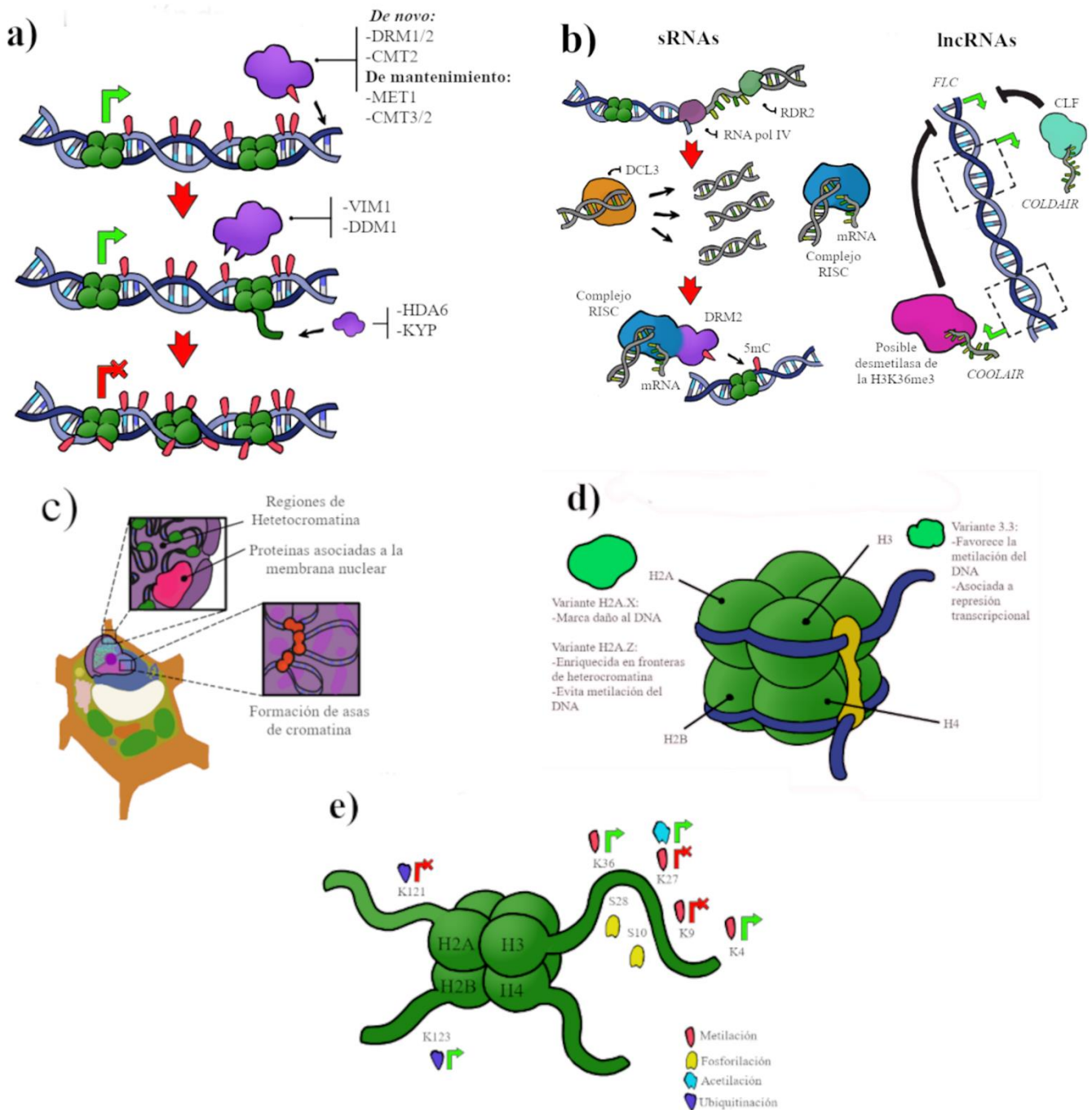


Figura 2. Mecanismos presentes en la regulación epigenética en plantas. a) Metilación del DNA. b) RNAs no codificantes. c) Dinámica de la estructura de la cromatina en el núcleo. d) Variantes de histonas. e) Modificaciones postraduccionales de histonas

Como es de esperarse, existe un repertorio de proteínas escritoras, lectoras, borradoras y proteínas de andamiaje específicas para cada lisina/histona específica

formando complejos proteicos involucrados en la regulación fina espaciotemporal de la expresión génica de los organismos. La función específica, mecanismos de acción e incluso los miembros de cada uno de estos complejos sigue siendo poco clara, sin embargo, existen algunos ejemplos de complejos caracterizados y conservados en la mayoría de los eucariontes. En particular, existen dos complejos proteicos de naturaleza antagónica que son importantes para mantener el estado de expresión o represión de los genes: el complejo *Trithorax* (TrxG) y el complejo *Polycomb* (PcG) de los cuales, se describirán más detalle a continuación.

2.5.1 Polycomb y Trithorax

Inicialmente, en *Drosophila melanogaster* fueron identificados varios genes cuyas mutaciones provocaban una desregulación en la expresión de genes homeóticos, estos genes posteriormente fueron clasificados como miembros del complejo Polycomb y el complejo Trithorax (Schuettengruber et al., 2017). Esta desregulación en los genes homeóticos despertó el interés científico provocando una investigación más a fondo, encontrando así a nuevos miembros de ambos complejos y poco a poco comprendiendo que su papel iba más allá de dichos genes *Hox*, participando en una plétora de procesos a lo largo de la vida de los organismos (Schuettengruber et al., 2017).

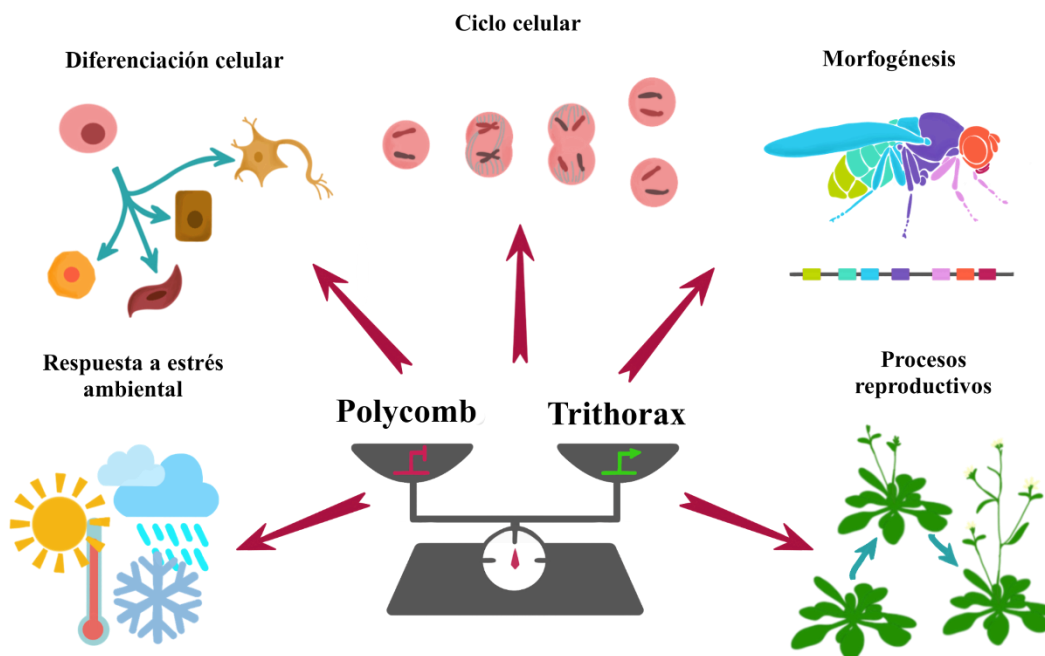


Figura 3. Diferentes papeles de Polycomb y Trithorax conservados en los eucariontes

Polycomb se encarga de depositar dos marcas: la H3K27me3 y la H2AKub (en plantas en la K121) asociadas a la represión, y Trithorax deposita la marca de H3K4me y la H3K36me3, que como ya se mencionó antes está asociada a la activación transcripcional. Procesos como la inactivación del cromosoma X, el cáncer, control del ciclo celular y

específicamente en plantas la diferenciación y mantenimiento de las células troncales en los meristemas, requieren de la acción de TrxG y PcG (Figura 3) (Sanchez et al., 2015; Schuettengruber et al., 2017). Cabe destacar que la relevancia de estudiar en conjunto a ambos complejos es que canónicamente sus actividades de represión-activación resultan ser antagónicas, es decir la presencia o ausencia de sus marcas en los mismos genes tiende a dictar el estado de expresión de un gran subconjunto de genes claves para la biología de los organismos.

2.5.1.1 El grupo Polycomb

Este grupo proteico está caracterizado por tener a su vez dos subcomplejos encargados de poner cada una de las dos marcas antes descritas: POLYCOMB REPRESSIVE COMPLEX 1 (PRC1) que coloca la marca de la H2AK121ub y POLYCOMB REPRESSIVE COMPLEX 2 (PRC2) que pone la marca de la H3K27me3 (Pien & Grossniklaus, 2007; Schuettengruber et al., 2017).

En *Arabidopsis thaliana*, PRC2 es por mucho el subcomplejo más estudiado. Puede estar conformado por una de las tres HKMT que colocan la marca de la H3K27me3: CURLY LEAF (CLF), SWINGER (SWN) expresadas a lo largo del desarrollo vegetativo o MEDEA (MEA), expresada específicamente durante el desarrollo de gametofito y semilla; asociadas siempre a las proteínas FERTILIZATION INDEPENDENT ENDOSPERM (FIE) y MULTICOPY SUPPRESSOR OF IRA 1 (MSI) (Sanchez et al., 2015; Xiao & Wagner, 2015). Adicionalmente, se ha identificado tres tipos de PRC2 dependiendo de que otra proteína interactuante esté presente, el tipo EMF-PRC2 cuando interactúa con EMBRYONIC FLOWER 2 (EMF2) con particular importancia en el desarrollo vegetativo, el VRN-PRC2 cuando está presente VERNALIZATION 2 (VRN2) importante en el proceso de vernalización y finalmente el FIS-PRC2 cuando interactúan con FERTILIZATION INDEPENDENT SEED 2 (FIS2), con papel en el desarrollo de gametofito y semilla (Figura 4a) (Derkacheva & Hennig, 2014; Sanchez et al., 2015).

En cuanto a PRC1, existía cierto debate respecto a si en *A. thaliana* se contaba con este complejo debido a la falta de homólogos al gen *Polycomb* (*Pc*), cuya proteína es indispensable para la función de PRC1 en metazoarios, sin embargo, se ha identificado distintas proteínas que cumplen la función del subcomplejo aún sin la presencia de *Pc* (Derkacheva & Hennig, 2014; Xiao & Wagner, 2015). Entre estas proteínas se tiene a dos familias de escritoras de la marca de ubiquitinación con dominio tipo REALLY INTERESTING NEW GENE (RING), las AtRING1A/B y las B CELL-SPECIFIC MOLONEY MURINE LEUKEMIA VIRUS INTEGRATION SITE 1 (AtBMI1) A/B/C, a la proteína LIKE HETEROCHROMATIN PROTEIN 1 (LHP1) que cumple la función de *Pc*, leyendo la marca de la H3K27me3 y además teniendo la capacidad de interactuar con PRC2 y finalmente se ha identificado a EMBRYONIC FLOWER 1 (EMF1) que en conjunto con las proteínas SHORT LIFE (SHL) y EARLY BOLTING IN SHORT DAYS (EBS), es también capaz de leer la marca de H3K27me3, teniendo una función importante en el mantenimiento del estado de represión (Figura 4b) (Pikaard & Scheid, 2014; Sanchez et al., 2015; Xiao & Wagner, 2015).

Cabe recalcar que para la acción de PcG en la represión génica, se ha identificado que la actividad de ambos subcomplejos puede ser dependiente e independiente, PRC2 puede reclutar a PRC1 pero también PRC2 puede ser reclutado por PRC1 o actuar de forma independiente y viceversa (Xiao & Wagner, 2015).

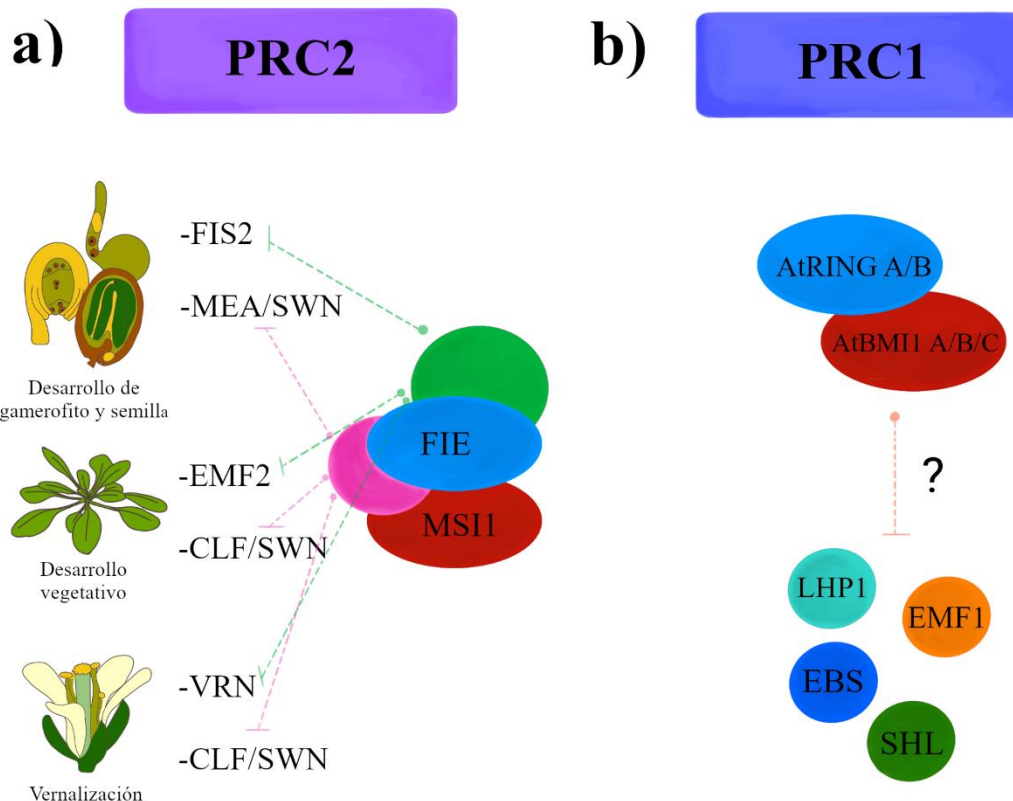


Figura 4. Proteínas miembros de Polycomb y algunas de sus interacciones identificadas. a) Componentes de PRC2 y los subcomplejos que forman entre ellos dependiendo del estadio y/o tejido. b) Componentes de PRC1

2.5.1.2 El grupo Trithorax

Curiosamente, muchas veces el complejo Trithorax ha sido reducido en los estudios a ser simplemente comparado a Polycomb. Además la heterogeneidad respecto a los componentes proteicos que lo forman tampoco contribuye a que se tenga una visión clara de su mecanismo de acción (Schuettengruber et al., 2017). A pesar de esto, se ha identificado el papel de elementos individuales del complejo en múltiples procesos del desarrollo vegetativo y reproductivo de *Arabidopsis*. Dentro de este grupo, los subcomplejos formados en cada contexto están mucho menos descritos y en lo general la definición de una proteína como miembro de TrxG tiene que ver simplemente con una asociación con la deposición de la H3K4me3 y con qué contrarresten la actividad de proteínas PcG (Sanchez et al., 2015).

Se han identificado dentro de Trithorax a cuatro grupos distintos de proteínas: las remodeladoras de la cromatina dependientes de ATP, proteínas núcleo AtCOMPASS, a las

HKMTs y proteínas accesorias asociadas a TrxG (Figura 5) (Fletcher, 2017; Schuettengruber et al., 2017). Dentro de las proteínas remodeladoras de la cromatina, en *Arabidopsis* se han identificado tres proteínas distintas con particular importancia: SPLAYED (SYD) con actividad en el meristemo aéreo y morfología floral, BRAHMA (BRM) con actividad en el meristemo aéreo, de raíz y represión floral y PICKLE (PKL) con un papel importante en el mantenimiento del meristemo de raíz (Fletcher, 2017). Respecto a las proteínas núcleo AtCOMPASS, estas tienen un papel importante al unir y reclutar a distintos elementos de Trithorax para llevar a cabo sus funciones, entre ellas destacan ARABIDOPSIS Ash2 RELATIVE (ASH2R), WD REPEAT-CONTAINING5 A y B (WDR5a/b) y RBBP5 LIKE (RBL) (Fletcher, 2017; Sanchez et al., 2015).

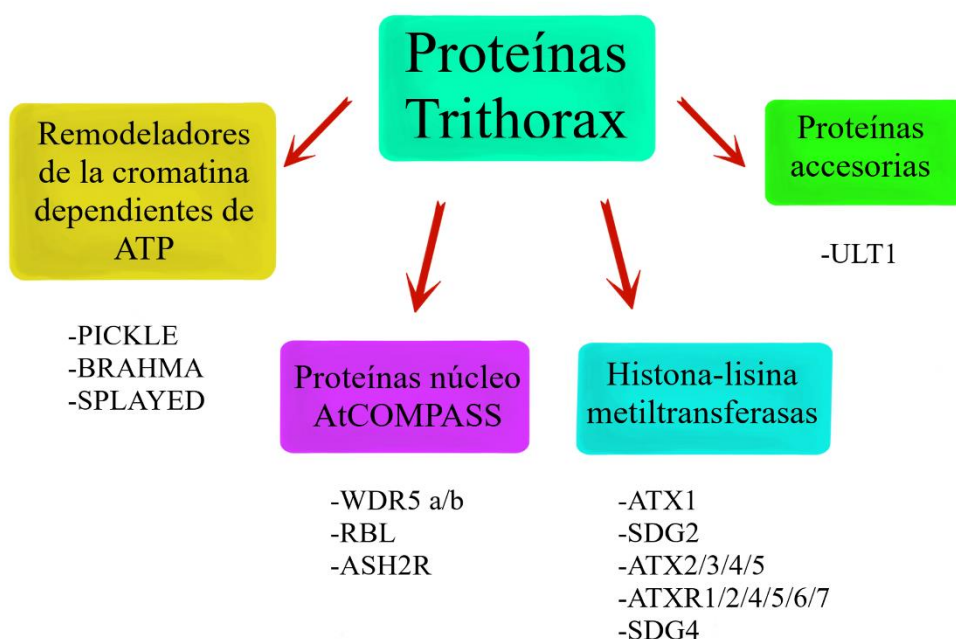


Figura 5. Distintos tipos de proteínas que integran al complejo Trithorax

Para las enzimas que ponen la marca, se ha identificado muchas proteínas con dominio SET y actividad de HKMT en *Arabidopsis*, existen alrededor de 47 genes identificados con esta actividad, si bien, no todos se encargan de catalizar la deposición de la marca de H3K4me3, aun así, se tiene 12 enzimas Clase III encargadas de dicha marca, 5 homólogas a las proteínas Trithorax en otros eucariontes denominadas ARABIDOPSIS TRITHORAX (ATX) de la 1-5 (ATX1-5) y 7 de tipo ATX-RELATED (ATXR) de la 1-7 (ATXR1-7) (Chen et al., 2017; Dong et al., 2015; Guo et al., 2010). Sorprendentemente de estas, muy pocas han sido caracterizadas y por mucho la más estudiada de todas es ATX1 de la cual se ha identificado su participación en múltiples procesos durante el desarrollo de *A. thaliana*. ATX1 participa en el desarrollo floral al activar genes homeóticos, participa en el mantenimiento de meristemo de raíz a través de regular la proliferación celular, en la formación de raíces laterales, respuesta a estrés hídrico y en morfogénesis foliar (Ding et al., 2011; Napsucialy-Mendivil et al., 2014; Pontvianne et al., 2010).

