



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

Análisis comparativo de las secuencias promotoras del
gen MADS-Box *APETALA3* en la planta homeótica
Lacandonia schismatica y en su taxón hermano, *Triuris*
brevistylis

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

BIÓLOGO

P R E S E N T A:

JOSÉ EMILIANO IGNACIO RODRÍGUEZ MEGA

DIRECTORA DE TESIS:
DRA. ADRIANA GARAY ARROYO

Cd. Universitaria, D .F 2015



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.



HOJA DE DATOS DEL JURADO

Datos del alumno

Rodríguez
Mega
José Emiliano Ignacio
55 59 44 54
Universidad Nacional Autónoma de México
Facultad de Ciencias
Biología
307644019

Datos del tutor

Dra.
Adriana
Garay
Arroyo

Datos del sinodal 1

Dra.
María Elena
Álvarez-Buylla
Roces

Datos del sinodal 2

Dr.
León Patricio
Martínez
Castilla

Datos del sinodal 3

Dr.
Carlos Alberto
Ochoa
Olmos

Datos del sinodal 4

Dra.
Mariana

Benítez
Keinrad

Datos del trabajo escrito

Análisis comparativo de las secuencias promotoras del gen MADS-Box *APETALA3* en la planta homeótica *Lacandonia schismatica* y en su taxón hermano, *Triuris brevistylis*

144 p
2015

AGRADECIMIENTOS

Esta tesis se realizó bajo la dirección de la Dra. Adriana Garay Arroyo en el Laboratorio de Genética Molecular, Desarrollo, Epigenética y Evolución de Plantas del Instituto de Ecología de la Universidad Nacional Autónoma de México. Gracias por tu enseñanza y amistad, Adriana. Me llevo la mejor experiencia. Todo fue padrísimo.

Agradezco la coordinación académica-científica de las Dras. Elena Álvarez-Buylla, Alma Piñeyro Nelson, Berenice García Ponce de León y María de la Paz Sánchez Jiménez, así como la (enorme) ayuda técnica del M. en C. Rigoberto Vicencio Pérez Ruiz y el gran apoyo administrativo y logístico de Diana Romo. Es justo mencionar que sin la constante guía y ayuda de Laura Rodríguez y Teresa Romero en la preparación de soluciones, medios y materiales diversos, la presente tesis habría sido muy distinta (en el peor sentido). Además, la paciencia infinita y accesibilidad del Dr. Juan Estévez lo convirtieron en uno de mis mejores y más queridos mentores. Por otro lado, el Dr. Félix Recillas-Targa, del Instituto de Fisiología Celular, fue muy amable al brindarme algunas enzimas para mis experimentos. ¡Muchas gracias!

También agradezco a mi comité tutorial, conformado por la Dra. Elena Álvarez-Buylla, la Dra. Mariana Benítez, el Dr. León Castilla y el Dr. Carlos Ochoa. Todos me han hecho crecer enormemente con sus comentarios y entusiasmo.

Por último, quiero dar gracias al Biól. Esteban Martínez, quien me enseñó con mano dura a manejar en autopista y me presentó un cosmos escondido en las profundidades de la Selva Lacandona.

Este trabajo fue financiado por Conacyt (180098, 180380, 152649, 105678), DGAPA-UNAM (IN204011-3, IN203113-3) y UC Mexus. Bien ahí, instituciones.

DEDICATORIA

Dedicar una tesis resulta más complicado de lo que uno podría suponer. Quizá sea más fácil comenzar por lo seguro.

La Universidad Nacional Autónoma de México y la (H.) Facultad de Ciencias han sido más que una segunda casa para mí. ¡Gracias! Qué orgullo haber vivido en sus aulas, jardines, cafeterías. Qué dicha haber conocido a una familia tan grande y unida. Qué alegría llamarme puma. Espero algún día poder pagarles de regreso todo lo que atesoro de ustedes.

Parece extraño nombrar propia una tesis que, sin duda, merece el crédito de muchas otras personas. Mis padres son un excelente ejemplo de ello. Mi esfuerzo, mis logros, son suyos también. Aunque pocas cosas puedo poner aquí que no les haya dicho ya en persona, puedo quizá parafrasear algunas.

Mamá, papá. Papá, mamá. Habría que inventar un nuevo idioma para expresar el cariño y admiración que les tengo. Gracias por el apoyo, la inspiración, las risas, las enseñanzas y el amor. Gracias, sobre todo, por creer en mí y alentarme en lo que creo. Son ustedes las personas más increíbles, más chingonas, que conozco.

Esta tesis también va dedicada a Martha y Ernesto. Porque sí. Porque puedo y quiero. No conozco mejor equipo que nosotros. Gracias por darle chispa a la vida. Mejores hermanos no pudieron darme sus padres.

A Matías y Matilde. Chaparros: gracias por recordarme que al final lo que cuenta es pasarla bien.

Acá caben de igual manera el resto de los Rodríguez y los Mega. Gracias por su cariño y apoyo, por darme lo mejor de ambas familias. Aun cuando no comprendan a fondo de qué trata este trabajo, sepan que se los dedico con todo mi amor. Eso debería hablar muy bien del mismo.

Sería injusto no incluir en estos agradecimientos a la institución responsable de haber formado –para bien o para mal– una parte importante de quien soy hoy, el Instituto Luis Vives. Agradezco a Gally, Tazzer, Laura, Erazo, Viqueira, Abraham, Marilú, Lupita, Mari y Connie por todo lo que les debo en formación humana y cultural. Gracias también por haberme reunido con los mejores amigos que he conocido. Sobre todo, aquellos que siguen

presentes de una forma u otra: Daniela, Adán, Rodrigo, Rafa, Florencia, Karla, Yael, Karen y Diego. Derivenmésta. ¡Vives! BIEN ¡Vives! BIEN ¡Vives! BIEN, COÑO, BIEN.

A mis profesores de la Facultad de Ciencias, en especial a Toño Lazcano, Luis Medrano, Ale Ortiz, Rodolfo Salas, Luis Robledo, Carlos Martorell, Vania Carmona, Angel Sánchez, Josefina Herrera, Martha Sánchez.

A mis amigos de la facultad: son ustedes lo mejor que me llevo de estos años. Sería necesario escribir otro anexo a esta tesis para poder incluirlos a todos, pero sepan que guardo todas las risas, chismes, fiestas, salidas de campo y abrazos en el corazón.

A mis compañeros de laboratorio, los que ya no están y los que siguen. A mis hermanos Triuridaceae, Ale, Karlita y Joel. A Aurora, Duhyadi, Andrea y Julio. A Morita, Petrone, Stefan, Pamela y Rosario. Pero sobre todo a Estephania, Valeria y Víctor. Encontré en ustedes la mejor de las amistades.

También dedico esta tesis a Vic, Sof, Pach y Alita. Gracias por enseñarme tantas cosas y mostrarme que la vida es verdaderamente ciencia. Son increíbles, muchachos.

A Diego. Por todo el amor que recibí de ti y todo el que aprendí a darte.

*“Es que ya no puedo, ya no puedo... ¡cómo no! ¡FUA! Y saco el carácter, la
fuerza, el poder”*

Video viral de YouTube

“Fuerza y voluntad”

Ignacio Rodríguez Reyna

ÍNDICE

I RESUMEN	p.1
II INTRODUCCIÓN	
Capítulo 1. La regulación transcripcional y la evolución del fenotipo	
1.1 El papel de la regulación genética en la evolución biológica: un enfoque histórico	p.2
1.2 Los genes eucarióticos, su estructura y su regulación	p.4
1.3 La hipótesis <i>cis</i> en la biología evolutiva del desarrollo	p.9
Capítulo 2. Bases morfogenéticas del desarrollo floral	
2.1 El ABC de la arquitectura floral	p.15
2.2 Los genes MADS-box	p.20
Capítulo 3. La esperanza de los monstruos lacandones	
3.1 La fascinación por <i>Lacandonia schismatica</i>	p.26
3.2 El encanto viene de familia	p.31
Capítulo 4. Antecedentes directos	
4.1 No es lo mismo decir ACB a decir ABC	p.34
4.2 Estructura y regulación del gen <i>APETALA3</i>	p.35
4.3 Estudios de genética molecular en <i>Lacandonia</i> y <i>Triuris</i>	p.39
III JUSTIFICACIÓN	p.42
IV OBJETIVOS	p.43
V HIPÓTESIS	p.43
VI MATERIAL Y MÉTODOS	
Material biológico	p.44
Extracción de ADN genómico	p.44
<i>Genome Walker</i> y construcción de librerías	p.45
Diseño de oligonucleótidos y amplificación de los promotores de <i>TbAP3</i>	p.46
Purificación, ligación, clonación y secuenciación de amplicones	p.49
.....	p.50

VII RESULTADOS Y DISCUSIÓN

Promotor parcial y primer intrón de <i>TbAP3v1</i> y <i>TbAP3v2</i>	p.53
Variación en las regiones promotoras de <i>APETALA3</i>	p.54
Elementos conservados en los promotores de <i>LsAP3</i> y <i>TbAP3</i>	p.58
Una visión compleja sobre la evolución del desarrollo en Triuridales	p.61

VIII CONCLUSIONES Y PERSPECTIVAS

p.64

IX REFERENCIAS

p.66

X APÉNDICES

Apéndice 1: <i>The role of transcriptional regulation in the evolution of plant phenotype</i> (revisión en la que soy primer autor)	p.88
Apéndice 2: Regiones promotoras de los ortólogos de <i>APETALA3</i>	p.116
Apéndice 3: Secuencias completas de <i>TbAP3v1</i> y <i>TbAP3v2</i>	p.122
Apéndice 4: Motivos detectados con FrameWorker	p.123

RESUMEN

La flor hermafrodita de *Lacandonia schismatica* ha provocado una fascinación particular, pues es la única de un cuarto de millón de especies donde se ha comprobado la existencia de carpelos centrales rodeados por estambres periféricos. Este fenotipo se encuentra fijado en las poblaciones naturales de *L. schismatica*, al sureste de Chiapas, donde brota la Selva Lacandona.

En años recientes se ha descubierto que el desplazamiento espacial de la expresión de un solo gen, *LsAP3*, es suficiente para dar lugar a la inversión de los órganos sexuales en la flor de esta planta. Cuál es el mecanismo molecular que provoca este cambio, sin embargo, permanece como una pregunta inquietante.

Hasta ahora se han planteado dos hipótesis que ayudarían a explicar los mecanismos moleculares necesarios para modificar la red de regulación genética de este gen: 1) alteraciones estructurales en la proteína *LsAP3* o en algunas de sus proteínas reguladoras; o 2) mutaciones en las secuencias reguladoras de *LsAP3*, específicamente en su promotor. Para probar esta última alternativa, analicé la región 5' de *LsAP3* y también cloné la de los genes ortólogos *TbAP3v1* y *TbAP3v2* que pertenecen a *Triuris brevistylis*, taxón hermano de *L. schismatica* en cuyas flores también se ha documentado cierto grado de heterotopía.

Mis resultados sugieren la presencia de diferencias importantes entre el promotor de estos tres genes y los de otras angiospermas incluidas en el estudio. Sin embargo, también demuestran que elegir entre una y otra hipótesis es un acercamiento erróneo pues ambas podrían formar parte de un escenario más complejo que explique el desarrollo y evolución de los monstruos lacandones.

INTRODUCCIÓN

CAPÍTULO 1. La regulación transcripcional y la evolución del fenotipo

1.1 El papel de la regulación genética en la evolución biológica: un enfoque histórico.

Estudiar el desarrollo de organismos multicelulares representa un acertijo mental: si el genoma de todas las células que los constituyen es idéntico, ¿por qué sólo sintetizan los productos genéticos que corresponden con el tejido al cual pertenecen? Esta duda permaneció en la comunidad científica hasta los años sesenta, cuando se llegó al consenso de que al menos una respuesta parcial podía recaer en la expresión diferencial de los genes, encargada de definir los tipos celulares desde el desarrollo temprano (Gilbert 2003).

De esta manera, en lugar de expresar todo su abanico genético, cada tipo celular activaría sólo un porcentaje de genes en un espacio definido y durante un tiempo específico del desarrollo. Ahora sabemos que los distintos perfiles de expresión emergen de las interacciones que ocurren dentro de complejas redes de regulación genética, donde colecciones de genes –o partes de genes– interactúan entre sí para controlar aspectos específicos del desarrollo, la diferenciación o la respuesta a señales ambientales (Dávila y Álvarez-Buylla 2014; Latchman 2008).

En los setenta comenzó a nacer una idea innovadora, y a la vez tentadora: los principales cambios anatómicos y funcionales en los organismos usualmente son consecuencia de mutaciones en las regiones reguladoras no codificantes que afectan la expresión de los genes, en tanto que las mutaciones en regiones codificantes que provocan cambios en las estructuras proteicas desempeñan un papel secundario.

Un buen ejemplo de ello se observó en los chimpancés (*Pan troglodytes*), que comparten alrededor de un 95% del genoma con nosotros (Britten 2002), lo cual los convierte en la especie más emparentada con los humanos desde un punto de vista filogenético. Además de la gran similitud genética y anatómica, los chimpancés son muy parecidos a los seres humanos en notables características conductuales, como la creación y utilización de herramientas, la formación de estructuras sociales muy especializadas, el intenso cuidado parental e incluso los comportamientos agresivos dentro y entre grupos (de Waal 2005).

Si humanos y chimpancés somos tan parecidos, entonces ¿qué es lo que nos hace distintos? Mary-Claire King y su tutor de doctorado, Alan Wilson, adoptaron en 1975 un enfoque totalmente nuevo para responder esa pregunta. Fueron los primeros en proponer que los cambios en la expresión de los genes (refiriéndose a las regiones reguladoras génicas) juegan un papel más importante en la evolución de los seres vivos que los genes mismos. Su conclusión se basa en un razonamiento simple: si la divergencia morfológica entre chimpancés y humanos no está correlacionada con un nivel correspondiente de divergencia proteica, la manera más armoniosa de reconciliar esto es bajo el supuesto de que son las mutaciones en las regiones reguladoras, y no en las codificantes, las que constituyen el principal mecanismo de cambio evolutivo (Britten y Davidson 1969; Doolittle et al. 1971; King y Wilson 1975; Stern y Orgogozo 2008).

Algunos estudios recientes han tratado de probar la hipótesis de King y Wilson (Gallego-Romero et al. 2012 hacen una revisión amplia sobre el tema). Se ha demostrado, por ejemplo, que ciertos genes presentan un nivel de expresión muy similar entre primates no humanos, pero su expresión se halla elevada o reducida de manera significativa en nuestra especie (Gilad et al. 2006; Khaitovich et al. 2006). Esto aporta un nuevo matiz en el estudio de la evolución de los seres vivos: si antes pensábamos que las secuencias codificantes constituían la única clave para entender el proceso evolutivo, ahora el papel que juegan las regiones reguladoras parece haber añadido otro grado de complejidad en los factores que participan durante el proceso evolutivo.

La investigación de King y Wilson como trabajo pionero sobre la importancia de las regiones reguladoras en la evolución morfológica, no sólo de chimpancés, sino también de aves (Prager y Wilson 1975), ranas (Wilson et al. 1974a) y otros mamíferos (Wilson et al. 1974b), así como investigaciones posteriores en otros animales y plantas (ver las revisiones de Wray et al. 2003 y Wray 2007 para más ejemplos), han constituido a lo largo de los años un sólido acervo de evidencia de que la adquisición, pérdida o modificación de los rasgos fenotípicos se debe en gran medida a alteraciones en la regulación genética durante el desarrollo de los organismos. Sin embargo, esto no significa que las modificaciones regulatorias sean la única fuente posible de cambio evolutivo (Rodríguez-Mega et al. 2015; ver **Apéndice 1**).

A su vez, otras investigaciones también han contribuido a tratar de responder cómo cada tipo celular alcanza un perfil de expresión particular. En los sesenta, por ejemplo,

comenzó a ganar fuerza el término ‘epigenética’ que Conrad Waddington había propuesto en 1953 para sugerir que el curso del desarrollo de los organismos está determinado no sólo por la interacción compleja de muchos genes entre sí, sino entre genes y ambiente (Haig 2004). Por otro lado, Brian Goodwin, quien impulsó el desarrollo de la biología teórica, reconoció que los genes no pueden explicar por sí solos la complejidad de los organismos. Para Goodwin (2000), existen fuerzas mucho más importantes que la selección natural que actúan sobre los genes, como las constricciones del desarrollo – físicoquímicas, morfológicas y filéticas–, las dinámicas no-lineales y la existencia de redes complejas de regulación genética.

Actualmente se ha comenzado a pensar en las secuencias reguladoras como algo más que ADN basura y, en varios casos, como agentes causales más importantes en la evolución del fenotipo. Si esto es verdad, aún quedan muchas más preguntas por resolver para entender cómo están constituidas, cómo se distribuye su variación en las poblaciones, qué consecuencias tiene el cambio en la expresión de algunos genes en la evolución fenotípica y cuáles son los mecanismos que subyacen tras el mismo (Rockman y Wray 2002).

1.2 Los genes eucarióticos, su estructura y su regulación

François Jacob y Jacques Monod desarrollaron a inicios de los sesentas las bases que explican la regulación de la síntesis proteica en bacterias (Jacob y Monod 1961). Los investigadores franceses postularon que la proteína en sí se encuentra determinada inicialmente por unidades genéticas que llamaron “genes estructurales”, cuya transcripción a ARN mensajero está regulada de manera positiva o negativa por otra proteína, a la cual nombraron “proteína reguladora”. Esta es, a su vez, producida por otros genes conocidos como “genes reguladores” y, una vez sintetizada, se puede unir al “operador”, una secuencia específica adyacente a los genes estructurales. Esta proteína es necesaria, mas no suficiente, para regular la tasa de expresión génica, puesto que está inactiva. Para activarse, necesita de una molécula más pequeña (el “inductor”) que actúa de dos maneras: permite la transcripción (llamándose entonces “activador”) o la bloquea (nombrándose “represor”). Al sistema completo se le llamó *operón* (Jacob et al. 1960), y su descubrimiento les valió a Jacob y Monod el premio Nobel en Fisiología o Medicina en 1965.

En la actualidad se conoce poco sobre los operones fuera de los procariontes: raras veces han sido descritos en otros organismos (Blumenthal et al. 2002). Las células eucariotas, a pesar de compartir varias características con los sistemas de regulación bacterianos, poseen muchas otras maneras de regular la expresión de sus genes y proteínas. Esto les ha permitido aumentar de cierta forma su complejidad genómica sin necesidad de incrementar enormemente el número de genes (Beckerman 2005).

Para el embriólogo Scott Gilbert (2003), existen dos características principales de los genes eucarióticos que los diferencian de los genes procarióticos: 1) se encuentran contenidos y empaquetados dentro de una envoltura nuclear y 2) son genes interrumpidos. Salvo excepciones, la mayoría de estos genes se compone de dos tipos de secuencias distinguibles: los exones, capaces de codificar un producto génico y, como tales, de formar parte del ARN maduro, y los intrones, secuencias no codificantes que se intercalan en el cuerpo del gen y que usualmente no forman parte del producto final (Gilbert 2003; Klug et al. 2006; Lewin 2008). Este rasgo, la alta densidad intrónica en el genoma, se encuentra tan conservado entre especies eucariotas, tanto basales como derivadas, que se ha propuesto la existencia de una invasión “catastrófica” de intrones durante las primeras etapas de la evolución de la célula eucariota (Rogozin et al. 2005; Lynch 2006).

Además de exones e intrones, los genes –entendidos en esta tesis como las unidades discretas del genoma capaces de sintetizar un producto, proteico o no– están compuestos por secuencias reguladoras que se encargan de controlar su expresión. Una de ellas, el promotor, se ubica en el extremo 5' y, de acuerdo con Beckerman (2005), se divide en promotor mínimo, región de control proximal y región de control distal (ver **Figura 1**).

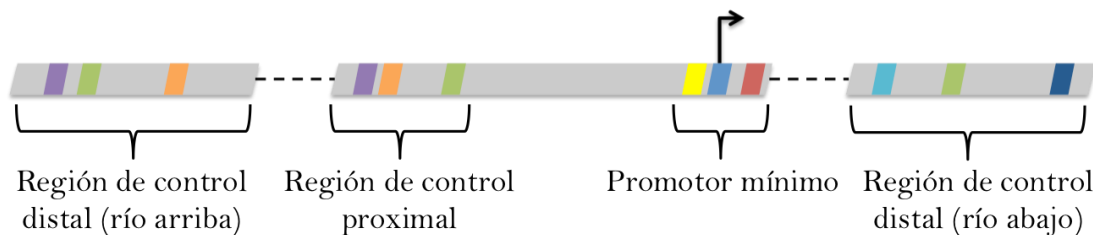


Figura 1. Estructura general de los promotores eucarióticos. La flecha ubicada en el promotor mínimo indica el sitio de inicio de la transcripción; los polígonos de colores representan secuencias o motivos que son reconocidos por distintos factores de transcripción. Las secuencias ubicadas en regiones de control distales (llamadas también *enhancers*) pueden estar involucradas en otros aspectos de la expresión genética. Imagen modificada de Beckerman (2005).

El *promotor mínimo* es una secuencia pequeña de unos 40pb y contiene sitios que son reconocidos por las proteínas de unión al ADN que forman parte del PIC (es decir, el complejo de preiniciación, que ayuda a posicionar a la ARN polimerasa II). En las regiones de control proximal y distal se encuentran otros motivos o secuencias de unión que son reconocidas de manera específica por proteínas reguladoras, conocidas como factores de transcripción, encargadas de controlar –al activar o reprimir– la tasa de transcripción de un gen en particular (Gilbert 2003). La *región de control proximal* se ubica río arriba del promotor mínimo y forma parte del mismo gen mientras que la *región de control distal* puede encontrarse ya sea río arriba o río abajo del promotor mínimo, en regiones intergénicas o incluso dentro de secuencias intrónicas (Beckerman 2005).

Estos componentes otorgan a los promotores una organización modular que, a diferencia de la de secuencias codificantes, es altamente variable porque los sitios de unión a factores de transcripción constituyen una minoría –del 10% al 20%– en el mar de nucleótidos que conforman los promotores (Wray 2003; Wray et al. 2003). En una revisión reciente, Spitz y Furlong (2012) mencionan que cada vez se acumula más evidencia sobre la posible inexistencia de reglas generalizables sobre la composición y posicionamiento –es decir, qué sitios de unión están presentes y cuál es su orden, orientación y espaciamiento– de los motivos que componen a regiones de control y promotores mínimos.

Dada esta naturaleza arquitectónica, las mutaciones en las vastas regiones no involucradas con la regulación transcripcional no tienen consecuencias funcionales ni restricciones evolutivas (Jordan y McDonald 1998; Jareborg et al. 1999) y, por tanto, las secuencias promotoras carecen de rasgos estructurales universales (Wray et al. 2003; Lewin 2008). Sin embargo, se ha demostrado que la variación en los módulos funcionales, cada uno de los cuales controla un aspecto discreto del perfil de expresión (Arnone y Davidson 1997), es capaz de alterar el fenotipo bioquímico u orgánico y tener consecuencias en un plano ecológico-evolutivo.

Desde una amplia perspectiva genética, se han identificado dos unidades básicas de la regulación transcripcional: los factores *trans* y los elementos *cis* (tanto *trans*- como *cis*- son prefijos latinos que significan “del otro lado” y “del mismo lado”, respectivamente). Ambas unidades coordinan la regulación espaciotemporal de la expresión genética, y pueden o no actuar de manera conjunta (Rodríguez Mega et al. 2015, ver en **Apéndice 1**).

Los factores *trans* abarcan cualquier producto génico móvil, como factores de transcripción y RNAs reguladores no codificantes, que actúe en genes blanco distantes a través del reconocimiento de pequeñas secuencias alelo-específicas de ADN llamadas elementos *cis*, que no son transponibles (es decir, que no pueden moverse a otras regiones del genoma) y, por tanto, sólo controlan la expresión de la secuencia adyacente a la que están físicamente ligados (Gilbert 2003; Lewin 2008). De acuerdo con Latchman (2008), los factores de transcripción poseen dominios en su estructura proteica que facilitan su unión al ADN¹. Una vez unidos, deben acoplarse a la ARN polimerasa II para controlar la expresión genética. Para ello, los factores de transcripción poseen otros dominios que les permiten interactuar con moléculas cercanas, sean proteínas reguladoras que interactúan con la ARN polimerasa II o directamente con alguna proteína de este complejo enzimático (Wray 2007).

Los elementos *cis* –a los que Jacob y Monod llamaron “operador”– se encuentran ubicados dentro de regiones no codificantes, como promotores, *enhancers*, silenciadores, UTRs e intrones, aunque también han sido descubiertos en exones (Neznanov 1997; Sandrelli et al. 2001). De acuerdo a su posición física en relación al gen que regulan, los elementos *cis* pueden ser divididos en elementos locales o distales (Emerson y Li 2010; ver **Figura 1**). Las unidades funcionales de los elementos *cis* son sitios de unión a factores de transcripción que mantienen la actividad transcripcional en un dominio de expresión definido y están involucrados ya sea en el proceso básico de transcripción o en la mediación de respuestas a estímulos particulares (Wray et al. 2003; Latchman 2008).

La expresión exitosa de un gen dado requiere de al menos un factor *trans* interactuando físicamente con los elementos *cis* de sus genes blanco en un patrón espaciotemporal específico (Rodríguez Mega et al. 2015, ver en **Apéndice 1**). La interacción entre factores *trans*, que reconocen y se unen a elementos *cis* específicos ubicados en regiones reguladoras, y posteriormente reclutan a otros intermediarios proteicos para formar *loops* o bucles moleculares que acercan todos los elementos de la maquinaria transcripcional, es necesaria para activar la expresión génica. Este tipo de interacciones son complejas, no lineales y, la mayoría de los casos, dependientes de contexto (Espinosa Soto et al. 2004; Álvarez-Buylla et al. 2007; Benítez et al. 2013; Wray 2007).

¹ Estos dominios han sido usados para clasificar a los distintos factores de transcripción en grupos llamados familias.

Algunas revisiones han identificado al menos cinco mecanismos capaces de alterar la actividad genética: 1) modificaciones epigenéticas que alteran el estado de la cromatina sin alterar la secuencia de ADN; 2) mutaciones codificantes no sinónimas que cambian la secuencias de aminoácidos de proteínas estructurales; 3) mutaciones codificantes no sinónimas que cambian la secuencias de aminoácidos de proteínas reguladoras (cambios en *trans*; ver **Figura 2**); 4) mutaciones reguladoras en *cis*, donde los cambios en nucleótidos de los dominios de unión modifican los patrones de expresión (cambios en *cis*, ver **Figura 3**); y 5) cambios en el número de copias génicas que ocurren por pérdidas o duplicaciones, y que involucran cambios en secuencias codificantes y reguladoras (Stern y Orgogozo 2008; Peter y Davidson 2010; Gruber et al. 2012; Rodríguez Mega et al. 2015, ver en **Apéndice 1**).