A pesar de lo anterior, la ausencia de ATX1 solamente afecta al 15% de la marca de H3K4me3, evidenciando que no posee un papel global tan importante en la deposición de la marca a lo largo del genoma (Alvarez-Venegas & Avramova, 2005). En contraste a lo anterior, se ha identificado a otras HKMTs con un papel más prominente en la deposición de la marca y regulación génica, SET DOMAIN GROUP 2 (SDG2/ATXR3) es una de ellas y parece ser de gran importancia para la deposición de la marca ya que en la mutante de pérdida de función existe una disminución significativa de alrededor de 34% de la H3K4me3 (Guo et al., 2010). Además, su ausencia provoca severas consecuencias fenotípicas como plantas pequeñas y débiles, floración temprana, órganos reproductivos aberrantes, esterilidad, raíces cortas, desarreglo en el nicho de células troncales, entre otras; lo cual contrasta con el efecto menor y menos pleiotrópico que tienen las mutantes de ATX1 (Berr et al., 2010; Chen et al., 2017; Pinon et al., 2017). Sin embargo, realmente es poco lo que se ha estudiado respecto a SDG2 debido a las complicaciones metodológicas que provoca el que las plantas se vean tan afectadas en la mutante.

Finalmente, describir a todas aquellas proteínas asociadas a Trithorax, pero que no entran en alguna de las categorías anteriormente mencionadas resulta complejo. Sin embargo, existen proteínas coactivadoras que interactúan con las HKMTs ayudando a la deposición de la marca, entre las cuales destaca ULTRAPETALA1 (ULT1). Ésta, es una proteína pequeña exclusiva de plantas de 268 aminoácidos, sin actividad catalítica, con un dominio tipo Spl00, AIRE-1, NucP41/75, DEAF-1 (SAND) de interacción con DNA, que se ha visto que tiene un papel importante en la regulación del meristemo apical actuando como un regulador negativo de la acumulación de células meristemáticas en meristemas florales y de inflorescencia (Carles et al., 2005; Fletcher, 2001). Fue descrita como un miembro de TrxG debido a su antagonismo con CLF ya que se observó que la planta sobre-expresora de ULT1 fenocopia a la mutante de CLF, la doble mutante *ult1 clf* recupera un fenotipo silvestre, su mutante presenta una disminución de la marca de la H3K4me3 en ciertos genes e *in vitro* puede interactuar físicamente con ATX1 (Carles & Fletcher, 2009). A pesar de esto, ULT1 resulta una proteína interesante debido a que puede presentar múltiples interacciones con factores de transcripción, interacciones dentro de TrxG independientes a ATX1 e inclusive interacciones que difieren en su antagonismo a CLF (Ornelas-Ayala et al., 2019, 2021).

La función de ULT1 independiente de ATX1 fue observada en la raíz, donde desde hace ya tiempo se identificó que ULT1 se expresa (Carles et al., 2005) y, estudios recientes demostraron su papel en este órgano, particularmente en el mantenimiento de la organización del nicho de células troncales (Ornelas-Ayala et al., 2019). Se identificó que es necesario para la correcta diferenciación de las células de la columela al inhibir una diferenciación prematura de las células iniciales, adicionalmente regula la tasa de división y estado indiferenciado del QC y participa en la regulación de la señalización de auxinas (hormonas vegetales importantes para el establecimiento del SCN) (Ornelas-Ayala et al., 2019). Lo más interesante de lo anterior, es que todos estos efectos que se observaron fueron independientes a la acción de ATX1. Las alteraciones observadas en la mutante *ult1-*

3 a diferencia de la mutante *atx1-3* fueron independientes a la proliferación celular del meristemo y la longitud de la raíz, así mismo el fenotipo de diferenciación prematura de las células iniciales de columela no está presente en *atx1* (Ornelas-Ayala et al., 2019).

Esta independencia funcional entre ATX1 y ULT1 en raíz, revela la versatilidad de ULT1 al presentar interacciones distintas dependiendo del tejido en que se encuentre, y por lo tanto surge el cuestionamiento sobre con qué otra HKMT de Trithorax puede estar interactuando para llevar a cabo su función (en este caso por comparativa) en la raíz (por ejemplo, SDG2). Además de ser tan importante para la deposición de la marca, SDG2 también presenta un fenotipo importante en raíz, se ha reportado que las mutantes de pérdida de función presentan raíces cortas con menos raíces laterales y, al igual que *ult1* presenta una diferenciación prematura de las células iniciales de la columela y una señalización reducida de auxinas (Yao et al., 2013).

Finalmente, se ha reportado que ULT1 puede tener actividad en conjunto a miembros del grupo PcG, particularmente con EMF1 (miembro de PRC1) en ciertas condiciones y además, la falta de ULT1 tiene un impacto en la baja de la H3K27me3 (Xu et al., 2018), lo cual da pie a cuestionamientos sobre la posibilidad de la existencia de interacciones que contra intuitivamente salen del modelo antagónico clásico de PcG vs TrxG (Ornelas-Ayala et al., 2021).

Debido a todo lo anterior, con el propósito de aportar más conocimiento al estudio de la regulación epigenética en plantas, este proyecto busca analizar el comportamiento de la expresión génica en *Arabidopsis thaliana* mediada por diferentes proteínas del grupo TrxG y su comparación con proteínas del grupo PcG. Esto a través de un enfoque bioinformático y experimental, el primero que permita de forma exploratoria a través de datos *in silico* observar comportamientos individuales y conjuntos de la expresión génica en mutantes de los factores epigenéticos ULT1, SDG2, ATX1 y CLF y de esta forma brindar información sobre la relación existente entre cada uno. En cuanto a la parte experimental, se planea generar las herramientas necesarias para el futuro estudio *in vivo* de los efectos de estas proteínas en la planta.

HIPÓTESIS

La gran mayoría de genes desregulados en las mutantes de TrxG analizadas se compartirán entre miembros de este mismo grupo en comparación a PcG. De la misma forma la mayor cantidad de genes desregulados compartidos entre PcG y TrxG serán aquellos que se desregulen en el sentido de su actividad de represión y activación respectivamente cumpliendo el papel canónico antagonista de ambos complejos.

OBJETIVOS

1. Objetivo general

Analizar el comportamiento de la expresión génica mediada por diferentes proteínas del grupo TrxG y su comparación con proteínas del grupo PcG.

2. Objetivos particulares

- Buscar datos de bioproyectos públicos de microarreglos obtenidos a partir de mutantes de pérdida de función en proteínas del complejo TrxG como ATX1, SDG2 y ULT1, así como de CLF, una HKMT miembro de PcG.
- Analizar la relación que existe entre ATX1 y SDG2 con ULT1 mediante un estudio *in silico* comparativo de los genes desregulados en las plantas mutantes de estos miembros de TrxG
- Estudiar la relación que existe entre ATX1, SDG2 y ULT1 con CLF, uno de los principales componentes de PcG
- Estudiar las vías de regulación en las que participan las proteínas de TrxG, mediante un análisis de ontología de genes y su comparación con PcG
- Generar cruza dobles y triples mutantes de *ult1 clf* y *ult1 clf atx1* como herramientas para describir la regulación genética que existe entre TrxG y PcG en la raíz de *Arabidopsis thaliana*.

MATERIALES Y MÉTODOS

1. Análisis Bioinformáticos

1.1. Obtención de datos

La búsqueda y obtención de bioproyectos públicos de microarreglos de mutantes de pérdida de función en proteínas del complejo TrxG como ATX1, SDG2 y ULT1 y CLF como representante de PcG se realizó en el NCBI (National Center for Biotechnology Information, <https://www.ncbi.nlm.nih.gov>) y el ArrayExpress database (<https://www.ebi.ac.uk/arrayexpress/>). Específicamente, para el procesamiento y manipulación de datos, se buscaron aquellos que fueran de la compañía *Affymetrix*, del tipo *Arabidopsis Genome ATH1 Array*

1.2. Procesamiento de datos y análisis bioinformáticos

Usando paqueterías para el software y lenguaje de programación R estos datos se procesaron para su comparación y análisis de la siguiente manera:

- Ocupando el paquete *GEOquery*, se descargaron y descomprimieron los archivos tipo “.CEL” obtenidos directamente de las bases de datos
- A través el paquete *affy*, se extrajeron los datos de los archivos con extensión .CEL, y fueron etiquetados respecto a cada condición del experimento (en este caso las plantas silvestres y las mutantes), así como las réplicas de cada uno. Esta base de datos se convirtió en un objeto de clase *Affybatch* manipulable para su análisis subsecuente.
- Dependiendo del caso, se usó el paquete *gcrma*, o la función del paquete *affy* de *rma* para la normalización de los datos de microarreglo.

- Ocupando el paquete *limma* se realizó el análisis para obtener todos los genes que se encuentran significativamente desregulados en las mutantes a comparar (ya sea de forma negativa o positiva) a través de una matriz de contraste. Posterior a eso se filtraron aquellos genes que tuvieron una *p* menor a 0.05 y que pasaron un límite de 1.5 o menores a -1.5 en cuanto a su *fold change* (tasa de cambio) para los genes desregulados de forma positiva y negativa respectivamente.
- Finalmente se generaron las tablas con los genes obtenidos y los gráficos pertinentes, ocupando los paquetes *VennDiagram* y *ggplot2*.

Para los análisis de ontología de genes, se ocupó la paquetería de *ClusterProfiler* para realizar los análisis de enriquecimiento y como referencia se usó la base de datos del paquete *org.At.tair.db* el cual contiene las anotaciones de los genes de *Arabidopsis thaliana*. Los resultados anteriores se graficaron con las paqueterías de *enrichplot*, *ggplot2* y *ggupset*.

2. Estudios experimentales

2.1. Material biológico

Para la parte experimental, se usaron plantas de la especie *Arabidopsis thaliana*, ecotipo Columbia (Col-0). Específicamente, en cuanto a las mutantes, se usaron:

- Mutantes de ULT1: *ult1-3* (SALK_074642), mutante de inserción de T-DNA localizada en el primer exón y de pérdida de función.
- Mutantes de ATX1: *atx1-3* (SALK_105214), mutante de inserción de T-DNA localizada en el intrón 19 y de pérdida de función
- Mutantes de CLF: *clf-29* (SALK_021003), mutante de inserción de T-DNA localizada en el tercer exón y de pérdida de función

Las semillas fueron desinfectadas lavándolas en una solución de hipoclorito de sodio comercial (Cloralex® al 3%) al 20% y polisorbato 20 (Tween® 20, Sigma Aldrich) al 0.01%, en agitación durante 15 min. Posteriormente se lavaron con agua bidestilada estéril 5 veces. Se pusieron en estratificación en oscuridad a 4°C por 72h, para sincronizar la germinación.

2.2. Condiciones de crecimiento

Después de la estratificación las semillas fueron sembradas en cajas petri cuadradas conteniendo medio Murashige y Skoog (MS) comercial al 0.2X (Murashige y Skoog 1962, MP Biomedicals), MES monohidratado al 0.05% (Sigma-Aldrich), sacarosa al 1% (Sigma-Aldrich) y agar al 1% (Becton, Dickinson y Company), pH de 5.6, ajustado con KOH 1M. Las semillas se incubaron a 22°C en condiciones de día largo (16 h luz/ 8 h oscuridad), con luz blanca, fría e intensidad aproximada de 110 $\mu\text{Em}^{-2}\text{s}^{-1}$. Las plantas para genotipificación y aquellas que se usaron en la fertilización cruzada se crecieron en tierra estéril a 22°C en condiciones de día largo.

2.3. Obtención de las plantas dobles y triples mutantes de *clf ult1* y *atx1 clf ult1*

Las plantas dobles mutantes *clf ult1* se obtuvieron mediante fertilización cruzada a partir de mutantes simples previamente identificadas para cada gen individual. Para la planta triple mutante *atx1 clf ult1*, se generó a partir de cruzar la mutante doble *clf ult1* previamente generada con una mutante sencilla de *atx1*. La cruce para la doble mutante se realizó primero tomando como madre a la mutante de *ult1* y como padre a la mutante de *clf*. Debido a que *A. thaliana* es una especie hermafrodita y cleistógama, los estambres de cada flor de la planta madre fueron retirados en una etapa temprana evitando su autopolinización, posteriormente se tomaron flores maduras de la planta padre y se frotaron con los pistilos de las flores. Se procedió a crecer la descendencia y genotipificar a las plantas hasta encontrar la planta doble mutante homocigota, esto a través de extracción de DNA y posteriores amplificaciones de PCR. Esta doble mutante *ult1 clf* se usó como madre y una mutante *atx1* como padre, repitiendo el proceso anterior para así generar la triple mutante *atx1 clf ult1*.

2.3.1 Genotipificación de las mutantes

Para comprobar la homocigocidad de las plantas obtenidas en cada paso de la generación de las mutantes, se realizaron extracciones de DNA a partir de hojas de las plantas. Este tejido fue triturado con nitrógeno líquido hasta obtener un polvo fino. A cada muestra se le agregaron 700µl de buffer de extracción (Tris-HCl pH 7.5 1M, NaCl 5M, EDTA pH 8 0.5M, SDS 10%), dejando incubar a temperatura ambiente por 10 min en agitación. Posteriormente se agregaron 700µl de Fenol-Cloroformo-Alcohol isoamílico (25:24:1), se incubaron 10 min a temperatura ambiente y se centrifugó a 13000 rpm por 15 min. El sobrenadante fue transferido a otro tubo y se precipitó con 1 volumen de isopropanol dejando incubar por 10 min a temperatura ambiente, para posteriormente centrifugar a 13000 rpm por 15 min. Se desechó el isopropanol y el pellet obtenido fue lavado con EtOH al 70% y después de dejar evaporar el EtOH, se resuspendió en 30 µl de agua miliQ agregando 1 µl de RNAsa.

La presencia o ausencia de cada alelo se corroboró realizando ensayos de PCR de punto final. Cada reacción contuvo 1x PCR Rnx Buffer, 1.5 mM MgSO₄, 0,2 mM de dNTPs, 0.4 de *primers* (oligonucleótidos), 200 ng de DNA y 1 U de Taq DNA polimerasa (Invitrogene) en un volumen final de 15 µl. Las condiciones y *primers* ocupados para cada gen e inserción de TDNA se especifican en la Tabla A1 del anexo.

2.4 Cinética de crecimiento y caracterización del fenotipo de diferenciación de columela

El crecimiento de las raíces de Col-0 (Wt), *clf ult1* y *atx1 clf ult1* fue monitoreado durante 12 días postsiembra. Se registró la longitud de la raíz cada 24 h con una marca sobre la caja de Petri donde se observó la punta. Finalmente, se escanearon las placas y se midió la distancia entre cada marca con el procesador de imágenes *Fiji*, siguiendo el eje principal de crecimiento.

Para la caracterización del fenotipo, se observó raíces al microscopio óptico con la finalidad de observar la morfología del nicho de células troncales. De la misma forma se usó tinción de Lugol al 1% en la raíz para posteriormente observarse al microscopio óptico y así poder identificar los gránulos de almidón característicos de células de columela diferenciadas, pudiendo así observar el estado de diferenciación de estos tejidos, particularmente el de la columela.

La observación de raíces al microscopio se realizó con un tratamiento previo usando una solución de hidrato de cloral al 80% y glicerol al 10%. Las raíces observadas fueron a los 5dps y se observaron bajo óptica Nomarski en un microscopio Olympus BX60, a un aumento de 40X.

RESULTADOS

1. Análisis del número de genes diferencialmente expresados en las mutantes *atx1*, *sdg2*, *ult1* y *clf*

Para poder observar el comportamiento global de los genes con relación a su desregulación en cada una de las mutantes de Trithorax (*sdg2*, *atx1*, *ult1*) y Polycomb (*clf*), a partir de datos de microarreglos de expresión, se generó un análisis de genes diferencialmente expresados para cada caso. Los datos usados fueron obtenidos de los artículos, material biológico y ecotipo de *Arabidopsis* descritos a continuación:

- ATX1 (Pien et al., 2008):
 - Número de accesión: E-MEXP-502, ArrayExpress
 - Material biológico: Ápice vegetativo, plántulas de 10 días post germinación
 - Ecotipo: Ws-0
- SDG2 (Guo et al., 2010):
 - Número de accesión: GSE23208, NCBI
 - Material biológico: Plántula completa de 10 días
 - Ecotipo: Col-0
- ULT1 (Tyler et al., 2019):
 - Número de accesión: GSE137976, NCBI
 - Material biológico: Ápices vegetativos, plántulas de 4 días post germinación
 - Ecotipo: Ler
- CLF (Tyler et al., 2019)
 - Número de accesión: GSE137976, NCBI
 - Material biológico: Ápices vegetativos, plántulas de 4 días post germinación
 - Ecotipo: Ler

Ocupando los datos de dichos experimentos, fueron normalizados los datos de las lecturas de las sondas para cada microarreglo usando el método *Robust Microarray Average* (RMA), el cual disminuye la cantidad de falsos positivos y provee estimados más consistentes para los análisis de expresión diferencial basados en el *Fold Change* (Irizarry et al., 2003), esto a través de la función con el mismo nombre dentro del paquete *Affy*.

Posteriormente, con el paquete *limma* (también en R), se obtuvieron los genes diferencialmente expresados para cada mutante. En este flujo de trabajo, de un inicio se requiere obtener el modelo lineal ajustado para cada gen, esto para estimar el logaritmo de las intensidades entre dos o más muestras de forma simultánea, adicionalmente se obtienen estadísticos de prueba para la clasificación de genes. Lo anterior se hace utilizando la función *lmFit*, que ajusta los modelos dependiendo del conjunto de datos de forma robusta o por mínimos cuadrados, una vez hecho esto, la función *makeContrasts* se usan para hacer la matriz de contraste entre las plantas silvestres (Wt) y las mutantes (Ritchie et al., 2015). Los datos de la matriz de contraste y del modelo lineal ajustado se ocupan para calcular los cambios de expresión basados el *Fold Change (FC)* que es el que se usó para obtener los cambios, también se obtuvieron los estadísticos necesarios como el valor *p* ajustado (Ritchie et al., 2015).