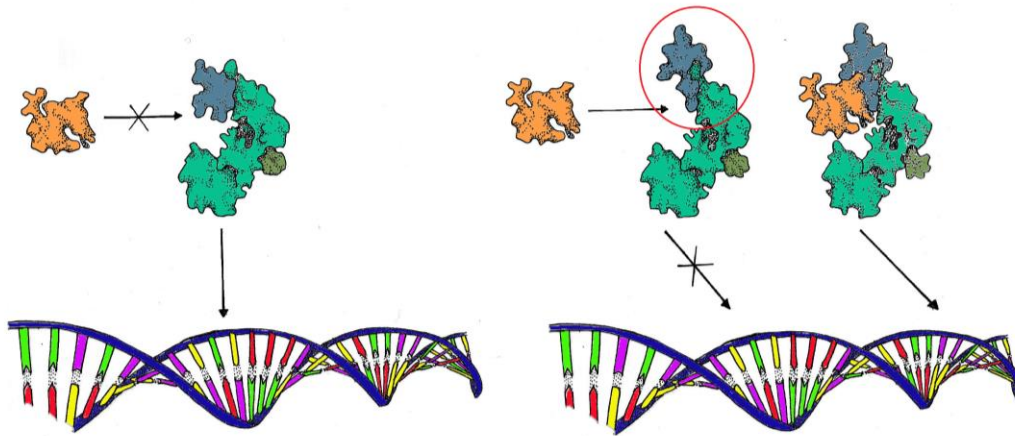


Figura 2. Alteración funcional de proteínas reguladoras (cambios en *trans*). Las modificaciones estructurales de los factores de transcripción pueden alterar su capacidad de unión al ADN o a otras proteínas. En este esquema, la proteína naranja, que antes no interactuaba con el complejo verde-azul, puede ahora acoplarse gracias al cambio estructural del dominio azul oscuro (señalado con un círculo rojo). El resultado de esta interacción varía, ya sea bloqueando o facilitando el reconocimiento del sitio de unión al ADN. En ambos casos, el resultado final es la alteración del patrón de expresión de algún gen particular. Ilustración de Estephania Zluhan, tomada de Rodríguez Mega y Hernández Marroquín (2013).

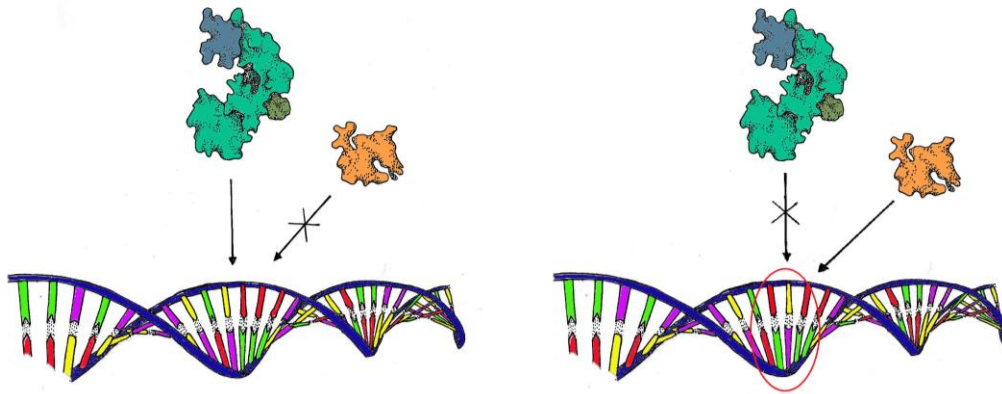


Figura 3. Alteración de los sitios de unión ubicados en secuencias reguladoras (cambios en *cis*). Las mutaciones de los sitios de unión ubicados sobre la secuencia de ADN (círculo rojo en la imagen derecha) puede ocasionar que algunos factores de transcripción, como el complejo verde-azul, dejen de reconocerlos o que otros más, como la proteína naranja, puedan unirse a una región que antes no reconocían. Ilustración de Estephania Zluhan, tomada de Rodríguez Mega y Hernández Marroquín (2013).

1.3 La hipótesis *cis* en la biología evolutiva del desarrollo

Un cuarto de siglo después de que Mary-Claire King y Alan Wilson sugirieran la importancia de la regulación genética –es decir, dónde, cuándo y cómo se activan los genes– para la evolución fenotípica, se inauguró la división de Biología Evolutiva del Desarrollo en la reunión anual de la Sociedad para la Biología Integrativa y Comparativa (SICB, por sus siglas en inglés). Así nació una nueva disciplina cuyo desarrollo conceptual llevaba gestándose desde antes que Charles Darwin publicara su afamado libro en 1859, y que, a la vez, no podría haber visto la luz de no ser por los explosivos avances en genética molecular de los noventa (Robert et al. 2001; Amundson 2005. Ver **Figura 4**).

La biología evolutiva del desarrollo² terminó de concebirse, de manera inevitable, como un esfuerzo por llenar los vacíos que había dejado la Teoría Sintética de la Evolución (Reydon 2011). Por ejemplo, al excluir disciplinas como la embriología, la morfología y la paleontología, la Síntesis Moderna no aclara enteramente cómo surge la variación fenotípica sobre la cual actúa la selección natural, pues sólo sugiere que la variación –considerada, además, como diferencias continuas entre organismos– surge por cambios cromosómicos. Por tanto, ignora otros mecanismos genéticos y de desarrollo que subyacen tras la generación de nuevos planes corporales y fenotipos (Müller 2007; Gilbert

² También llamada “evo-devo” por muchos autores.

y Epel 2009; Willmore 2012). Parafraseando al genetista neerlandés Hugo de Vries (1905), si bien la Síntesis Moderna puede explicar la supervivencia del más apto, no puede explicar la *llegada* del más apto³.

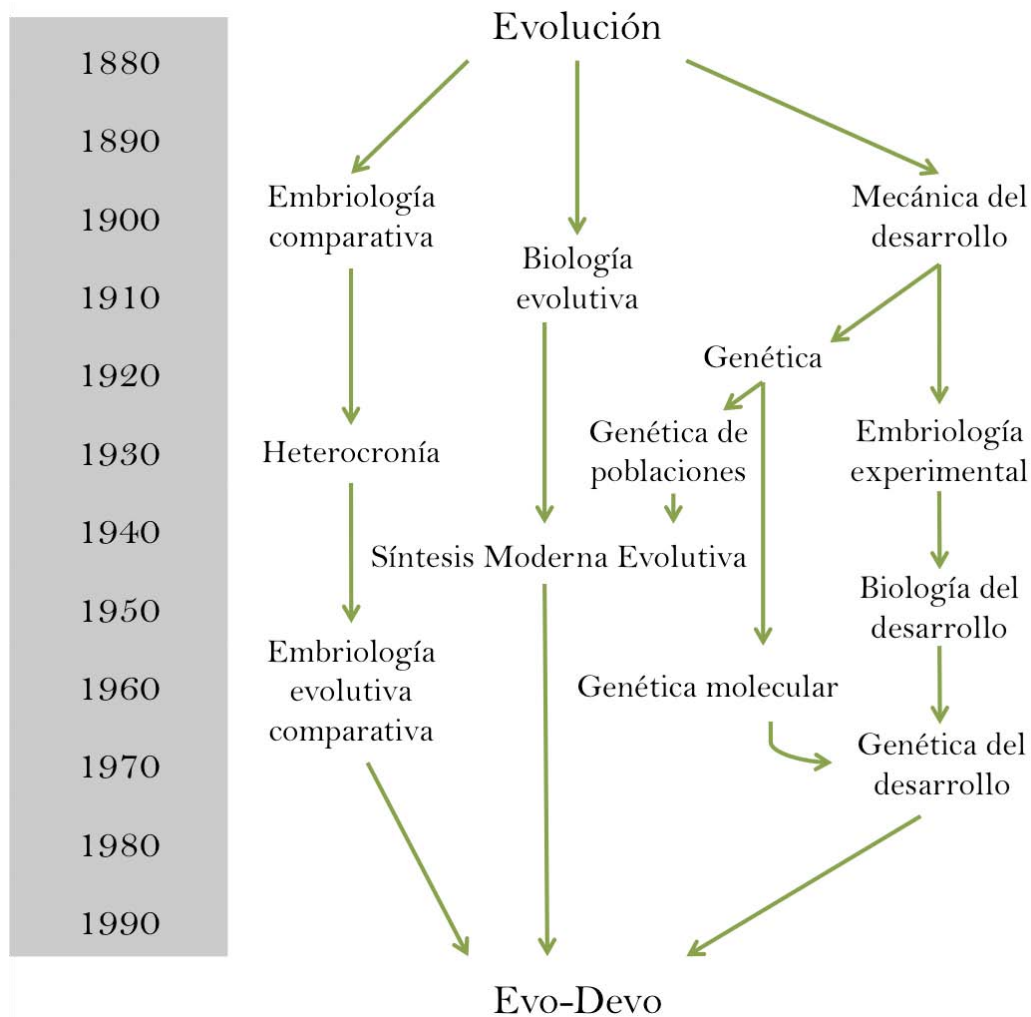


Figura 4. Breve historia de la biología evolutiva del desarrollo. Se observan los principales componentes sobre los cuales se funda la evo-devo contemporánea. Nótese la importancia de la heterocronía –cambio cronológico del desarrollo con consecuencias morfológicas–, que fue vista como un mecanismo universal durante gran parte del siglo XX y que impulsó la unión entre embriología y evolución (Hall 2012). Imagen modificada de Love y Raff (2003).

³ La frase de De Vries ha sido criticada, principalmente, por Ernst Mayr, uno de los fundadores de la Síntesis Moderna. Mayr, quien separó de tajo el desarrollo de la evolución, dice que la recombinación genética y el éxito reproductivo son parte del proceso de selección natural y, como tal, pueden ser fuerza creativa de variación (Mayr 1982, p. 519). Reydon (2011) hace un análisis detallado sobre este desacuerdo.

Como resultado de su creación colectiva, los principios de la biología evolutiva del desarrollo aún no terminan de definirse, pero se ha llegado al consenso general de que el principal objetivo de esta disciplina consiste en encontrar la relación entre el desarrollo y la evolución. Esto es, explorar cómo los mecanismos del desarrollo generan y modulan la variación fenotípica, cómo el desarrollo afecta el cambio evolutivo y cómo ha evolucionado el desarrollo embrionario (Robert et al. 2001; Baguña y Garcia Fernández 2003; Laubichler 2007; Müller 2007; Carroll 2008; Gilbert y Epel 2009; Hall 2012; Willmore 2012).

En biología evolutiva del desarrollo no existe un frente único desde donde se aborden los problemas que le atañen, sino una pluralidad de programas de investigación con distintas aproximaciones y metodologías enfocados a situar el estudio del desarrollo en un contexto evolutivo (Carroll 2008; Hall 2012). La presente tesis se enmarca dentro de, al menos, dos de ellos: el programa comparativo, dedicado a estudiar las diferencias morfogénicas –usando herramientas como la comparación de secuencias y patrones de expresión– y el programa de genética evolutiva del desarrollo, que estudia la evolución de la maquinaria genética del desarrollo y su regulación (Müller 2007; Laubichler 2007; Carroll 2008).

A inicios de los ochenta se describió el segundo elemento más importante que terminaría por unir desarrollo con evolución (el primero había sido la heterocronía, como menciono en la **Figura 4**): la secuencia *homeobox* (Garber et al. 1983; Scott et al. 1983). Esta secuencia forma parte de la estructura de genes reguladores del desarrollo que sintetizan factores de transcripción, y codifica su dominio de unión al ADN⁴.

Los genes *homeobox* ocupan sitios clave dentro de las redes de regulación genética, de tal manera que sus mutantes dan lugar a cambios en la polaridad, número e identidad de los segmentos corporales que componen a los animales (Gehring y Hiromi 1986; Hall 2012). Por otro lado, se han diversificado en distintas familias de genes, de los cuales los más famosos probablemente sean los genes Hox, pues fue en ellos donde se encontró la secuencia *homeobox* por primera vez⁵.

⁴ En la proteína, la secuencia *homeobox* tiene el nombre de dominio *homeobox* o, de manera alternativa, homeodominio.

⁵ El hallazgo de la secuencia *homeobox* se hizo gracias al estudio de animales con genes Hox mutados, específicamente de moscas de la fruta con mutaciones en el gen *Antennapedia*. Dado que las moscas mutantes mostraban partes del cuerpo en un lugar equivocado –las antenas de su cabeza se transformaban en un par

Tras su descubrimiento, pronto se supo que no sólo la secuencia *homeobox* es prácticamente idéntica entre especies animales –sean cercanas o no–, sino que el resto de la región codificante de los genes Hox, sus patrones de expresión y, por tanto, su función también están muy conservados (McGinnis et al. 1984; Gehring y Hiromi 1986). El hallazgo de otros genes *homeobox*, todos ellos relacionados e involucrados en la regulación del desarrollo embrionario, dio paso a pensar que existe un juego de herramientas moleculares, un *toolkit* genético, que subyace tras el desarrollo de todos los animales.

Pero de este descubrimiento surge un problema evidente: ¿Cómo es posible que fenotipos tan distintos de animales como insectos o mamíferos sean delineados por los mismos genes del desarrollo, conservados entre especies y expresados en patrones similares? Esta es la paradoja Hox (Wray 2002), y resolverla implicaría reconocer, al menos en parte, que ocurren cambios en las reglas constantes de ensamblaje de las redes regulatorias del desarrollo. La reorganización de redes genéticas de regulación explicaría, por ejemplo, cómo ha cambiado la interacción entre genes ortólogos y proteínas durante la evolución de linajes, de manera que regulen procesos divergentes del desarrollo y generen distintos fenotipos resultantes o incluso estructuras no homólogas (Wray y Lowe 2000; Davidson 2001; Wray 2003; Wray et al. 2003; Dávila y Álvarez-Buylla 2014).

Dado que el desarrollo se ve controlado por redes genéticas de regulación, la evolución fenotípica se verá afectada de acuerdo a los cambios que ocurran entre sus interacciones (Wray 2002; Carroll 2008). Pero, ¿qué tipo de mutaciones son necesarias para reorganizar estas redes complejas de regulación? Aquí menciono la participación de sólo dos: la alteración estructural y funcional de factores *trans* o la modificación de elementos *cis*.

Durante el curso de la evolución, algunas proteínas pueden comenzar a participar en distintos aspectos independientes del desarrollo. Este efecto pleiotrópico, adjudicada a factores de transcripción por autores como Carroll (2006) o Gilbert y Epel (2009), funcionaría entonces como una restricción evolutiva al cambio de los factores *trans*, pues es más probable que su alteración estructural y funcional tenga consecuencias negativas sobre el desarrollo y el éxito reproductivo (Carroll 2006; Gilbert y Epel 2009). Ejemplos de esta pleiotropía abundan en la literatura. Es el caso de la proteína Decapentaplegic (Dpp), que moldea diversos aspectos del desarrollo embrionario de la mosca de la fruta,

de patas adicionales– se les llamó “mutantes homeóticos”, término acuñado por William Bateson en 1894 y que describe la transformación de una estructura del cuerpo en otro segmento corporal (Gehring y Hiromi 1986).

como la polaridad dorsoventral, la morfogénesis intestinal y la formación de alas, patas y otros apéndices (Gelbart 1989). Por otro lado, también se ha comprobado que la delección de uno de los motivos peptídicos que componen a la proteína Ultrabithorax (Ubx), conservada a través del linaje de los insectos, tiene efectos distribuidos en muchos tejidos (Hittinger et al. 2005). Este tipo de ejemplos no sólo se ha observado en insectos; también se han documentado en aves (McMahon et al. 2003) y mamíferos (Galis 1999). Sin embargo, esto puede no ser más que un reflejo de cuán poco sabemos sobre la participación de los factores *trans* en cambios regulatorios, pues se ha demostrado que algunos de ellos pueden ser tan modulares como los elementos *cis* (Rodríguez-Mega et al. 2015; ver **Apéndice 1**).

En contraste, los elementos *cis* se caracterizan por su organización modular donde cada módulo representa una subunidad genética discreta y funcional involucrada en regular la transcripción de un locus cercano de manera conjunta o independiente (Arnone y Davidson 1997; Bolker 2000; Wray 2003). Esta modularidad reduce el efecto pleiotrópico de los elementos *cis*, al tiempo que les permite mutar y modificar el patrón espaciotemporal de expresión de los genes que regulan⁶ (Vergara Silva 2009; Gilbert y Epel 2009). No resulta difícil, pues, entender porqué durante los últimos años de la biología evolutiva del desarrollo se ha considerado a los elementos *cis* como el sustento genético de la evolución morfológica, novedades evolutivas y modificación del plan corporal de los organismos.

Desde cambios en sistemas reproductivos de plantas como el tomate (Chen et al. 2007), la aparición de novedades morfológicas en la flor de otras solanáceas (Khan et al. 2012), las diferencias en morfología larval, en distribución de pelos torácicos y en patrones de pigmentación alar y corporal entre especies de *Drosophila* (Sucena y Stern 2000; McGregor et al. 2007; Gompel et al. 2005; Jeong et al. 2008) o la distribución de escamas y la reducción pélvica en peces espinosos (Colosimo et al. 2005; Shapiro et al. 2006) hasta la adquisición de rasgos como la resistencia a la malaria (Tournamille et al. 1995) o la persistencia de lactasa (Olds y Sibley 2003) en algunas poblaciones humanas, e incluso algunas diferencias cognitivas y conductuales entre chimpancés y humanos (Rockman et

⁶ La modularidad no sólo está presente en los *enhancers*, sino en otros aspectos distintos: los módulos, por ejemplo, también pueden ser tejidos o regiones de dichos tejidos. Estos módulos discretos que interactúan entre sí son los que permiten que haya lugar para cambios en un proceso tan complejo como lo es el desarrollo.

al. 2005), la mutación y modularidad de los elementos *cis* parece ser una base importante de la variación biológica. Esto no es trivial: además de ser capaces de alterar las redes de regulación genética que controlan aspectos de la morfología, fisiología y comportamiento de los organismos, las mutaciones de elementos *cis* también ofrecen una base sólida para explicar cómo evoluciona el fenotipo de los sistemas biológicos (Wray 2003; Wray et al. 2003; Wray 2007; Carroll 2008; Wittkopp y Kalay 2011).

Terminar de ligar esta creciente evidencia con sus consecuencias moleculares, rasgos fenotípicos y contextos ecológicos ayudará a contribuir a uno de los principios más importantes de la evo-devo: el desarrollo es construcción, producto y causa del proceso evolutivo.

CAPÍTULO 2. Bases morfogenéticas del desarrollo floral

2.1 El ABC de la arquitectura floral

Existe una pandemia en biología evolutiva que ha logrado contagiar a otras disciplinas, incluida la evo-devo: los resultados de la investigación proliferan principalmente gracias al estudio de sistemas animales. Esto nos obliga a enfrentarnos con preguntas nada triviales. ¿Hasta qué punto estos resultados pueden ser extrapolados a otros linajes o reproducibles los mismos experimentos en distintos grupos? ¿Hasta qué punto el desarrollo de animales es comparable con, digamos, el de plantas?

Aunque muchos estudios sobre el desarrollo de plantas se han realizado a la sombra de la biología animal, el escenario cambia a paso lento desde hace ya un par de décadas. Ahora sabemos, por ejemplo, que los mecanismos del desarrollo entre plantas y animales son bastante distintos (Meyerowitz 2002; Friedman y Diggle 2011), como se esperaría de linajes con historias de vida tan diferentes que han evolucionado de manera independiente durante millones de años.

En la ausencia de cualquier tipo de comportamiento migratorio, respuesta inmune o sistema nervioso, las plantas exhiben un abanico de fenotipos cuando se enfrentan a agentes externos bióticos o abióticos, y ajustan continuamente su desarrollo y morfología a las condiciones que las rodean (Schlichting 1986). Quizá la sofisticación de los sistemas de regulación genética en plantas sea justamente una consecuencia evolutiva de ser organismos sésiles, no móviles, cuyo desarrollo responde de manera plástica al ambiente. Algunos análisis comparativos de genómica han sugerido que la regulación transcripcional podría interpretar un papel más importante en plantas que en animales, pues las plantas dedican un mayor porcentaje de su genoma para codificar reguladores transcripcionales (Riechmann et al. 2000; Gong et al. 2004; Xiong et al. 2005).

Hombro con hombro junto a mecanismos ambientales, genéticos y fisiológicos, el desarrollo ha participado en la creación de la impresionante complejidad estructural que ostenta el reino vegetal. Mientras más comprendamos su papel en la evolución de las plantas, más razones tendremos para ubicar y defender su estudio a un nivel igual de

importante que el de la biología animal (Steeves y Sussex 1989; Kaplan y Hagemann 1991; Bateman y DiMichele 2002; Meyerowitz 2002; Friedman et al. 2004).

Este esfuerzo no es cosa nueva: del siglo XVIII al XX, la curiosidad de figuras como Johann Wolfgang von Goethe, Wilhelm Hofmeister, Karl von Goebel y Wilhelm Troll contribuyeron a fortalecer el estudio de las formas vegetales, principalmente gracias a su intento de construir un puente entre la estructura y el desarrollo de las plantas al tratar de descubrir la base causal de su diversidad morfológica (Kaplan 2001).

De todos ellos, Goethe interesa de manera particular. Poeta, novelista y dramaturgo, también se sumergió en el estudio de la botánica debido a una aparente crisis vocacional y publicó un libro llamado *La metamorfosis de las plantas*, donde propone que estos seres se construyen a partir de una unidad básica: la hoja. Todos los demás órganos –a excepción de la raíz, que omite por completo en su escrito– serían transformaciones, metamorfosis originadas a partir de cambios estructurales de la hoja, y la diversidad organográfica de las angiospermas podría ser explicada de esta manera (Coen 2001; Kaplan 2001).

Hoy sabemos que la mayoría de las flores hermafroditas está integrada por cuatro tipos de órganos ordenados en círculos concéntricos, los verticilos, que forman las partes básicas de su arquitectura (ver **Figura 5**): el perianto (formado por sépalos y pétalos, ambos visualmente similares a las hojas), el androceo (la parte masculina, formada por estambres) y el gineceo (la parte femenina de las flores, compuesta por carpelos).

Ocasionalmente, y bajo ciertas circunstancias, los estambres y carpelos pueden transformarse en órganos del perianto. En las rosas cultivadas, por ejemplo, estos han sido reemplazados por un mayor número de pétalos durante un proceso de domesticación, y fue en ellas Goethe vio la comprobación de su teoría porque los órganos sexuales se han transformado en órganos más parecidos a hojas: los pétalos (Coen 2001; Álvarez-Buylla 2002). Podría decirse, entonces, que las distintas partes de una flor son equivalentes entre sí como variaciones de un mismo tema.

Aunque Goethe no incluyó explícitamente una visión evolutiva o del desarrollo en su hipótesis, sí proveyó las bases para que otros lo hicieran siglos después (Coen 2001; Theissen y Saedler 2001; Friedman y Diggle 2011). A inicios de los noventa se publicaron tres artículos que compartían el objetivo de examinar las bases genéticas del desarrollo floral al examinar flores anormales encontradas una década antes en las angiospermas modelo *Antirrhinum majus* (Carpenter y Coen 1990; Schwarz-Sommer et al.

1990) y *Arabidopsis thaliana* (Coen y Meyerowitz 1991). Tras analizar muchas variantes florales, los estudios sugirieron la existencia de tres tipos principales de mutantes: *a*, *b* y *c* (ver **Figura 6**).



Figura 5. Arquitectura básica de una flor. Las flores hermafroditas están formadas, en general, por cuatro verticilos ordenados de manera concéntrica. En la imagen se observa una foto y el diagrama floral de la planta modelo *Arabidopsis thaliana*, donde el verticilo más externo contiene a los sépalos (S), que protegen a la flor (en verde); el siguiente, a los pétalos (P), usualmente la parte más llamativa de la flor (en blanco); después siguen los estambres (E), órganos masculinos productores de polen (en gris); y ubicados en el centro se encuentran los carpelos (C), órganos femeninos de la flor. El diagrama floral ejemplifica la distribución espacial de cada verticilo. La fotografía es propiedad de Jürgen Berger, de la Unidad de Microscopía Electrónica del Instituto Max Planck; el diagrama floral es de mi autoría.

Los primeros, los mutantes *a*, no poseen pétalos ni sépalos; su lugar ha sido ocupado por más estambres y carpelos. Los mutantes *b* carecen de estambres y pétalos, pero tienen sépalos y carpelos extra. Por último, los mutantes *c* se parecen mucho a las rosas cultivadas, pues sus órganos sexuales han desaparecido a expensas de un mayor número de sépalos y pétalos.

Estas observaciones ayudaron a construir la propuesta de un modelo capaz de explicar las bases genéticas del desarrollo floral: el modelo ABC (Coen y Meyerowitz 1991), que dio un fuerte apoyo a la idea de que todas las flores comparten mecanismos básicos para definir la identidad de sus órganos.

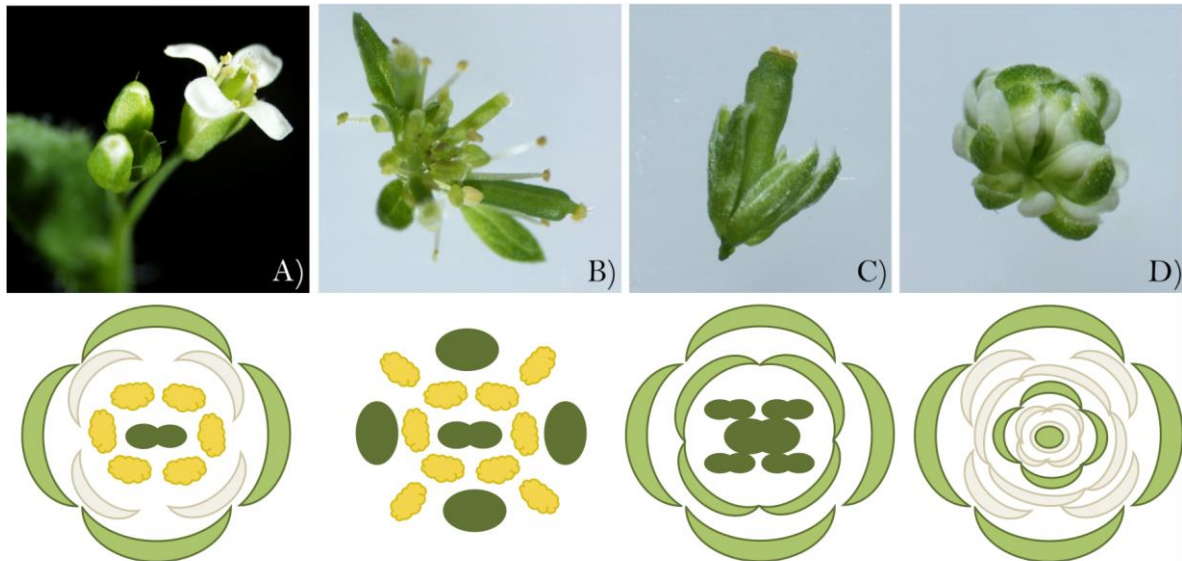


Figura 6. Las flores mutantes a, b y c. (A) El fenotipo silvestre de *Arabidopsis* comparado con sus variantes encontradas y analizadas a principios de los noventa: mutante a (B), mutante b (C) y mutante c (D). Ver la descripción de cada una en el texto. Este tipo de mutaciones son claramente homeóticas, y cada una se ejemplifica con su respectivo diagrama floral. Imágenes de *Arabidopsis* tomadas de Wikimedia Commons y Flickr; los diagramas florales son de mi autoría.

El modelo ABC plantea la existencia de tres regiones dentro del meristemo floral (A, B y C), cada una de las cuales coincide con el dominio espacial de expresión de tres clases distintas de genes homeóticos. En la región A se desarrollan los sépalos y pétalos; en la región B, los pétalos y estambres; en la región C, los estambres y carpelos (Coe y Meyerowitz 1991). Una combinación específica de “funciones” (llamadas también A, B y C) contribuyen al desarrollo de un órgano floral particular (ver **Figura 7**). La función A está representada por los genes *APETALA1* y *APETALA2*; la función B, por los genes *APETALA3* y *PISTILLATA*; y la función C, por el gen *AGAMOUS*. Por si sola, la función A puede dar origen a sépalos, pero junto con la función B puede otorgar identidad a pétalos. De la misma forma, la función C es capaz de promover el desarrollo de carpelos, pero si se combina con la función B puede dar inicio a la caracterización de estambres. El modelo sólo funciona si también se asume que las funciones A y C son mutuamente antagonistas: cuando una desaparece, la otra se expande y ocupa su lugar. Drews et al. (1991) comprobaron este supuesto a nivel genético.

La importancia del modelo ABC, y la razón por la cual sigue siendo válido para estudiar la estructura de las flores (Bowman et al. 2012), es que no sólo logra explicar de manera

sencilla el desarrollo de una flor normal y sus variantes homeóticas naturales (como las observadas en la **Figura 6**), sino que también permite hacer predicciones bastante cercanas sobre las modificaciones florales. Por otro lado, el modelo ABC ha sido utilizado para predecir cómo cambia la arquitectura floral en diversas especies de angiospermas (Rijpkema et al. 2010) –aunque Kramer y Irish (1999) han cuestionado la conservación funcional de los genes B en angiospermas basales–, y parece ser que puede darnos pistas sobre la evolución de la flor misma: la hipótesis de Goethe, por ejemplo, se vio parcialmente comprobada de alguna manera en mutantes que carecen de las tres funciones y cuyas flores sólo poseen órganos muy parecidos a hojas (Bowman et al. 1991).

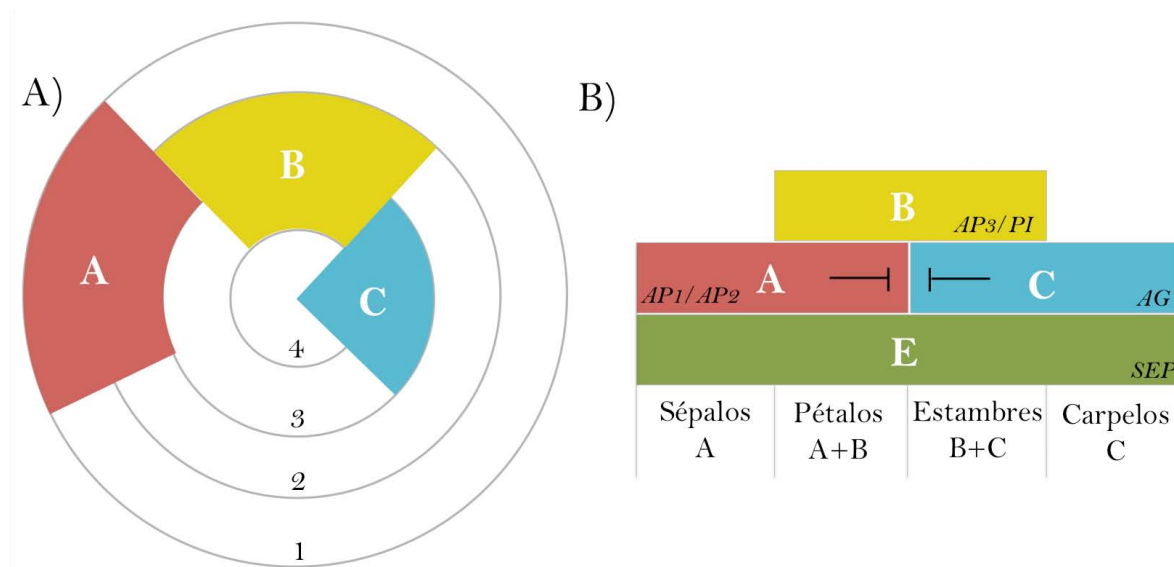


Figura 7. El ABC de la arquitectura floral en *Arabidopsis*. (A) El meristemo floral se divide en tres regiones que funcionan como dominios de acción de distintos genes homeóticos. La combinación de dichos genes otorga la identidad de los órganos florales que ocupan cada verticilo (de afuera hacia adentro, sépalos, pétalos, estambres y carpelos). (B) El modelo ABC –modificado aquí en ABCE– propone que, en el primer verticilo, los genes de función A (*APETALA1* [*AP1*] y *APETALA2* [*AP2*]) especifican la identidad de sépalos; en el segundo verticilo, los genes A y los de función B (*APETALA3* [*AP3*] y *PISTILLATA* [*PI*]) especifican para pétalos; en el tercer verticilo, los genes B y el de función C (*AGAMOUS* [*AG*]) otorgan la identidad de estambres; y, en el cuarto verticilo, el gen C dirige el desarrollo de carpelos. Existe también la función E, representada por la familia de genes *SEPALLATA* [*SEP*], necesaria para el desarrollo de todos los órganos florales. Diagramas modificados de Coen y Meyerowitz (1991).

De acuerdo con Theissen y Melzer (2007), los mutantes florales A, B y C –así como el modelo al que dieron nombre– resultaron ser muy reveladores, pues indicaron que el

desarrollo floral se rige por la acción conjunta de genes homeóticos capaces de controlar un “programa genético necesario para el diseño de cada órgano particular”. Pero a pesar de sus bondades, el modelo ABC pronto comenzó a poner en duda lo que Goethe dijo hace más de 200 años.

Si todas las partes de la planta fueran simples variaciones de un sólo tema, entonces podríamos asumir que las hojas pueden transformarse en órganos florales. Sin embargo, la expresión ectópica de genes B y C no logró alterar la identidad de los órganos vegetativos (revisado en Causier et al. 2010), lo cual dejó claro que las tres funciones florales son necesarias, pero no suficientes, para imponer el programa genético floral en un contexto molecular distinto (Theissen y Melzer 2007; Causier et al. 2010; Wellmer et al. 2014).

En 2000 llegó la primera demostración sobre la existencia de una cuarta función que podría ayudar a completar el modelo ABC. Pelaz et al. (2000) observaron que los triples mutantes florales de genes de la familia SEPALLATA (SEP1-SEP3) exhibían órganos muy parecidos a sépalos y demostraron que los genes de función B y C requieren de los SEP para especificar las identidades florales. Meses más tarde, descubrieron que los SEP otorgaban el contexto molecular necesario para convertir hojas en pétalos, estambres y carpelos (Pelaz et al. 2001). Con sus experimentos, corroboraron lo que Goethe alguna vez pensó y modificaron el modelo ABC en uno más preciso: el modelo ABCE, donde la función E resulta imprescindible para la formación de todos los órganos de la flor (ver **Figura 7**).

2.2 Los genes MADS-box

El modelo extendido ABCE fue un gran avance teórico que dio validez a ideas concebidas desde el siglo XVIII sobre la naturaleza de las flores, al tiempo que brindó un apoyo sólido para el estudio de la genética del desarrollo floral. Pero, al igual que su antecesor, era incapaz de proponer un mecanismo molecular que explicara la interacción de las proteínas capaces de especificar la identidad de sépalos, pétalos, estambres y carpelos (Theissen & Melzer 2007). Nuevas aproximaciones comenzaron a tomar forma con el entendimiento de que todos los genes involucrados en el modelo ABCE —a excepción de *AP2*— son integrantes de una sola familia de genes: los MADS-box (Coen y Meyerowitz 1991; Álvarez-Buylla et al. 2010a)

Los genes MADS-box se encuentran en prácticamente todos los linajes eucariontes (Friedman et al. 2004; Melzer y Theissen 2011), y sintetizan factores de transcripción encargados de regular la expresión de otros genes. El acrónimo ‘MADS’ se deriva de los nombres de los primeros genes caracterizados en esta familia: *MINICHROMOSOME MAINTENANCE 1*, de la levadura *Saccharomyces cerevisiae*; *AGAMOUS*, de la planta *Arabidopsis thaliana*; *DEFICIENS*, de la planta *Antirrhinum majus*; y *SERUM RESPONSE FACTOR*, del ser humano (Shwarz-Sommer et al. 1990). El término “box”, por su parte, se refiere a la secuencia conservada de unos 180 nucleótidos que agrupa como familia a estos genes y que codifica el dominio de unión al ADN en sus proteínas (Vergara-Silva et al. 2000). Este dominio evolucionó como resultado de la duplicación de una topoisomerasa *TOPOIIA-A* ancestral, una de cuyas copias divergió y adquirió una especificidad única de unión al ADN (Gramzow et al. 2010; ver **Figura 8a**). Hoy en día, todas las proteínas MADS-box reconocen secuencias en las regiones reguladoras de sus genes blanco denominadas cajas CArG con la estructura consenso CC(A/T)₆GG (Taylor et al. 1989; Riechmann et al. 1996), aunque se han reportado numerosas variantes con algunos cambios permitidos (de Folter y Angenent 2006).

Dada su amplia distribución, los genes MADS cumplen funciones muy distintas. En animales y hongos, por ejemplo, están involucrados en la proliferación y diferenciación celular, metabolismo de aminoácidos, desarrollo muscular y respuesta a feromonas y estrés osmótico (Messenguy y Dubois 2003). En las plantas, sin embargo, los MADS se han diversificado de manera impresionante a través de duplicaciones génicas y están involucrados en el control de los principales aspectos de sus ciclos de vida, incluido un abanico de procesos morfogénicos que dirige el desarrollo de estructuras vegetativas y reproductivas (Puruggannan et al. 1995; Friedman et al. 2004; Soltis et al. 2007; Gramzow y Theissen 2010). De hecho, es en el reino vegetal –sobre todo en espermatofitas, las plantas con semilla– donde los genes MADS se han expandido de manera impresionante gracias a duplicaciones genómicas y divergencia de parálogos: a diferencia de hongos o metazoarios, los genomas de las angiospermas más representativas incluyen un promedio de 100 genes MADS⁷ (Smaczniak et al. 2012).

⁷ Tan sólo en el genoma de *Arabidopsis* se han encontrado 107 genes MADS-box (Pařenicová et al. 2003) mientras que animales como la mosca de la fruta o el ser humano poseen sólo dos y cinco, respectivamente (Gramzow et al. 2010). En hongos, como *Saccharomyces*, el número es igual de bajo.

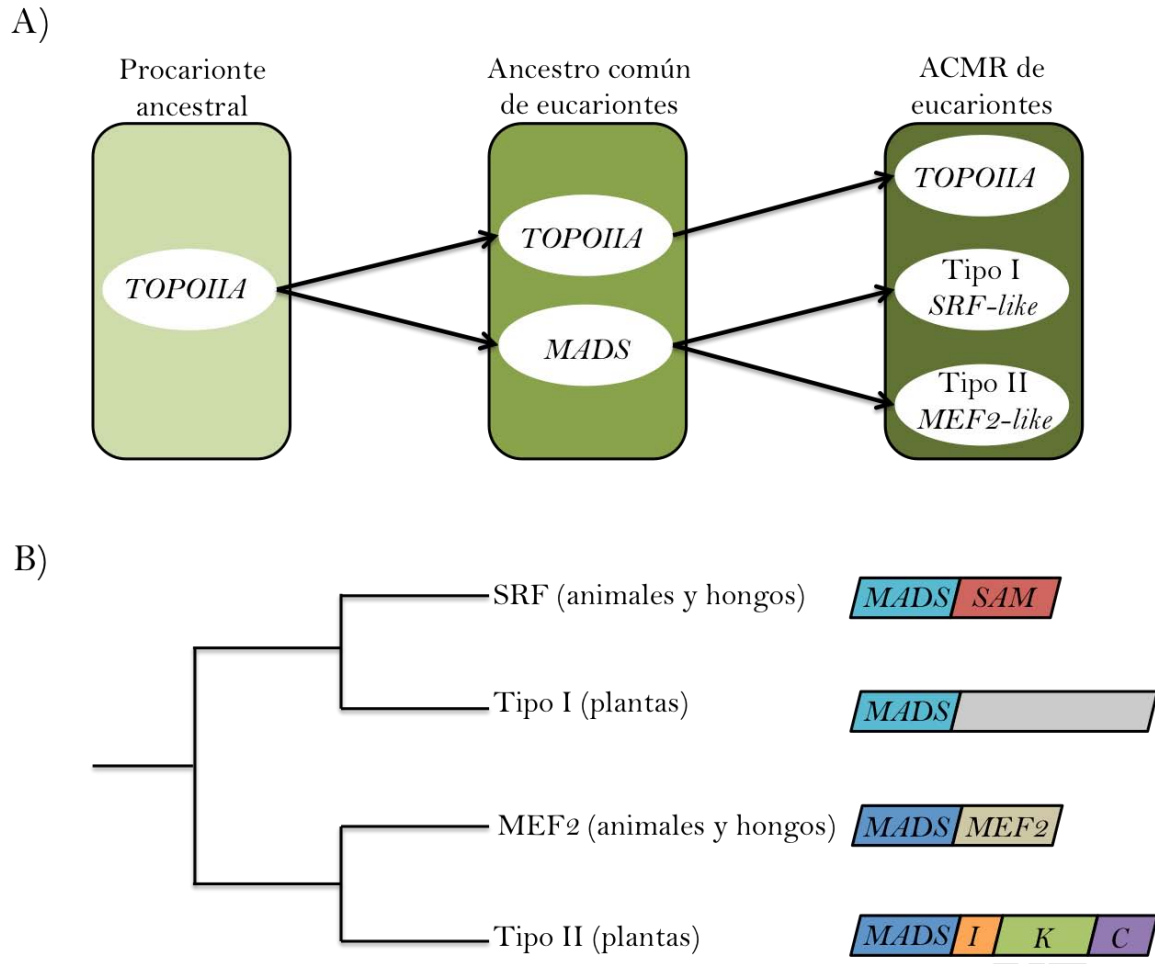


Figura 8. Origen y evolución de la familia MADS. (A) La duplicación de un gen de topoisomerasa codificado por un procarionte ancestral dio origen a un gen *TOPOIIA* y un gen con dominio MADS en el ancestro común de los eucariontes. Otra duplicación posterior del gen con dominio MADS concluyó en la evolución de genes con dos dominios MADS distintos en el ancestro común más reciente (ACMR) de plantas, animales y hongos. (B) Estos eventos resultaron en la división de ambos linajes de genes MADS que se encuentran en plantas: el tipo I y el tipo II. A excepción del dominio MADS *SRF-like*, los genes tipo I no presentan ningún otro dominio conservado. En contraste, los genes tipo II (MIKC), que son específicos de plantas, comparten un dominio MADS *MEF2-like* seguido de los dominios *intervening* (I), *keratin-like* (K) y *C-terminal* (C). Las tres subunidades del dominio K están resaltadas por líneas negras. Imagen modificada de Gramzow et al. (2010) y de Folter y Angenent (2006).

La familia de factores de transcripción MADS-box puede dividirse en dos linajes, según su estructura proteica: los tipo I, o *SRF-like*, y tipo II, o *MEF2-like* (ver **Figura 8b**). De acuerdo con Álvarez-Buylla et al. (2000), ambos linajes se originaron durante la evolución temprana de los eucariontes a partir de un evento de duplicación que ocurrió antes de la divergencia de plantas y animales. Otras investigaciones, como las de Pařenicová et al.

(2003), Nam et al. (2003) y Gramzow et al. (2010), han comprobado los resultados de este estudio y sugerido la existencia de linajes MADS en el ancestro común más reciente de plantas, animales y hongos.

En los genes MADS tipo I de plantas sólo se encuentra conservado el dominio de unión al ADN, es decir, el que les da nombre. No fue sino hasta hace pocos años que se comenzó a descubrir que los genes *SRF-like* tienen importantes funciones durante la fase reproductiva, pues participan en el desarrollo del endospermo y gametofito femenino, así como en la regulación de dosis génicas y compatibilidad postcigótica (Masiero et al. 2011).

Por su parte, los genes MADS tipo II de plantas exhiben una estructura modular bastante definida, reflejada también en sus proteínas (ver **Figura 8b**): del extremo 5' al 3', un dominio **M**ADS muy conservado que facilita la unión a las cajas CArG distribuidas en el ADN, aunque a veces también se ve involucrado en la dimerización proteica y la interacción con factores accesorios (Shore y Sharrocks 1995; Pellegrini et al. 1995), un dominio **I**ntervening medianamente conservado que contribuye a especificar la formación del dímero de unión al ADN y la interacción proteína-proteína (Gramzow y Theissen 2010), un dominio **K**eratin-like, el segundo más conservado de todos, dividido en tres subdominios –K1, K2 y K3– que regulan la interacción proteína-proteína y ayudan en la formación de dímeros y multímeros (Fan et al. 1997; Kaufmann et al. 2005) y, por último, en el extremo 3' se encuentra el dominio **C**-terminal, muy poco conservado entre especies, con un papel importante en la transactivación de genes blanco y formación de complejos multiméricos (Cho et al. 1999; Egea-Cortines et al. 1999; Honma y Goto 2001). Debido al orden de dominios, a los productos de los genes MADS tipo II se les conoce también como proteínas MIKC.

De todos los genes MADS⁸, aquellos del modelo ABCE, involucrados en esculpir la arquitectura de la flor, han recibido mayor atención en estudios sobre evolución y desarrollo (Puruggannan et al. 1995; Ng y Yanofsky 2001; Friedman et al. 2004). Fue gracias a ellos que se logró esbozar el modelo –hasta ahora más aceptado– que explica el mecanismo molecular del desarrollo floral (Álvarez-Buylla et al. 2010a): el cuarteto floral

⁸ Algunos análisis filogenéticos, como el de Rounsley et al. (1995), han encontrado que muchos otros genes MADS –fuera del grupo de los genes homeóticos florales– representan los loci más basales de la familia, lo cual sugiere que quizá la función de los primeros MADS-box tenía que ver con el desarrollo vegetativo y embrionario. Tras numerosas duplicaciones, algunos de estos genes podrían haber diversificado y adquirido nuevas funciones dentro del desarrollo floral.

(Theissen y Saedler 2001). Este sugiere que las proteínas MADS-box MIKC forman tetrámeros –en específico, un dímero de dímeros– que se unen a dos cajas CARG ubicadas en las regiones reguladoras de sus genes blanco, a los que activan o reprimen durante el periodo de determinación de órganos florales, es decir, durante la transición de inflorescencia a flor. El modelo también predice la existencia de un complejo proteico específico que controla el desarrollo de cada tipo de órgano (ver **Figura 9**). De esta manera, un cuarteto de proteínas homeóticas A+E da identidad a sépalos; uno de proteínas A+B+E, a pétalos; otro de proteínas B+C+E, a estambres; y finalmente uno de proteínas C+E, a carpelos.

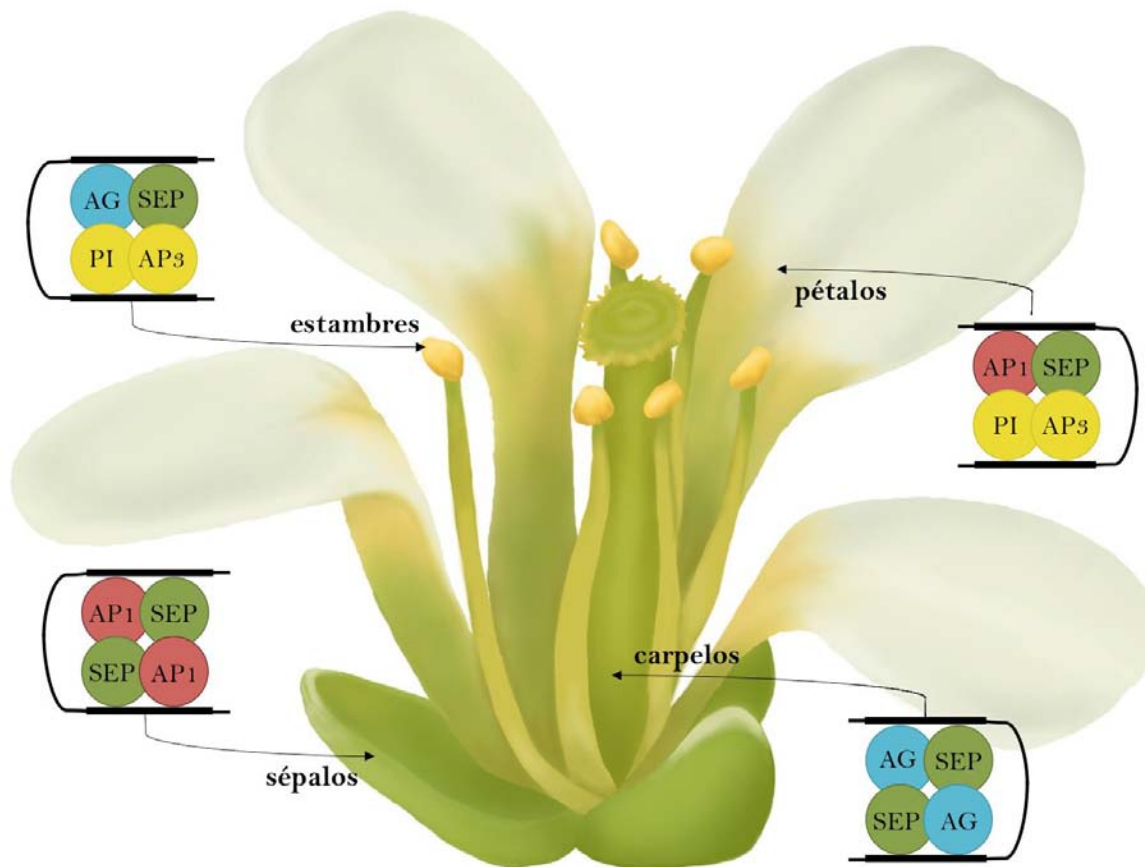


Figura 9. El cuarteto floral. La identidad de los órganos florales se ve determinada por cuatro combinaciones distintas de tetrámeros conformados por proteínas MADS-box. Dos dímeros de cada tetrámero reconocen dos cajas CARG distintas (rectángulos negros), ubicadas en regiones reguladoras, y se unen a través de un *loop* en el ADN. Uno de los dímeros podría funcionar como el dominio transactivador del complejo mientras que el otro podría ser importante para la regulación genética órgano-específica. Imagen modificada de Theissen y Saedler (2001) y Wellmer et al. (2014). Ilustración de *Arabidopsis* realizada por Erin Dolan y tomada de la página del Partnership for Research and Education in Plants.

El cuarteto floral conoció en paralelo el apoyo experimental de Honma y Goto (2001), quienes evaluaron a través de ensayos de tres y cuatro híbridos en levadura la formación de los tetrámeros propuestos por Theissen y Saedler (2001) para el desarrollo de pétalos y estambres. Además, lograron convertir hojas vegetativas en órganos petaloides y estaminoides al coexpresar de manera ectópica los genes MADS necesarios para modificar el contexto molecular durante el desarrollo, comprobando así “la base molecular de la interacción combinatoria de los genes homeóticos florales” (Theissen y Saedler 2001).

Desde un punto de vista de la evo-devo, los avances teóricos y análisis genéticos sobre los que he hablado en este capítulo resultan fundamentales para comprender los cambios corporales de las plantas, pues estudiar los mecanismos moleculares que subyacen tras la morfogénesis floral puede ayudarnos a descubrir cómo la variación genética se ve reflejada en variación fenotípica (Carroll 2008). Al respecto, los genes MADS-box representan, sin duda, algunos de los nodos más importantes de la red de regulación genética que gobierna el desarrollo de los órganos florales, y que provee una explicación dinámica del modelo ABC (Espinosa-Soto et al. 2004; Álvarez-Buylla et al. 2007; ver **Figura 14c** del Capítulo 4). Investigar su función y regulación es la clave para explicar las bases evolutivas de la diversificación estructural de la flor.

CAPÍTULO 3. La esperanza de los monstruos lacandonos

3.1 La fascinación por *Lacandonia schismatica*

Richard Goldschmidt fue un genetista alemán que ganó reconocimiento por su extenso trabajo en la determinación sexual de la polilla *Lymantria* y sus observaciones sobre la íntima relación entre el gen y los procesos fisiológicos del desarrollo en *Drosophila* (Dietrich 2003). Sin embargo, no es por eso que se le recuerda. Poco tiempo después de que el régimen nazi diera inicio a la purga de científicos judíos en Alemania, Goldschmidt salió de su país en 1935 y se instaló en la Universidad de Berkeley, Estados Unidos, donde comenzó a provocar controversia en la academia por sus propuestas “heréticas” que iban en contra de la teoría sintética de la evolución. Estas propuestas finalmente vieron la luz con la publicación de su libro *Las bases materiales de la evolución* (Goldschmidt 1940).

Con la idea de que la selección natural no es el medio de producción de nuevas especies, pues sólo actúa sobre pequeñas diferencias intrínsecas dentro de las poblaciones, Goldschmidt rechazó las suposiciones neosintéticas de que la macroevolución⁹ no es más que la acumulación de cambios microevolutivos extendidos en el tiempo y de que la divergencia morfológica siempre ocurre a pasos pequeños y graduales¹⁰ (Goldschmidt 1933; Goldschmidt 1940).

En cambio, propuso que las especies divergían entre sí gracias a la acción de mutaciones que serían desastrosas la mayoría de las veces para la supervivencia de los organismos, ahora monstruosos, pero hipotetizó que en raras ocasiones uno de estos individuos sobreviviría y, con él, la esperanza de adaptarse a un estilo de vida diferente y dar inicio a un nuevo linaje evolutivo (Dietrich 2003; Theissen 2009b). El papel de la selección natural entraría, más bien, cuando actuara sobre las monstruosidades viables o no viables, de acuerdo a un criterio embriológico y ambiental. Así nació el concepto de *monstruo esperanzado*, también conocido como monstruo prometedor (Goldschmidt 1933).

⁹ Definida como la evolución *en y por arriba* del nivel de especie, aunque algunos autores como Theissen (2006) prefieren pensar en ella como innovaciones en el plan corporal.

¹⁰ Richard Goldschmidt nunca negó la existencia de una evolución gradual, pero estaba convencido de que ésta se encontraba confinada a un plano microevolutivo y no era suficiente para explicar el aparente salto entre una especie y otra (Goldschmidt 1933; Dietrich 2003; Theissen 2006).

En su libro, Goldschmidt (1940) sugirió la existencia de dos mecanismos principales capaces de explicar el origen de los monstruos esperanzados: las mutaciones sistémicas y las macromutaciones del desarrollo. Las primeras, que dañaron más gravemente su reputación por la negación del concepto clásico de gen¹¹, consisten en rearrreglos cromosómicos que provocan notables cambios genotípicos y efectos fenotípicos inmediatos capaces de crear una nueva especie (Dietrich 2003; Theissen 2009a). Por otro lado, también pensó que si las macromutaciones de genes clave durante el desarrollo temprano de los organismos fueran viables, podrían modificar en un sólo paso los procesos de diferenciación morfológica y conducir a una divergencia evolutiva abrupta (Goldschmidt 1940).

Sin embargo, las ideas de Goldschmidt fueron despreciadas y denigradas por los padres de la Teoría Sintética. Una de sus críticas fue la improbabilidad de que se fijara en condiciones naturales y a una escala poblacional el fenotipo de un solo individuo monstruoso (Charlesworth 1982; Schwartz 1999). Esto sigue siendo un tema de debate, pero se han propuesto algunas explicaciones complementarias que respaldan la existencia de los monstruos esperanzados.