Cabe mencionar que antes de obtener las tablas finales de los genes sobre- y subexpresados, los datos de la matriz de contraste y los producidos por *lmFit* se pasan por la función *contrasts.fit* para estimación de coeficientes y errores estándar. Posteriormente los datos resultantes se pasan por la función *eBayes*, que computa estadísticos t y F moderados y el logaritmo de las probabilidades de los datos de expresión diferencial a través de moderación empírica bayesiana de los errores estándar.

Primeramente, para la obtención de los genes diferencialmente expresados en cada mutante respecto a las plantas silvestres, se realizó un corte de aquellos genes que contaran con un valor de *p* ajustado menor o igual a 0.05 y de aquel subconjunto, el umbral que se eligió para considerar un gen como sobreexpresado (Up) o subexpresado (Down), fue ocupando el *FC* y que éste fuera mayor a 1.5 y menor a -1.5 respectivamente.

De los 22746 genes contenidos en un microarreglo, los genes significativamente desregulados para *atx1* fueron un total de 2168, para *clf* 1380, para *ult1* 333 y el que presentó un mayor número fue *sdg2* con un total de 3956 genes. Para tener una idea del impacto general que la falta de estas proteínas tiene sobre la expresión en *A. thaliana*, y debido a que todos los conjuntos de datos procedían de microarreglos de la misma compañía (*Affymetrix*), se graficó que porcentaje del total de los genes presentes en dicho microarreglo (los 22746 genes) que se desregulaban a la baja y al alza en cada una de las mutantes (Figura 6). De esto se observó que la mutante que presenta un mayor número de genes desregulados es *sdg2* (17.6%) y después en orden decreciente *atx1* (9.7%), *clf* (6.1%) y por último *ult1* (1.5%).

Respecto a los sentidos de las desregulaciones, para *atx1* se presentaron un total de 1295 genes desregulados al alza y 873 a la baja, representando un 59.7% de sus genes los sobreexpresados y 40.3% de estos desregulados a la baja. En el caso de *clf* se presentaron 850 genes desregulados arriba y 530 desregulados a la baja, representado un 61.6% del total de los genes desregulados en esta mutante, al alza y 38.4% a la baja. Para *ult1* se presentó un total de 223 genes sobreexpresados y 110 subexpresados, presentando una proporción del 66.9% de los genes al alza y 33.1% a la baja. Finalmente, con *sdg2* se presentaron 1973 genes sobreexpresados y 1983 subexpresados; lo cual, del total de genes desregulados en

esta mutante, los sobrepresados representan el 49.9% y los desregulados a la baja el 50.1%.

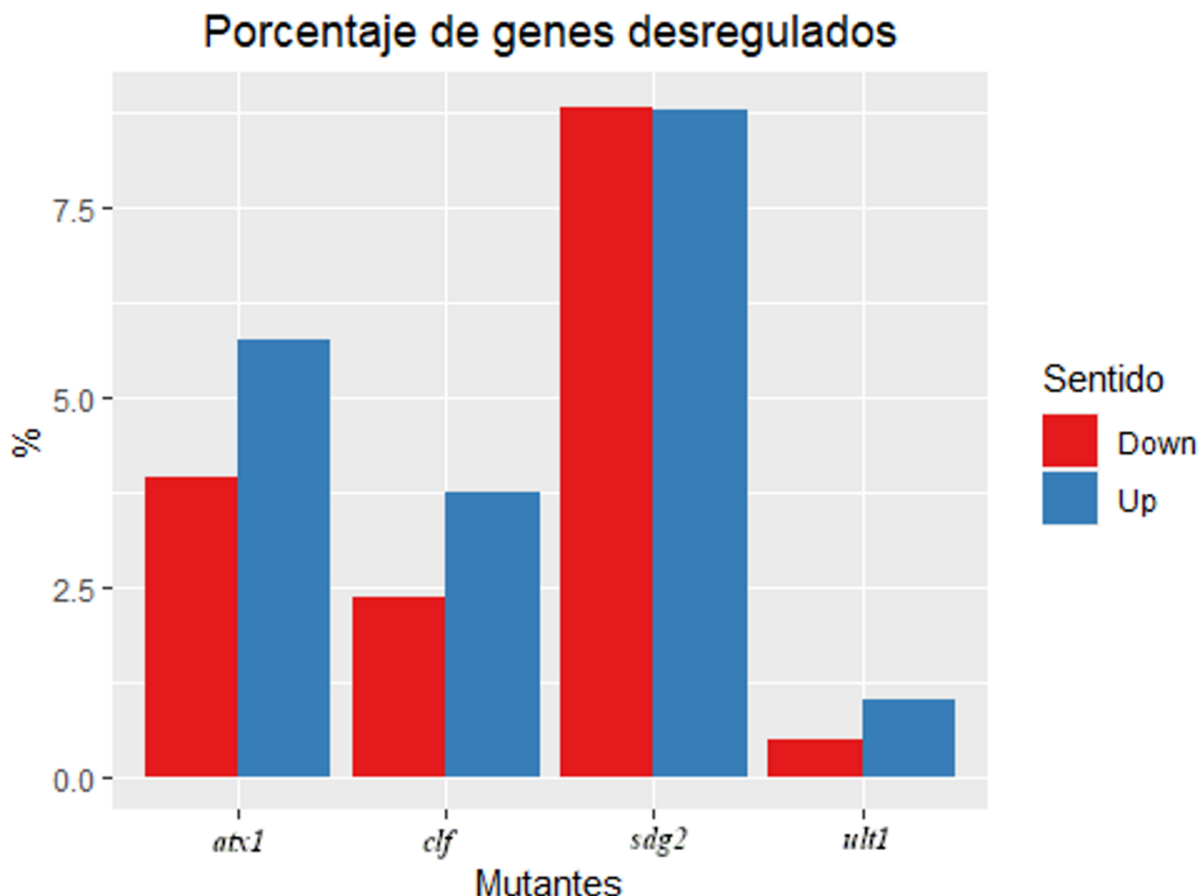


Figura 6. Genes diferencialmente expresados en mutantes de proteínas del grupo Trithorax y Polycomb. Porcentaje de genes desregulados al alza (Up) en rojo y a la baja (Down) en azul en cada una de las mutantes (*atx1*, *clf*, *sdg2* y *ult1*) respecto al total de genes contenidos en el microarreglo de *Affymetrix*

De manera interesante se esperaría observar que las proteínas del grupo Trithorax (con función de activación) al no estar presentes, la desregulación de genes se viera sesgada a la baja y en contraste, proteínas de Polycomb (con función de represión) con un sesgo al alza, sin embargo, en ningún caso se ve el sesgo claro y como se puede observar por los porcentajes inclusive en el caso de *atx1* y *ult1* la poca diferencia existente entre los genes a sub- y sobrepresados está al revés de lo esperado.

Cabe destacar que, para validar los resultados obtenidos en este trabajo, en primera instancia se buscó, para cada conjunto de datos de cada mutante, el gen mutado dentro de los genes pertenecientes al subconjunto de los subexpresados. Es decir, que *ATX1*, se encontrara subexpresado en la mutante de *atx1*, *CLF* en la mutante de *clf* y así sucesivamente. Contrario a lo esperado, únicamente *CLF* y *ATX1* cumplieron con este requisito de encontrarse en la tabla de los genes que más bajaban su expresión, *CLF*

presentó un *logFC* de -1.2 en la mutante *clf*, mientras que *ATX1* presentó un *logFC* de -1.8 en la mutante de *atx* (Figura S1a).

Extrañamente para *ULT1*, no se encontró la expresión de este gen entre las que cumplían las características de aquellos subexpresados ($FC \leq -1.5$, $p \leq 0.04$), por lo que la validación de este análisis fue hecha a partir de la reproducibilidad de los datos obtenidos en este trabajo respecto a los genes encontrados como diferencialmente expresados en el artículo donde provienen el microarreglo (Tyler et al., 2019). Se seleccionaron 20 genes de los reportados como diferencialmente expresados y se graficaron para contrastar su *logFC*, con lo obtenido en este trabajo, y como se observa en la Figura S1b en todos los casos el cambio es prácticamente idéntico, tanto en magnitud como en sentido. Cabe destacar que gracias a análisis transcriptómicos no publicados del laboratorio, tenemos información de que cuando se tiene la mutación homocigota *ult1-3*, hay un enriquecimiento de RNAs no codificantes provenientes del locus de *ULT1*, que, si bien no son funcionales, en este tipo de análisis de expresión, pueden ser confundidos con mensajeros de este gen.

Finalmente, para la validación de *sdg2*, ocurrió algo aún más anómalo: el análisis devolvió la expresión de *SDG2*, en el conjunto de genes sobreexpresados, con un *logFC* de 2.06 (Fig. S1a). Esto evidentemente destaca bastante, porque la mutante de donde se extrajo el RNA, esta identificada como de pérdida de función y en el artículo fuente no mencionan algo al respecto (Guo et al., 2010). Sin embargo, en dicho artículo de origen hicieron la validación de 17 genes contrastando su microarreglo con PCRs cuantitativas para ver si el cambio observado era el mismo, y para este trabajo se ocuparon esos mismos 17 loci, logrando reproducir dichos resultados, en magnitud y sentido (Figura S1c). Esto puede indicar dos cosas, la primera es que este ocurriendo un fenómeno similar al explicado con *ult1*, en el que la mutación provoca la transcripción de RNAs no codificantes no funcionales, que además tendría sentido si se toma en cuenta que *SDG2* es un gen muy grande y con muchos intrones (la secuencia completa tiene 11182 pb con 20 exones y 20 intrones). La segunda opción, es que el enriquecimiento de *SDG2* en la mutante *sdg2*, responda a errores metodológicos realizados por Guo y colaboradores. Sin embargo, debido a que este conjunto de datos representa el único microarreglo disponible para esta mutante, y que la evidencia de la gran cantidad de genes desregulados soporta fuertemente el que todo esté bien, se procedió a ocupar este experimento y apoyar a la primera hipótesis de la presencia de transcritos no funcionales.

2. Genes regulados de forma conjunta e independiente entre proteínas del grupo Trithorax

Entendiendo que cada conjunto de datos proviene de tejidos distintos y plantas de edades distintas, cabe resaltar que la comparabilidad entre los datos obtenidos debe manejarse con cuidado. Sin embargo, lo que se busca es obtener (más que genes específicos coregulados por las proteínas a analizar), comportamientos que reflejen una regulación génica general por las proteínas epigenéticas ya mencionadas. A partir de los conjuntos de datos de genes diferencialmente expresados, se generaron diagramas de Venn para observar las concordancias en cuanto a regulación entre las proteínas del grupo Trithorax, los genes

que se encuentran desregulados de forma compartida entre más de una mutante podrían indicar redundancia o acción conjunta para que se dé la correcta expresión de éstos (Figura 7a). En conjunto, *atx1*, *sdg2* y *ult1* comparten un total de 40 genes desregulados, *ult1* a pesar de tener significativamente menos genes desregulados que las otras dos proteínas, sí comparte un gran número de genes con ambas mutantes: 97 genes con *atx1* (el 29.1% de sus genes desregulados) y 136 con *sdg2* (40.8% de sus genes desregulados), de los cuales 57 solo son compartidos con *atx1* y 96 con *sdg2*. Adicionalmente, para ahondar, no solo en la cantidad de genes compartidos, sino en el sentido de la regulación, se generó un *Heatmap* ocupando el *logFC* del cambio en la expresión dentro de las mutantes, para cada uno de los 40 genes compartidos entre *atx1*, *sdg2* y *ult1*, y como se puede observar en la Figura 7b, resulta contraintuitivo observar que no todos estos genes compartidos se desregulan en el mismo sentido en las tres mutantes, mostrando que a pesar de existir coregulación, esta no siempre es sinérgica. También cabe destacar, que no parecen coincidir más *sdg2* con *atx1*, *ult1* con *atx1* o *ult1* con *sdg2* en el sentido de regulación de dichos 40 genes compartidos, es decir, no existe un sesgo claro compartido entre estas tres proteínas Trithorax.

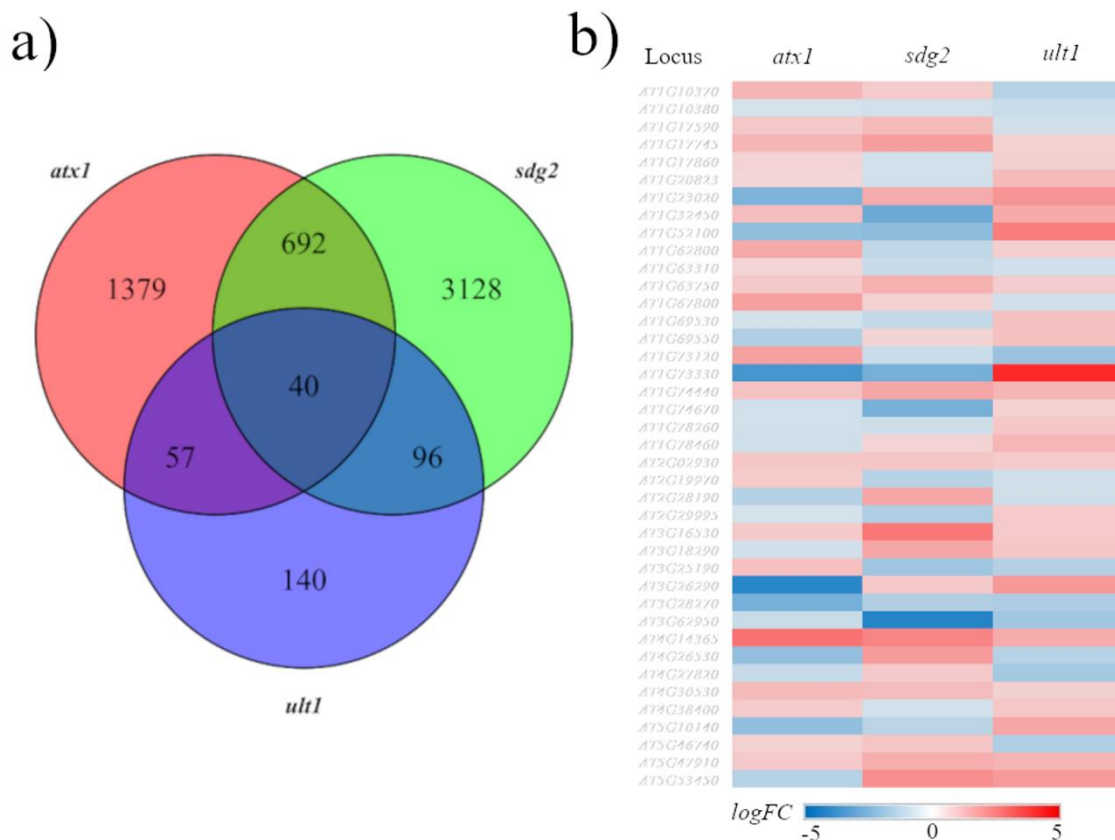


Figura 7. Genes desregulados compartidos entre los miembros de Trithorax. a) Diagrama de Venn mostrando los genes compartidos y exclusivos en las tres proteínas de Trithorax analizadas. b) *Heatmap* de los cuarenta genes compartidos entre las *atx1*, *sdg2* y *ult1*

Curiosamente, a pesar de que tanto ATX1 como SDG2 comparten la misma acción enzimática de colocar la marca de la H3K4me3, los subconjuntos de genes desregulados no tienen dominios tan sobrelapados como se esperaría, ya que únicamente comparten 732 genes (el 33.8% de los genes desregulados en *atx1* y el 18.5% de los desregulados en *sdg2*) (Fig. 7a). También cabe destacar que *sdg2* a pesar de ser reportada como la más pleiotrópica de las HKMTs de la K4, no absorbe totalmente el conjunto de genes desregulados en *atx1*, conservando esta última 1379 genes que no cambian en *sdg2*, pero *sdg2* presenta una cantidad muy superior de genes desregulados de forma exclusiva: 3128 genes.

3. Relación de los genes regulados de forma conjunta e independiente entre proteínas del grupo Trithorax y CLF como representante de Polycomb

3.1 Interacciones con *ult1*

Por su parte, *ult1* como ya se mencionó al principio representa una proteína importante debido a las múltiples posibles interacciones que puede tener en su papel como reguladora, no solo actúa a través de Trithorax, sino que también puede actuar con factores de transcripción (Ornelas-Ayala et al., 2021) y con Polycomb (Ornelas-Ayala et al., 2021; Xu et al., 2018). Es por esto que es relevante analizar no solo los genes que comparte con *atx1* y *sdg2*, sino también con *clf*. Como podemos ver en la Figura 8, en lo general contrario a lo que se pudiera pensar, no hay un sesgo preferencial de genes coregulados entre ULT1

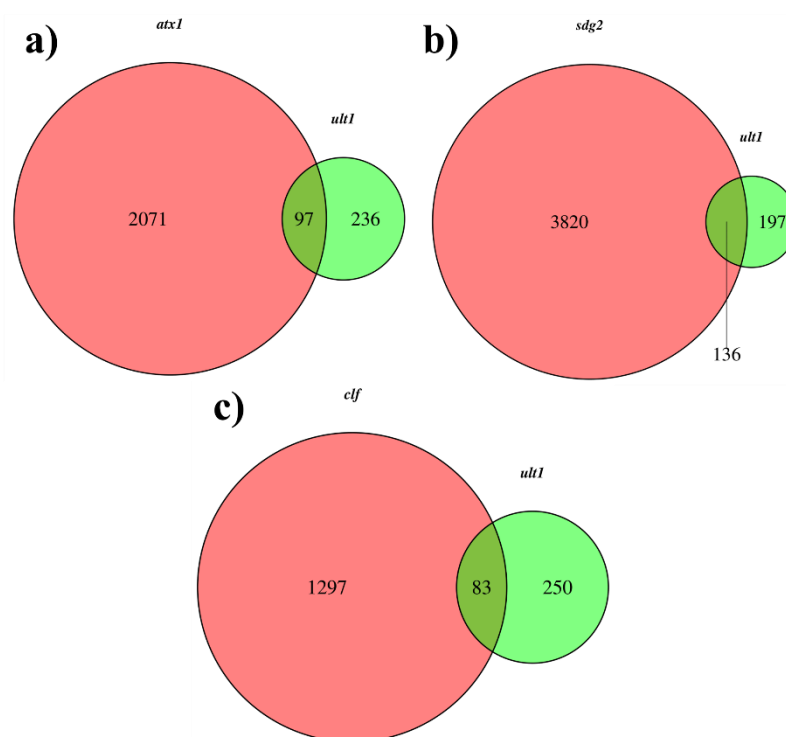


Figura 8. ULT1 comparte genes desregulados en proporciones similares con SDG2, ATX1 y CLF. Diagramas de Venn mostrando los genes compartidos y exclusivos entre *ult1* y a) *atx1*, b) *sdg2* y c) *clf*.

y las dos HKMTs de Trithorax sobre CLF. Sin embargo, con esta proteína comparte un total de 83 genes, representando el 24.9% del total de genes desregulados en *ult1* (Fig. 8c) lo cual es comparable con el 29.1% compartidos con *atx1* (Fig. 8a), destacando únicamente lo compartido con *sdg2* que como ya se mencionó anteriormente, representa el 40.8% de los genes desregulados en *ult1* (Fig. 8b), evidenciando que a pesar de que lo compartido con las dos proteínas Trithorax es mayor a lo que ocurre con CLF, no es un sesgo tan obvio si se toma en cuenta que la diferencia mayor es con *sdg2* y de por sí esta proteína tiene un número mucho mayor de genes desregulados.

Continuando con la exploración de esta posible interacción de *ult1* independiente de Trithorax, particularmente con *clf*, se realizaron los diagramas de Venn de los genes co-desregulados con los conjuntos correspondientes al sentido esperado acorde a la actividad canónica de ambas proteínas: al alza en *clf* (al tratarse de un represor) y a la baja en *ult1* (al ser miembro de Trithorax y ayudar a la activación). Adicionalmente se probó el subconjunto compartido entre los desregulados en sentido para *clf* y al alza también en *ult1*, pensando en la posibilidad de que pudiera haber blancos que este último regulase de manera negativa. Como podemos ver en la Figura 9 son realmente muy pocos los genes compartidos entre ambas mutantes en el sentido esperado tomando el modelo canónico de un antagonismo entre ambas proteínas (únicamente 13 genes, el 11.8% de los genes desregulados a la baja en *ult1*). De igual forma, aunque en una proporción mayor, son pocos los genes compartidos desregulados de forma positiva en ambas mutantes (un total de 48 genes, el 21.5% de todos los desregulados al alza en *ult1*), esto abre la posibilidad de que exista una coregulación de al menos algunos blancos entre CLF y ULT1.

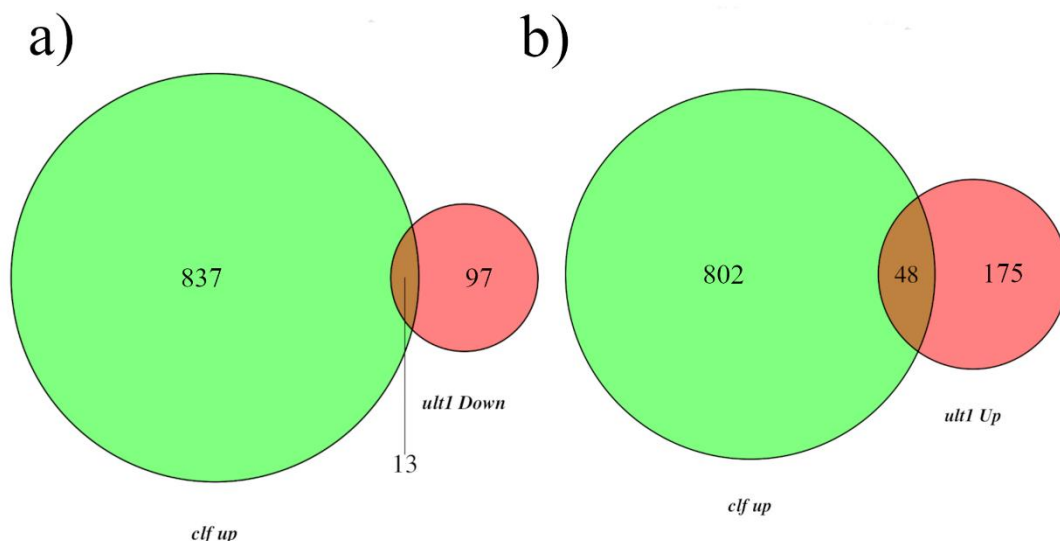


Figura 9. Genes desregulados compartidos entre *ult1* y *clf*. En a) se presenta el diagrama de los genes desregulados en los sentidos esperados respecto a su acción canónica: en *ult1* genes subexpresados al tratarse de un activador y en *clf* genes sobreexpresados al tratarse de un represor y en b) los genes desregulados en el sentido esperado para *clf* (sobreexpresados) y el contrario para *ult1* (sobreexpresados).

3.2 Interacciones entre CLF y las HKMTs de Trithorax

La hipótesis intuitiva sería pensar que los genes desregulados a la alza en *clf* estarían compartidos en gran medida con los genes desregulados a la baja en *sdg2* o *atx1*, debido a que las tres proteínas tienen la misma función enzimática de catalizar la deposición de la trimetilación de lisinas de la histona H3, y particularmente de poner marcas que señalizan regulaciones opuestas: CLF pone, como ya se mencionó, la H3K27me3 de represión transcripcional y SDG2/ATX1 ponen la marca de H3K4me3 que antagónicamente señala activación transcripcional. Sin embargo, al realizar la comparativa como podemos ver a través de los diagramas de Venn presentados en la Figura 10, no es esto lo que ocurre.

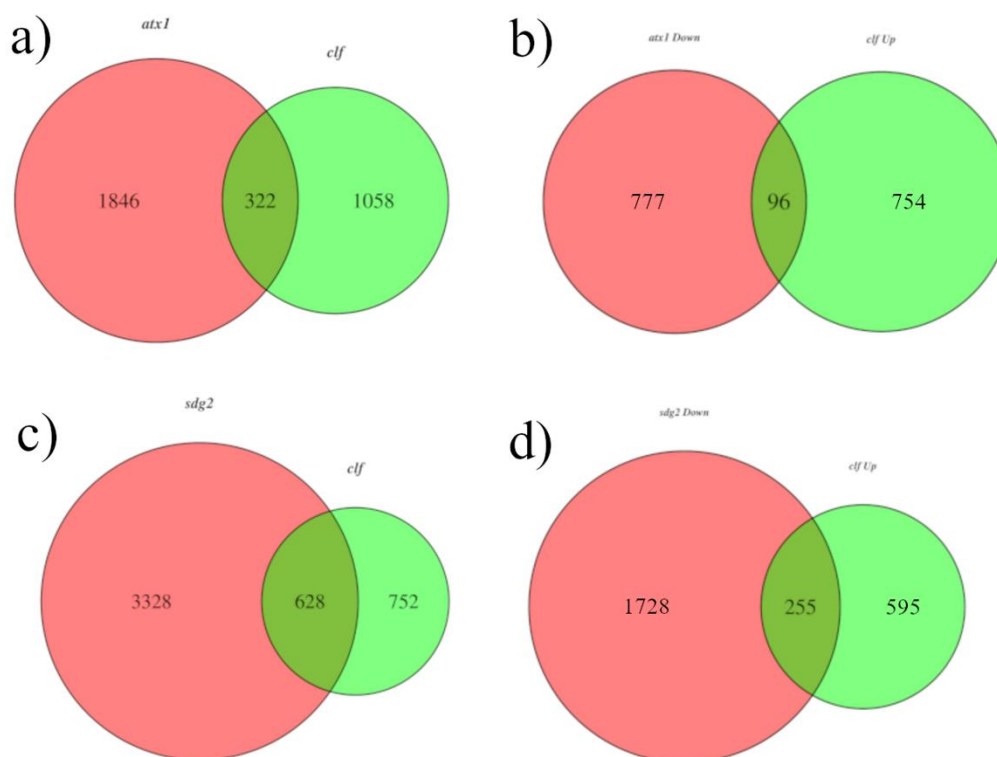


Figura 10. Genes desregulados compartidos entre las HKMTs de Trithorax (*atx1* y *sdg2*) y *clf*. En a) y c) se presenta el diagrama de los genes desregulados en cualquier sentido (a la baja y al alza) entre *clf* y (a) *atx1* o c) *sdg2* y en b) y d) los genes desregulados en los sentidos esperados respecto a su función canónica entre *clf* al alza al ser un represor y b) *atx1* o d) *sdg2* hacia la baja al tratarse de activadores transcripcionales.

Teniendo la comparativa entre el total de genes desregulados en las mutantes de *sdg2/atx1* y *clf* en cualquier sentido y la comparativa en los sentidos esperados acorde a su función, se esperaba que se compartiera un conjunto mayor de genes en el segundo caso (si el comportamiento de las proteínas fuera predominantemente antagónico). Sin embargo, en ambos casos no se observó esto, *atx1* comparado con *clf* en cualquier sentido compartieron únicamente 322 genes (representando el 14.9% del total de genes desregulados en *atx1* y el 23.3% de los de *clf*), número que se redujo al analizarse las desregulaciones en sentido a su función, a solo 255 genes compartidos (representando el 10.9% del total de genes

desregulados a la baja en *atx1* y el 11.3% de los desregulados al alza en *clf*). Lo mismo ocurre en *sdg2* en donde la comparación de los genes con *clf* únicamente presento un subconjunto de 628 genes compartidos entre ambos (siendo el 15.9% del total de genes desregulados en *sdg2* y el 45.5% de los de *clf*). Al igual que con *atx1*, el número se reduce en la comparativa con los genes en sentido de la regulación llevada a cabo por cada proteína: únicamente se desregulan 255 genes (representando el 12.9% del total de genes desregulados a la baja en *sdg2* y el 30% de los desregulados al alza en *clf*). En principio, además de los pocos genes compartidos en ambos casos, pareciera que SDG2 regula más genes de forma conjunta y antagónica a CLF que ATX1.

4. Procesos biológicos asociados a los genes desregulados en las mutantes de *atx1*, *sdg2*, *ult1* y *clf*

Para tener una idea acerca de los procesos biológicos en los cuales se encuentran involucradas las proteínas analizadas de Trithorax y Polycomb se realizó una ontología de genes ocupando los conjuntos de aquellos genes desregulados en el sentido esperado acorde a la función de cada proteína. Este análisis se realizó ocupando el paquete de R de *ClusterProfiler*, específicamente la función *enrichGO* y ocupando la base de datos del genoma de *Arabidopsis* con sus anotaciones contenida en el paquete *org.At.tair.gb*.

Este análisis mostró que para los genes desregulados a la baja en *atx1* (Fig. 11a), se tiene un enriquecimiento importante en genes relacionados a respuesta a estímulos externos, pertenecientes a las categorías de: *Respuesta a frío*, *respuesta celular a químicos ácidos*, *procesos del sistema inmune*, *respuesta a estrés abiótico*, *respuesta a estímulos ambientales*; también presentó enriquecimiento a procesos relacionados a la respuesta a luz, de las categorías: *Respuesta a luz roja o roja lejana*, *respuesta a intensidad de luz*, *respuesta a estímulos de luz*, *respuesta a radiación* y *respuesta a luz roja*; a procesos relacionados a la respuesta a carbohidratos, de las categorías: *Respuesta a karrikina*, *respuesta a carbohidratos*, *respuesta a disacáridos*, *respuesta a sacarosa* y *respuesta a fructosa* y finalmente también tuvo dos categorías enriquecidas relacionadas a el ciclo circadiano: *Ritmo circadiano* y *Procesos rítmicos*.

Respecto a los genes desregulados a la baja en *sdg2* (Fig. 11b), tiene un muy importante enriquecimiento en procesos relacionados al oxígeno, presentando las categorías de: *Respuesta a estrés oxidativo*, *respuesta a hipoxia*, *respuesta celular a hipoxia*, *respuesta a bajos niveles de oxígeno*, *respuesta celular a bajos niveles*, *respuesta a niveles de oxígeno*, *de oxígeno* y *respuesta celular a niveles de oxígeno*. También presentó categorías relacionadas a la respuesta a agua: *Respuesta a agua* y *respuesta a falta de agua*, relacionadas a respuesta hormonal: *Respuesta a auxinas*, *respuesta a estímulos de citoquininas* y *respuesta a la vía de señalización activada por citoquininas*, relacionado al ciclo circadiano: *Ciclo circadiano* y *Regulación del ciclo circadiano*, a respuestas a luz: *Respuesta a luz roja o luz roja lejana* y *respuesta a luz roja* y finalmente presentó enriquecidas las categorías de: *Respuesta a lípidos* y *tropismo*.

Por último, los genes desregulados al alza en la mutante *clf* (Fig. 11c), también presentaron un enriquecimiento importante en genes relacionados al oxígeno: *Respuesta a*

hipoxia, respuesta celular a hipoxia, respuesta a bajos niveles de oxígeno, respuesta celular a bajos niveles, respuesta a niveles de oxígeno, de oxígeno y respuesta celular a niveles de oxígeno. Adicionalmente presentó un gran número de categorías relacionadas a respuesta a patógenos, específicamente: *Respuesta de defensa a bacterias, procesos del sistema inmune, respuesta inmune, respuesta inmune innata, respuesta a hongos, respuesta a ácido salicílico, respuesta de defensa a hongos, respuesta de defensa a interacciones incompatibles, regulación de la respuesta de defensa y resistencia sistémica adquirida*. Por último, el análisis arrojó otro grupo de categorías de los genes desregulados relacionadas con la respuesta a compuestos químicos: *Respuesta a fármacos, respuesta a compuestos*

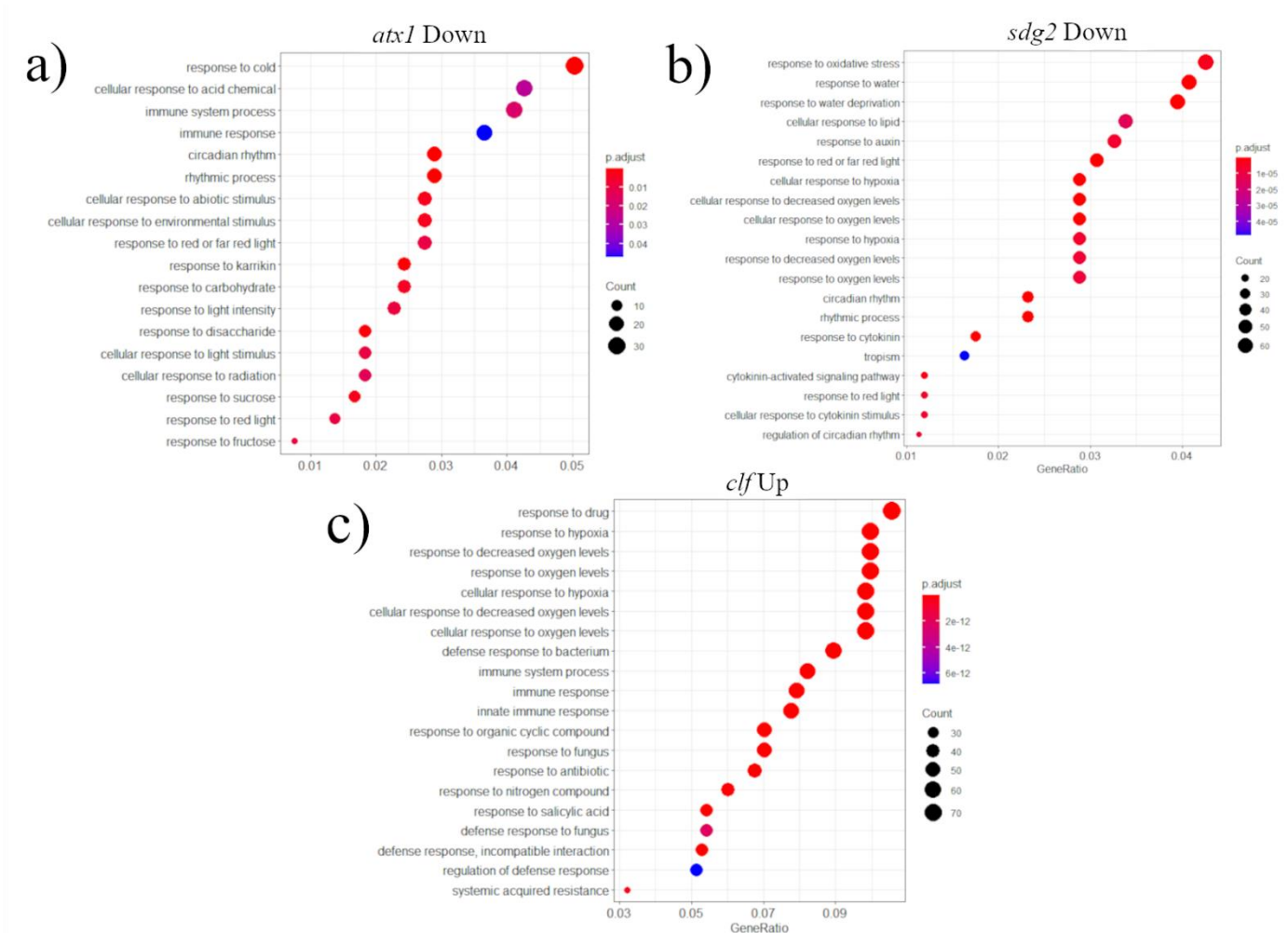


Figura 11. Categorías enriquecidas de los procesos biológicos desregulados entre las HKMTs de *Trithorax* (*atx1* y *sdg2*) y *clf*. En a) se presenta el diagrama de puntos con las categorías enriquecidas de los genes desregulados a la baja para la mutante *atx1*, en b) para *sdg2*, de igual manera para los genes a la baja y en c) los genes desregulados al alza en *clf*.

orgánicos cíclicos, respuesta a antibióticos y respuesta a compuestos nitrogenados.

Es interesante puntualizar que de las tres mutantes, *clf* fue la que tuvo categorías más relacionadas entre sí. A pesar de tener la misma función enzimática *atx1* y *sdg2* únicamente comparten las categorías relacionadas al *ciclo circadiano* y de *respuesta a luz y luz roja*. También destaca el hecho que *atx1* y *clf* en estos conjuntos de genes desregulados

en el sentido esperado acorde a su función antagonónica, no comparten más que la categoría de *procesos asociados al sistema inmune*, mientras que con *sdg2*, *clf* comparte todos los procesos relacionados a respuesta a oxígeno (*respuesta a hipoxia*, *respuesta celular a hipoxia*, *respuesta a bajos niveles de oxígeno*, *respuesta celular a bajos niveles*, *respuesta a niveles de oxígeno*, *de oxígeno* y *respuesta celular a niveles de oxígeno*) sugiriendo la existencia de más antagonismo entre estas dos últimas, que entre *clf* y *atx1*.

5. Procesos biológicos asociados a los genes desregulados en la mutante de *ult1*

Consecuentemente se realizó este mismo análisis de ontología de genes para la mutante de *ult1*. Sin embargo, en contraste a las anteriores donde sólo se realizó con los conjuntos de genes desregulados en los sentidos esperados acorde a su función, en este caso se realizó tanto para los que se desregulaban a la baja como al alza siguiendo la línea de que ULT1 puede tener interacciones diversas con proteínas tanto de Trithorax como de Polycomb y que, como se observó en los diagramas de Venn anteriormente descritos, comparte genes desregulados con las otras tres proteínas analizadas en este trabajo.

Del análisis realizado para el conjunto de genes que se desregula a la baja en esta mutante (en sentido a su actividad conjunta con Trithorax) (Fig. 12a), se obtuvo un enriquecimiento importante en categorías relacionadas con la homeostasis y respuesta a iones metálicos: *Homeostasis iónica*, *homeostasis de iones metálicos*, *homeostasis celular de iones metálicos*, *homeostasis celular de cationes*, *homeostasis celular de iones*, *homeostasis celular de hierro*, *respuesta a iones de hierro*, *homeostasis de metales de transición*, *secuestro intracelular de iones de hierro*, *secuestro de iones de hierro*, *secuestro de iones metálicos*. También presentó categorías relacionadas con la respuesta a compuestos químicos diversos: *Homeostasis química*, *respuesta a sustancias tóxicas*, *respuesta celular a sustancias tóxicas*, *respuesta a especies reactivas de oxígeno*, *detoxificación*, categorías relacionadas a la a la localización celular: *mantenimiento de la posición en la célula* y *mantenimiento de la localización* y finalmente, la categoría más enriquecida fue la de *homeostasis celular*.