West-Eberhard (2003; 2005; 2008) sugiere que un estímulo, sea un estrés ambiental o una mutación genética, puede afectar a algunos individuos al alterar rutas reguladoras del desarrollo que den lugar de manera simultánea a un cambio en el fenotipo. Este cambio fenotípico viene de un fenotipo ancestral cuyo desarrollo es plástico y se reorganiza para producir nuevas variantes cuando es estimulado (West-Eberhard 2005). El origen de nuevos rasgos a través de patrones alterados de expresión de rasgos ancestrales es algo a lo que West-Eberhard (2003) llama *recombinación del desarrollo*. El nuevo fenotipo puede comenzar a expandirse en una subpoblación, y si tiene un efecto en la adecuación entonces se verá sometido a selección natural y fijado en la población por *acomodación genética*, es decir, por cambios en la frecuencia de genes que afectan la regulación y forma del nuevo rasgo (West-Eberhard 2005).

Por su parte, Schwartz (1999) propone un modelo de cambio evolutivo basado en la herencia mendeliana donde las mutaciones de genes reguladores aparecerían en un principio como alelos recesivos. A través de varias generaciones, y gracias al proceso de

¹¹ Goldschmidt comenzó a cuestionar el concepto clásico de gen cuando se vio convencido del efecto de posición —es decir, cuando un fenotipo cambia debido a un cambio en la localización de un gen—, pues el fenómeno cuestionaba la existencia de unidades estructural y funcionalmente independientes.

heterocigosis, estos alelos recesivos se distribuirían de manera silenciosa en la población hasta que la reproducción azarosa produjera descendencia homocigota para los mismos (Schwartz 1999). De esta manera, en los homocigotos se expresaría abruptamente el fenotipo monstruoso, sin ningún rastro de modificación gradual.

Una de las premisas de la evo-devo es que las novedades evolutivas son resultado de cambios en genes clave del desarrollo temprano de los organismos (Theissen 2009b; ver primer capítulo de esta tesis). No es coincidencia que Goldschmidt reconozca la homeosis como evidencia de sus macromutaciones del desarrollo, capaces de producir un efecto importante en el fenotipo de manera súbita (Vergara Silva 2009). Durante sus últimos años de vida, incluso, comenzó a ver a algunas moscas con mutaciones homeóticas como la prueba irrefutable de sus monstruos esperanzados (Dietrich 2003).



Figura 10. Las flores homeóticas de *Lacandonia*. *Lacandonia schismatica* es una planta herbácea que pertenece al grupo de las monocotiledóneas. Se caracteriza por su ausencia de clorofila, su pequeño tamaño (hasta de nueve centímetros de longitud), su fecundación cleistogámica (es decir, antes de la antesis floral) y sus inflorescencias racemosas donde se desarrollan flores cuyos órganos sexuales se encuentran invertidos (Martínez y Ramos 1989; Ambrose et al. 2006). A la izquierda se muestra la arquitectura racemosa de un individuo de *Lacandonia* con estructuras florales antéticas y preantéticas; a la derecha, un detalle de su flor donde se distinguen tépalos (T), carpelos (C) y estambres (E). Ambas fotografías son propiedad de Joel McNeal.

Más de 50 años después de su muerte, en 1958, los monstruos esperanzados parecen reclamar su lugar dentro de la biología evolutiva. Algunos estudios han demostrado la aparición de grandes diferencias fenotípicas debido a cambios sutiles en el genoma de poblaciones naturales de plantas (Bateman y DiMichele 2003; Kanno et al. 2003; Kramer et al. 2003; Hintz et al. 2006; Theissen 2006), animales (Rieppel 2001; Peichel et al. 2001; Dietrich 2003; Theissen 2006; Chan et al. 2010; Page et al. 2010) y, discutiblemente, bacterias (Blount et al. 2008). Pero sin duda uno de los ejemplos que más llaman la atención es el de la planta mexicana *Lacandonia schismatica* (familia Triuridaceae; Vergara-Silva et al. 2000).

Desde su descubrimiento en 1985 por Esteban Martínez y Clara Ramos (1989), cerca de la localidad de Crucero Corozal en la Selva Lacandona, Chiapas, *Lacandonia schismatica* (ver **Figura 10**) fue motivo de polémica entre la comunidad científica, pues ninguno de sus rasgos –ni su apariencia blanquecina, pequeño tamaño, saprotrofia o fecundación cleistogámica– resulta tan extraño como su flor.

La diversidad floral de las más de 350 mil especies aceptadas de angiospermas (The Plant List 2013) es alucinante pero, hasta donde sabemos, los cimientos básicos de su arquitectura están bastante conservados; al menos en flores hermafroditas, del verticilo más externo al más interno, el orden espacial de los órganos florales es siempre el mismo (ver **Figura 7** del capítulo anterior). Hasta hace poco, se consideraba que *Lacandonia schismatica* representaba la única especie cuyo arreglo floral era diferente, pero en 2012 se reportó el descubrimiento de *Lacandonia brasiliiana* en la Reserva Biológica de Guaribás ubicada en el estado de Paraíba, en Brasil (Melo y Alves 2012). Sin embargo, a la fecha no se han realizado análisis filogenéticos, moleculares o morfogenéticos para distinguir una especie de otra ni para conocer más a fondo la biología de *L. brasiliiana*.

La flor de *Lacandonia schismatica* difiere de todas las demás porque sus órganos sexuales se encuentran invertidos: de 60 a 80 carpelos periféricos rodean tres estambres centrales (Martínez y Ramos 1989; Márquez Guzmán et al. 1989; Ambrose et al. 2006; ver **Figura 10**). Este fenotipo homeótico –fijado en poblaciones naturales, lo cual reafirma su calificativo de monstruo prometedor y su estudio como un caso de evolución saltacionista (Vergara-Silva et al. 2003)– ha sido el estandarte de discusiones sobre qué interpretación se merecen sus estructuras reproductivas: ¿inflorescencias comprimidas con órganos

abortados o flores verdaderas? (Stevens 1991; Rudall 2003; Ronse de Craene 2003). La evidencia histológica, embriológica, estructural, molecular y poblacional de investigaciones como las de Márquez Guzmán et al. (1989), Vergara-Silva et al. (2003), Ambrose et al. (2006) y Álvarez-Buylla et al. (2010b) apoya la visión euantial de que la de *Lacandonia schismatica* es una flor verdadera con una inversión homeótica.

¿Será entonces *Lacandonia* un ejemplo de una innovación evolutiva? Hace más de dos décadas, Müller y Wagner (1991) redefinieron el concepto de “novedad morfológica” en un intento por homogeneizar las propuestas previas y aclarar confusiones generalizadas y, en años más recientes, West-Eberhard (2003) se dio a la tarea de hacer lo propio. Sus definiciones integran el desarrollo al concepto de innovación evolutiva y dan un peso importante al surgimiento de estructuras no homólogas. Más tarde, Pigliucci (2008) logró combinar ambas propuestas: “una novedad evolutiva es un nuevo rasgo, o combinación de rasgos previamente existentes, que surge durante la evolución de un linaje y que realiza una nueva función dentro de su ecología”.

Si por homología morfológica pensamos en aquellas partes corporales casi independientes que divergieron a partir de una misma parte en un ancestro común (Wagner y Lynch 2010), parece difícil pensar en la flor de *Lacandonia* como una novedad evolutiva por el hecho de que “la homología a nivel de la identidad de sus órganos florales es inequívoca” (Vergara-Silva et al. 2003). Además, según la definición de Pigliucci (2008), la homeosis observada en los órganos sexuales de la flor de esta triuridácea tampoco parece implicar un cambio en su polinización (i.e. la ecología), pues *L. schismatica* es una especie cleistogámica (**Figura 10**). Sin embargo, Müller (2010) defiende la clasificación de tres tipos de innovaciones morfológicas. El primer tipo serían cambios anatómicos en el plan corporal, el segundo tipo sería la adición de un nuevo elemento a un plan corporal existente, y el tercer tipo sería un cambio importante en un carácter existente del plan corporal (Müller 2010). De acuerdo a esta clasificación, las flores homeóticas de *Lacandonia schismatica* bien podrían considerarse una innovación morfológica de tipo III.

Bajo el supuesto de que los sistemas modelo, tanto en plantas como en animales, no son suficientes para entender la evolución de la biodiversidad o los mecanismos generadores de variación fenotípica, la biología evolutiva del desarrollo ha salido en busca de especies nuevas para estudiar (Jenner y Wills 2007; Abzhanov et al. 2008; Kramer 2009; Sommer 2009). Considerando que la urgencia ahora recae en un acercamiento pragmático para

realizar investigaciones experimentales, no pienso que *L. schismatica* represente un modelo biológico alternativo, pues es una especie endémica en peligro de extinción que hasta ahora no ha logrado ser cultivada de manera exitosa en condiciones de laboratorio. Pero sí creo que *Lacandonia* representa una ventana única desde donde asomarse hacia la comprensión de la diversidad de los mecanismos morfogénéticos del desarrollo floral y, probablemente, hacia el papel de la homeosis en la evolución fenotípica de las plantas.

3.2 El encanto viene de familia

Escondida bajo la hojarasca de la Selva Lacandona se encuentra también *Triuris brevistylis*, que ha sido señalada como el taxón más cercanamente relacionado a *Lacandonia schismatica* y la única otra especie de triuridácea encontrada en México (Vergara-Silva 2003). Como la mayoría de los géneros agrupados dentro de la familia Triuridaceae, *T. brevistylis* es una especie dioica, esto es, sus poblaciones se ven integradas por individuos que presentan flores masculinas o femeninas (ver **Figura 11**), lo cual sugiere una reproducción biparental –aunque nunca se ha documentado en la literatura–. Sin embargo, algunos estudios poblacionales y comparativos han demostrado la existencia individuos monoicos que poseen ambos tipos de flores (Espinosa Matías et al. 2012).

En años recientes, *Triuris brevistylis* se ha visto incluida en estudios más profundos sobre su biología dada su relación filogenética con *L. schismatica*. Vergara-Silva et al. (2003) fueron los primeros en hacer un análisis poblacional de *T. brevistylis* en busca de variación morfológica en sus flores. Después de un trabajo de campo exhaustivo, de 1995 a 2001, se recolectaron más de mil individuos, de los cuales sólo el 1% eran masculinos, 0.5% eran bisexuales (**Figura 12**) y el resto eran femeninos (Vergara-Silva et al. 2003). Una investigación posterior de Ambrose et al. (2006), donde se estudiaron las series comparativas de desarrollo de las estructuras reproductivas de *T. brevistylis* y *L. schismatica*, comprobó que los ejes reproductivos de ambas plantas son flores verdaderas.

Estos resultados sugieren distintas conclusiones. Por ejemplo, la proporción de individuos masculinos y femeninos en poblaciones naturales de *T. brevistylis* no descarta la existencia de doble fecundación en la especie, pero sí indica fuertemente que esta planta también podría reproducirse gracias otro mecanismo de reproducción. La hipótesis fue

comprobada hace poco por Espinosa Matías et al. (2012), quienes observaron por primera vez la existencia de cleistogamia en flores bisexuales de *T. brevistylis*.



Figura 11. *Triuris brevistylis* es una especie dioica. Como los demás miembros de Triuridaceae, *T. brevistylis* también es una planta micoheterótrofa con una evidente ausencia de clorofila. A pesar de tener una distribución geográfica más amplia y extenderse a otras regiones neotropicales, *Triuris* vive en proximidad con *L. schismatica*, aunque en altitudes más altas. En la imagen se observan fotos de individuos femeninos (izquierda) y masculinos (derecha). Ambas fotografías son propiedad de Joel McNeal.

Por otro lado, Vergara-Silva et al. (2003) postularon que la observación de individuos hermafroditas en poblaciones naturales de *T. brevistylis* con cierto grado de homeosis en su arreglo floral podría indicar que las flores del ancestro común entre *L. schismatica* y su taxón hermano ya presentaban una tendencia hacia la inversión espacial de sus órganos sexuales. Comprender los mecanismos morfogenéticos del desarrollo floral en *T. brevistylis* podría iluminar mucho más el estudio del fenotipo homeótico en *L. schismatica* y quizá también podría darnos pistas de que los patrones morfogenéticos y reproductivos en *T. brevistylis* son, en efecto, incluso más complejos que los de *L. schismatica* (Espinosa Matías et al. 2012).

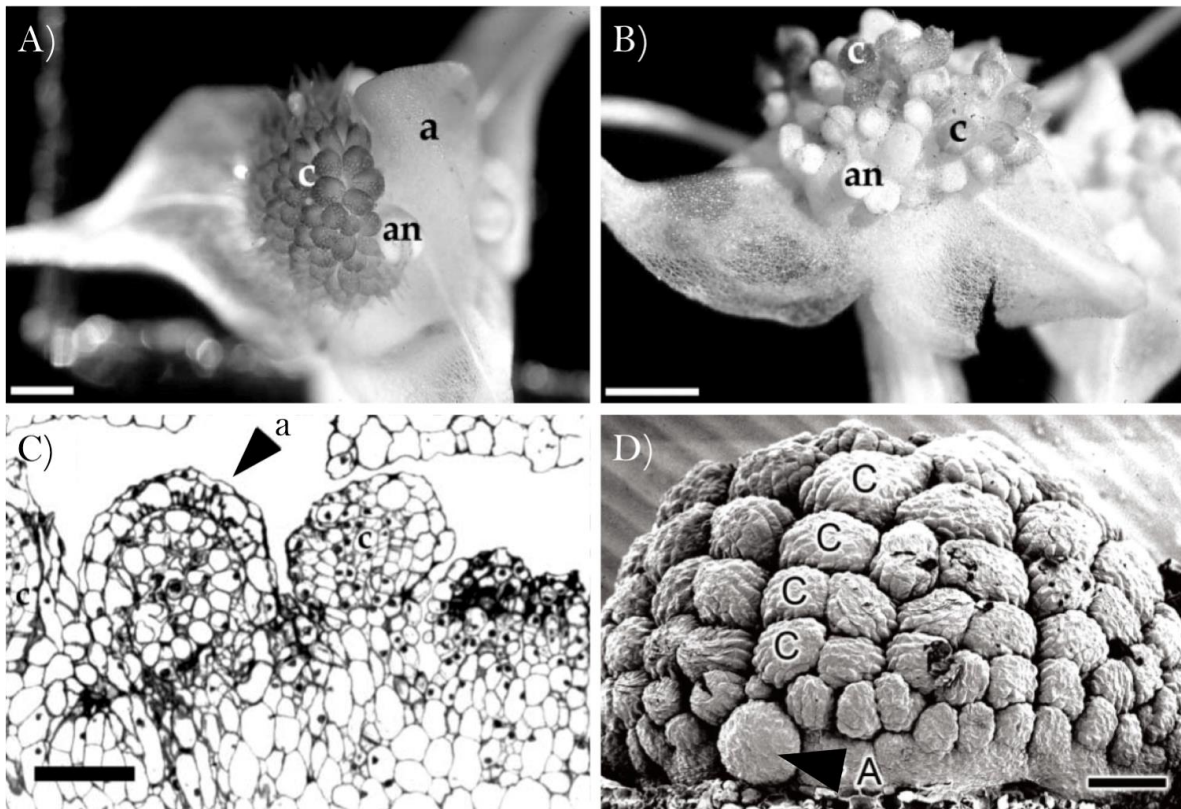


Figura 12. La homeosis de las flores bisexuales de *T. brevistylis*. Vergara-Silva et al. (2003) y Espinosa Matías et al. (2012) han identificado al menos cuatro tipos de arreglos florales homeóticos en los ejes bisexuales de *T. brevistylis*. (A) Flores masculinas con carpelos heterotópicos, escala = 0.5cm. (B) Flores femeninas con estambres entremezclados con carpelos, escala = 1cm. (C) Flores bisexuales con un arreglo tipo *L. schismatica*, donde los estambres se ven rodeados por carpelos, escala = 100μm. (D) Flores femeninas con estambres ubicados en la base del receptáculo, escala = 90μm. Abreviaturas: *an*, andróforo; *a*, antera (parte terminal del estambre); *c*, carpelo. Fotografías tomadas de Vergara-Silva et al. (2003) y Espinosa Matías et al. (2012).

CAPÍTULO 4. Antecedentes

4.1 No es lo mismo decir ACB a decir ABC

El modelo ABC del desarrollo floral representa una herramienta útil para hacer predicciones teóricas sobre los cambios observados en la arquitectura básica de las flores. Uno de los estudios más reveladores sobre las bases morfogenéticas del fenotipo homeótico de *Lacandonia schismatica* fue justamente concebido a partir de una simple hipótesis sobre la modificación del modelo ABC: el desarrollo de estambres en el centro del meristemo floral es consecuencia del desplazamiento espacial de la función B (Vergara-Silva et al. 2003; Álvarez-Buylla et al. 2010b; **Figura 13**).

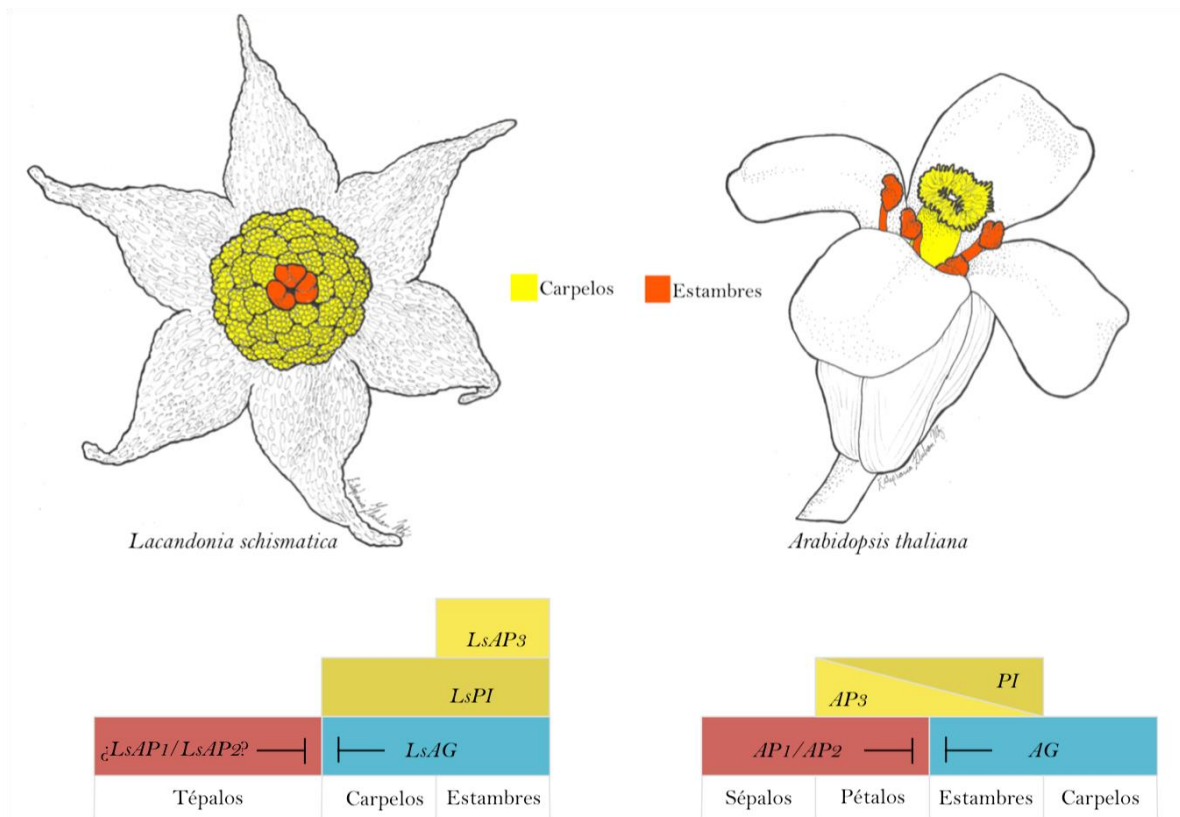


Figura 13. Cuando ABC se convierte en ACB. La inversión homeótica de estambres y carpelos en la flor de *L. schismatica* se explica por el desplazamiento espacial de la función B hacia el centro del meristemo floral, donde coexiste la expresión de los genes *LsAP3* y *LsPI*. Este cambio relativamente pequeño es suficiente para convertir al modelo ABC en un modelo ACB. Ilustraciones de Estephania Zluhan, tomadas de Rodríguez Mega y Hernández Marroquín (2013).

Tras la clonación y caracterización de los genes ortólogos de la función B y C en *Lacandonia*, Álvarez-Buylla et al. (2010b) lograron determinar su expresión espaciotemporal con experimentos de hibridación *in situ*. Sus resultados muestran una conservación funcional de los genes *LsPI* y *LsAG*, cuyos dominios de expresión siguen el patrón observado en otras angiospermas mientras que *LsAP3* se expresa desde etapas muy tempranas del desarrollo en la región central de la flor, zona donde más tarde se formarán los primordios de estambres.

Aunque el patrón de expresión del gen B *LsPI* se extiende hacia el tercer verticilo, se traslapa con la expresión de *LsAP3* en el centro del meristemo floral. Es entonces sólo en esta región donde se ubica la función B que, junto con la expresión de la función C (representada por el gen *LsAG*), contribuye al desarrollo de estambres (Álvarez-Buylla et al. 2010b; **Figura 13**). Esta observación ofrece una explicación relativamente sencilla a la inversión homeótica de las flores de *Lacandonia* a la vez que resalta el interés por estudiar de manera más profunda al gen MADS-box *APETALA3*, cuyo papel en la evolución morfológica floral ha sido documentado previamente en distintos estudios (Ainsworth et al. 1995; Kanno et al. 2003; Kramer et al. 2003; Tsai et al. 2004; Kim et al. 2005; de Martino et al. 2006; Tsaftaris et al. 2006; Xu et al. 2006; Drea et al. 2007).

4.2 Estructura y regulación del gen *APETALA3*

Los primeros análisis sobre *AP3* en la década de los noventa están relacionados con el interés por desmenuzar las bases genéticas de los mutantes homeóticos florales observados en la planta modelo *Arabidopsis thaliana* (Bowman et al. 1989; ver **Figura 6** en el segundo capítulo). Con ellos se descubrió que *AP3* está involucrado en el desarrollo de estambres y pétalos, órganos donde sus niveles de ARNm se encuentran más elevados (Jack et al. 1992). Tras comprobarse que *AP3*, como otros genes homeóticos de plantas, presenta en su estructura un dominio MADS, fue catalogado como un factor de transcripción (Jack et al. 1992). Esto sería corroborado más tarde por Jack et al. (1994), quienes detectaron niveles importantes de la proteína AP3 en el núcleo celular, lo cual sugería una interacción directa con el ADN.

Después de haberse descrito la estructura genética de *AP3* (Jack et al. 1992; **Figura 14a**), otros estudios comenzaron a investigar la regulación de este gen floral. Jack et al. (1994), por ejemplo, fueron los primeros en ofrecer evidencia experimental sobre la autoregulación positiva de *AP3* y en sugerir que *PI* también juega un papel en la activación del promotor de *AP3* en tejido floral. Esto fue comprobado poco tiempo después por Krizek y Meyerowitz (1996), quienes también demostraron que la activación de *AP3* y *PI* es suficiente para proveer la función B durante el desarrollo floral y especificar la identidad de pétalos y estambres en conjunto con los genes de función A y C, respectivamente. Ese mismo año, Davies et al. (1996) y Riechmann et al. (1996) detectaron que *AP3* y *PI* forman un heterodímero específico de unión al ADN, facilitado en gran medida por la interacción de sus dominios K, capaz de reconocer secuencias reguladoras llamadas cajas CArG. El modelo del cuarteto floral, explicado en el segundo capítulo de esta tesis, sería propuesto más tarde (ver **Figura 9**).

No fue sino hasta 1998 cuando dos publicaciones paralelas escudriñaron a fondo la secuencia promotora de *AP3* con el fin de disectarla y examinar a detalle cómo la expresión de este gen se establece y mantiene en la flor. A partir de la fusión de fragmentos del promotor de *AP3* con genes reporteros y su expresión en plantas transgénicas de *Arabidopsis thaliana*, Hill et al. (1998) y Tilly et al. (1998) definieron algunos de los elementos *cis* requeridos para su expresión espacial y temporal (ver **Figura 14a**). Además, Tilly et al. (1998) demostraron que sólo 496 pb del promotor –y no 1.7 kb, como habían reportado Irish y Yamamoto (1995)– son suficientes para dirigir la expresión de genes reporteros en el mismo patrón espaciotemporal que se observa con *AP3* en plantas silvestres. La evidencia de ambos estudios sugiere que, al menos en *Arabidopsis*, las secuencias reguladoras encargadas de controlar la expresión de *AP3* se encuentran exclusivamente en la región 5' del ATG.

Tilly et al. (1998) también observaron que un fragmento del promotor de *AP3* donde se encuentran las tres cajas CArG podía recuperar aspectos clave de la expresión normal del gen. Interesados por analizar el papel de las cajas CArG en *AP3*, realizaron experimentos de mutagénesis dirigida y documentaron los efectos mediante construcciones con sistemas reporteros. Cuando mutaron la CArG1 descubrieron una ausencia de la expresión génica en todos los estadios florales del desarrollo; cuando mutaron la CArG2 sólo pudieron medir actividad en los primordios de estambres, pero no en los de pétalos; cuando mutaron

la CArG3 observaron un aumento en la intensidad de la expresión durante las etapas tempranas del desarrollo de pétalos y estambres (Tilly et al. 1998; **Figura 14b**).