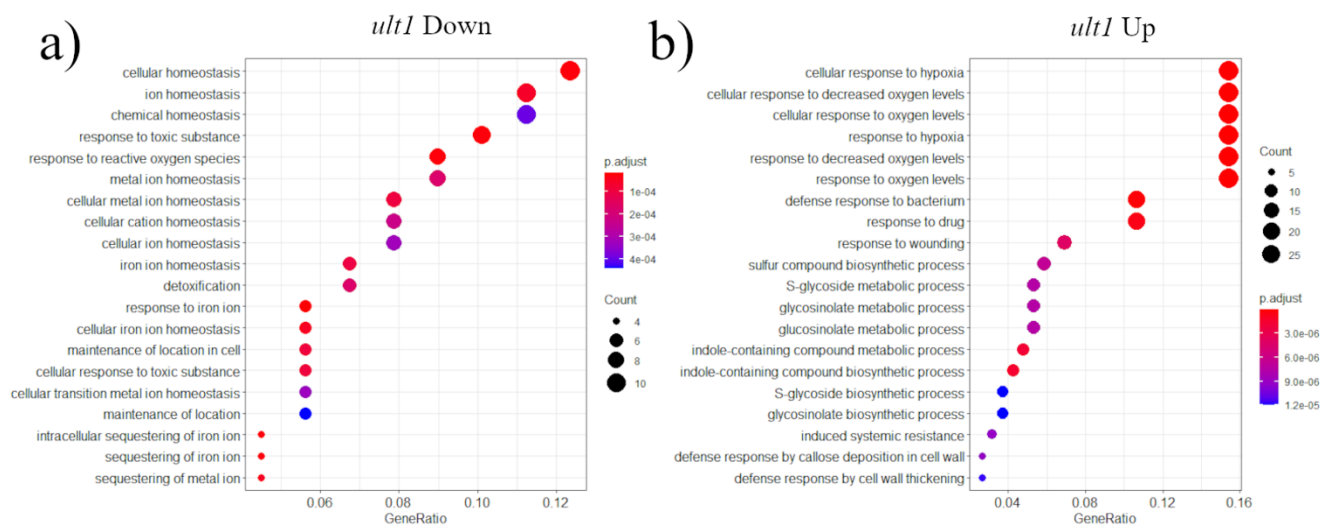


Figura 12. Categorías enriquecidas de los procesos biológicos desregulados *ult1*. En a) se presenta el diagrama de puntos con las categorías enriquecidas de los genes desregulados a la baja para la mutante de *ult1* y en b) al alza para esta misma mutante

Respecto a las categorías enriquecidas en aquellos genes que se desregulan al alza en la mutante de *ult1* (Fig. 12b), se tiene categorías enriquecidas relacionadas con el oxígeno: *Respuesta celular a hipoxia*, *respuesta a hipoxia*, *respuesta celular a bajos niveles de oxígeno*, *respuesta a bajos niveles de oxígeno* y *respuesta a niveles de oxígeno*. También, de manera importante se observa un enriquecimiento a cuestiones relacionadas a compuestos con azufre: *Procesos biosintéticos de compuestos de azufre*, *procesos metabólicos de S-glucósidos*, *procesos metabólicos del glucosinolato*, *procesos biosintéticos de S-glucósidos* y *Procesos biosintéticos del glucosinolato*. Presenta también categorías relacionadas a la defensa contra patógenos: *Respuesta de defensa a bacterias*, *respuesta a heridas*, *resistencia sistémica inducida*, *respuesta de defensa por deposición de calosa en la pared celular* y *respuesta de defensa por engrosamiento de la pared celular* y finalmente categorías relacionadas a otro tipo de químicos: *Respuesta a fármacos*, *procesos metabólicos de compuestos con indoles* y *procesos biosintéticos de compuestos con indoles*.

Es interesante destacar que las categorías a las que pertenecen los genes desregulados a la baja (en el sentido esperado si ULT1 actuase principalmente como un coactivador del grupo Trithorax) no son compartidas por las otras mutantes analizadas. Contrastantemente, dentro de las categorías enriquecidas presentes en los genes desregulados al alza, sí existen varias compartidas con los desregulados a la baja en *sdg2* y al alza con *clf*, con el *sdg2* principalmente todas aquellas categorías relacionadas al oxígeno y con *clf* las de *respuesta de defensa a bacterias* y *respuesta a fármacos*. Lo anterior sugiere la existencia de una regulación de forma antagónica de algunos genes entre SDG2 y ULT1, asimismo, la existencia de una coregulación de otro subconjunto de genes entre CLF y ULT1.

6. Generación de la doble y triple mutante de *clf ult1* y *ult1 clf atx1* como herramientas para describir la interacción entre TrxG y PcG

Con el propósito de generar herramientas para continuar el estudio a futuro de las posibles interacciones entre Trithorax y Polycomb, se realizó, una doble mutante *clf ult1* y una triple mutante *atx1 ult1 clf*, a través de cruza y posterior genotipificación de la descendencia. Lo anterior para que a futuro se puedan hacer distintos ensayos como lo puede ser el análisis de expresión génica comparando las mutantes sencillas o las distintas combinaciones de las dobles. Para este proyecto en particular serán usadas para observar fenotipos específicos que se describirán a más adelante.

Después de la fertilización cruzada entre mutantes previamente identificadas como homocigotas para las mutaciones sencillas de *ult1* y *clf*, las semillas producto de esta cruza se crecieron. De estas plantas a través de un ensayo PCR se probó la homocigocidad de cada mutación, como podemos ver en la Figura 13, la planta marcada como *clf-29 ult1-3* 4.4.4.13.1 no amplificó para el gen endógeno de *CLF* (1277 pb) (Fig. 13a) y sí amplificó para la inserción de T-DNA *clf-29* (± 950 pb) (Fig. 13b). Además, no amplificó para el gen endógeno de *ULT1* (900 pb) (Fig. 13c) y sí amplificó para la inserción de T-DNA *ult1-3*

(±400 pb) (Fig. 13d), indicando así que la planta 4.4.4.13.1 es mutante homocigota para *clf* y *ult1*.

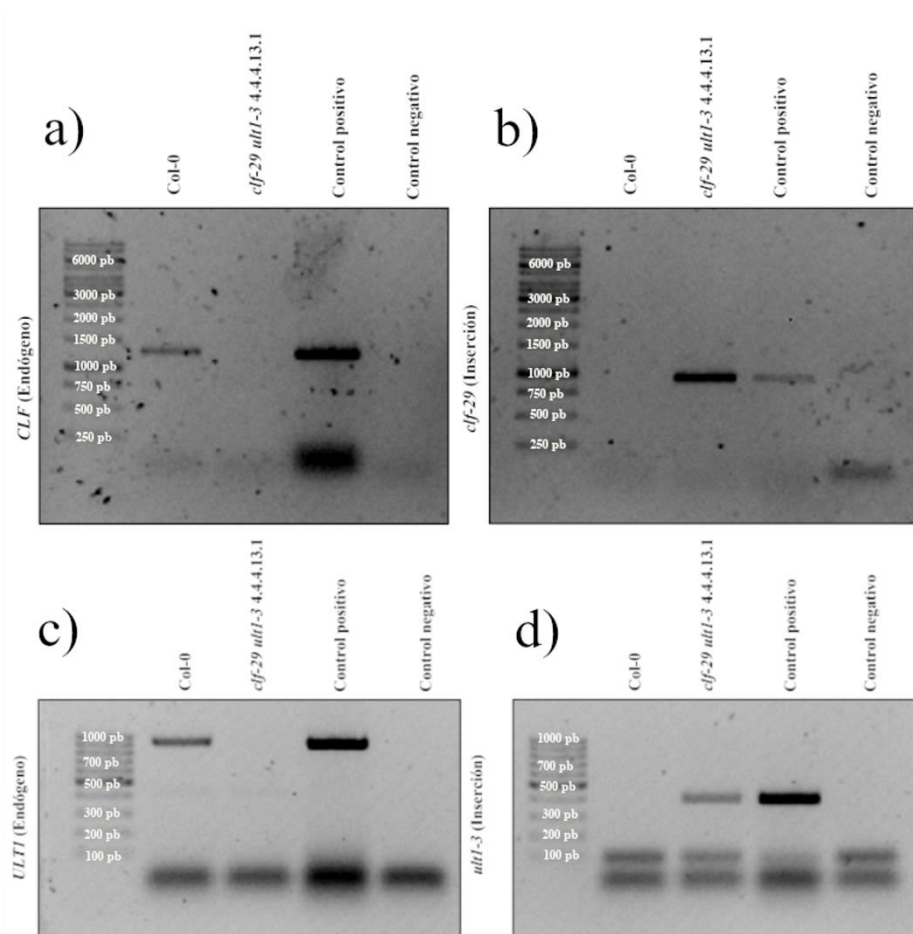


Figura 13. Genotipificación de la doble mutante homocigota *clf ult1*. En a) el gel de electroforesis para la amplificación del gen endógeno de *CLF*, donde solo se observa en la planta Col-0 (Wt) y el control positivo, en b) la amplificación de la inserción de TDNA *clf-29*, donde solo la mutante lo amplifica y su respectivo control positivo. c) Amplificación del endógeno de *ULT1*, donde solo la planta Col-0 y su respectivo control positivo lo amplifica, b) Amplificación de la inserción de TDNA *ult1-3*, donde solo la mutante y su respectivo control positivo presentan banda. Marcador de peso molecular para a) y b): *Gene Ruler 1kb DNA Ladder* de ThermoScientific y para c) y d): *100 bp ladder* de NEB.

Respecto a la generación de la triple mutante *atx1 clf ult1*, una vez identificadas la planta doble mutante homocigota *clf ult1* mencionada anteriormente, esta fue cruzada con una mutante simple homocigota a *atx1* y fue seguida la descendencia producto de esta cruce hasta identificar la planta homocigota para las tres mutaciones. Como se puede observar en la Figura 14, la planta marcada como *atx1-3 clf-29 ult1-3 2.3.1.7.3* no amplificó para el gen endógeno de *ATXI* (937 pb) (Fig. 14a) y sí amplificó para la inserción de T-DNA *atx1-3* (±610 pb) (Fig. 14b). Además, no amplificó para el endógeno de *CLF* (1277 pb) (Fig. 14c) y sí amplificó para la inserción de T-DNA *clf-29* (±950 pb) (Fig. 14d) y no amplificó para

el gen endógeno de *ULT1* (900 pb) (Fig. 14e) pero sí amplificó la inserción de T-DNA *ult1-3* (± 400 pb) (Fig. 14f), obteniendo así la triple mutante homocigota *atx1 clf ult1*.

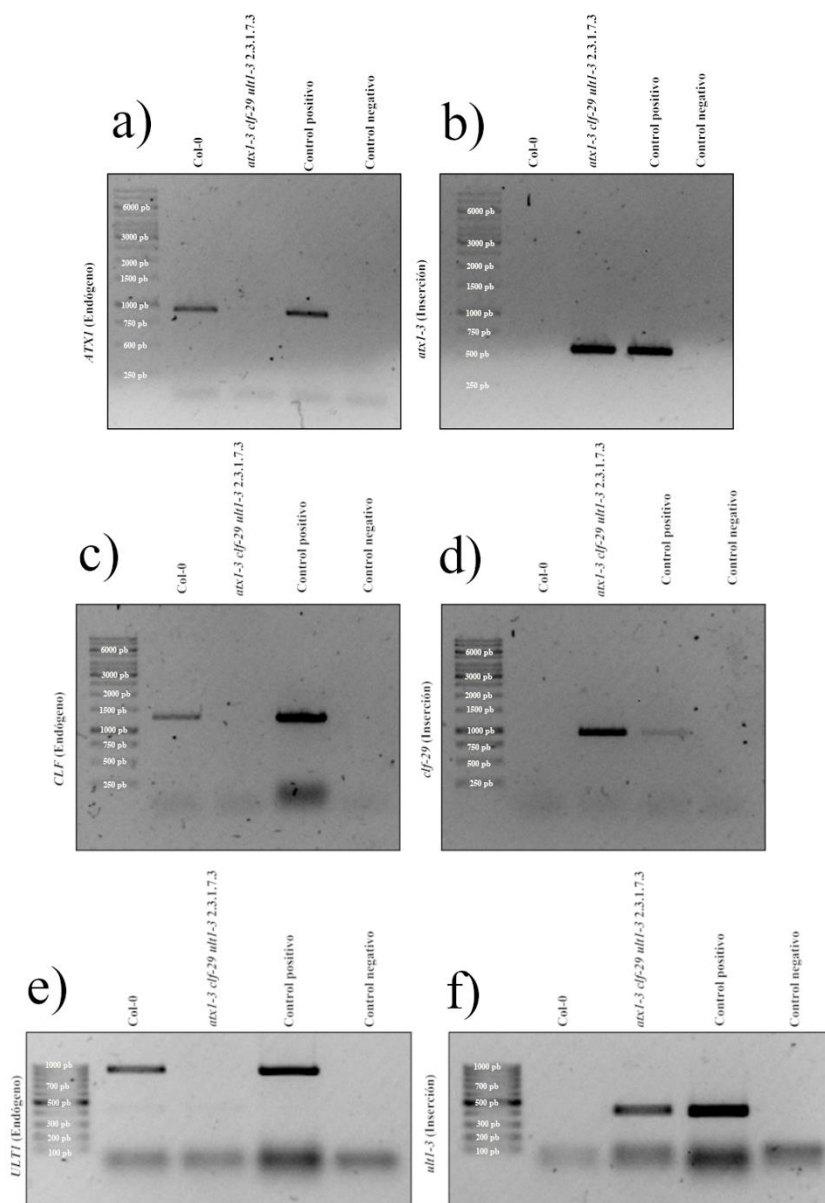


Figura 14. Genotipificación de la triple mutante homocigota *atx1 clf ult1*. En a) el gel de electroforesis donde se muestra la amplificación del gen endógeno de *ATXI*, solo en la planta Col-0 (Wt) y el control positivo. b) Amplificación de la inserción de TDNA de *atx1-3*, donde solo la mutante presenta banda. c) Gel donde se observa el gen endógeno de *CLF* solo en la planta Col-0. d) Amplificación de la inserción de TDNA *clf-29*, donde solo la mutante lo amplifica. e) Amplificación del endógeno de *ULT1*, donde solo la planta Col-0 lo tiene. e) Amplificación de la inserción de TDNA *ult1-3*, donde solo la mutante presenta banda. Marcador de peso molecular para a) y b): *Gene Ruler 1kb DNA Ladder* de ThermoScientific y para c) y d): *100 bp ladder* de NEB.

6.1 Efectos de la mutación de *CLF*, *ULT1* y *ATX1* en el fenotipo: usando a la raíz como modelo

Como ya se mencionó, la regulación epigenética y específicamente de *Trithorax* y *Polycomb* son importantes para los procesos de morfogénesis y diferenciación, por lo tanto, el que las plantas posean una organogénesis continua, resulta bastante útil para estudiar la regulación epigenética en dichos procesos.

En este sentido, es importante destacar que las plantas poseen tejidos específicos cuya función es renovar el reservorio de células para este desarrollo continuo de sus órganos (Dinneny & Benfey, 2008). Estos tejidos se denominan meristemas, existiendo en la planta dos principales: el aéreo y el de raíz. Dentro de los meristemas, se encuentran las llamadas células troncales, de carácter pluripotencial y que, a través de su división y posterior diferenciación, mantienen dicha organogénesis (Aichinger et al., 2012; Pi et al., 2015); estas células se encuentran organizadas en pequeños reservorios denominados nichos de células troncales (SCN) (Dinneny & Benfey, 2008).

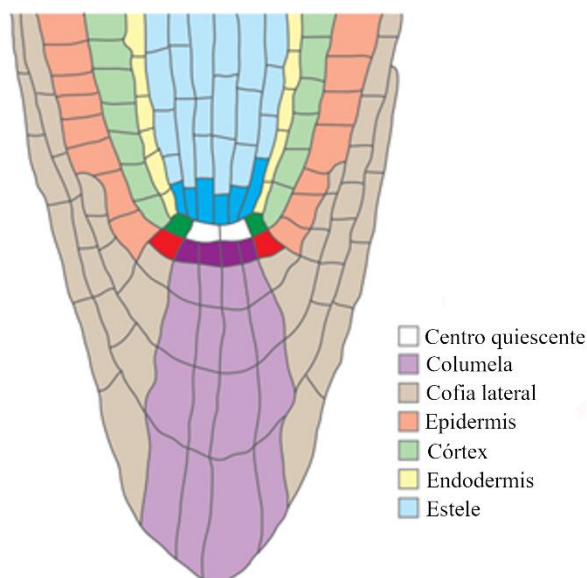


Figura 15. Tipos celulares del nicho de células troncales de la raíz de *Arabidopsis thaliana* (Aichinger et al., 2012)

Como se puede observar en la Figura 15, el SCN de la raíz se conforma por las células del centro quiescente (QC) rodeado de las células iniciales que dan origen a todos los tejidos de la raíz, en la región superior al QC y del centro de la raíz a la periferia se encuentran el estele, la endodermis, el córtex y la epidermis; mientras que en la parte inferior al QC se encuentran las células de la columela y la cofia lateral (Aichinger et al., 2012; Cederholm et al., 2012).

Por todo lo anterior, la raíz representa un modelo de estudio ideal para ahondar en el efecto que tiene la falta de las proteínas epigenéticas de *Trithorax* (*ATX1* y *ULT1*), y *Polycomb* (*CLF*) en fenotipos de diferenciación y proliferación de este órgano. En este

La raíz de *Arabidopsis* no solo está compuesta por múltiples linajes celulares, sino que, a su vez, a lo largo del órgano estos tipos celulares se encuentran en estadios de desarrollo distintos, permitiendo estudiar la proliferación, maduración y diferenciación celular. Además, importante la raíz posee una arquitectura estereotípica constante que permite en muchas ocasiones, por simple morfología y posición, determinar qué tipo celular corresponde con solo observar una imagen (Cederholm et al., 2012; Ivanov & Dubrovsky, 2013).

trabajo se decidió observar dos fenotipos particulares fáciles de medir y para los que hay datos observados en otras mutantes distintas a las generadas en este trabajo (tanto simples como dobles de los genes en cuestión): uno un tanto más macroscópico y relativo al órgano completo: la longitud total de la raíz primaria y el segundo de carácter más específico y relativo a la diferenciación celular: el fenotipo de diferenciación de las células iniciales de columela (CSCs) en células diferenciadas de columela (DCCs).