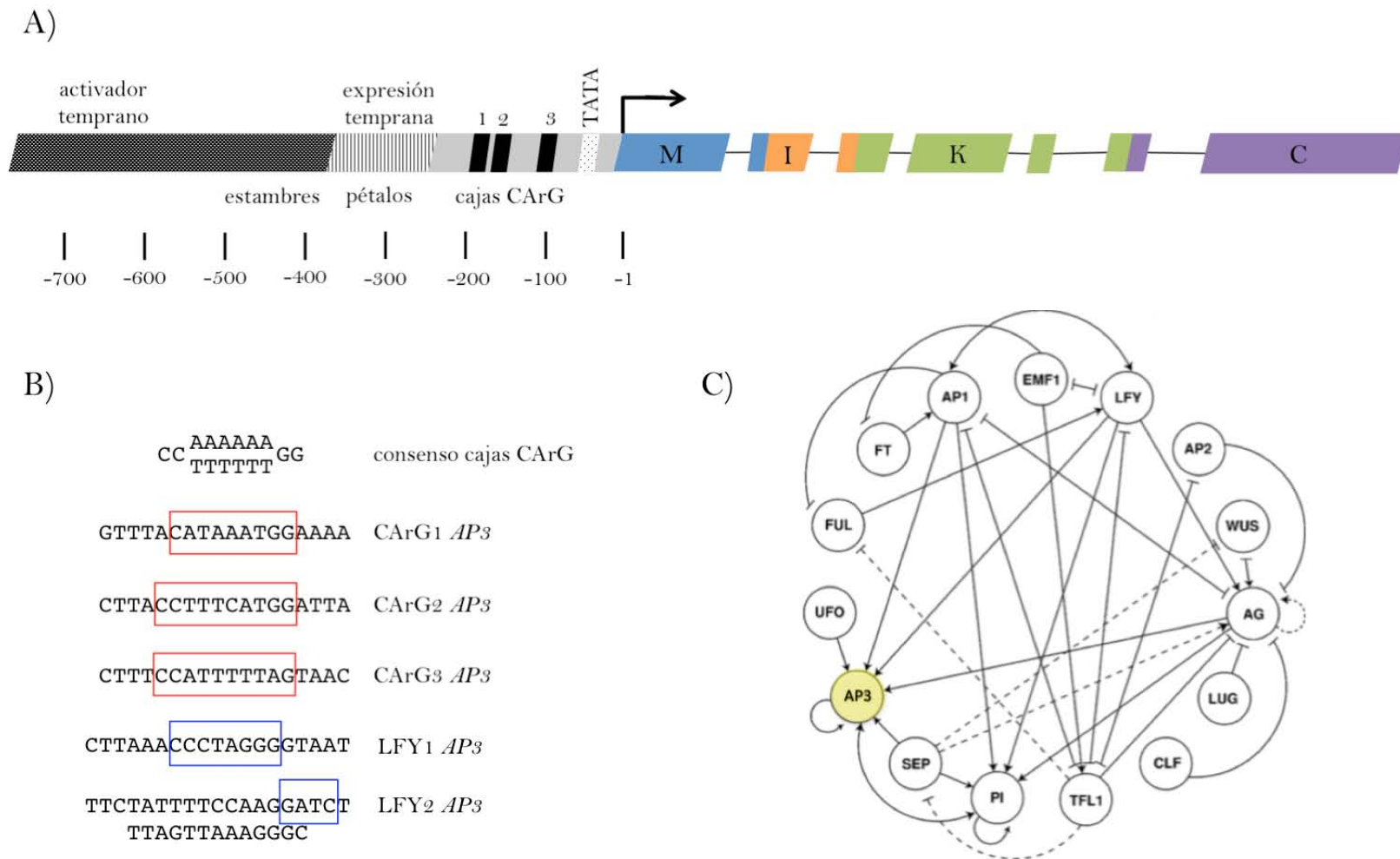


Figura 14. Estructura y análisis del gen MADS-box *APETALA3*. (A) Estructura génica de *AP3* y sus principales secuencias reguladoras. Fiel a la composición clásica de las proteínas MIKC, la región codificante de *AP3* está compuesta por siete exones de longitud variable y seis exones intermediarios. Al menos en el ortólogo de *Arabidopsis*, toda la expresión genética está controlada por motivos ubicados en el promotor. El polígono blanco con puntos representa la caja TATA y los polígonos negros simbolizan las tres cajas CARG reportadas; los demás motivos representan secuencias de regulación que dirigen la expresión de *AP3* en pétalos y estambres o que controlan la expresión temprana del gen durante los estadios 3-5 del desarrollo floral (se han identificado dos motivos de activación temprana, uno proximal y otro distal, llamados PEE y DEE, respectivamente). Por razones de espacio, la región codificante no está a escala. (B) Motivos de unión de proteínas MADS-box encontrados en el promotor de *AP3*. Se han identificado tres cajas CARG: CARG1, a la que se une un activador de la expresión temprana y tardía, CARG2, un sitio de unión para un activador específico de pétalos, y CARG3, reconocida por un represor de la expresión en etapas tempranas del desarrollo floral. También se conocen dos sitios de unión de LFY; en azul, se enmarcan sus secuencias palindrómicas. (C) La red de regulación genética para la determinación de órganos florales en *Arabidopsis*. Algunos reguladores transcripcionales como UFO, LFY, PI, AP1, AG y SEP interactúan con AP3, en amarillo, que también se autoregula. Ver explicación en el texto. Datos resumidos de Jack et al. (1992), Tilly et al. (1998), Hill et al. (1998), Lamb et al. (2002), Espinosa-Soto et al. (2004).

Por otra parte, aunque *AP3* puede autoregularse de manera positiva con ayuda de *PI*, Hill et al. (1998) probaron a través de ensayos de cambio en la corrida electroforética (EMSA, por sus siglas en inglés) que otras proteínas MADS juegan un papel importante en su regulación transcripcional, pues *AG* y *AP1* reconocen dos de las tres cajas *CArG* distribuidas en el promotor de *AP3* (*CArG1* y *CArG3*). El heterodímero de *AP3/PI* parece tener, a su vez, mayor afinidad por la *CArG3* (Hill et al. 1998).

Pero *AG* y *AP1* no son los únicos reguladores que interactúan con el promotor de *AP3* (**Figura 14c**). Distintos análisis de mutantes han argumentado que el establecimiento de la expresión inicial de este gen en *Arabidopsis* se debe a la actividad conjunta de los genes *LEAFY (LFY)* y *UNUSUAL FLORAL ORGANS (UFO)*¹² (Weigel y Meyerowitz 1993; Levin y Meyerowitz 1995; Wilkinson y Haughn 1995). Tras demostrarse su naturaleza de activador transcripcional, Percy et al. (1998) observaron que *LFY* no sólo participa en la especificación de la identidad del meristemo floral sino también en la activación de genes homeóticos que dirigen el desarrollo de la flor –como *AP3*, *AG* y *AP1*–, en cuyas regiones reguladoras reconoce la secuencia consenso (T/A)NNNNCCANTG(T/G)NNNN(T/A), donde N representa cualquier base (Chae et al. 2008; Hamès et al. 2008).

Para el caso concreto del promotor de *AP3*, Lamb et al. (2002) se dedicaron a investigar cuál de los dos elementos de activación temprana identificados por Tilly et al. (1998) y Hill et al. (1998) podría funcionar como sitio de unión a *LFY* (**Figura 14a**). A través de pruebas como EMSA fueron capaces de encontrar que *LFY* sólo se une a un fragmento de 52 pb del promotor que contiene al elemento distal DEE, y lo dividieron en dos secciones: el sitio de unión I, necesario y suficiente para la unión de *LFY*, y el sitio de unión II, que no es suficiente para la unión de *LFY* pero sí necesario, pues actúa de forma cooperativa con el sitio de unión I (Lamb et al. 2002; **Figura 14b**).

Sin embargo, *LFY* parece no poder unirse al promotor de *AP3* por sí solo y activar su expresión en los primordios de pétalos y estambres. En cambio requiere de la actividad de *UFO*, una proteína F-box cuyo dominio de expresión coincide con el de *AP3* y que actúa como cofactor al interactuar físicamente con *LFY* y estabilizar su unión al ADN (Chae et al. 2008; **Figura 14c**), proceso conservado en los ortólogos de varios linajes de plantas (Ingram et al. 1997; Taylor et al. 2001; Zhang et al. 2003; Ikeda et al. 2007).

¹² Ni *LFY* ni *UFO* son genes MADS y, por tanto, no reconocen ni se unen a cajas *CArG*.

4.3 Estudios de genética molecular en *Lacandonia* y *Triuris*

Los experimentos de Álvarez-Buylla et al. (2010b) ofrecieron datos suficientes para demostrar que la homeosis conservada en el arreglo floral de las poblaciones naturales de *L. schismatica* es resultado del desplazamiento espacial del patrón de expresión del gen *LsAP3* hacia el centro del meristemo floral desde etapas muy tempranas en el desarrollo. Sin embargo, poco se sabía sobre qué tipo de cambios eran necesarios para alterar de tal manera la red de regulación genética en control de la morfogénesis de *Lacandonia*.

Fue así como surgieron dos hipótesis alternativas: la expresión atípica de *LsAP3* es causada por cambios en *cis* o cambios en *trans* (Piñeyro-Nelson et al. 2010; Garay-Arroyo et al. 2012; ver Capítulo 1 de esta tesis).

Los primeros involucran mutaciones en las secuencias reguladoras del gen, como el promotor o los intrones, alterando así la afinidad de unión al ADN de reguladores transcripcionales (ver **Figura 3**). Estos cambios en *cis* podrían también explicar de manera parcial la existencia de homeosis en las flores bisexuales de *T. brevistylis* o, al menos, ayudar a entender la regulación de *AP3* en la familia Triuridaceae (Garay-Arroyo et al. 2012).

En sus respectivas revisiones, Piñeyro-Nelson et al. (2010) y Garay-Arroyo et al. (2012) proponen que los cambios en *trans* podrían constituir otra explicación a la inversión de órganos sexuales en flores de *Lacandonia*: si la estructura proteica de los reguladores transcripcionales de *LsAP3* se ve afectada, modificaría su afinidad de unión al ADN o su capacidad de interactuar con otros factores *trans*, cambiando así el patrón de expresión de *LsAP3* (ver **Figura 2**). Algunas proteínas candidatas de estos cambios en *trans* son LFY y UFO, necesarias para activar la expresión del *AP3* en el meristemo floral.

Como una aproximación para tratar de abordar ambas hipótesis, el Laboratorio de Genética Molecular, Epigenética, Desarrollo y Evolución de Plantas del Instituto de Ecología de la UNAM, en colaboración con la iniciativa Monocot Tree of Life, secuenció el transcriptoma de *L. schismatica* y ambos morfos (masculino y femenino) de *T. brevistylis* (Garay-Arroyo et al. 2012; Capello et al. no publicado). Estos datos demostraron la presencia de una sola copia del gen *AP3* en *L. schismatica* (*LsAP3*) y dos copias en *T. brevistylis* (*TbAP3v1* y *TbAP3v2*), cuyas secuencias pueden consultarse en el **Apéndice I**.

Pero fue hasta 2013 cuando los primeros resultados relacionados con las bases causales de la morfogénesis floral en *Lacandonia* y *Triuris* se dieron a conocer.

Quizá una de las aportaciones más extensas al tema fue la de Piñeyro Nelson (2013), quien presentó datos preliminares sobre la regulación en *cis* de *LsAP3*. Después de clonar 1.1 kb de su promotor, Piñeyro Nelson (2013) comparó la secuencia con la de otros promotores de ortólogos de *AP3* en dicotiledóneas (*Arabidopsis thaliana*, tomate, papa, petunia) y monocotiledóneas (azucena) y reconoció una divergencia notable en la estructura de las tres cajas CArG, aunque otros motivos –como las cajas TATA, GATA, MYB y otras relacionadas con regulación hormonal y respuesta a estrés osmótico– sí estaban presentes. Además, con el fin de analizar la función del promotor de *LsAP3*, generó líneas reporteras con la construcción pro*LsAP3*::GUS en plantas de *A. thaliana* y observó que la expresión observada, si bien era localizable en estambres de etapas tardías del desarrollo floral, no recuperaba por completo el patrón espacio-temporal de *LsAP3* reportado en estudios previos (Álvarez-Buylla et al. 2010b; Piñeyro-Nelson 2013).

En un intento por comenzar a explorar qué cambios en *trans* serían necesarios para explicar la alteración del arreglo floral en triuridáceas, Sánchez Lara (2013) clonó la secuencia de ADN y caracterizó la secuencia proteica de los ortólogos de LFY y UFO tanto en *L. schismatica* como en *T. brevistylis*: LsLFY y LsUFO; TbLFYv1, TbLFYv2 y TbUFO. Sánchez Lara (2013) demostró que estos genes comparten la estructura característica de los genes MADS tipo II (MIKC, ver **Figura 8**) y, como resultado de análisis filogenéticos, observó que estos genes son ortólogos de los correspondientes en *Arabidopsis*. Además, reportó que la mayoría de las sustituciones en los ortólogos de LFY fueron encontradas en el dominio C-terminal, lo cual podría provocar un cambio en la actividad bioquímica de este factor transcripcional (Sánchez Lara 2013). Finalmente, en cuanto a LsUFO, Sánchez Lara (2013) identificó mutaciones novedosas en el dominio F que podrían alterar su capacidad de interacción con proteínas blanco, como reguladores negativos de *LsAP3*, con el fin de marcarlas para su degradación por el proteosoma.

En su tesis de licenciatura Dolores Fuentes (2013) clonó, ensambló e identificó los genes *TbAP3* (dos parálogos), *TbPI* y *TbAG*, incluida la región 3' UTR de cada uno. Con la ayuda de alineamientos a nivel de aminoácidos, también descubrió que el motivo K1 – parte del dominio K– y el dominio C-terminal de las proteínas TbAP3v1, TbAP3v2 y

LsAP3 exhiben sustituciones respecto a los ortólogos de otras monocotiledóneas (Dolores Fuentes 2013).

Esto sugiere que la capacidad de heterodimerización entre AP3 y PI podría ser distinta en *L. schismatica* y *T. brevistylis*. Además, Dolores Fuentes (2013) utilizó matrices de identidad de aminoácidos y análisis filogenéticos para estudiar la relación evolutiva entre los ortólogos de AP3 de ambas triuridáceas y encontró que *TbAP3v1* y *LsAP3* comparten un mayor grado de identidad entre sí que ambos parálogos de *TbAP3*. Ella propone que, dado su mayor parecido a la secuencia de *LsAP3*, *TbAP3v1* podría estar involucrado en la determinación de estambres mientras que *TbAP3v2* sería responsable de otorgar identidad a los tépalos de *Triuris* –sin embargo, los genes de función A en *Lacandonia* podrían ser suficientes para el desarrollo de tépalos sepaloideos–. Otra opción recae en la posibilidad de que *TbAP3v2* sea una de las copias de AP3 presentes en las triuridáceas y que por alguna razón se perdió en el genoma de *Lacandonia* (Dolores Fuentes 2013).

La presente tesis se concibió de manera paralela a los estudios mencionados, dirigidos a investigar los cambios necesarios para desplazar la expresión espacial de *LsAP3*. Durante el desarrollo de este trabajo, me he enfocado a buscar más a detalle posibles mutaciones en *cis* ubicadas en las secuencias promotoras de este gen y sus parálogos *TbAP3v1* y *TbAP3v2*.

JUSTIFICACIÓN

La identificación de sitios de unión de factores de transcripción significa una parte importante para definir las redes de regulación genética en control de diversos aspectos del desarrollo de los organismos. Sin estudiar la contribución de la variación en motivos de regulación, no es posible comprender el contexto completo de la función biológica de dichas redes.

El análisis bioinformático de secuencias promotoras ha demostrado ser una herramienta útil para identificar posibles elementos *cis* de distintos genes, incluyendo a *AP3* y sus ortólogos. Un estudio más profundo sobre la regulación transcripcional de *LsAP3* es necesario para comprender a fondo su regulación anómala, que puede explicar la rara heterotopía floral predominante en las flores de *Lacandonia schismatica*.

Además, una aproximación comparativa con los promotores de *TbAP3v1* y *TbAP3v2* de *Triuris brevistylis*, taxón hermano de *Lacandonia*, resulta fundamental para encontrar motivos compartidos y divergentes entre ambas especies de triuridáceas mexicanas. Esto resulta especialmente interesante si se toman en cuenta las investigaciones de Vergara-Silva et al. (2003) y Espinosa-Matías et al. (2012), quienes demuestran la existencia de flores atípicas de *Triuris brevistylis* con cierto grado de homeosis en poblaciones naturales, fenómeno que bien podría deberse, como en el caso de *Lacandonia*, a la alteración de la expresión espacial de *AP3*.

Finalmente, los datos obtenidos del transcriptoma de *Triuris brevistylis* demuestran la presencia de dos copias de *AP3* en ambos sexos. Esto, sumado al estudio de Espinosa-Matías et al. (2012), donde no se reporta ningún tipo de aborto de órganos sexuales, refuerza el interés por hacer un análisis preliminar de la regulación en *cis* de estos tres genes mediante análisis bioinformáticos de *phylogenetic footprinting*.

OBJETIVOS

Estudiar la regulación genética en *cis* de *LsAP3*, *TbAP3v1* y *TbAP3v2* mediante la búsqueda de posibles motivos de regulación en sus promotores.

- Clonar y secuenciar los promotores de *TbAP3v1* y *TbAP3v2*.
- Comparar las secuencias regulatorias de *LsAP3*, *TbAP3v1* y *TbAP3v2* y diferenciar sus motivos compartidos y divergentes.
- Evaluar la divergencia de las secuencias promotoras de estos genes respecto a las de otras angiospermas.

Hacer un análisis conceptual sobre *Lacandonia* y *Triuris* enmarcado dentro de la biología evolutiva del desarrollo.

HIPÓTESIS

Las secuencias promotoras de *LsAP3* y ambas copias de *TbAP3* comparten motivos de regulación exclusivos de *Lacandonia schismatica* y *Triuris brevistylis* que están involucrados en los mecanismos moleculares responsables de los eventos homeóticos que ocurren durante su desarrollo floral.

MATERIALES Y MÉTODOS

Material biológico

Se recolectaron flores femeninas de *Triuris brevistylis* en la localidad de Nahá, Ocuisingo, en el sitio Jaguar. La recolección se realizó el 15 de septiembre de 2011. Inmediatamente después de la colecta, el material se congeló en nitrógeno líquido.

Extracción de ADN genómico

Utilicé un protocolo de extracción modificado a partir de la versión del Centro Internacional de Maíz y Trigo (CIMMYT 2006), basada a su vez en el método de Saghai-Marooft et al. (1984).

Con el fin de evitar la contaminación y la degradación del tejido, todo el material de laboratorio fue esterilizado, cubierto con aluminio y congelado en nitrógeno líquido. En un tubo de 2ml, se puso el equivalente a 0.25ml de tejido molido de individuos de *Triuris brevistylis*. Posteriormente, se agregó 1ml de buffer de extracción (Tris-7.5 1M, NaCl 5M, EDTA-8.0 0.5 M, CTAB 1%) previamente calentado a 65°C y 10 µl de β-mercaptoetanol (**Tabla 1**).

El CTAB es un detergente catiónico que es muy usado para precipitar ADN genómico de organismos que producen una alta cantidad de polisacáridos, como las plantas, porque puede formar complejos con éstos en soluciones de alta concentración iónica. Además, los ácidos nucleicos son solubles en CTAB pero precipitan cuando la concentración iónica disminuye (Murray y Thompson 1980; Valadez y Kahl 2000). Durante la extracción es necesario proteger al ADN de nucleasas endógenas que lo puedan degradar; para este fin, el protocolo requiere del uso de EDTA, un compuesto quelante que inactiva estas enzimas al unirse a los cationes metálicos que funcionan como sus cofactores (Velasco Mosquera 2005). Por su parte, el β-mercaptoetanol es un agente reductor que inhibe la oxidación de los compuestos fenólicos de la planta (Velasco Mosquera 2005).

Después de agregar el buffer de extracción y el β-mercaptoetanol, se incubaron las muestras a 65°C en baño de agua con agitación durante una hora y media. Posteriormente,

se añadieron 500 µl de una solución de cloroformo:octanol (24:1) para separar el ADN de las proteínas. La fase acuosa se transfirió a otro tubo, donde había 0.7 vol. de isopropanol frío (-20°C), un alcohol selectivo en la precipitación de ADN en presencia de ARN (Velasco Mosquera 2005). Los tubos se dejaron a -20°C durante 16-24 horas. Tras centrifugarlos y decantar el sobrenadante, la precipitación y el lavado prosiguieron con la adición de 1 ml de una primera solución compuesta por EtOH 76% y NaOAc 0.2 M. Realicé otro ciclo de centrifugación y añadí 500 µl de una segunda solución compuesta por EtOH 76% y NH₄OAc 10 mM. Finalmente, después de centrifugar y decantar nuevamente, se agregó EtOH 70% frío (4°C) para terminar el lavado. Resuspendí la pastilla de ADN resultante en agua destilada.

Tabla 1. Solución amortiguadora con CTAB para extracción de ADN genómico (RXN = reacción) ¹

		1RXN	5RXN	10RXN	20RXN	50RXN
Solución	[Final]	10ml	50ml	100ml	200ml	500ml
dH ₂ O		6.5 ml	32.5 ml	65.0 ml	130.0 ml	325.0 ml
1M Tris-7.5	100 mM	1.0 ml	5.0 ml	10.0 ml	20.0 ml	50.0 ml
5M NaCl	700 mM	1.4 ml	7.0 ml	14.0 ml	28.0 ml	70.0 ml
0.5M EDTA-8.0	50 mM	1.0 ml	5.0 ml	10.0 ml	20.0 ml	50.0 ml
CTAB ²	1%	0.1 g	0.5 g	1.0 g	2.0 g	5.0 g
14 M BME ³	140 mM	0.1 ml	0.5 ml	1.0 ml	2.0 ml	5.0 ml

¹ Utilizar la solución recién hecha; calentar a 65°C antes de agregar CTAB y BME

² CTAB = bromuro mixto de alquiltrimetil-amonio

³ Agregar BME (β-mercaptoetanol) bajo campana extractora

GenomeWalker™ y construcción de librerías

Para clonar los promotores de TbAP3v1 y TbAP3v2, usé el sistema GenomeWalker™ (Clontech Laboratories, Inc.), basado en la técnica de *genome walking* (“paseo genómico”) que sirve para identificar secuencias no conocidas de ADN genómico adyacentes a alguna

secuencia conocida y que, en eucariontes, ha sido utilizada principalmente para encontrar regiones regulatorias (Leoni et al. 2011).

Aunque la estrategia experimental del *genome walking* sea la reacción en cadena de la polimerasa (PCR, por sus siglas en inglés), existe un amplio abanico de estrategias cuyas modificaciones intentan mejorar la insuficiente especificidad de sus semejantes anteriores. El sistema de Clontech es uno de los más usados, y está basado en el método descrito por Siebert et al. (1995) que requiere de la previa digestión del ADN genómico, seguida de una ligación con oligonucleótidos adaptadores. También está diseñado para prevenir la amplificación exponencial de secuencias no específicas gracias a un efecto ‘supresor’ durante la PCR (Siebert et al. 1995).

Siguiendo los pasos del manual del GenomeWalker™ (protocolo PT3042-1, versión PR742239), construí cuatro librerías distintas del ADN genómico de *T. brevistylis* previamente purificado por el método CTAB modificado (CIMMYT 2006), y digerido por las enzimas de restricción DraI, EcoRV, PvuII y StuI (para cada digestión, se usaron 2.5 µg de ADN y 80 unidades de enzima; la incubación fue de dos horas a 37°C). De manera adicional, preparé una digestión control de ADN humano con DraI. La ligación de la secuencia adaptadora se hizo de acuerdo al protocolo, con la enzima ligasa T4 (se ligan cuatro µl de ADN digerido con 1.9 µl del adaptador y tres unidades de enzima; la incubación fue de 18 horas a 16°C). Todas las endonucleasas seleccionadas dejan extremos romos, facilitando la ligación del adaptador –que tiene un extremo romo– a ambos lados de los fragmentos de ADN genómico generados durante la digestión.

El GenomeWalker™ implica la combinación de distintas variantes de la PCR, principalmente la PCR a larga distancia (LD-PCR), la PCR touchdown y la PCR anidada. Para favorecer la LD-PCR, utilicé el Advantage 2 PCR Kit (Clontech), que incluye dos polimerasas termoestables que aumentan el alcance y la precisión de la amplificación.

Diseño de oligonucleótidos y amplificación de los promotores de *TbAP3*

A partir de la secuencia codificante conocida a partir del transcriptoma de *TbAP3v1* y *TbAP3v2*, diseñé oligos para avanzar hacia la región 5' (**Tabla 2; Figura 15**). En este caso, las secuencias promotoras pertenecen a una misma familia génica, la de *APETALA3*, cuya región codificante se encuentra altamente conservada (ver **Antecedentes**). Por ello,

adopté la estrategia que propusieron Leoni et al. (2009) y escogí regiones conservadas para diseñar oligos comunes y específicos.

Se usaron los programas de ciclaje sugeridos en el manual del GenomeWalker™ para la amplificación por PCR de ocho tubos distintos (**Tabla 3**). La **Tabla 4** muestra los parámetros y condiciones de amplificación sugeridas por el Genome Walker™. El tamaño de los fragmentos obtenidos se observó mediante electroforesis en geles de agarosa 1.5% junto con marcadores de peso molecular (**Figura 16**).

>TbAP3v1 (690nt)

ATGGGCAGGGGAAGATCGAGATCAAGAAGATCGAGAACCCACCAACCGGCAGGTGACCTATTCCAAGCGCCGG
CTGGGGCTGATGAAGAAGCCAAGGAGCTCACCGTCTCTGCGACGCCAGGTCTCCATCGTCATGGTCTCCTCCA
CCAACAAGTTCGCCGAGTACTGCAGCCCTCCACCGATCTGAAGACCATCTACGACCGCTACCAAGAA**GCCTCCGG**
GATCAACCTCTGGACGGCTCAATA**CGAGAAGATGCAGAACCGCGCTGAGCC**ATCTGAAGGAGATCAACCGCAACCTG
AGGAGAGAGATCAGGCAGAGGATGGGGGAGGAGCTGGAGGAGATGGAGGCGCAGGAGCTGCGCGGTCTTGAGCAAA
CCTTGACGAGTCACTGAAGCTCGTCCGCCAGAGGAAATATCATGTGATCACCACACAGACAGATACTTACAAGAA
GAAGATAAAGAATTCTCATGAAGCCTACAAGACTTTGCTGCGAGAGCTGGAAATCCGAGATGGGGGAGGCGATATG
GGTTATGTTATCGATGAGGAGGCTGTACCCGTTATGATGGAGGTACGCTGGCGATGGGGAACGGAGGTAGCCCTC
ACATGTTACATTCCGCGTTCAGCCCAGCCAGCCCAATCTCCACCTCATGGGCAATGGTTTCCATGATCTTCGTCT
CGCC**TAG**

>TbAP3v2 (666nt)

ATGGGCAGGGGAAGATCGAGATCAAGAAGATCGAGAACCCGACCAGCCGCCAGGTGACCTACTCGAAGCGCCGC
CTCGGCCTCATGAAGAAAGCCAAGGAGCTCACCGTCTCTGCGACGCCAGGTGGGCGTCATCATGATCTCCAGCA
CCAAGAAGCTCGCGAATACTGCAGCCCTCCACCAACCTCAAGTCGATGTACGACCGCTACCAGGAA**CCAACGGG**
CATCTATCTCTGGGGCTCCCAATA**CGAGAAGATGCAGAACCGCGCTGAGCC**ACCTGAAGGAGAACAACATGAATCTC
AGCCATGAAATCAAGCTGAGGATGGGGGAGGAGCTCGAGGGGATGGACGTGCAGGAATTGCGTGGTCTTGAGCAAA
CAATGGATGAGGCACTCAAGATCGTGCGCCAGAGGAAGTATCATGTGATCACTACACAGACAGAGACATACAAGAA
GAAGGTTAAGAACTCCAACGAGGCGTACAAGAACCTCGTCCGGGAGCTGGAGATGCGCAACGAACATGTCGACATT
GTCTACGCAGACGTGATTGGTGCGGGCGGCCCTATGCTGGGCCGGTGCCGGTTCGAGATGCTCACCTTCCTTGTGG
CGCCCGGGCCACCCTTCGACAATGAACTATGCTATTACGACCTCCGCTGGCC**TAA**

Figura 15. Ubicación física de los oligonucleótidos usados para la amplificación de los dos promotores de ambos parálogos de *TbAP3*. El oligo GSP1 compartido (en naranja) se encuentra sobre los dominios I y K en una zona altamente conservada. Los oligos GSP2 específicos (en verde para TbAP3v1 y en azul para TbAP3v2; ver Tabla 2) están diseñados en una zona del dominio I variable entre ambos parálogos. Secuencia obtenida a partir del transcriptoma de *T. brevistylis*.