Con el objetivo de observar el efecto de las mutaciones al fenotipo de longitud total de la raíz, se midió cada 24 horas el tamaño total desde el primer día después de haber sido sembradas y durante el transcurso de 12 días en total, esto tanto para la planta silvestre (Col-0) como para la doble mutante *clf-29 ult1-3* y la triple mutante *atx1-3 clf-29 ult1-3*. De esta cinética de crecimiento, se obtuvo que para la doble mutante *clf-29 ult1-3* no existieron

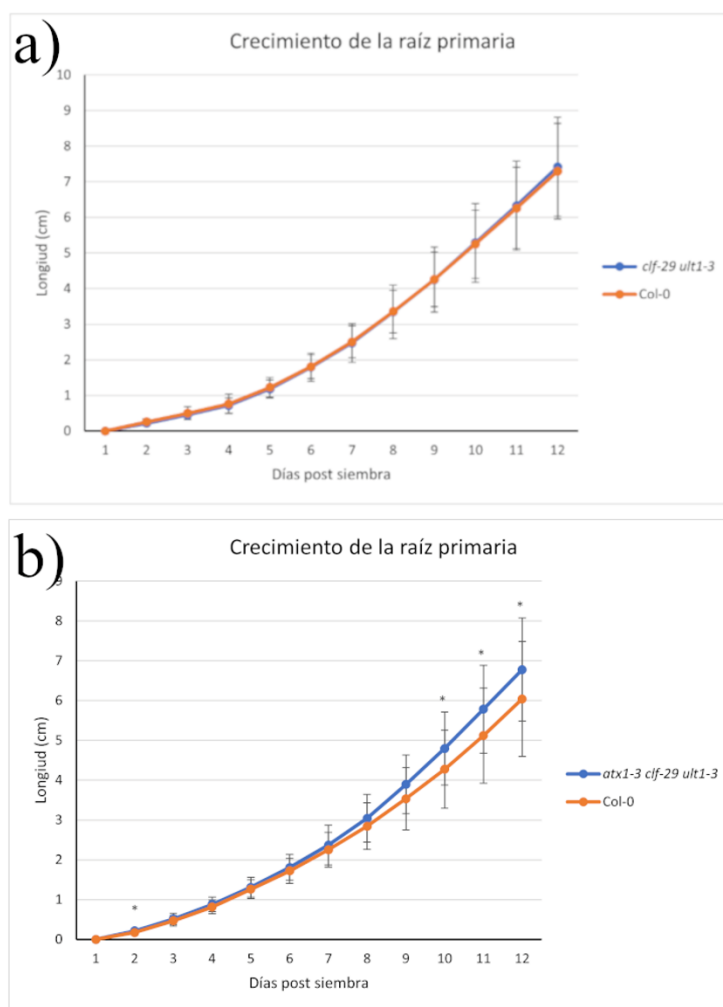


Figura 16. Cinética de crecimiento de la raíz primaria. Arriba, en a) las curvas de crecimiento para las plantas silvestres Col-0 (naranja) y para la doble mutante *clf-29 ult1-3* (azul) durante 12 días después de la siembra ($n \geq 33$). Abajo, en b) las curvas de crecimiento para las plantas silvestres Col-0 (naranja) y para la triple mutante *atx1-3 clf-29 ult1-3* (azul), también durante 12 días después de la siembra ($n \geq 28$). Los asteriscos marcan diferencias significativas del tamaño, acorde a una prueba de T de Student ($p \leq 0.05$)

diferencias de tamaño en ningún momento de los 12 días en comparación a la planta silvestre (Fig. 16a). Contrastantemente, las plantas mutantes *atx1-3 clf-29 ult1-3* sí presentaron diferencias estadísticamente significativas en los tres últimos días de crecimiento (Fig. 16b), con una tendencia a ser más largas; en el día dos (aunque realmente no es observable en la gráfica) y del día 10-12. Esto último sugiere que la triple mutación de los genes estudiados tiene un efecto ligero más que en la velocidad de crecimiento de los primeros días del desarrollo del órgano, en el tamaño total final de ésta.

Un análisis relativamente sencillo de los procesos de diferenciación celular del nicho de células troncales de la raíz, es verlo a través de las células iniciales y diferenciadas de columela. Estas últimas, como parte de su fenotipo de diferenciación terminal, generan gránulos de almidón con papel importante en el gravitropismo, estos gránulos son sencillamente teñidos de color morado oscuro cuando se les aplica Lugol. De esta forma, las células troncales que dan origen al tejido de la columela no poseen los gránulos y cuando se diferencian los adquieren, pudiendo distinguir a la célula troncal parental de la

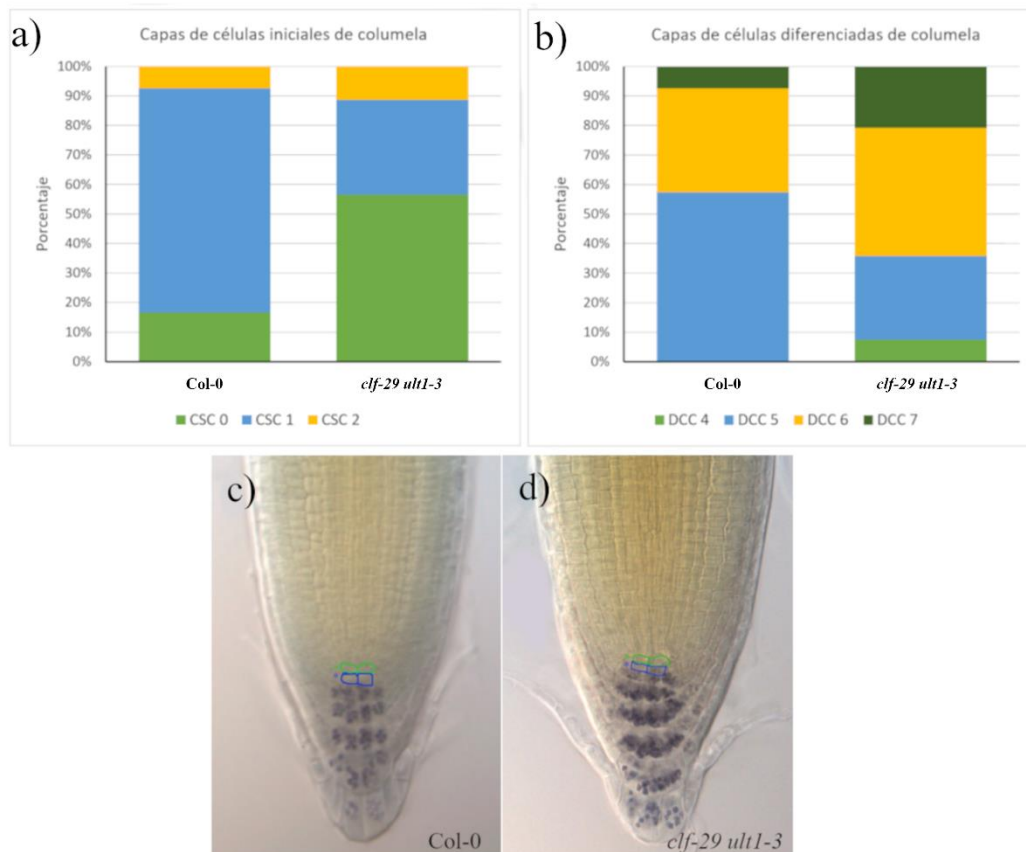


Figura 17. Impacto de la doble mutante *clf-29 ult1-3* en la diferenciación de células de columela. En a) se muestra la cuantificación del número de raíces que presentaban 0, 1 o 2 capas de células iniciales de columela (CSC) entre las plantas Col-0 (izquierda) y las *clf-29 ult1-3* (derecha). En b) se muestra la cuantificación del número de raíces que presentaban 4, 5, 6 o 7 capas de células de columela diferenciadas (DCC) entre las plantas Col-0 (izquierda) y las *clf-29 ult1-3* (derecha). En c) y d) imágenes representativas de microscopía óptica de raíces teñidas con Lugol de los fenotipos más prevalentes en las plantas Col-0 (c) y *clf-29 ult1-3*. Centro quiescente encerrado en color verde y las CSC en color azul. Todas las raíces observadas a 5 días post-siembra, a 40x (n≥53).

diferenciada únicamente con una tinción. Además, como ya se mencionó anteriormente, el patrón estereotípico del nicho de células troncales permite distinguir al centro quiescente y debajo de este la posición donde en teoría deberían estar estas células iniciales de la columela y si en esa posición se observan gránulos posteriormente a la tinción, se puede considerar la existencia de una diferenciación prematura de las CSC.

Resultado de este análisis para la doble mutante *clf-29 ult1-3* se obtuvo que sí presenta un fenotipo de diferenciación temprana de las CSC, en una proporción mucho mayor a las de las plantas Col-0 (Fig. 17c, Fig. 17d). Las plantas silvestres presentaron una capa de células iniciales en el 75.93% de las raíces observadas, cero capas en el 16.67% y por último 2 capas en el 7.41% (Fig. 17a). En contraste, en la doble mutante el fenotipo más prevalente fue el de cero capas (56.60% de las raíces), seguido de una capa (32.08%) y por último dos capas (11.32%) (Fig. 17b).

Respecto al número de capas de células de columela diferenciadas, se obtuvo que lo más común para las plantas Col-0 fue el presentar 5 capas (representado el 57.41%), seguido de 6 y 7 (35.19% y 7.41% respectivamente); mientras que para la doble mutante

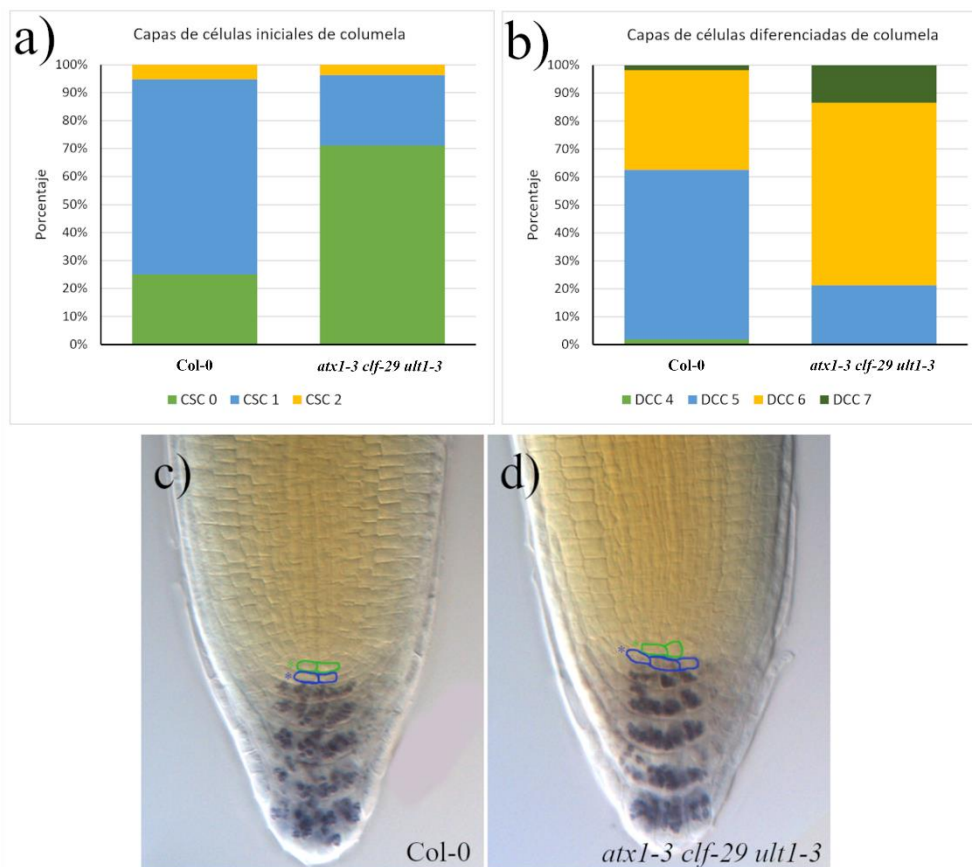


Figura 18. Impacto de la triple mutante *atx1-3 clf-29 ult1-3* en la diferenciación de células de columela. En a) se muestra la cuantificación del número de raíces que presentaban 0, 1 o 2 capas de células iniciales de columela (CSC) entre las plantas Col-0 (izquierda) y las *atx1-3 clf-29 ult1-3* (derecha). En b) se muestra la cuantificación del número de raíces que presentaban 4, 5, 6 o 7 capas de células de columela diferenciadas (DCC) entre las plantas Col-0 (izquierda) y las *atx1-3 clf-29 ult1-3* (derecha). En c) y d) imágenes representativas de microscopía óptica de raíces teñidas con Lugol de los fenotipos más prevalentes en las plantas Col-0 (c) y *atx1-3 clf-29 ult1-3*. Centro quiescente encerrado en color verde y las CSC en color azul. Todas las raíces observadas a 5 días post-siembra, a 40x (n≥52).

existió una mayor variación presentado una mayor proporción de raíces con 6 capas (43.40%), seguido de 5 (28.30%), 7 (20.75%) e inclusive una pequeña población de raíces con 4 capas (7.55%).

Interesantemente, la triple mutante *atx1-3 clf-29 ult1-3* también presenta como fenotipo más prevalente el tener una diferenciación temprana de las células iniciales de columela (CSC=0) con un 71.15% de todas las raíces analizadas y al igual que la doble mutante, seguido de presentar una sola capa de células iniciales (25%) y por último presentar dos capas (3.85%), en este ensayo las plantas Col-0 presentaron prácticamente el mismo patrón que en las observadas para la doble mutante, con un 69.64% de raíces con 1 capa de CSCs, 25% con 0 capas y 5.36% con 2 capas.

Y finalmente, para las capas de células diferenciadas de columela, las plantas Col-0 presentaron un 60.71% de plantas con 5 capas, un 35.71% con 6 capas y muy pocas con 7 y 4 capas (1.79% en ambos casos). Para la triple mutante, se obtuvo (similar nuevamente a la doble mutante) que la proporción mayor (65.38%) fue el fenotipo de 6 capas de células diferenciadas, seguido de aquellas que presentaron 5 capas (21.15%), y por último 7 capas (13.46%).

Estos resultados muestran la importancia de la regulación epigenética en el mantenimiento del nicho de células troncales y en la regulación de los procesos de diferenciación en el mismo.

DISCUSIÓN

1. La desregulación génica en la mutante *sdg2* corresponde a su fenotipo

Un razonamiento directo que soporta el hecho de que, de todas las mutantes analizadas en este trabajo, *sdg2* presentó un mayor número de genes desregulados, teniendo un 7.9% más que *atx1* (la segunda con más cantidad) y 16.1% más que *ult1* (la de menor desregulación). Lo anterior está directamente asociado a un efecto más pleiotrópico y por lo tanto un fenotipo más drástico en *sdg2* respecto al de *atx1* y *ult1*. Como se mencionó anteriormente, la falta de SDG2 provoca plantas débiles y susceptibles a estímulos ambientales y mucho más pequeñas en comparación a las plantas Wt, tanto en parte aérea como en la raíz (Guo et al., 2010), lo cual contrasta con ATX1, cuya mutante de pérdida de función provoca un fenotipo macroscópico mucho más discreto, con plantas ligeramente pequeñas en relación a la Wt, raíces cortas sin llegar al tamaño de *sdg2* y defectos sutiles en el tiempo de floración y morfología de las flores (Alvarez-Venegas et al., 2003). Incluso, el fenotipo de *sdg2* es más drástico que el de la mutante de CLF, que tiene una redundancia funcional menor debido a la existencia solo de otras dos proteínas encargadas de la misma función (SWN y MEA), y aunque una de ellas (SWN) comparte expresión espacio-temporal con CLF, no tiene un gran impacto sobre la deposición de la H3K27me3 (Sanchez et al., 2015; Xiao & Wagner, 2015). A pesar de que algunos fenotipos de *clf* y *sdg2*, como el tamaño de la planta y su sensibilidad a estímulos ambientales, son muy parecidos entre sí (Guo et al., 2010; G. Kim et al., 1998), las deficiencias reproductivas presentes en *sdg2*, (que son además independientes al ocurrir en estadios distintos del desarrollo del

gametofito masculino y femenino (Berr et al., 2010; Pinon et al., 2017)) hacen de su actividad más necesaria para la compleción del ciclo de vida de *Arabidopsis*.

Lo anterior pone a SDG2 sobre ATX1 no solo en relación con la importancia en su papel biológico sino respecto a su papel más generalizado en la regulación génica a lo largo del genoma. Esto además es interesante debido a que los dominios de expresión de SDG2 y ATX1 no son mutuamente excluyentes ni espacial ni temporalmente, ya que ambos se expresan a lo largo de la planta durante todo el desarrollo vegetativo y reproductivo (Guo et al., 2010; Saleh et al., 2008), lo cual abre la puerta para sugerir que a pesar de tener la misma actividad enzimática, *in vivo* la planta tiene bien diferenciadas las funciones de cada una. Con los datos disponibles, es difícil especular cómo es que se logra esta especificidad de cada proteína, sin embargo, no sería tan absurdo sugerir que es a través de la formación de complejos diferenciales dependientes de otras proteínas accesorias que actúan de acuerdo con el tejido en que se encuentren. Idea soportada adicionalmente por la evidencia presentada, por ejemplo, de que en tejido aéreo ULT1 actúa con ATX1 en la activación de genes, mientras que en la raíz la función de ULT1 es independiente a la de ATX1 (Carles et al., 2005; Ornelas-Ayala et al., 2019).

2. La proporción entre los subconjuntos de genes sub- y sobreexpresados en las mutantes no corresponden a su actividad de represión o activación

La idea canónica que siempre se ha tenido respecto a Trithorax y Polycomb es bastante dicotómica: dos complejos proteicos que depositan marcas específicas y antagónicas en el efecto que tienen sobre la expresión génica, siendo el primero de activación y el segundo de represión transcripcional. El análisis general que se realizó en este trabajo no muestra un sesgo en los genes sobre o subregulados que correlacione con la pertenencia a TrxG o PcG. Es decir, uno esperaría encontrar que el subconjunto de los genes que se subexpresan en las mutantes fuera mucho mayor en las pertenecientes a Trithorax (al participar en la activación transcripcional) y lo contrario en CLF. Más bien lo que se obtuvo, como se puede observar en la Figura 6, fue una repartición aproximada de sus totales de genes desregulados del 60% de genes desregulados al alza y un 40% a la baja en *atx1*, *clf* y *ult1*, y para *sdg2* prácticamente del 50% para cada subconjunto.

La conclusión más obvia de lo anterior es pensar que aquellos genes desregulados en sentido opuesto a la acción canónica de la proteína tienen que ver más con regulación indirecta que con el hecho de que sean blancos. Sin embargo, es importante indicar que cada vez existe más evidencia de proteínas PcG o TrxG participando en procesos contrarios a su función canónica, es decir, proteínas TrxG involucradas en procesos de represión y proteínas PcG con funciones de activación transcripcional. Por ejemplo, como ya se mencionó brevemente, ULT1 tiene la capacidad de asociarse físicamente con EMF1 (proteína del grupo Polycomb) para mantener el estado de represión e integridad de la cromatina durante la germinación; adicionalmente, ATX1 participa en este proceso y la triple mutante *atx1 emf1 ult1* sorprendentemente, tiene menos H3K27me3 que la mutante sencilla *emf1* (Xu et al., 2018). Respecto a Polycomb, también hay reportes de proteínas participando en la activación transcripcional de ciertos genes, tal es el caso de LHP1, cuya

mutante baja la expresión de varios genes *YUCCA* (relacionados con la biosíntesis de auxinas) y adicionalmente algunos de estos genes presentan unión de la proteína a su cromatina, soportando así que son blancos de LHP1 (Rizzardí et al., 2011). Además, en nuestro grupo de investigación, se tiene evidencia de otros genes relacionados a la señalización de auxinas que son regulados positivamente por LHP1 (Alejo-Vinogradova, 2020).