Tabla 2. Oligos usados para amplificar los promotores de *TbAP3*

Gen	Nombre de oligo	Secuencia (5'-3')	Dirección	Tamaño (nt)
<i>TbAP3</i> *	TbAP3Pro2-GSP1	GGCTCAGCGCGTTCTGCATCTTCTCG	Antisentido	26
	AP1	GTAATACGACTCACTATAGGGC	Sentido	22
	AP2	ACTATAGGGCACGCGTGGT	Sentido	19
<i>TbAP3v1</i>	TbAP3.1Pro-GSP2	AGCCGTCCAGAGGTTGATCCCGGAGGC	Antisentido	27
<i>TbAP3v2</i>	TbAP3.2Pro-GSP2	GGAGCCCCAGAGATAGATGCCCGTTGC	Antisentido	27

* Los oligos TbAP3Pro2-GSP1, AP1 y AP2 fueron usados para realizar la primera PCR de ambas copias de *TbAP3*. TbAP3.1Pro-GSP2 y TbAP3.2Pro-GSP2 se usaron en la PCR anidada de cada gen individual.

Tabla 3. Librerías construidas a partir del protocolo del GenomeWalker™

Librería de ADN	Tubo n°	Oligos PCR primaria	Tubo n°	Oligos PCR anidada
DraI	1A	GSP1 y AP1 ^a	1B	GSP2 y AP2 ^b
EcoRV	2A	“	2B	“
PvuII	3A	“	3B	“
StuI	4A	“	4B	“
1^{er} control negativo				
H ₂ O	5A	“	5B	“
1^{er} control positivo				
Librería de control ^c	6A	PCP1 y AP1 ^a	6B	PCP2 y AP2 ^b
2^{do} control negativo				
H ₂ O	7A	“	7B	“
2^{do} control positivo				
Librería de control pre-construida ^d	8A	“	8B	“

^a Oligo usado sólo en la PCR primaria

^b Oligo usado sólo en la PCR anidada

^c Yo construí esta librería con ADN genómico humano provisto por el kit

^d Esta librería pre-construida ya está incluida en el kit

Tabla 4. Parámetros para la amplificación de fragmentos			
PCR primaria		PCR anidada	
1 ciclo:	94°C 3 minutos (HOT START)	1 ciclo:	94°C 3 minutos (HOT START)
7 ciclos:	94°C 25 segundos	5 ciclos:	94°C 25 segundos
	72°C 3 minutos		72°C 3 minutos
32 ciclos:	94°C 25 segundos	20 ciclos:	94°C 25 segundos
	67°C 3 minutos		67°C 3 minutos
1 ciclo:	67°C 7 minutos	1 ciclo:	67°C 7 minutos

Purificación, ligación, clonación y secuenciación de amplicones

Una vez finalizadas las PCR anidadas de ambas copias de *TbAP3*, identifiqué los insertos aislados que fueran mayores a 300pb. Se cortaron y purificaron dos bandas de librerías distintas para cada parólogo (una banda de ~750pb con DraI y otra de ~500pb con EcoRV para *TbAP3v1*; una banda de ~750pb con DraI y otra de ~600pb con PvuII para *TbAP3v2*). Para purificar las bandas del gel de agarosa, las coloqué en tubos con malla y un orificio inferior por donde pueda salir el líquido, donde está soluble el ADN, que son retenidos en otros tubos más grandes. De esta manera, la agarosa se queda atrapada en la malla. Centrifugué a 8000 rpm por 5' y se añadieron 2.5 µl de glucógeno a 25mg/ml, que favorece la precipitación de ADN.

Después a cada tubo se agregaron 3.5 vol. De AcNH₄⁺/EtOH (1:6) para degradar la fracción proteica asociada al ADN y precipitar el mismo, pues es insoluble en alcohol. Los tubos se enfriaron a -20°C durante 20' y, tras una centrifugación, lavé la pastilla de ADN con 500µl EtOH al 70% y la resuspendí en agua destilada.

Posteriormente, se llevó a cabo la ligación de amplicones a vectores de clonación con el sistema pGEM®-T Easy (Promega Corporation), incubando las reacciones durante 24 horas a 4°C. Para continuar con la transformación de bacterias y seleccionar las colonias transformantes, se prepararon cajas con LB/Ampicilina/IPTG/X-Gal. Las bacterias

usadas fueron de la cepa TOP10, y las transformé mediante choque térmico a 42°C durante 45-50". Las cajas fueron incubadas por 24 horas a 37°C y, después, realicé una purificación de plásmidos en pequeña escala (o miniprep) a partir de un protocolo modificado del laboratorio.

El ADN plasmídico se mandó a secuenciar a la Unidad de Síntesis y Secuenciación de DNA del Instituto de Biotecnología de la UNAM, en Cuernavaca, a cargo del Dr. Paul Gaytán y el M. en C. Jorge Arturo Yañez.

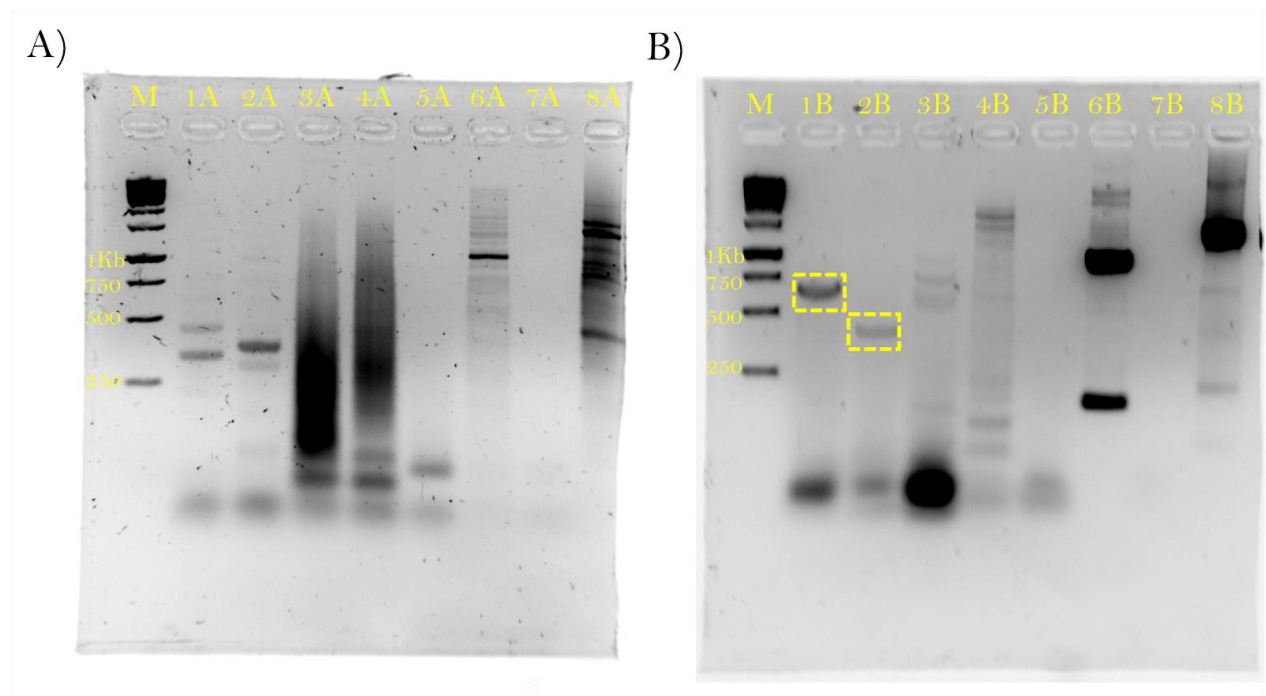


Figura 16. Resultados de la PCR primaria y la PCR anidada del GenomeWalker™. (A) En general, las PCR primarias produjeron múltiples fragmentos de tamaños distintos en cada librería o barridos de ADN, como los que se observan en los carriles 3A y 4A. Cada carril del gel está etiquetado con los nombres asignados en la Tabla 3. (B) La PCR anidada, sin embargo, aumenta la especificidad de amplificación. Las bandas aisladas de al menos 400pb, como las que se observan en los carriles de las librerías 1B y 2B, pueden ser cortadas y purificadas. Cada carril del gel está etiquetado con los nombres asignados en la Tabla 3. M representa el marcador molecular de 1Kb.

Análisis bioinformáticos

Para corroborar que las secuencias obtenidas forman parte de la región promotora de los parálogos de *TbAP3*, se alineó el cDNA de cada copia –conseguido a partir del transcriptoma de *Triuris brevistylis* (ver **Antecedentes**)– y el ADN genómico clonado con

los oligos presentados en la **Tabla 2**. El alineamiento pareado fue realizado con el programa MultAlin (Corpet 1988) usando parámetros de identidad 1-0.

Las secuencias consideradas en esta tesis fueron elegidas de acuerdo a dos criterios principales: su disponibilidad en bases de datos y la existencia de información experimental sobre sus patrones de expresión. La identificación de los genes ortólogos de *Apetala3* fue realizada a partir de búsquedas en las bases de datos del National Center for Biotechnology Information (acceso de GenBank: AB158291.1, AB180962.1, AY496050.1, CM000782.3, AP005463.3, U30729.1, X62810.1, X67511.1, AF057708.2, HM192809.1) y artículos experimentales donde se demuestra su conservación funcional en algunas monocotiledóneas (Tsuchiya et al. no publicado; Ambrose et al. 2000; Nagasawa et al. 2003; Tsai et al. 2004; Ito et al. 2005) y dicotiledóneas (Rui y Zhen no publicado, Schwarz-Sommer et al. 1992; Garcia-Maroto et al. 1993, Irish y Yamamoto 1995; Sheppard et al. 2000).

Las regiones promotoras (de la posición -1,000 a la 0, en relación al sitio de inicio de traducción¹³) de estos genes ortólogos, a excepción de los de maíz y arroz, fue extraída de manera directa en el NCBI. Para encontrar los promotores de *SILKY1* y *SPW1/OsMADS16*, utilicé el Genome Browser del Maize Genetics and Genomics Database y del Rice Genome Annotation Project, respectivamente. Una vez localizado el cromosoma donde se ubican estos genes (cromosoma 6, para ambos), rastreeé su locus y descargué 1Kb río arriba de su sitio de inicio de traducción. El **Apéndice 2** contiene las secuencias reguladoras usadas en esta tesis. Todos los promotores, incluidos los de *LsAP3* y *TbAP3*, fueron sometidos a una búsqueda restringida a secuencias de angiospermas en BLAST del NCBI para comprobar que formaran parte de su gen correspondiente y asegurarme de que su región intergénica no fuera más corta de 1Kb.

Para obtener estadísticas sobre el contenido de GC de cada región promotora, utilicé el DNA Sequence Toolbox del software Genomatix. La presencia de islas CpG se cuantificó con la herramienta Newcpgreport del software EMBOSS con parámetros de longitud mínima de 200pb, porcentaje mínimo de 50% y tamaño de espaciamiento de 100pb.

¹³ Idealmente, las regiones reguladoras 5' deberían comenzar en el sitio de inicio de la transcripción, no de la traducción. Sin embargo, identificarlo en el caso de todos los genes demostró ser difícil. Otros estudios, como el de De Bodt et al. (2006), se han encontrado con obstáculos similares y han optado por usar como referencia el sitio de inicio de la traducción.

Con el objetivo de extraer patrones comunes de elementos *cis*, escogí únicamente las regiones promotoras de monocotiledóneas y, para comparación con dicotiledóneas, la de *A. thaliana*. Las demás secuencias no fueron incluidas en el análisis, pues se supuso que su regulación genética ha divergido demasiado. Después usé la herramienta FrameWorker del software Genomatix, diseñada para el análisis comparativo de regiones promotoras. Este programa no sólo sirve para la búsqueda de elementos compartidos entre secuencias reguladoras de distintas especies sino que también recupera patrones o modelos, es decir, todos los sitios de unión que se distribuyen en un mismo orden y a una cierta distancia dentro de al menos dos secuencias. El análisis se llevó a cabo con los grupos de matrices para plantas, una similitud core de 0.75 y una similitud de matrices optimizada.

FrameWorker trabaja en conjunto con la herramienta MatInspector (Cartharius et al. 2005), que utiliza una gran colección de elementos *cis* reportados en la literatura y los localiza en secuencias de ADN. Los motivos reguladores similares son agrupados en familias de matrices. Todos los elementos encontrados con FrameWorker enseñan un valor de similitud a estas matrices. En esta tesis sólo se tomaron en cuenta aquellos motivos con similitudes superiores a 0.80.

Los motivos encontrados fueron después buscados en las bases de datos PLACE (Higo et al. 1999) y PlantCARE (Lescot et al. 2002). Fue posible identificar otros sitios de unión a factores de transcripción, como los de LFY, con alineamientos realizados en DIALIGN (Morgenstern et al. 1998), un programa que busca similitudes locales entre distintas secuencias.

El análisis filogenético se realizó con las secuencias de aminoácidos de TbAP3v1 (Dolores Fuentes 2013), TbAP3v2 (Dolores Fuentes 2013), LsAP3 (GQ214161.1), SPW1 (AF424549.1), SILKY1 (NM_001111481.1), AoDEF (AB094964.1), LrDEF (AB071378.1), PeMADS5 (AY378148.1) y AP3 (D21125.1). El alineamiento se realizó con el programa MUSCLE (Edgar 2004) y se curó en GBlocks (Talavera y Castresana 2007). Los programas PhyML (Guindon et al. 2010) y TreeDyn (Chevenet et al. 2006) fueron usados para construir el árbol y visualizarlo, respectivamente.

RESULTADOS Y DISCUSIÓN

Promotor parcial y primer intrón de *TbAP3v1* y *TbAP3v2*

Las secuencias parciales de la región reguladora 5' que cloné pertenecen a su respectivo parólogo de *TbAP3*, como lo demuestran los alineamientos pareados entre el cDNA y el ADN genómico de cada copia (ver **Figura 17** para *TbAP3v1*; ver **Figura 18** para *TbAP3v2*). Estos también sugieren la clonación exitosa del primer intrón de cada parólogo, de una longitud promedio de 74 pb, que interrumpe la secuencia del dominio I. Las secuencias completas de *TbAP3v1* y *TbAP3v2* pueden encontrarse en el **Apéndice 3**.

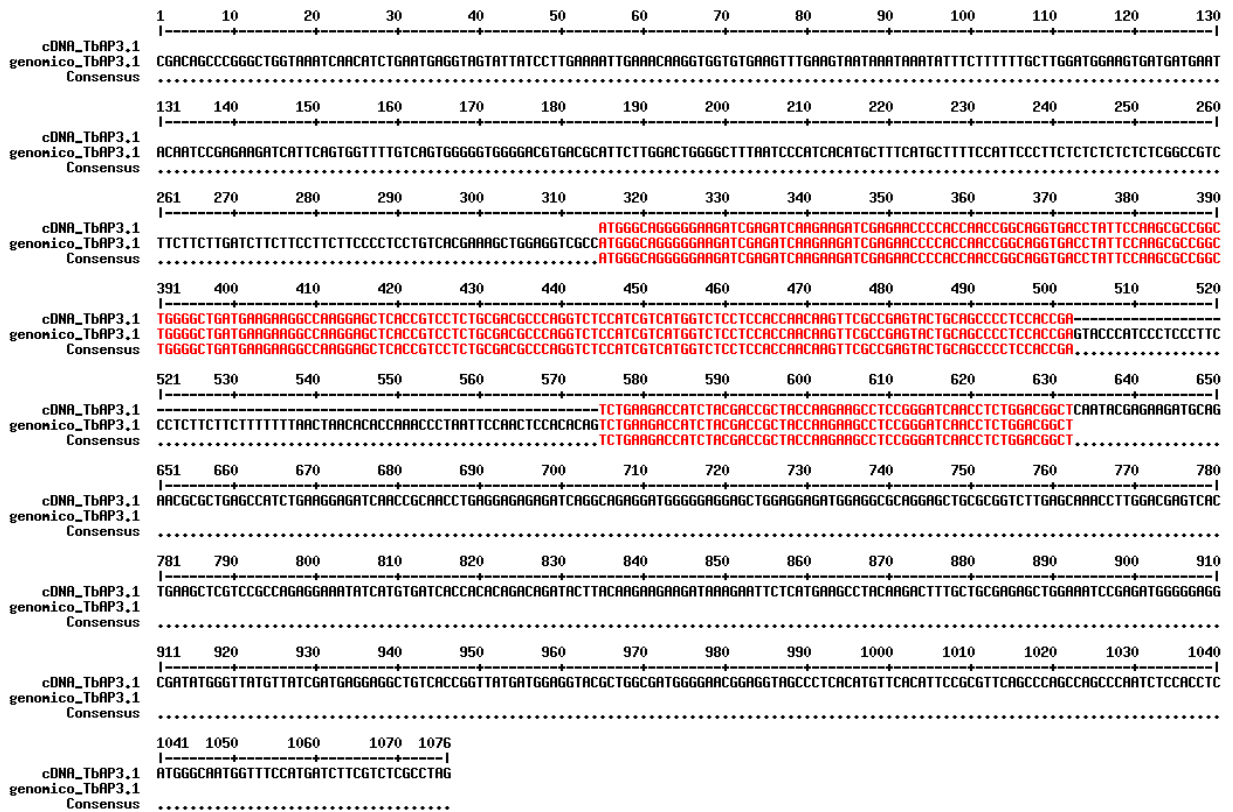


Figura 17. Regiones reguladoras de *TbAP3v1*. El alineamiento pareado entre el cDNA y el ADN genómico de *TbAP3v1* muestra la conservación de los dominios M-I, así como la existencia de un primer intrón de 72 pb que interrumpe el dominio I. En esta tesis sólo fui capaz de clonar 297 pb de la región promotora de *TbAP3v1*. Alineamiento realizado en el programa MultAlin.

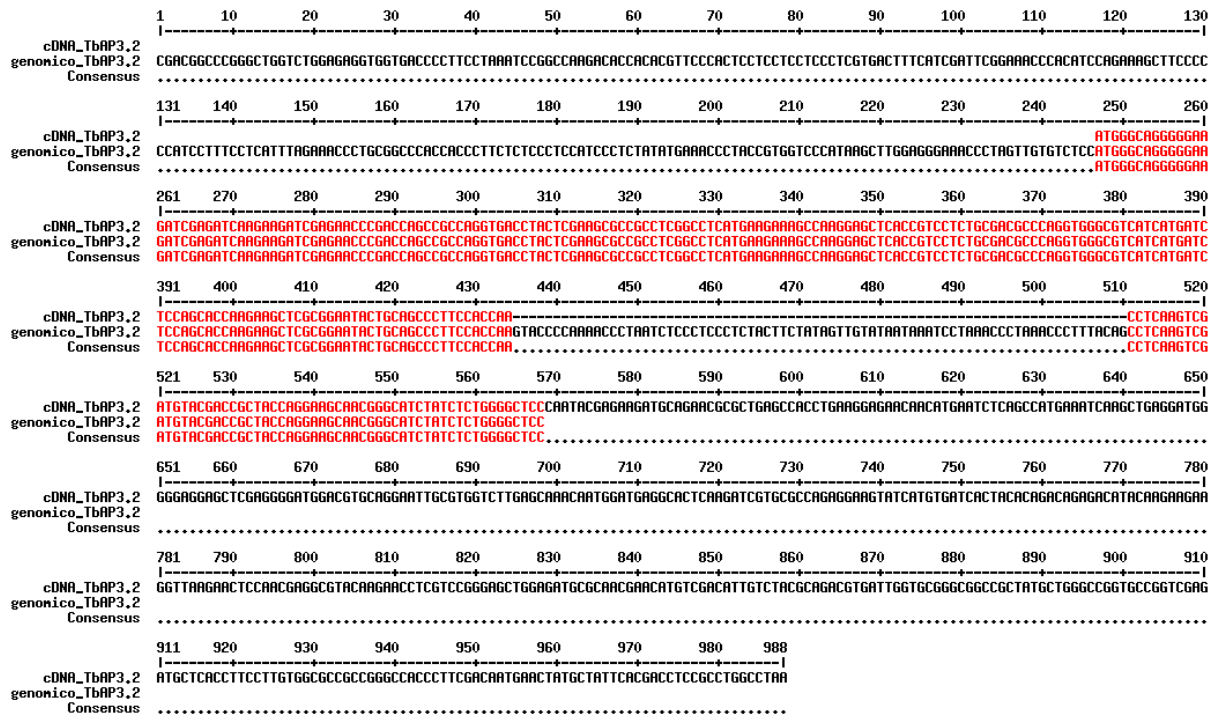


Figura 18. Regiones reguladoras de *TbAP3v2*. El alineamiento pareado entre el cDNA y el ADN genómico de *TbAP3v2* muestra la conservación de los dominios M-I, así como la existencia de un primer intrón de 76 pb que interrumpe el dominio I. En esta tesis sólo fui capaz de clonar 229 pb de la región promotora de *TbAP3v2*. Alineamiento realizado en el programa MultAlin.

A pesar de haber modificado los parámetros de ciclaje, diseñado varios oligos ubicados río arriba del ATG de ambos parálogos de *TbAP3* y secuenciado múltiples veces, no me fue posible amplificar más de 297 pb del promotor de *TbAP3v1* ni más de 229 pb del promotor de *TbAP3v2*. Esto podría influir en el número de motivos reguladores que encuentre para ambas copias, sobre todo suponiendo que la regulación de *TbAP3*, como en *Arabidopsis*, se dirige de manera primordial desde el promotor (Jack et al. 1992; Tilly et al. 1998; Hill et al. 1998). Sin embargo, es también cierto que sólo 469 pb del promotor de *AP3* son capaces de recuperar un patrón de expresión silvestre en esta planta modelo (Tilly et al. 1998), lo cual sugiere la presencia de al menos unos pocos sitios de unión en los promotores de *TbAP3* capaces de controlar algunos aspectos de su expresión.

Variación en las regiones promotoras de *APETALA3*

A pesar de su importancia en las investigaciones modernas de evo-devo de plantas, es muy poco lo que sabemos sobre las características básicas de los promotores de genes de

angiospermas (de Bruijn et al. 2012). El contenido GC es un parámetro general usado en estudios comparativos para revelar patrones sobre la composición de genomas, y ha sido incluido en algunos análisis de promotores de plantas (Koch et al. 2001; Liu et al. 2013). En esta tesis analicé la región promotora de los dos principales linajes de angiospermas, monocotiledóneas (*TbAP3* de *Triuris brevistylis*, *LsAP3* de *Lacandonia schismatica*, *LRDEF* de *Lilium regale*, *AODEF* de *Asparagus officinalis*, *MADS5* de *Phalaenopsis equestris*, *SILKY1* de *Zea mays*, *SPW1* de *Oriza sativa*) y dicotiledóneas (*AP3* de *Arabidopsis thaliana*, *DEFA* de *Antirrhinum majus*, *DEF4* de *Solanum tuberosum*, *AP3-like* de *Vitis x lambruscana*, *PTD* de *Populus trichocarpa*), con el fin de detectar un patrones similares o divergentes entre ambos.

Mis resultados muestran que el contenido GC de todos los promotores analizados varía de 27.72% a 59.12%, con un promedio de 40.97%. El contenido GC de los promotores de dicotiledóneas varía de 27.2% en el árbol *P. trichocarpa* a 39.78% en la vid *Vitis x lambruscana*, con un promedio de 32.74%, mientras que el contenido GC de los promotores de monocotiledóneas varía de 39.68% en el espárrago *A. officinalis* a 59.12% en el maíz *Z. mays*, con un promedio de 46.11%. Estos datos concuerdan con las observaciones de Liu et al. (2013), quienes reportaron que el contenido GC en promotores de monocotiledóneas es mayor al de promotores de dicotiledóneas. Dado que las regiones AT mutan más frecuentemente, es razonable pensar que los genes de plantas dicotiledóneas cuya regulación se encuentra dirigida desde el promotor se regulan de manera más compleja.

Otro parámetro que utilicé para evaluar la variación en las regiones promotoras de los ortólogos de *AP3* fue la presencia de islas CpG, en general relacionadas a los promotores de genes codificantes en otros organismos, como los mamíferos, y cuya función parece radicar en el control de la transcripción al constituir secuencias blanco de modificaciones epigenéticas (Bird 2002). Pero de todos los promotores sólo encontré una isla CpG de 403pb en el gen *SILKY1* del maíz (resultado no mostrado), lo cual indica que los genes de plantas no conservan los mismos patrones de regulación ni la misma estructura de las regiones promotoras de genes animales, como han señalado otros análisis similares (Yamamoto et al. 2007a).

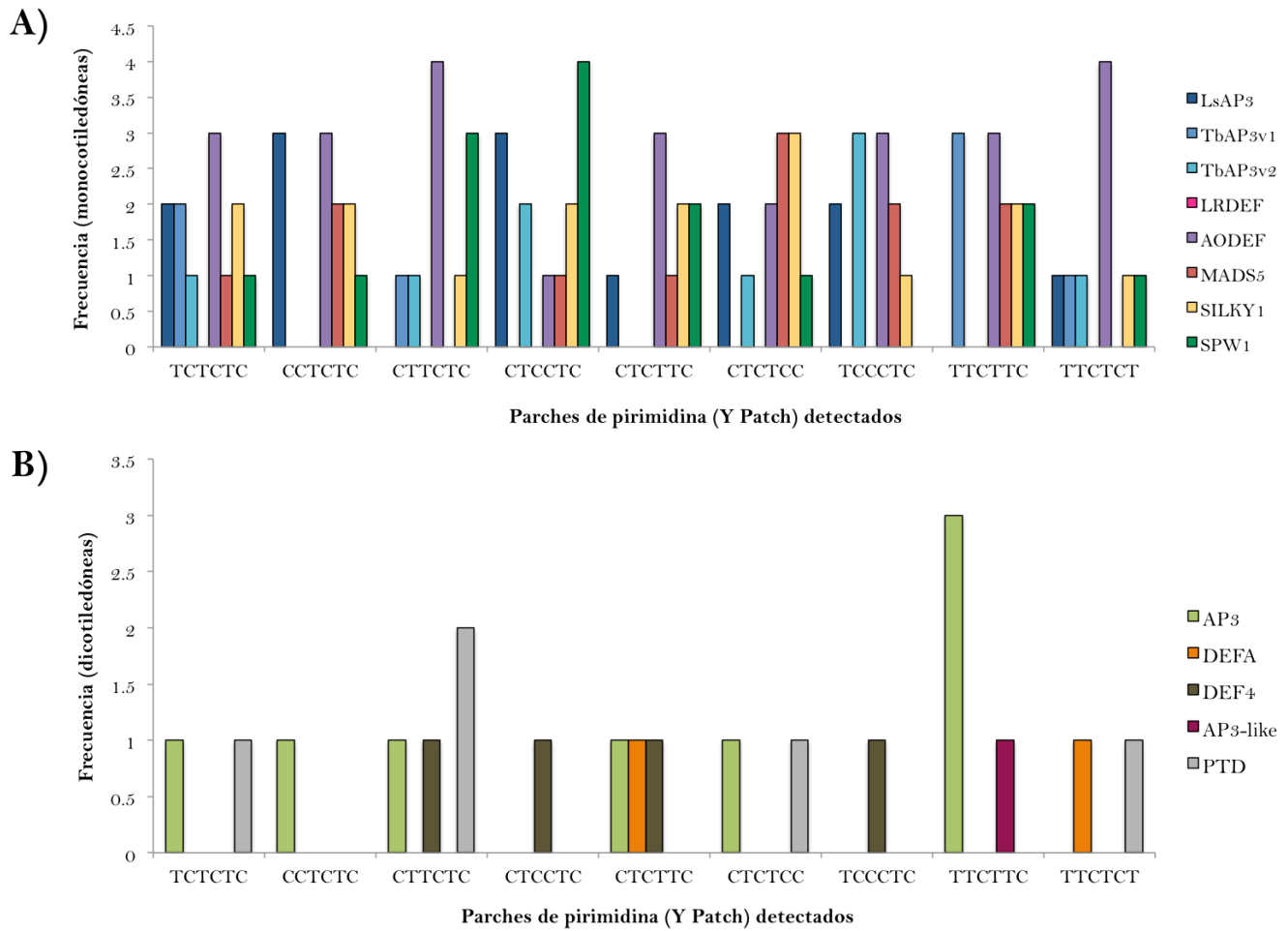


Figura 19. Distribución de distintos parches de pirimidina (Y Patch) en promotores de genes ortólogos de *AP3*. La presencia de varios motivos Y Patch parece ser una característica conservada en monocotiledóneas (A), al menos entre los ortólogos de *AP3*, indicando una posible función en el establecimiento o mantenimiento de la expresión génica. De manera interesante, *LsAP3* y *TbAP3v2* comparten diversos motivos que no están presentes en *TbAP3v1*. Salvo excepciones llamativas, como *LRDEF*, las monocotiledóneas exhiben la mayoría de las variantes encontradas por Yamamoto et al. (2007b). Esto no ocurre en dicotiledóneas (B), cuya distribución de motivos Y Patch es más bien dispersa.