Todo lo anterior puede sugerir que existen, al menos en *Arabidopsis*, límites difusos respecto a las actividades de algunas de las proteínas de cada complejo en relación con su regulación de la expresión. No obstante, esta no es la única conclusión que puede surgir de este comportamiento inesperado, también se puede hablar sobre la contundencia de los efectos que tienen ambas marcas en el estado de la cromatina y, ultimadamente en la expresión de los genes que las poseen. Resulta un poco simplista el visualizar este sistema como una cuestión de *todo o nada* en la que necesariamente la presencia de la marca de H3K4me3 implica activación y la presencia de H3K27me3, represión transcripcional. Estudios recientes relativos a Trithorax, específicamente para ATX1, revelan que del total de genes desregulados a la baja en la mutante muy pocos presentaron una pérdida de la marca de la H3K4me3 (Wang et al., 2021). En este estudio, en *atx1* disminuyó la expresión de 3221 genes y de esos, únicamente 107 (el 3.32%) presentaron también una pérdida de la marca de la H3K4me3, rompiendo con la idea de una relación directa entre la ausencia de la marca y la desregulación, aun cuando se trate de una desregulación en el sentido esperado. También hay evidencia de este comportamiento para Polycomb, en donde si bien, las marcas de la H3K27me3 y la H2AK121ub y sus distintas combinaciones tienen un impacto en el estado de la cromatina y su capacidad de respuesta; la pérdida de función de Polycomb y consecuente apertura de la cromatina, no necesariamente tiene como efecto último un cambio en la expresión génica (Yin et al., 2021).

Por lo tanto, no es arriesgado sugerir que el sistema de regulación Polycomb-Trithorax es altamente dinámico, actuando de forma diferencial entre estadios de desarrollo, tejidos e inclusive hasta en genes en específico. Esto complica bastante el estudio de la regulación mediada por estos complejos, sugiriendo cambiar la idea de un modelo binario de prendido y apagado y realizando más estudios de carácter específico espaciotemporales.

2.1 Polycomb y Trithorax pueden regular genes de forma independiente el uno del otro

Pensar en el antagonismo entre Polycomb y Trithorax compitiendo para regular la activación de la expresión o represión de blancos comunes resulta útil para explicar su papel en los diferentes procesos biológicos en los que participan a lo largo del desarrollo de la planta. No obstante, dogmatizarlo representa un impedimento para el progreso de su estudio. Por ejemplo, particularmente con Trithorax esto ha representado el que la descripción de miembros al complejo se vea en muchas ocasiones realizada sólo por la capacidad de antagonizar la actividad a Polycomb, y como podemos observar en los resultados de este estudio, en la Figura 8c, 9 y 10 el traslape de los genes desregulados no

es total entre *clf* y las otras tres mutantes de TrxG, ni siquiera es predominante: con *clf*, *ult1* comparte únicamente el 24.9% de sus genes, *atx1* el 14.9% y *sdg2* el 15.9%.

Estos resultados sugieren dos cosas, la más directa es que, como reportan Wang *et al.* (2021), existe una gran cantidad de genes regulados por elementos de Trithorax o en este caso Polycomb cuya expresión es controlada de alguna forma por estas proteínas, de forma independiente a la marca, o en este caso independiente a su antagonismo con el complejo contraparte, es decir, TrxG promueve la transcripción en algunos genes que no son reprimidos por PcG y PcG reprime genes cuya expresión no es promovida por TrxG.

La segunda conclusión derivada de los resultados del presente trabajo sería una vez más la posible existencia de complejos diferenciales que actúan en la regulación de ciertos genes en ciertas etapas. Es decir, el que el traslape de genes corregulados por CLF y algún Trithorax sea mucho mayor, depende importantemente de la proteína con la que se compare. Al existir múltiples enzimas que ponen la marca de la H3K4me3, dependerá mucho de cuál de ellas se compare y en que etapa lo realice para observar esta corregulación, dejando a un lado la idea de la existencia de un solo complejo TrxG generalista que actúe en todo el desarrollo vegetativo de *A. thaliana*.

3. Trithorax presenta interacciones intracomplejo distintas a las esperadas

3.1 ULT1: un caso especial dentro de Trithorax

En lo que respecta a ULT1, como ya se mencionó, puede presentar interacciones con factores de transcripción, estar asociado a proteínas Polycomb, actuar como Trithorax y dentro del complejo TrxG, actuar de forma independiente a ATX1 (la única HKMT con la que se ha visto interacción) (Ornelas-Ayala *et al.*, 2021). En este trabajo los datos obtenidos apoyan a estas actividades contraintuitivas de ULT1, para empezar, no existe preferencia hacia *atx1* en la desregulación génica y es realmente *sdg2* quien comparte más, exactamente un 11.7% más genes desregulados e interesantemente comparte cantidades similares entre *clf* y *atx1* (24.9% y 29.1% respectivamente). Esto pone a SDG2 como un mejor candidato para formar un mayor número de complejos con ULT1, por ejemplo en la raíz, donde se ha visto que su papel es independiente al de ATX1 (Ornelas-Ayala *et al.*, 2019).

Por otra parte, la gran cantidad de genes que se comparten entre ULT1 y CLF evidencia la existencia de una corregulación con CLF, en donde ULT1 podría estar funcionando en represión, como se ha reportado para el factor de transcripción *FLC*, cuya represión es inducida por PcG. Al respecto, la pérdida de función de ULT1 incrementa la expresión de este gen, en lugar de disminuirla, como se esperaría para el caso de un regulador positivo de la transcripción, además la unión de ULT1 al gen indica una represión directa de ULT1 sobre *FLC* (Ornelas-Ayala *et al.*, 2021).

Al analizar los procesos biológicos a los que pertenecen los genes desregulados en *ult1* y compararlos con los de las otras tres mutantes (Figuras 11 y 12), resulta interesante que cuando se observan aquellos genes que se subexpresan (lo cual iría en concordancia

respecto a la función Trithorax de ULT1), son únicos para esta mutante. Sin embargo, el subconjunto de genes que aumentan su expresión comparte categorías de procesos biológicos con SDG2, y al menos en todos aquellos genes relacionados a la respuesta a oxígeno esta corregulación parece ser antagonista entre ULT1 y SDG2 (el primero actuando en la represión de los genes y el segundo en activación). También en este mismo sentido (sobrexpresados en *ult1*), comparte las categorías de defensa a bacterias y respuesta a fármacos con *clf*, abriendo la posibilidad de una regulación conjunta con esta proteína y consecuentemente con PcG.

3.2 SDG2 y ATX1 no comparten la totalidad de sus blancos de regulación

Sumando a todo lo ya mencionado respecto a SDG2, que tiene una mayor cantidad de genes blancos y su amplio efecto en el fenotipo, además de la evidencia de que ATX1 no actúa con las mismas proteínas en raíz, este trabajo apoya al hecho de que no existe una redundancia importante entre ambas proteínas y que, por el contrario, regulan distintos conjuntos de genes. Solo el 33.8% de los genes desregulados en *atx1* y el 18.5% de los desregulados en *sdg2* son compartidos por ambas (Figura 8). Afirmando el papel más generalizado de SDG2, pero también revelando que ATX1 tiene funciones específicas e independientes a SDG2.

4. Los defectos en la vía de las auxinas presentes en *ult1* pueden deberse a problemas en la síntesis de la hormona

Una de las explicaciones por las cuales la mutante *ult1* presenta defectos en el SCN de la raíz, es el presentar problemas en la regulación de la vía de las auxinas (Ornelas-Ayala et al., 2019). Estas hormonas tienen un impacto importante en la regulación de varios factores de transcripción importantes para la homeostasis del SCN, particularmente para el mantenimiento de la troncalidad y la proliferación (Dric Finet & Jaillais, 2012). Sin embargo, no se ha logrado identificar específicamente en qué parte de la vía se encuentra la afectación presente en *ult1*, debido a que, a pesar de que la respuesta a auxinas (medida con reporteros indirectos) sí cambia, los cambios de expresión en genes de transporte, conjugación y respuesta a auxinas son muy discretas (Ornelas-Ayala et al., 2019).

Este trabajo, puede presentar indicios respecto a qué es lo que se encuentra afectado. Por un lado, el *Gene Ontology* realizado, muestra que entre los genes desregulados al alza en la mutante de *ult1*, se encuentran enriquecidos en dos categorías relacionadas al metabolismo y síntesis de anillos indoles (Figura 12), los cuales son componentes principales de estas hormonas (Dric Finet & Jaillais, 2012), además tendría sentido si se toma en cuenta que de los únicos *pasos* en la vía de señalización de las auxinas que no fue medido, justo fue la biosíntesis de las mismas. Y, por otra parte, en ese mismo *Gene Ontology*, en los genes sobrexpresados, encontramos múltiples categorías relacionadas al metabolismo y síntesis de glucosinolatos, y existe evidencia que desregulaciones en componentes de estas vías afectan a la homeostasis de las auxinas, generando acumulación de las mismas (Mikkelsen et al., 2004). Respecto a esto último, datos no publicados de mi grupo de investigación, indican que, en concordancia a lo anterior la mutante de *ult1* sí presenta una acumulación de auxinas en la raíz.

5. ULT1 tiene el fenotipo más prevalente respecto a la diferenciación de las células de columela

Los análisis fenotípicos de longitud de raíz de las mutantes en *CLF*, *ATX1* y *ULT1* concluyen que, en este caso, *ULT1* no participa en este fenotipo, ya que la mutante sencilla no presenta cambios respecto a la *Wt*, la doble mutante *atx1 ult1* tiene el fenotipo de *atx1* de tener una raíz corta (Ornelas-Ayala et al., 2019) y la doble generada en este trabajo (*clf ult1*) simplemente presenta un fenotipo *Wt*. En cuanto a la triple mutante generada (*atx1 clf ult1*), a pesar de ser ligeramente más larga en estadios más tardíos del desarrollo de la raíz analizados en este trabajo, no presenta un fenotipo exacerbado o sesgado al de alguna de las tres mutantes sencillas (ni siquiera a la de *clf* cuya longitud es más larga a lo largo de los 12 días (Aichinger et al., 2011)).

Para los análisis experimentales sobre el impacto que tuvieron las mutaciones en *CLF*, *ATX1* y *ULT1* sobre el fenotipo de raíz, específicamente en la diferenciación de las células de columela, primero es importante describir los ya reportados previamente a este trabajo. En primera instancia, la mutante sencilla para *atx1* es prácticamente igual al *Wt*, con la excepción de que posee ligeramente más proporción de poseer 2 células iniciales (Ornelas-Ayala et al., 2019); la mutante sencilla de *clf* presenta un fenotipo de tener el 50% de las raíces con 1 sola célula inicial de columela y el otro 50% tener 2 (Aichinger et al., 2011) y finalmente la mutante sencilla de *ult1* presenta un fenotipo predominante de presentar 0 capas de iniciales (es decir una diferenciación prematura de las mismas) alrededor del 70% del total (Ornelas-Ayala et al., 2019)

Al generar la combinación de dobles mutantes, la doble mutante *atx1 ult1* presenta el fenotipo prácticamente idéntico al reportado de *ult1* (Ornelas-Ayala et al., 2019). Mientras que la doble mutante *atx1 clf* presenta un fenotipo más parecido al de *clf* (datos no publicados del laboratorio). Recordando lo obtenido en este trabajo, la doble mutante generada (*clf ult1*) tiene una mayor proporción de raíces sin capas de células troncales de columela, al igual que la mutante sencilla de *ult1*. De manera similar, el fenotipo de la triple mutante (*atx1 clf ult1*) es aún más parecido a la mutante sencilla de *ult1* teniendo también un 71% de raíces sin capas de células iniciales de columela.

Por lo tanto, se puede concluir que es realmente *ULT1* el que tiene el papel más importante en el proceso de diferenciación de células de columela y que el fenotipo de su mutante es epistático al de *atx1* y *clf*. Además, *atx1* presenta el fenotipo más fácil de enmascarar por cualquiera de las otras dos mutantes (ya que la mutante sencilla no presenta fenotipo y en combinación, es el fenotipo de la(s) otra(s) proteína(s) mutadas el observado) y por ende se puede decir que es la proteína cuya participación es menor en esta diferenciación celular.

Relativo a los fenotipos de las células diferenciadas de columela, es muy difícil concluir debido a que existe mucha variación entre las mutantes. Tanto las sencillas, dobles (en cualquiera de sus combinaciones) o triples mutantes no presentan un patrón que sea sencillamente explicable. Inclusive la ausencia de células iniciales o la presencia de una capa extra no correlaciona a un mayor o menor número de capas de células diferenciadas

que sea estable y constante. Por lo tanto, no existe conclusión sobre el efecto de la presencia/ausencia de ATX1, CLF o ULT1 en este fenotipo. Una explicación a este comportamiento es que el proceso de formación de células troncales de columela y la diferenciación de estas, está desacoplado.

También, otra cosa que se puede evidenciar es una vez más el hecho que no en todos los fenotipos, ATX1 con ULT1 y CLF actúan de forma antagonista. Esto debido a que en las diferentes combinaciones de dobles mutantes PcG + TrxG y en la triple mutante se esperaría recuperar un fenotipo Wt, lo cual no ocurre. Por lo anterior, sería interesante realizar análisis moleculares de expresión génica con esta triple mutante para observar comportamientos de regulación más específicos que nos revelen qué conjunto de genes sí está siendo corregulado por estas proteínas. Además, será importante definir los blancos específicos de cada una de estas proteínas en estos tejidos de la raíz para determinar sus comportamientos en un contexto de doble o triple mutante.

CONCLUSIONES

- SET DOMAIN GROUP2 (SDG2) es la Histona lisina metiltransferasa de Trithorax que afecta la expresión de un gran número de genes, lo cual está directamente asociado con que su pérdida de función genere mayor afectación fenotípica en diferentes etapas del desarrollo de *Arabidopsis*, en comparación con las otras proteínas de Trithorax e incluso en comparación con CULRY LEAF (CLF), la enzima que coloca la marca de Polycomb.
- Dentro de las enzimas que colocan la marca de la trimetilación de la lisina 4 en la histona H3 (H3K4me3) analizadas en este trabajo, SDG2 parece ser la que tiene un papel más generalizado. Adicionalmente, ARABIDOPSIS TRITHORAX 1 (ATX1) y SDG2 no comparten la totalidad de sus blancos de regulación génica a pesar de tener la misma actividad enzimática.
- De las proteínas de Trithorax analizadas, ULTRAPETALA1 (ULT1), SDG2 y ATX1 se observa que comparten muy pocos genes blancos de regulación, lo cual predice funciones diferenciales dependiendo del tejido, estadio de desarrollo y blancos específicos.
- ULT1 puede tener interacciones distintas tanto dentro de Trithorax como con proteínas del grupo Polycomb
- Existen subconjuntos de genes regulados de forma antagónica por Trithorax y Polycomb, pero también existen aquellos que solo son blancos de Trithorax y otros que solo por Polycomb.
- La función de ULT1 es epistática a la de ATX1 y CLF en el proceso de diferenciación de células de columela de la raíz de *Arabidopsis*. Esta epistásis se evidencia gracias a que el fenotipo de presentar una diferenciación temprana de estas células iniciales de columela que posee la mutante *ult1-3* es el que predomina en las dobles mutantes *atx1 ult1*, *clf ult1* y la triple mutante *atx1 clf ult1*
- ATX1 no participa en el proceso de diferenciación de células de columela de la raíz de *Arabidopsis thaliana*.

ANEXO

Nombre del primer	Secuencia	Condiciones	Notas
ATX1-F1	CGGACAGGTATGCAGC TACTG	Tm: 55 C, 35 ciclos	Para amplificar gen endógeno <i>ATX1</i> e inserción
ATX1-R1	CAGCTATTGACGGTCTA ACAAG	Tm: 55 C, 35 ciclos	Para amplificar gen endógeno <i>ATX1</i>
CLF_TDAR	ATGGCGTCAGAAGCTT CGC	Tm: 55 C, 35 ciclos	Para amplificar gen endógeno <i>CLF</i> e inserción
CLF_TDNA_F	ACTTGCTAGTTCCGCCA GGAC	Tm: 55 C, 35 ciclos	Para amplificar gen endógeno <i>CLF</i>
LBC1	CCGAGATAGGGTTGAG TGTTGTTCC	Tm: 55 C, 35 ciclos	Para amplificar inserción
ULT1_CC_pGG2_ F1b	GGCTACGGTCTCC GGCT ATGGCGAAC AATGAGGGAGAG	Tm: 55 C, 35 ciclos	Para amplificar gen endógeno <i>ULT1</i> e inserción
ULT1_alelo1_DO R	GGTGCCAAAGTTTGCC ATAGA	Tm: 55 C, 35 ciclos	Para amplificar gen endógeno <i>ULT1</i>

Tabla A1. Lista de todos los *primers* y condiciones de PCR usados en el genotipificado de las mutantes generadas

a)

Locus	Gen	logFC	P.Value	adj.P.Val
AT2G23380	CLF	-1.207409	3.73E-06	0.0001236
AT2G31630	ATX1	-1.805867	7.72E-08	1.43E-05
AT4G15180	SDG2	2.0672696	2.10E-09	7.46E-08

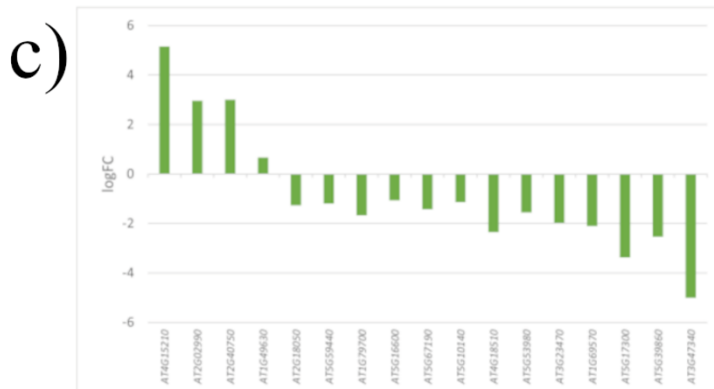
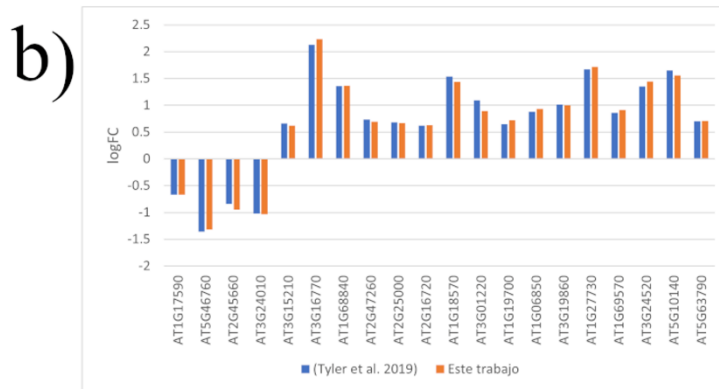


Figura S1. Validación de los datos obtenidos en el análisis de expresión diferencial. a) Datos de la expresión de los genes mutados en sus mutantes. b) Comparación de los genes diferencialmente expresados para *ult1* en el artículo de Tyler et al. (2019) y los obtenidos en este trabajo. c) Cambio en la expresión de los 17 genes probados por Guo et al. (2010) para validar los datos de *sdg2*.