Sin embargo, Yamamoto et al. (2007b) han identificado, a través de perfiles de distribución local de secuencias cortas, grupos de hexámeros distribuidos en los promotores de *Arabidopsis* y arroz: cajas TATA y parches de pirimidina, llamados también Y Patch. Aunque la función de estos últimos no ha sido esclarecida se sabe que son motivos específicos de los promotores de plantas (Yamamoto et al. 2007a).

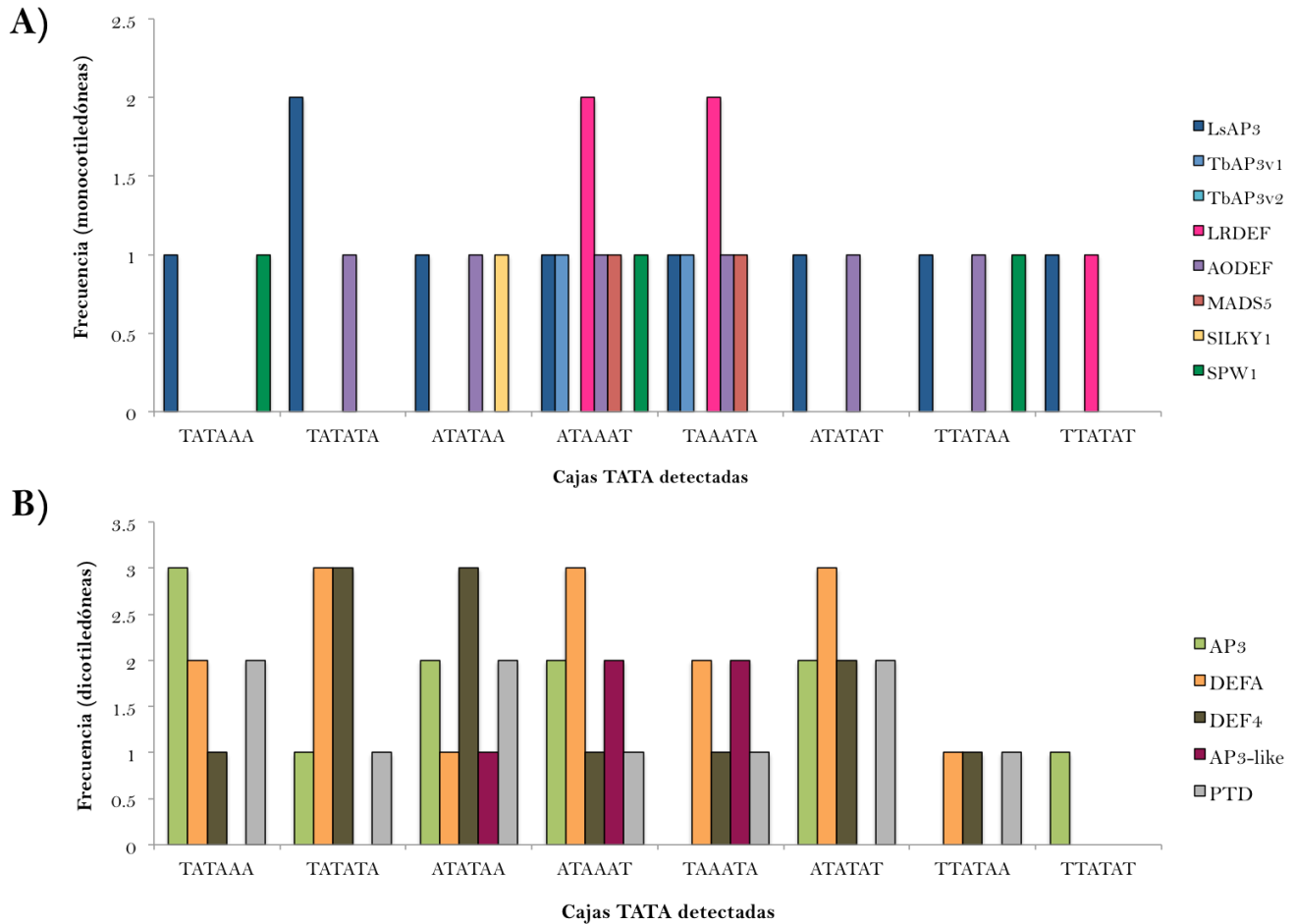


Figura 20. Distribución de distintas cajas TATA en promotores de genes ortólogos de *AP3*. El patrón de distribución cajas TATA en monocotiledóneas (**A**) no es tan claro, aunque puede distinguirse una conservación general. Es de notar, eso sí, que todas las variantes identificadas por Yamamoto et al. (2007b) están presentes en el promotor de *LsAP3*. (**B**) Por otro lado, se observa una mayor presencia de cajas TATA en los promotores de dicotiledóneas como *Arabidopsis*, *Antirrhinum* y *Solanum*, así como una conservación más evidente de estos motivos en este linaje.

La **Figura 19** demuestra que la distribución de motivos Y Patch es desigual entre los linajes de monocotiledóneas y dicotiledóneas, lo cual puede interpretarse en términos de posibles diferencias en la regulación de *AP3*. Más aún, dentro del grupo de las monocotiledóneas es posible identificar algunos patrones. Resalta el hecho, por ejemplo, de que *LsAP3* y *TbAP3v2* comparten un número importante de motivos Y Patch a pesar de que *LsAP3* se encuentra más cercanamente relacionado a *TbAP3v1* (Dolores Fuentes 2013). Además, resulta interesante que las secuencias TCTCTC, CTTCTC, TTCTCT conservan la misma posición física dentro de los promotores de *TbAP3v1* y *TbAP3v2* (resultado no mostrado), pues se ha sugerido que los patrones de distribución espacial de

elementos *cis* proveen una medida a grandes rasgos del grado de conservación en patrones de regulación tras la divergencia de genes o especies (Xu et al. 2012).

Algunos genes de plantas carecen de cajas TATA, e inician la transcripción con la ayuda de elementos iniciadores (YYANWYY; Nakamura et al. 2002). Para indagar al respecto, en la **Figura 20** representé la frecuencia de distintas cajas TATA previamente identificadas en regiones reguladoras de plantas y encontradas en los promotores de los ortólogos de *AP3*. Aunque no resulta tan sencillo distinguir un patrón de distribución entre linajes, llama la atención que el promotor de *LsAP3* es el único donde se ven representadas todas las variantes de este motivo –todos los promotores poseen al menos una variante de la caja TATA–. Es también notable que la secuencia ATAAAT está incorporada de manera predominante en los promotores de angiospermas.

Elementos conservados en los promotores de *LsAP3* y *TbAP3*

Tagle et al. (1994) escogieron el término *phylogenetic footprinting* para describir las comparaciones filogenéticas que detectan elementos *cis* conservados entre regiones reguladoras de genes ortólogos. Muchos estudios han aprovechado este tipo de análisis para comprender a detalle la regulación genética de distintos grupos de organismos. En plantas, estos han demostrado que los elementos *cis* de genes involucrados en el desarrollo (como *APETALA3*) pueden ser disectados gracias a la comparación de promotores u otras regiones no codificantes, como los intrones (Koch et al. 2001; Hong et al. 2003; Lee et al. 2005; De Bodt et al. 2006; Xu et al. 2012).

La premisa fundamental del *phylogenetic footprinting* es que los motivos reguladores funcionales –usualmente pequeños, de 4 pb a 20 pb– evolucionan a una tasa más baja que otras regiones no codificantes. Sin embargo, si las especies usadas en el análisis han divergido lo suficiente, es probable que el ruido de fondo oculte la señal de los elementos conservados (Blanchette y Tompa 2002). Además, el número de secuencias analizadas también puede influir en los resultados del análisis: mientras más se incluyan aumenta la probabilidad de que la regulación de algunas de ellas haya cambiado por completo (Blanchette y Tompa 2002).

Para este trabajo decidí entonces sólo elegir aquellos promotores de especies relativamente cercanas a *Lacandonia* y *Triuris*, es decir, los de monocotiledóneas. El promotor de *AP3* de *Arabidopsis* se analizó también a manera de comparación con los

mecanismos de regulación en dicotiledóneas y, además, porque la mayor cantidad de información experimental proviene del mismo (Jack et al. 1992; Tilly et al. 1998; Hill et al. 1998; Lamb et al. 2002).

Los patrones generales de conservación de elementos *cis* entre ortólogos de *AP3* se resumen en la **Figura 21**. Llama la atención su distribución espacial. Xu et al. (2012) proponen que el patrón de distancia y orden entre sitios de unión en el ADN podría ofrecer una medida a grandes rasgos del grado de conservación entre secuencias durante su divergencia evolutiva.

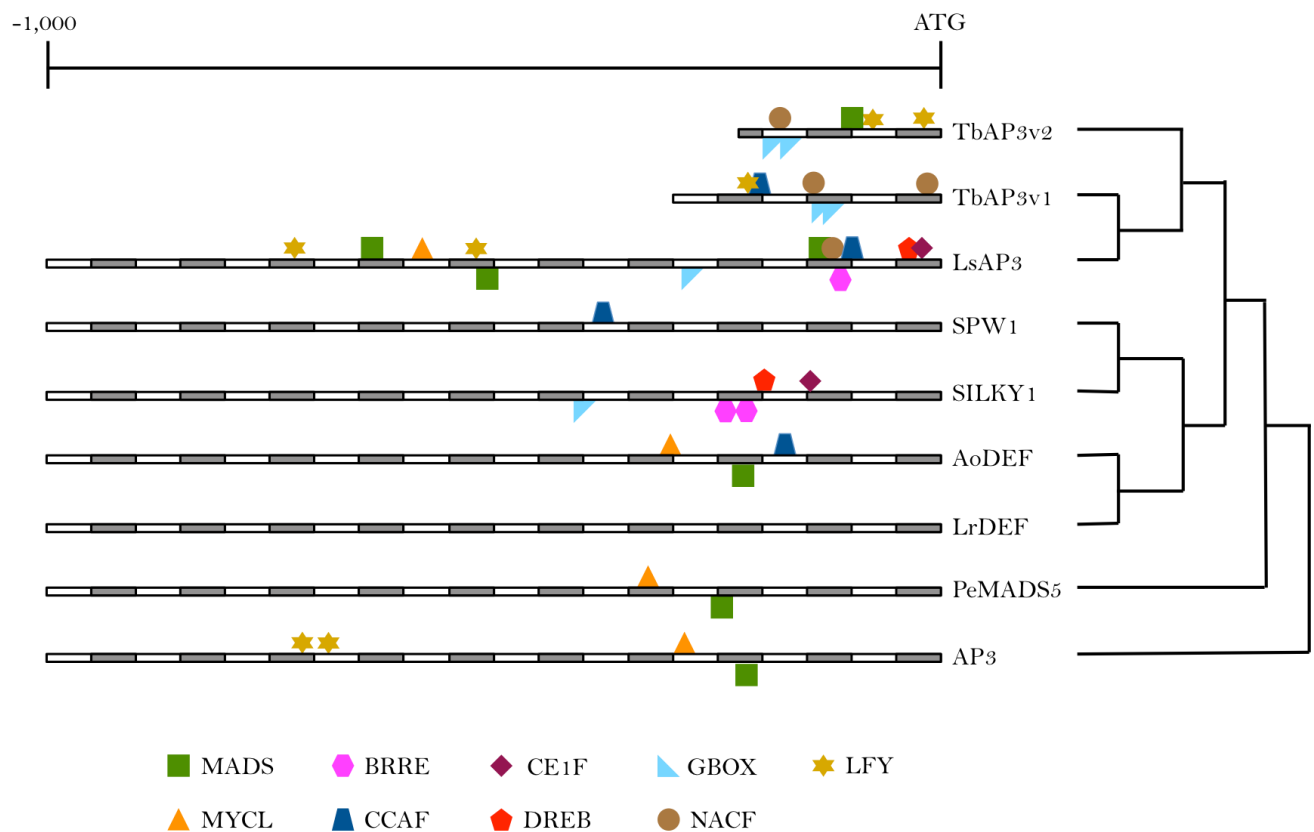


Figura 21. Representación visual de los elementos *cis* detectados entre los promotores de varios ortólogos de *APETALA3*. Cada símbolo representa el sitio de unión de una familia de factores de transcripción; las figuras invertidas indican que la ubicación del motivo se encuentra en la hebra complementaria de la región promotora. Los elementos *cis* aquí presentados forman parte de los patrones elaborados en FrameWorker a excepción de los sitios de unión a LFY, que encontré con alineamientos locales en DIALIGN. No todas las cajas CArG ubicadas en los promotores de *LsAP3* y *TbAP3v2* forman un patrón definido pero las incluí por su importancia en esta tesis. Para una lista completa de todos los sitios de unión compartidos entre secuencias, ver el **Apéndice 4**. Valor de p de todos los patrones < 0.01.

Como se ve en mis resultados, predomina una mayor conservación de patrones entre Pandanales (*Triuris* y *Lacandonia*), Asparagales (espárrago) y Poales (arroz y maíz) y, en general, una alta similitud entre los promotores de monocotiledóneas, pues comparten un mayor número de motivos y patrones de regulación que no están presentes en *A. thaliana*, su contraparte dicotiledónea. La misma tendencia se refleja en la filogenia construida a partir de secuencias de aminoácidos.

También es evidente, sin embargo, el pequeño número de patrones de elementos *cis* identificados en los promotores de monocotiledóneas (no encontré ninguno en la región reguladora del ortólogo de *Lilium regale*, por ejemplo, y el promotor del gen *PeMADS5* de la orquídea *Phalaenopsis equestris* exhibe uno solo). Esta variación podría indicar que el contexto de regulación de *APETALA3* no sólo ha evolucionado tras la divergencia entre ambos linajes de angiospermas, definiendo así un perfil único de expresión espaciotemporal en monocotiledóneas y dicotiledóneas, sino incluso dentro de cada uno.

Para intentar comprender la función biológica de los supuestos sitios de unión recuperados a partir de la identificación de patrones en FrameWorker los busqué en las bases de datos PLACE y PlantCARE. Según lo que encontré, estos participan en procesos del desarrollo vegetativo y reproductivo, incluyendo crecimiento de la raíz, especificación de identidad de órganos florales, respuesta a estrés biótico y abiótico, ritmos circadianos, regulación del ácido abscísico y respuesta a luz. En el **Apéndice 4** se enlistan todos los elementos *cis* compartidos entre al menos una de las regiones promotoras de *LsAP3* o *TbAP3* y las de otros ortólogos.

Uno de los sitios de unión encontrados de mayor interés es la caja CArG, a la que reconocen y se unen las proteínas MADS (Taylor et al. 1989; Riechmann et al. 1996). Aunque no lo muestra la **Figura 21**, ubiqué cajas CArG en casi todas las regiones promotoras analizadas, incluidas las tres cajas canónicas de *A. thaliana* (Hill et al. 1998; Tilly et al. 1998). Este es un resultado novedoso para las triuridáceas. Si bien Piñeyro Nelson (2013) realizó alineamientos múltiples de la región 5' de algunos pocos ortólogos de *AP3* y sugirió la existencia de cajas CArG en el promotor de *LsAP3*, no pudo hallarlas con exactitud y concluyó que quizá habían perdido su secuencia consenso. En esta tesis yo demuestro con ayuda de análisis de *phylogenetic footprinting* que *LsAP3* posee al menos tres cajas CArG localizadas en las posiciones -645 (CCAAGATTAG), -511 (CCTTTATTTGG) y -146 (CTTTTTAAAG). Además, encontré otra presunta caja

CAR_G en la posición -107 (CCTCATTTAG) del promotor parcial de *TbAP3v2*. Ninguna de estas cajas CAR_G ha sido reportada previamente aunque todas ellas conservan estructuras consenso de algunas variantes (de Folter y Angenent 2006).

Mientras que el heterodímero AP3/PI regula la expresión tardía de *AP3*, se ha observado en *Arabidopsis* que LFY y UFO controlan el establecimiento de su dominio de expresión en etapas tempranas del desarrollo floral. Si ocurriera lo mismo en *L. schismatica* y *T. brevistylis*, se esperaría encontrar sitios conservados de unión a LFY en las regiones promotoras tanto de *LsAP3* como de *TbAP3*. Basándome en esto, detecté cinco posibles sitios de unión a LFY en todas las triuridáceas basándome en alineamientos con los motivos identificados por Lamb et al. (2002) en *Arabidopsis* (ver **Figura 22**).

En resumen, en esta tesis logré descubrir motivos de regulación conservados y divergentes dentro de los promotores de *LsAP3*, *TbAP3v1* y *TbAP3v2*. Sin embargo, quizá valga la pena recordar que sólo la validación experimental podrá dar sustento a la función biológica predicha aquí.

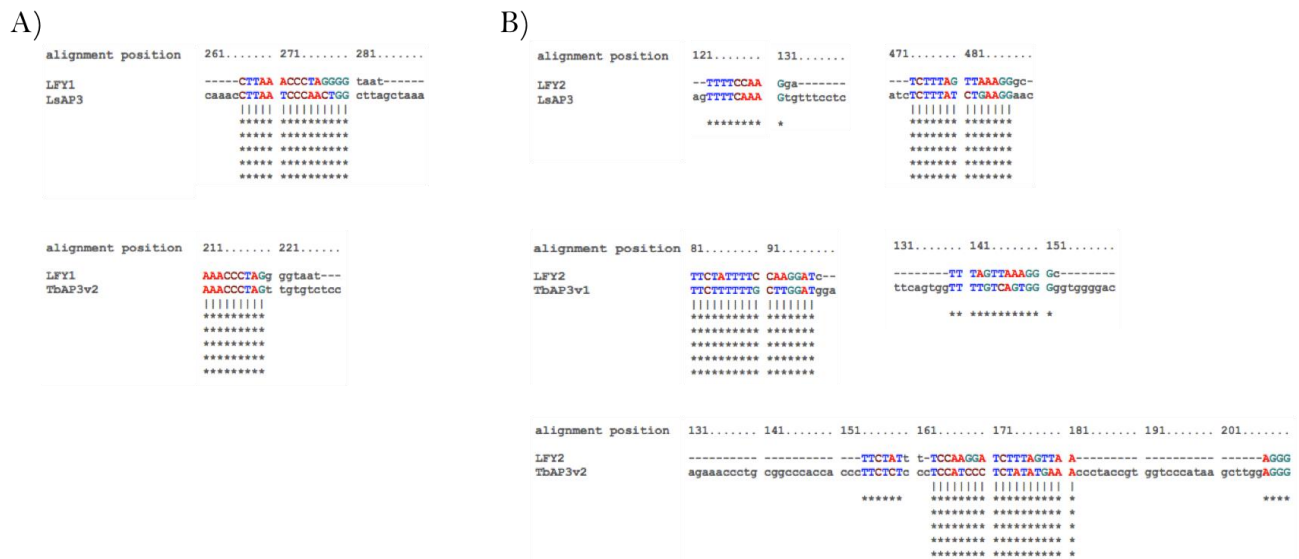


Figura 22. Alineamientos locales de ambos sitios de unión a LFY y secuencias promotoras de *LsAP3* y *TbAP3*. (A) El primer sitio de unión a LFY, LFY1, se alinea a las regiones -735 de *LsAP3* y -19 de *TbAP3v2* con un porcentaje de identidad de 57% y 47%, respectivamente. El alineamiento con *TbAP3v1* no se muestra, pues el porcentaje de identidad fue muy bajo (36%). (B) LFY2 se alinea a las posiciones -527 de *LsAP3*, -217 de *TbAP3v1* y -67 de *TbAP3v2* con un porcentaje de identidad de 59%, 68% y 50%, respectivamente. Hice todos los alineamientos en DIALIGN, que busca regiones cortas conservadas. Sólo se muestran los segmentos de alineamiento; el resto de las secuencias se omite por razones de espacio.

Una visión compleja sobre la evolución del desarrollo en Triuridales

Las diferencias entre las copias parálogas de *TbAP3* y la mayor variabilidad entre ambas con respecto a *LsAP3* (Dolores Fuentes 2013) hacen de su estudio un caso interesante. Existen tres escenarios posibles que podrían representar la relación entre ambos parálogos de *TbAP3*: 1) ambos genes podrían haber divergido tanto en estructura de sus promotores como en sus patrones de expresión debido a eventos de duplicación y subfuncionalización (Doebley y Lukens 1998; Force et al. 1999; Xu et al. 2012; Dolores Fuentes 2013), o bien, 2) podrían haber conservado cierto grado de redundancia en sus promotores como una forma de refuerzo genético durante el desarrollo (Nadeau y Sankoff 1997; Kafri et al. 2005). También hay lugar para la posibilidad, por supuesto, de que alguna de las copias de *TbAP3* haya perdido funcionalidad. Hasta ahora, no me es posible discernir entre estos escenarios. La clonación completa de los promotores de *TbAP3v1* y *TbAP3v2* será necesaria si queremos comprender a fondo el papel de ambos parálogos durante la evolución y el desarrollo de *T. brevistylis* y sus mutantes homeóticas.

Por otro lado, el hallazgo de cajas CARG modificadas y sitios de unión a LFY variables en el promotor de *LsAP3* permite pensar en alternativas posibles sobre cómo se desplazó su dominio espacial de expresión. El fenotipo homeótico observado en las flores de *L. schismatica* podrá ser resultado de un cambio relativamente sencillo en el patrón de expresión de este gen, pero quizá la red de regulación genética dentro de la cual interactúa el mismo haya sufrido cambios más complejos para dar lugar a esta innovación morfológica.

Un debate continuo en evo devo versa sobre cuál es la contribución más importante para la evolución fenotípica de los organismos: la de mutaciones reguladoras en elementos *cis* o la de mutaciones codificantes en factores *trans* (Wray 2007; Hoekstra y Coyne 2007; Carroll 2008; Stern y Orgogozo 2008). Tomando en cuenta el contexto molecular completo de la regulación genética de *LsAP3*, propongo que la evolución del arreglo floral en *Lacandonia* no se debe a modificaciones en los elementos *cis* de su promotor o a alteraciones en las capacidades de unión o interacción de las proteínas que lo regulan. Sugiero, más bien, que ambos mecanismos han contribuido a la evolución del desarrollo alterado de esta planta mexicana.

Los experimentos de Piñeyro Nelson (2013), donde fusionó el promotor de *LsAP3* al gen reportero GUS y luego lo expresó en plantas de *A. thaliana*, no logran reproducir el

patrón homeótico observado en *Lacandonia*. Sin embargo, sí recuperan algunos aspectos como la expresión en estambres y la expresión en etapas tardías del desarrollo floral. Esto demuestra que las proteínas AP3 y LFY de *Arabidopsis* son capaces de reconocer e interactuar, al menos débilmente, con algunos de los elementos *cis* presentes en el promotor de *LsAP3*. También quiere decir que los cambios de secuencias reguladoras encontrados en esta tesis no son suficientes para explicar la posición atípica de los órganos sexuales de *Lacandonia*. Sin embargo, es muy probable que sean necesarios.

Según los estudios realizados por Sánchez Lara (2013) y Dolores Fuentes (2013), las proteínas LsUFO, LsLFY y LsAP3 poseen mutaciones importantes en su secuencia de aminoácidos capaces de modificar su actividad. Dichos reguladores proteicos alterados (principalmente LsUFO y LsAP3, pues en *Arabidopsis* LFY interactúa con otros genes como AG y, de haber cambiado su función en *Lacandonia* veríamos con seguridad una alteración en el patrón de expresión de *LsAG*, que no ocurre) podrían unirse a los sitios de unión modificados que se encuentran en el promotor de *LsAP3* para dar como resultado un desplazamiento espacial de la expresión de este gen.

Es verdad que este no es el escenario más parsimonioso posible, pues involucra mutaciones a distintos niveles de la red de regulación genética que controla el desarrollo de estambres en las flores de *Lacandonia*, pero es verdad también que la parsimonia no siempre refleja la realidad biológica. Será necesario realizar más análisis con acercamientos prácticos y teóricos antes de poder descubrir las causas de la promesa de los monstruos lacandones.

CONCLUSIONES Y PERSPECTIVAS

A lo largo de esta tesis, he tratado de explorar las bases causales de los arreglos florales homeóticos presentes en dos triuridáceas mexicanas: *Lacandonia schismatica* y *Triuris brevistylis*. Desde un punto de vista de la biología evolutiva del desarrollo, el drástico fenotipo que resulta de la inversión espacial de órganos sexuales puede ser consecuencia de un cambio relativamente sencillo, la modificación del dominio de expresión de un gen MADS-box que participa en el desarrollo de estambres: *LsAP3*.

Esto se ha demostrado de manera experimental en estudios pasados, pero aún resta indagar sobre los mecanismos que subyacen tras la alteración de la regulación de este gen. Mis resultados y algunos antecedentes directos de este proyecto sugieren que en el desplazamiento espacial de la expresión de *LsAP3* hacia el centro del meristemo floral podría estar dibujado un escenario un poco más complejo donde distintos niveles de la red de regulación genética se han visto alterados. No sólo podrían estar involucrados cambios en las proteínas reguladoras de *LsAP3* –e incluso cambios en el mismo producto proteico de este gen– sino que también podrían ser necesaria la presencia de elementos *cis* modificados en el promotor.

Esto supone un obstáculo metodológico para aquellos estudios que deseen recuperar un arreglo floral tipo *Lacandonia* en plantas modelo, pero también significa que el desarrollo de las plantas no puede ser concebido como un proceso lineal.

Por otro lado, esta tesis demostró que la búsqueda de huellas filogenéticas entre secuencias reguladoras de genes ortólogos puede ser una herramienta útil para encontrar pistas sobre la evolución y divergencia de especies. Sin embargo, sería un engaño pensar que los genes no son únicos en un contexto regulador que define los aspectos más generales y finos de su patrón espaciotemporal de expresión. Sólo una combinación de enfoques prácticos y teóricos podrá permitirnos comprender más a detalle el papel de la regulación genética en la evolución fenotípica de los organismos y, en este caso, en la evolución de la flor atípica de *Lacandonia* y la aparición de fenotipos parecidos en las flores de su taxón hermano.

Aquí me aventuro a proponer una hipótesis sobre el origen de la homeosis en estas plantas, pero todavía falta recorrer un largo y sinuoso camino antes de poder aceptarla.