REFERENCIAS

- Aichinger, E., Kornet, N., Friedrich, T., & Laux, T. (2012). Plant Stem Cell Niches. *Annu. Rev. Plant Biol.* *63*, 615–636.
- Aichinger, E., Villar, C. B. R., di Mambro, R., Sabatini, S., & Köhler, C. (2011). The CHD3 chromatin remodeler PICKLE and polycomb group proteins antagonistically regulate meristem activity in the Arabidopsis root. *Plant Cell*, *23*(3), 1047–1060.
- Alejo-Vinogradova, M. T. (2020). *El desarrollo de la raíz de Arabidopsis descrito a través de LHP1*. [Tesis de Maestría no publicada]. Universidad Nacional Autónoma de México.
- Allis, C. D., & Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nature Publishing*

Group, 17(8), 487–500.

- Alvarez-Venegas, R., & Avramova, Z. (2005). Methylation patterns of histone H3 Lys 4, Lys 9 and Lys 27 in transcriptionally active and inactive Arabidopsis genes and in atx1 mutants. *Nucleic Acids Research*, 33(16), 5199–5207.
- Alvarez-Venegas, R., Pien, S., Sadler, M., Witmer, X., Grossniklaus, U., & Avramova, Z. (2003). ATX-1, an Arabidopsis Homolog of Trithorax, Activates Flower Homeotic Genes. *Current Biology*, 13(8), 627–637.
- Ariel, F., Lucero, L., Christ, A., Liu, C., Benhamed, M., Crespi, M., Mammarella, M. F., Jegu, T., Veluchamy, A., Mariappan, K., Latrasse, D., & Blein, T. (2020). R-Loop Mediated trans Action of the APOLO Long Noncoding RNA Article R-Loop Mediated trans Action of the APOLO Long Noncoding RNA. *Molecular Cell*, 77.
- Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Nature Publishing Group*, 21(3), 381–395.
- Berr, A., McCallum, E. J., Ménard, R., Meyer, D., Fuchs, J., Dong, A., & Shen, W. H. (2010). Arabidopsis SET DOMAIN GROUP2 is required for H3K4 trimethylation and is crucial for both sporophyte and gametophyte development. *Plant Cell*.
- Berr, A., Shafiq, S., & Shen, W. H. (2011). Histone modifications in transcriptional activation during plant development. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1809(10), 567–576.
- Berry, S., & Dean, C. (2015). Environmental perception and epigenetic memory: Mechanistic insight through FLC. *Plant Journal*, 83(1), 133–148.
- Caballero, L., Benítez, M., Alvarez-Buylla, E. R., Hernández, S., Arzola, A. V., & Cocho, G. (2012). An Epigenetic Model for Pigment Patterning Based on Mechanical and Cellular Interactions. *Journal of Experimental Zoology*, 3, 209–223.
- Carles, C. C., Choffnes-Inada, D., Reville, K., Lertpiriyapong, K., & Fletcher, J. C. (2005). ULTRAPETALA1 encodes a SAND domain putative transcriptional regulator that controls shoot and floral meristem activity in Arabidopsis. *Development*, 132(5), 897–911.
- Carles, C. C., & Fletcher, J. C. (2009). *The SAND domain protein ULTRAPETALA1 acts as a trithorax group factor to regulate cell fate in plants.*
- Cederholm, H. M., Iyer-Pascuzzi, A. S., & Benfey, P. N. (2012). Patterning the primary root in Arabidopsis. *Wiley Interdisciplinary Reviews: Developmental Biology*, 1(5), 675–691.
- Chen, L.-Q., Luo, J.-H., Cui, Z.-H., Xue, M., Wang, L., Zhang, X.-Y., Pawlowski, W. P., & He, Y. (2017). ATX3, ATX4, and ATX5 Encode Putative H3K4 Methyltransferases and Are Critical for Plant Development. 174, 1795–1806.
- Chen, X., Hu, Y., & Zhou, D.-X. (2011). *Epigenetic gene regulation by plant Jumonji group of histone demethylase.*
- Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J. J., Goldberg, R. B., Jacobsen, S. E., & Fischer, R. L. (2002). DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in Arabidopsis. *Cell*, 110(1), 33–42.

- Costa, F. F. (2010). Non-coding RNAs: Meet thy masters. *BioEssays*, 32(7), 599–608.
- Darwin, C. (1859). *El origen de las especies*.
- Deal, R. B., & Henikoff, S. (2011). Histone variants and modifications in plant gene regulation. *Current Opinion in Plant Biology*, 14(2), 116–122.
- Deichmann, U. (2016a). Epigenetics: The origins and evolution of a fashionable topic. *Developmental Biology*, 416(1), 249–254.
- Deichmann, U. (2016b). Why epigenetics is not a vindication of Lamarckism - and why that matters. *Studies in History and Philosophy of Science Part C :Studies in History and Philosophy of Biological and Biomedical Sciences*, 57, 80–82.
- Derkacheva, M., & Hennig, L. (2014). Variations on a theme: Polycomb group proteins in plants. *Journal of Experimental Botany*, 65(10), 2769–2784.
- Ding, Y., Avramova, Z., & Fromm, M. (2011). The Arabidopsis trithorax-like factor ATX1 functions in dehydration stress responses via ABA-dependent and ABA-independent pathways. *The Plant Journal*, 66(5), 735–744.
- Dinneny, J. R., & Benfey, P. N. (2008). *Plant Stem Cell Niches : Standing the Test of Time*. D, 553–557.
- Dong, H., Liu, D., Han, T., Zhao, Y., Sun, J., Lin, S., Cao, J., Chen, Z.-H., & Huang, L. (2015). *Diversification and evolution of the SDG gene family in Brassica rapa after the whole genome triplication*.
- Dric Finet, C., & Jaillais, Y. (2012). *AUXOLOGY: When auxin meets plant evo-devo*.
- Du, Z., Li, H., Wei, Q., Zhao, X., Wang, C., Zhu, Q., Yi, X., Xu, W., Liu, X. S., Jin, W., & Su, Z. (2013). Genome-wide analysis of histone modifications: H3K4me2, H3K4me3, H3K9ac, and H3K27ac in *Oryza sativa* L. *Molecular Plant*, 6(5), 1463–1472.
- Fletcher, J. C. (2001). The ULTRAPETALA gene controls shoot and floral meristem size in Arabidopsis. *Development*, 128(8), 1323–1333.
- Fletcher, J. C. (2017). State of the art: TrxG factor regulation of post-embryonic plant development. *Frontiers in Plant Science*, 8(November), 1–8.
- Gong, Z., Morales-Ruiz, T., Ariza, R. R., Roldán-Arjona, T., David, L., & Zhu, J. K. (2002). ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. *Cell*, 111(6), 803–814.
- Grant-Downton, R. T., & Dickinson, H. G. (2005). Epigenetics and its implications for plant biology. 1. The epigenetic network in plants. *Annals of Botany*, 96(7), 1143–1164.
- Guo, L., Yu, Y., Law, J. A., & Zhang, X. (2010). SET DOMAIN GROUP2 is the major histone H3 lysine 4 trimethyltransferase in Arabidopsis. *Proc Natl Acad Sci*, 107(51), 3919–3924.
- Haig, D. (2004). The (Dual) Origin of Epigenetics. *Cold Spring Harbor Symposia on Quantitative Biology*, LXIX, 67–70.
- Heo, J. B., Yong-Suk, L., & Sibum, S. (2013). Epigenetic regulation by long noncoding RNAs. *Chromosome Res.*, 21, 685–693.

- Holliday, R. (2006). Epigenetics: A historical overview. *Epigenetics*, 1(2), 76–80.
- Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., & Speed, T. P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, 4(2), 249–264.
- Ivanov, V. B., & Dubrovsky, J. G. (2013). Longitudinal zonation pattern in plant roots: Conflicts and solutions. *Trends in Plant Science*, 18(5), 237–243.
- Jablonka. (2009). Epigenetic Inheritance: Prevalence, Mechanisms, and Implications for the Study of Heredity and Evolution. *QRevBio*, 84(2), 131–176.
- Jablonka, E. (2016). Cultural Epigenetics. *The Sociological Review*, 64(1_suppl), 42–60.
- Jablonka, E., & Lamb, M. J. (2006). The Changing Concept of Epigenetics. *Annals of the New York Academy of Sciences*, 981(1), 82–96.
- Jones, A. L., & Sung, S. (2014). Mechanisms underlying epigenetic regulation in arabidopsis thaliana. *Integrative and Comparative Biology*, 54(1), 61–67.
- Karp, G. (2013). *Cell and Molecular Biology: Concepts and Experiments, 7th Edition* (7th ed.). Wiley Global Education.
- Kim, D. H., & Sung, S. (2017). Vernalization-Triggered Intragenic Chromatin Loop Formation by Long Noncoding RNAs. *Developmental Cell*, 40(3), 302-312.e4.
- Kim, G., Tsukaya, H., & Uchimiya, H. (1998). The CURLY LEAF gene controls both division and elongation of cells during the expansion of the leaf blade in Arabidopsis thaliana. *Planta*, 206(2), 175–183.
- Klemm, S. L., Shipony, Z., & Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nature Reviews Genetics*, 20(4), 207–220.
- Knight, K. (2015). A comparative perspective on epigenetics. *Journal of Experimental Biology*, 218(1), 1–5.
- Lauria, M., & Rossi, V. (2011). Epigenetic control of gene regulation in plants. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1809(8), 369–378.
- Marije Kooistra, S., & Helin, K. (2012). *Molecular mechanisms and potential functions of histone demethylases*.
- Mikkelsen, M. D., Naur, P., & Halkier, B. A. (2004). Arabidopsis mutants in the C-S lyase of glucosinolate biosynthesis establish a critical role for indole-3-acetaldoxime in auxin homeostasis. *Plant Journal*, 37(5), 770–777.
- Napsucially-Mendivil, S., Alvarez-Venegas, R., Shishkova, S., & Dubrovsky, J. G. (2014). ARABIDOPSIS HOMOLOG of TRITHORAX1 (ATX1) is required for cell production, patterning, and morphogenesis in root development. *Journal of Experimental Botany*, 65(22), 6373–6384.
- Nicoglou, A., & Merlin, F. (2017). Epigenetics: A way to bridge the gap between biological fields. *Studies in History and Philosophy of Science Part C :Studies in History and Philosophy of Biological and Biomedical Sciences*, 66(March), 73–82.

- Noble, D. (2015). Conrad Waddington and the origin of epigenetics. *Journal of Experimental Biology*, 218(6), 816–818.
- Ornelas-Ayala, D., Garay-arroyo, A., García-ponce, B., Álvarez-buylla, E. R., De, M., & Sanchez, P. (2021). *The Epigenetic Faces of ULTRAPETALA1*. 12(February), 1–8.
- Ornelas-Ayala, D., Vega-Leon, R., Petrone-Mendoza, E., Garay-Arroyo, A., Garcia-Ponce, B., Alvarez-Buylla, E. R., Sanchez, P., Junto, C. E., & Bot, J. (2019). *ULTRAPETALA1 maintains Arabidopsis root stem cell niche independently of ARABIDOPSIS TRITHORAX1*.
- Pfluger, J., & Wagner, D. (2007). Histone modifications and dynamic regulation of genome accessibility in plants. *Current Opinion in Plant Biology*, 10(6), 645–652.
- Pi, L., Aichinger, E., Groot, E., Laux, T., Pi, L., Aichinger, E., Graaff, E. Van Der, Llavata-peris, C. I., Weijers, D., Hennig, L., & Groot, E. (2015). Organizer-Derived WOX5 Signal Maintains Root Columella Stem Cells through Chromatin-Mediated Repression of CDF4 Expression Article Organizer-Derived WOX5 Signal Maintains Root Columella Stem Cells through Chromatin-Mediated Repression of CDF4 Expression. *Developmental Cell*, 1–13.
- Pien, S., Fleury, D., Mylne, J. S., Crevillen, P., Inzé, D., Avramova, Z., Dean, C., & Grossniklaus, U. (2008). ARABIDOPSIS TRITHORAX1 dynamically regulates FLOWERING LOCUS C activation via histone 3 lysine 4 trimethylation. *The Plant Cell*, 20(3), 580–588.
- Pien, S., & Grossniklaus, U. (2007). Polycomb group and trithorax group proteins in Arabidopsis. *Biochimica et Biophysica Acta - Gene Structure and Expression*, 1769(5–6), 375–382.
- Pikaard, C. S., & Scheid, O. M. (2014). Epigenetic Regulation in Plants. *Cold Spring Harbor Perspectives in Biology*, 6(12).
- Pinon, V., Yao, X., Dong, A., & Shen, W.-H. (2017). *SDG2-Mediated H3K4me3 Is Crucial for Chromatin Condensation and Mitotic Division during Male Gametogenesis in Arabidopsis*.
- Pontvianne, F., Blevins, T., & Pikaard, C. S. (2010). *Arabidopsis Histone Lysine Methyltransferases*. 53(10), 1–22.
- Pontvianne, F., & Grob, S. (2020). Three-dimensional nuclear organization in Arabidopsis thaliana. *Journal of Plant Research*, 133(4), 479–488.
- Poulet, A., Duc, C., Voisin, M., Desset, S., Tutois, S., Vanrobays, E., Benoit, M., Evans, D. E., Probst, A. V., & Tatout, C. (2017). The LINC complex contributes to heterochromatin organisation and transcriptional gene silencing in plants. *Journal of Cell Science*, 130(3), 590–601.
- Probst, A. V., Desvoyes, B., & Gutierrez, C. (2020). Similar yet critically different: The distribution, dynamics and function of histone variants. *Journal of Experimental Botany*, 71(17), 5191–5204.
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43(7), e47.
- Rizzardi, K., Landberg, K., Nilsson, L., Ljung, K., & Sundås-Larsson, A. (2011). TFL2/LHP1 is involved in auxin biosynthesis through positive regulation of YUCCA genes. *Plant Journal*, 65(6), 897–906.

- Saleh, A., Alvarez-Venegas, R., Yilmaz, M., Oahn-Le, Hou, G., Sadler, M., Al-Abdallat, A., Xia, Y., Lu, G., Ladunga, I., & Avramovaa, Z. (2008). The highly similar Arabidopsis homologs of trithorax ATX1 and ATX2 encode proteins with divergent biochemical functions. *Plant Cell*, *20*(3), 568–579.
- Sanchez, M. de la P., Aceves-García, P., Petrone, E., Steckenborn, S., Vega-León, R., Álvarez-Buylla, E. R., Garay-Arroyo, A., & García-Ponce, B. (2015). The impact of Polycomb group (PcG) and Trithorax group (TrxG) epigenetic factors in plant plasticity. *New Phytologist*.
- Satyaki, P. R. V., & Gehring, M. (2017). DNA methylation and imprinting in plants: machinery and mechanisms. *Critical Reviews in Biochemistry and Molecular Biology*, *52*(2), 163–175.
- Schuettengruber, B., Bourbon, H. M., Di Croce, L., & Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell*, *171*(1), 34–57.
- Sun, L., Jing, Y., Liu, X., Li, Q., Xue, Z., Cheng, Z., Wang, D., He, H., & Qian, W. (2020). Heat stress-induced transposon activation correlates with 3D chromatin organization rearrangement in Arabidopsis. *Nature Communications*, *11*(1).
- Sung, S., & Amasino, R. M. (2004). Vernalization and epigenetics: how plants remember winter. *Current Opinion in Plant Biology*, *7*(1), 4–10.
- Tyler, L., Miller, M. J., & Fletcher, J. C. (2019). The Trithorax group factor ULTRAPETALA1 regulates developmental as well as biotic and abiotic stress response genes in Arabidopsis. *G3: Genes, Genomes, Genetics*, *9*(12), 4029–4043.
- Veluchamy, A., Jégu, T., Ariel, F., Latrasse, D., Mariappan, K. G., Kim, S.-K., Crespi, M., Hirt, H., Bergounioux, C., Raynaud, C., & Benhamed, M. (2016). LHP1 Regulates H3K27me3 Spreading and Shapes the Three-Dimensional Conformation of the Arabidopsis Genome. *PLoS One*, *11*(7), e0158936.
- Wang, X., Wang, D., Xu, W., Kong, L., Ye, X., Zhuang, Q., Fan, D., & Luo, K. (2021). Histone methyltransferase ATX1 dynamically regulates fiber secondary cell wall biosynthesis in Arabidopsis inflorescence stem. *Nucleic Acids Research*, *49*(1), 190–205.
- Wollmann, H., Stroud, H., Yelagandula, R., Tarutani, Y., Jiang, D., Jing, L., Jamge, B., Takeuchi, H., Holec, S., Nie, X., Kakutani, T., Jacobsen, S. E., & Berger, F. (2017). The histone H3 variant H3.3 regulates gene body DNA methylation in Arabidopsis thaliana. *Genome Biology*, *18*(1), 94.
- Xiao, J., & Wagner, D. (2015). Polycomb repression in the regulation of growth and development in Arabidopsis. *Current Opinion in Plant Biology*, *23*, 15–24.
- Xu, F., Kuo, T., Rosli, Y., Liu, M.-S., Wu, L., Oliver Chen, L.-F., Fletcher, J. C., Renee Sung, Z., & Pu, L. (2018). Trithorax Group Proteins Act Together with a Polycomb Group Protein to Maintain Chromatin Integrity for Epigenetic Silencing during Seed Germination in Arabidopsis. *Mol. Plant*, *11*, 659–677.
- Yao, X., Feng, H., Yu, Y., Dong, A., & Shen, W. H. (2013). SDG2-Mediated H3K4 Methylation Is Required for Proper Arabidopsis Root Growth and Development. *PLoS ONE*.
- Yin, X., Romero-Campero, F. J., de Los Reyes, P., Yan, P., Yang, J., Tian, G., Yang, X. Z., Mo, X., Zhao, S., Calonje, M., & Zhou, Y. (2021). H2AK121ub in Arabidopsis associates with a less accessible

chromatin state at transcriptional regulation hotspots. *Nature Communications*, 12(1).

Yun, M., Wu, J., Workman, J. L., & Li, B. (2011). Readers of histone modifications. *Cell Research*, 21(4), 564–578.

Zhang, H., Lang, Z., & Zhu, J.-K. (2018). Dynamics and function of DNA methylation in plants. *Nature Reviews Molecular Cell Biology*, 19(8), 489–506.

Zhang, M., Kimatu, J. N., Xu, K., & Liu, B. (2010). DNA cytosine methylation in plant development. *Journal of Genetics and Genomics*, 37(1), 1–12.