Por un lado, será necesario terminar de clonar los promotores de *TbAP3v1* y *TbAP3v2* con el objetivo de poder realizar análisis bioinformáticos mucho más completos de los que yo presento aquí. Para este fin, la técnica de *genome walking* podrá ser útil pero la secuenciación del genoma completo de *Triuris* –y, de paso, el de *Lacandonia*– ofrecerá datos mucho más ricos. Sería increíble también clonar y secuenciar los ortólogos de *LsAP3*, *LsUFO* y *LsLFY* de otras triuridáceas, e incluso de especies que forman parte de las Pandanales, orden al que pertenece la familia Triuridaceae, para poder analizar con profundidad la evolución de estos genes.

Además, la funcionalidad biológica de los motivos más importantes hallados en esta tesis, como las cajas CArG y los motivos de unión a LFY, deberá ser comprobada experimentalmente –por ejemplo, a través de mutagénesis dirigidas–.

Comenzar a explorar las bases conceptuales de evo-devo en el estudio del caso *Lacandonia* y *Triuris* ayudará a consolidar y complementar los avances experimentales que se hagan. Sería crucial, por ejemplo, comprender cómo se fijó el fenotipo monstruoso de las flores de *Lacandonia* en sus poblaciones naturales: ¿por la distribución de alelos recesivos en sus poblaciones ancestrales o por acomodación genética?

Por otro lado, idear enfoques teóricos como la reconstrucción de redes de regulación genética que consideren la acción de factores genéticos y no genéticos en el desarrollo de *Lacandonia* y *Triuris* darán un panorama más completo sobre qué tipo de cambios han ocurrido durante la divergencia de ambas especies.

REFERENCIAS

- Abzhanov A, Extavour CG, Groover A, Hodges SA, Hoekstra HE, Kramer EM, Monteiro A. 2008. Are we there yet? Tracking the development of new model systems. *Trends Genet* 24: 352-360.
- Ainsworth C, Thangavelu M, Crossley S, Buchanan-Wollaston V, Parker J. 1995. Male and female flowers from the dioecious plant *Rumex acetosa* show different patterns of MADS-box gene expression. *Plant Cell* 7: 1583-1598.
- Álvarez-Buylla ER, Pelaz S, Liljegren SJ, Gold SE, Burgeff C, Ditta GS, Ribas de Pouplana L, Martínez-Castilla L, Yanofsky MF. 2000. An ancestral MADS-box gene duplication occurred before the divergence between plants and animals. *Proceedings of the National Academy of Sciences* 97: 5328-5333.
- Álvarez-Buylla ER. 2002. La diversidad de las formas vegetales: variaciones sobre un mismo tema. *Ciencias* 65: 18-28.
- Álvarez-Buylla ER, Benítez M, Balleza Dávila E, Chaos A, Espinosa-Soto C, Padilla-Longoria P. 2007. Gene regulatory network models for plant development. *Curr Op Plant Biol* 10: 83-91.
- Álvarez-Buylla ER, Benítez M, Corvera-Poiré A, Chaos Cador A, de Folter S, Gamboa de Buen A, Garay-Arroyo A, García-Ponce B, Jaimes-Miranda F, Pérez-Ruiz RV, Piñeyro-Nelson A, Sánchez-Corrales YE. 2010a. Flower development. *The Arabidopsis Book* 8: e0127. doi:10.1199/tab.0127
- Álvarez-Buylla ER, Ambrose BA, Flores-Sandoval E, Englund M, Garay-Arroyo A, García-Ponce B, de la Torre-Bárcena E, Espinosa-Matías S, Martínez E, Piñeyro-Nelson A, Engström P, Meyerowitz EM. 2010b. B-function expression in the flower center underlies the homeotic phenotype of *Lacandonia schismatica* (Triuridaceae). *The Plant Cell* 22: 3543-3559.
- Ambrose BA, Lerner DR, Ciceri P, Padilla CM, Yanofsky MF, Schmidt RJ. 2000. Molecular and genetic analyses of the *silky1* gene reveal conservation in floral organ specification between eudicots and monocots. *Molecular Cell* 5: 569-579.
- Ambrose BA, Espinosa-Matías S, Vázquez-Santana S, Vergara-Silva F, Martínez E, Márquez-Guzmán J, Álvarez-Buylla ER. 2006. Comparative developmental series

- of the mexican triurids support an euanthial interpretation for the unusual reproductive axes of *Lacandonia schismatica* (Triuridaceae). *American Journal of Botany* 93: 15-35.
- Amundson R. 2005. *The changing role of the embryo in evolutionary thought: roots of evo-devo*. Cambridge University Press, Cambridge. 280 p.
- Arnone MI, Davidson EH. 1997. The hardwiring of development: organization and function of genomic regulatory systems. *Development* 124: 1851-1864.
- Baguñà J y Garcia-Fernández J. 2003. Evo-devo: the long and winding road. *International Journal of Developmental Biology* 47: 705-713.
- Bateman RM, DiMichele WA. 2002. Generating and filtering major phenotypic novelties: neoGoldschmidtian saltation revisited. En: Cronck QCB, Bateman RM, Hawkins JA (eds.) *Developmental Genetics and Plant Evolution*. Taylor & Francis, Londres, pp. 109-159.
- Bateman RM, DiMichele WA. 2003. Genesis of phenotypic and genotypic diversity in land plants: the present as key to the past. *Syst Bio* 1: 13-28.
- Bäumlein H, Nagy I, Villarroel R, Inzé D, Wobus U. 1992. *Cis*-analysis of a seed protein gene promoter. The conservative RY repeat CATGCATG within the legumin box is essential for tissue-specific expression of a legumin gene. *Plant J* 2: 233-239.
- Beckerman M. 2005. *Molecular and Cellular Signaling*. Springer Science. New York, USA. 574 p.
- Benítez M, Azpeitia E, Álvarez-Buylla ER. 2013. Dynamic models of epidermal patterning as an approach to plant eco-evo-devo. *Curr Opin Plant Biol*. 16:11-8. doi: 10.1016/j.pbi.2012.11.005.
- Bird A. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev* 16: 6-21.
- Blanchette M, Tompa M. 2002. Discovery of regulatory elements by a computational method for phylogenetic footprinting. *Genome Research* 12: 739-748.
- Blount ZD, Borland CZ, Lenski RE. 2008. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc Nat Acad Sci* 105: 7899-7906.
- Blumenthal T, Evans D, Link CD, Guffanti A, Lawson D, Thierry-Mieg J, Thierry-Mieg D, Chiu WL, Duke K, Kiraly M, Kim SK. 2002. A global analysis of *Caenorhabditis elegans* operons. *Nature* 417: 851-854.

- Bolker JA. 2000. Modularity in development and why it matters to evo-devo. *American Zoologist* 40: 770-776.
- Borneman AR, Gianoulis TA, Zhang ZD, Yu H, Rozowsky J, Seringhaus MR, Wang LY, Gerstein M, Snyder M. 2007. Divergence of transcription factor binding sites across related yeast species. *Science* 317: 815-819.
- Bowman JL, Smyth DR, Meyerowitz EM. 1989. Genes directing flower development in *Arabidopsis*. *The Plant Cell* 1: 37-52.
- Bowman JL, Smyth DR, Meyerowitz EM. 1991. Genetic interactions among floral homeotic genes in *Arabidopsis*. *Development* 112: 1-20.
- Bowman JL, Smyth DR, Meyerowitz EM. 2012. The ABC model of flower development: then and now. *Development* 139: 4095-4098.
- Bradley RK, Li XY, Trapnell C, Davidson S, Pachter L, Chu HC, Tonkin LA, Biggin MD, Eisen MB. 2010. Binding site turnover produces pervasive quantitative changes in transcription factor binding between closely related *Drosophila* species. *PLoS Biology* 8: e1000343.
- Britten R, Davidson EH. 1969. Gene regulation for higher cells: a theory. *Science* 165: 349-357.
- Britten RJ. 2002. Divergence between samples of chimpanzee and human DNA sequences is 5%, counting indels. *Proceedings of the National Academy of Sciences* 99: 13633-13635.
- Busk K, Pages M. 1998. Regulation of abscisic acid-induced transcription. *Plant Mol Biol* 37: 425-435.
- Carpenter R, Coen E. 1990. Floral homeotic mutations produced by transposon-mutagenesis in *Antirrhinum majus*. *Genes & Development* 4: 1483-1493.
- Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134: 25-36.
- Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M, Bayerlein M, Werner T. 2005. MatInspector and beyond: promoter analysis based on transcription binding sites. *Bioinformatics* 21: 2933-2942.
- Causier B, Schwarz-Sommer Z, Davies B. 2010. Floral organ identity: 20 years of ABCs. *Seminars in Cell & Developmental Biology* 21: 73-79.

- Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT). 2006. *Protocolos de laboratorio: laboratorio de genética molecular aplicada del CIMMYT*. Tercera edición. México, D.F.
- Chae E, Tan QKG, Hill TA, Irish VF. 2008. An *Arabidopsis* F-box protein acts as a transcriptional co-factor to regulate floral development. *Development* 135: 1235-1245.
- Chan YF, Marks ME, Jones FC, Villareal GJ, Shapiro MD, Brady SD, Southwick AM, Absher DM, Grimwood J, Schmutz J, Myers RM, Petrov D, Jónsson B, Schluter D, Bell MA, Kingsley DM. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* 327: 302-305.
- Charlesworth B. 1982. Hopeful monsters cannot fly. *Paleobiology* 8: 469-474.
- Chen KY, Cong B, Wing R, Vrebalov J, Tanksley SD. 2007. Changes in regulation of a transcription factor lead to autogamy in cultivated tomatoes. *Science* 318: 643-645.
- Chevenet F, Brun C, Banuls AL, Jacq B, Christen R. 2006. TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* 7: 439.
- Cho S, Jang S, Chae S, Chung KM, Moon Y, An G, Jang SK. 1999. Analysis of the C-terminal region of *Arabidopsis thaliana* APETALA1 as a transcription activation domain. *Plant Molecular Biology* 40: 419-429.
- Coen E. 2001. Goethe and the ABC model of flower development. *Comptes rendus de l'Academie des Sciences* 324: 523-530.
- Coen E, Meyerowitz E. 1991. The war of the whorls: genetic interactions controlling flower development. *Nature* 353: 31-37.
- Cole MD, Cowling VH. 2008. Transcription-independent functions of MYC: regulation of translation and DNA replication. *Nat Rev Mol Cell Biol* 9: 810-815.
- Colosimo PF, Hosemann KE, Balabhadra S, Villareal G, Dickson M, Grimwood J, Schmutz J, Myers RM, Schluter D, Kingsley DM. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307: 1928-1933.
- Corpet F. 1998. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research* 16: 10881-10890.
- Davidson EH. 2001. *Genomic regulatory systems: development and evolution*. Academic Press, Sand Diego, California.

- Dávila J, Álvarez-Buylla E. 2014. Bridging the genotype and the phenotype: towards an epigenetic landscape approach to evolutionary systems biology. *BioRxiv* 1-15.
- De Bodt S, Theissen G, Van de Peer Y. 2006. Promoter analysis of MADS-Box genes in eudicots through phylogenetic footprinting. *Mol Biol Evol* 23: 1293-1303.
- de Bruijn S, Angenent GC, Kaufmann K. 2012. Plant 'evo-devo' goes genomic: from candidate genes to regulatory networks. *Trends in Plant Science* 17: 441-447.
- de Folter S, Angenent GC. 2006. *Trans* meets *cis* in MADS science. *TRENDS in Plant Science* 11(5): 224-231.
- de Martino G, Pan I, Emmanuel E, Levy A, Irish VF. 2006. Functional analyses of two tomato *APETALA3* genes demonstrate diversification in their roles in regulating floral development. *The Plant Cell* 18: 1833-1845.
- De Vries H. 1905. *Species and varieties: their origin by mutation*. The Open Court Publishing Company, Londres. 826 p.
- De Waal FB. 2005. A century of getting to know the chimpanzee. *Nature* 437: 56-59.
- Dietrich MR. 2003. Richard Goldschmidt: hopeful monsters and other 'heresies'. *Nature Reviews Genetics* 4: 68-74.
- Doebley J, Lukens L. 1998. Transcriptional regulators and the evolution of plant form. *The Plant Cell* 10: 1075-1082.
- Dolores Fuentes AC. 2013. Clonación, caracterización molecular y análisis filogenético de los genes MADS-box de desarrollo floral *APETALA3*, *PISTILLATA* y *AGAMOUS* de *Triuris brevistylis* (Triuridaceae). Tesis de licenciatura. Universidad Nacional Autónoma de México.
- Doolittle RF, Wooding GL, Lin Y, Riley M. 1971. Hominoid evolution as judged by fibrinopeptide structures. *Journal of Molecular Evolution* 1: 74-83.
- Drea S, Hileman LC, de Martino G, Irish VF. 2007. Functional analyses of genetic pathways controlling petal specification in poppy. *Development* 134: 4157-4166.
- Drews GN, Bowman JL, Meyerowitz EM. 1991. Negative regulation of the *Arabidopsis* homeotic gene *AGAMOUS* by the *APETALA2* product. *Cell* 65: 991-1002.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792-1797.
- Egea-Cortines M, Saedler H, Sommer H. 1999. Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in

- the control of floral architecture in *Antirrhinum majus*. The EMBO Journal 18: 5370-5379.
- Espinosa-Matías S, Vergara-Silva F, Vázquez-Santana S, Martínez-Zurita E y J Márquez-Guzmán. 2012. Complex patterns of morphogenesis, embryology and reproduction in *Triuris brevistylis*, a species of Triuridaceae (Pandanales) closely related to *Lacandonia schismatica*. Botany 90: 1133- 1151.
- Espinosa-Soto C, Padilla-Longoria P, Álvarez-Buylla ER. 2004. A gene regulatory network model for cell-fate determination during *Arabidopsis thaliana* flower development that is robust and recovers experimental gene expression profiles. Plant Cell 16: 2923-2939.
- Eyal Y, Curie C, McCormick S. 1995. Pollen specificity elements reside in 30 bp of the proximal promoters of two pollen expressed genes. Plant Cell 7: 373-384.
- Fan HY, Hu Y, Tudor M, Ma H. 1997. Specific interactions between the K domains of *AG* and *AGLs*, members of the MADS domain family of DNA binding proteins. The Plant Journal 12: 999-1010.
- Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151: 1531-1545.
- Friedman WE, Moore RC, Purugganan MD. 2004. The evolution of plant development. American Journal of Botany 91: 1726-1741.
- Friedman WE, Diggle PK. 2011. Charles Darwin and the origins of plant evolutionary development. The Plant Cell 23: 1194-1207.
- Galis F. 1999. Why do almost all mammals have seven cervical vertebrae? Developmental constraints, Hox genes and cancer. Journal of Experimental Zoology 285: 19-26.
- Gallego-Romero I, Ruvinsky I, Gilad Y. 2012. Comparative studies of gene expression and the evolution of gene regulation. Nat Rev Genet 18: 505-516.
- Garay-Arroyo A, Piñeyro-Nelson A, García-Ponce B, de la Paz Sánchez M, Álvarez-Buylla ER. 2012. When ABC becomes ACB. Journal of Experimental Botany 63: 2377-2395.
- Garber RL, Kuroiwa A, Gehring WJ. 1983. Genomic and cDNA clones of the homeotic locus *Antennapedia* in *Drosophila*. The EMBO Journal 2: 2027-2036.

- Garcia-Maroto F, Salamini F, Rohde W. 1993. Molecular cloning and expression patterns of three alleles of the *Deficiens* homologous gene *St-Deficiens* from *Solanum tuberosum*. *The Plant Journal* 4: 771-780.
- Gehring WJ, Hiromi Y. 1986. Homeotic genes and the homeobox. *Annual Review of Genetics* 20: 147-173.
- Gelbart WM. 1989. The decapentaplegic gene: a TGF-beta homologue controlling pattern formation in *Drosophila*. *Development* 107: 65-74.
- Gilad Y, Oshlack A, Smyth GK, Speed TP, White KP. 2006. Expression profiling in primates reveals a rapid evolution of human transcription factors. *Nature* 440: 242-245.
- Gilbert SF, Epel D. 2009. *Ecological developmental biology. Integrating epigenetics, medicine, and evolution*. Sinauer Associates, Inc. Sunderland, Massachusetts. 480 p.
- Gilbert SF. 2003. *Developmental Biology*. 7^a edición. Sinauer Associates Inc. Sunderland, Massachusetts. 838 p.
- Goldschmidt R. 1933. Some aspects of evolution. *Science* 78: 539-547.
- Goldschmidt R. 1940. *The material basis of evolution*. Yale University Press, New Haven.
- Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB. 2005. Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* 433: 481-487.
- Goodwin BC. 2000. The life of form. Emergent patterns of morphological transformation. *Comptes rendus de l'academie des sciences III* 323: 15-21.
- Gramzow L, Theissen G. 2010. A hitchhiker's guide to the MADS world of plants. *Genome Biology* 11: 214-225.
- Gramzow L, Ritz MS, Theissen G. 2010. On the origin of MADS-domain transcription factors. *Trends in Genetics* 26: 149-153.
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* 59: 307-321.
- Guo AY, Zhu QH, Gu X, Ge S, Yang J, Luo J. 2008. Genome-wide identification and evolutionary analysis of the plant specific SBP-box transcription factor family. *Gene* 418: 1-8.

- Haig D. 2004. The (dual) origin of epigenetics. *Cold Spring Harbor Symposia on Quantitative Biology* 69: 67-70.
- Hall BK. 2012. Evolutionary developmental biology (evo-devo): past, present and future. *Evolution: Education & Outreach* 5: 184-193.
- Harmer SL, Hogenesch JB, Straume M, Chang HS, Han B, Zhu T, Wang X, Kreps JA, Kay SA. 2000. Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. *Science* 290: 2110-2113.
- Higo K, Ugawa Y, Iwamoto M, Korenaga T. 1999. Plant *cis*-acting regulatory DNA elements (PLACE) database. *Nucleic Acids Research* 27: 297-300.
- Hill TA, Day CD, Zondlo SC, Thackeray AG, Irish VF. 1998. Discrete spacial and temporal *cis*-acting elements regulate transcription of the *Arabidopsis* floral homeotic gene *APETALA3*. *Development* 125: 1711-1721.
- Hintz M, Bartholmes C, Nutt P, Ziermann J, Hameister S, Neuffer B, Theissen G. 2006. Catching a 'hopeful monster': sheperd's purse (*Capsella bursa-pastoris*) as a model system to study the evolution of flower development. *Journal of Experimental Botany* 57(13): 3531-3542.
- Hittinger CT, Stern DL, Carroll SB. 2005. Pleiotropic functions of a conserved insect-specific Hox peptide motif. *Development* 132: 5261-5270.
- Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61: 995-1016.
- Hong RL, Hamaguchi L, Busch MA, Weigel D. 2003. Regulatory elements of the floral homeotic gene *AGAMOUS* identified by phylogenetic footprinting and shadowing. *The Plant Cell* 15: 1296-1309.
- Honma T, Goto K. 2001. Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature* 409: 525-529.
- Ikeda K, Ito M, Nagasawa N, Kyojuka J, Nagato Y. 2007. Rice ABERRANT PANICLE ORGANIZATION1, encoding an F-box, regulates meristem fate. *Plant J* 51: 1030-1040.
- Ingram GC, Doyle S, Carpenter R, Schultz EA, Simon R, Coen ES. 1997. Dual role for *fmbriata* in regulating floral homeotic genes and cell division in *Antirrhinum*. *EMBO J* 16: 6521-6534.

- Irish VF, Yamamoto YT. 1995. Conservation of floral homeotic gene function between *Arabidopsis* and *Antirrhinum*. *The Plant Cell* 7: 1635-1644.
- Ito T, Suzuki G, Nakada M, Kameya T, Kanno A. 2005. Genomic organization of the *AODEF* gene in *Asparagus officinalis* L. *Genes Genet Syst* 80: 95-103.
- Jack T, Brockman LL, Meyerowitz EM. 1992. The homeotic gene *APETALA3* of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. *Cell* 68: 683-697.
- Jack T, Fox GL, Meyerowitz EM. 1994. *Arabidopsis* homeotic gene *APETALA3* ectopic expression: transcriptional and posttranscriptional regulation determine floral organ identity. *Cell* 76: 703-716.
- Jacob F, Monod J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* 3: 318-389.
- Jacob F, Perrin D, Sanchez C, Monod J. 1960. The operon: a group of genes whose expression is coordinated by an operator. *Comptes rendus des scéances de l'Academie des Sciences* 250: 1727-1729.
- Jareborg N, Birney E, Durbin R. 1999. Comparative analysis of non-coding regions of 77 orthologous mouse and human gene pairs. *Genome Research* 9: 815-824.
- Jenner RA, Wills MA. 2007. The choice of model organisms in evo-devo. *Nat Rev Genet* 8: 311-319.
- Jeong S, Rokas A, Carroll SB. 2006. Regulation of body pigmentation by the abdominal-b hox protein and its gain and loss in *Drosophila* evolution. *Cell* 125: 1387-1399.
- Jeong S, Rebeiz M, Andolfatto P, Werner T, True J, Carroll SB. 2008. The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. *Cell* 132: 783-793.
- Jordan IK, McDonald JF. 1998. Interelement selection in the regulator region of the *copia* retrotransposon. *Journal of Molecular Evolution* 47: 670-676.
- Kafri RA, Bar-Even A, Pilpel Y. 2005. Transcription control reprogramming in genetic backup circuits. *Nat Genet* 37: 295-299.
- Kang JY, Choi HI, Im MI, Kim SY. 2002. *Arabidopsis* basic leucine zipper proteins that mediate stress-responsive abscisic acid signaling. *Plant Cell* 14: 343-357.

- Kanno A, Saeki H, Kameya T, Saedler H, Theissen G. 2003. Heterotopic expression of class B floral homeotic genes supports a modified ABC model for tulip (*Tulipa gesneriana*). *Plant Molecular Biology* 52: 831-841.
- Kaplan DR, Hagemann W. 1991. The relationship of cell and organism in vascular plants: are cells the building-blocks of plant form? *BioScience* 41: 693-703.
- Kaplan DR. 2001. The science of plant morphology: definition, history, and role in modern biology. *American Journal of Botany* 88: 1711-1741.
- Kasowski M, Grubert F, Heffelfinger C, Hariharan M, Asabere A, Waszak SM, Habegger L, Rozowsky J, Shi M, Urban AE, Hong MY, Karczewski KJ, Huber W, Weissman SM, Gerstein MB, Korbel JO, Snyder M. 2010. Variation in transcription factor binding among humans. *Science* 328: 232-235.
- Kaufmann K, Melzer R, Theissen G. 2005. MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. *Gene* 347: 183-198.
- Khaitovich P, Enard W, Lachmann M, Pääbo S. 2006. Evolution of primate gene expression. *Nat Rev Genet* 7: 693-702.
- Khan MR, Hu J, Ali GM. 2012. Reciprocal loss of CArG-boxes and Auxin Response Elements drives expression divergence of *MPF2*-like MADS-Box genes controlling calyx inflation. *PLOS ONE* 7(8): e42781.
- Kim S, Koh J, Yoo MJ, Kong H, Hu Y, Ma H, Soltis PS, Soltis DE. 2005. Expression of floral MADS-box genes in basal angiosperms: implications for the evolution of floral regulators. *The Plant Journal* 43: 724-744.
- King MC, Wilson AC. 1975. Evolution at two levels in humans and chimpanzees. *Science* 188: 107-116.
- Klug W, Cummings M, Spencer C. 2006. *Conceptos de Genética*. Pearson Educación, S.A. Madrid, España. 884 p.
- Kobayashi F, Maeta E, Terashima A, Kawaura K, Ogihara Y, Takumi S. 2008. Development of abiotic stress tolerance via bZIP-type transcription factor LIP19 in common wheat. *J Exp Bot* 59: 891-905.
- Koch MA, Weisshaar B, Kroymann J, Haubold B, Mitchell-Olds T. 2001. Comparative genomics and regulatory evolution: conservation and function of the *Chs* and *Apetal3* promoters. *Mol Biol Evol* 18: 1882-1891.

- Kozaki A, Hake S, Colasanti J. 2004. The maize ID1 flowering time regulator is a zinc finger protein with novel DNA binding properties. *Nucleic Acids Res* 12: 1710-1720.
- Kramer EM, Irish VF. 1999. Evolution of genetic mechanisms controlling petal development. *Nature* 399: 144-148.
- Kramer EM, Di Stilio VS, Schluter PM. 2003. Complex patterns of gene duplication in the *APETALA3* and *PISTILLATA* lineages of the Ranunculaceae. *Int J Plant Sci* 164: 1-11.
- Kramer EM. 2009. New model systems for the study of developmental evolution in plants. *Curr Top Dev Biol* 86: 67-105.
- Krizek BA, Meyerowitz EM. 1996. The *Arabidopsis* homeotic genes *APETALA3* and *PISTILLATA* are sufficient to provide the B class organ identity function. *Development* 122: 11-22.
- Lamb RS, Hill TA, Tan QKG, Irish VF. 2002. Regulation of *APETALA3* floral homeotic gene expression by meristem identity genes. *Development* 129: 2079-2086.
- Latchman DS. 2008. *Eukaryotic Transcription Factors*. Academic Press. Londres, Reino Unido. 448 p.
- Laubichler MD. 2007. *Evolutionary developmental biology*. En: Hull DL, Ruse M (eds.) *The Cambridge companion to the philosophy of biology* (pp. 342-360). Cambridge University Press. Cambridge, Inglaterra.
- Lee JY, Baum S, Alvarez J, Patel A, Chitwood DH, Bowman JL. 2005. Activation of CRABS CLAW in the nectaries and carpels of *Arabidopsis*. *The Plant Cell* 17: 25-36.
- Leoni C, Volpicella M, Placido A, Gallerani R, Ceci LR. 2009. Application of a genome walking method for the study of the spinach *Lhcb 1* multigene family. *Journal of Plant Physiology* 167: 138-143.
- Leoni C, Volpicella M, De Leo F, Gallerani R, Ceci LR. 2011. Genome walking in eukaryotes. *FEBS Journal* 278: 3953-3977.
- Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S. 2002. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Res* 30: 325-327.

- Levin JZ, Meyerowitz EM. 1995. UFO: an *Arabidopsis* gene involved in both floral meristem and floral organ development. *The Plant Cell* 7: 529-548.
- Lewin B. 2008. *Genes IX*. Jones & Bartlett Publishers. Sudbury, Massachusetts. 892 p.
- Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K. 1998. Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell* 10: 1391-1406.
- Liu Y, Yin J, Xiao M, Mason AS, Gao C, Liu H, Li J, Fu D. 2013. Characterization of structure, divergence and regulation patterns of plant promoters. *Journal of Molecular Biology Research* 3: 23-36.
- Love AC, Raff RA. 2003. Knowing your ancestors: themes in the history of evo-devo. *Evolution & Development* 5: 327-330.
- Lynch M. 2006. The origins of eukaryotic gene structure. *Mol Biol Evol* 23: 450-468.
- Márquez Guzmán J, Engleman M, Martínez Mena A, Martínez E, Ramos CH. 1989. Anatomía reproductiva de *Lacandonia schismatica* (Lacandoniaceae). *Annals of the Missouri Botanical Garden* 76: 124-127.
- Martínez ER, Ramos CH. 1989. Lacandoniaceae (Triuridaceae): una nueva familia de México. *Annals of the Missouri Botanical Garden* 76: 128-135.
- Masiero S, Colombo L, Grini PE, Schnittger A, Kater MM. 2011. The emerging importance of type I MADS box transcription factors for plant reproduction. *The Plant Cell* 23: 865-872.
- Mayr E. 1982. *The growth of biological thought: diversity, evolution and inheritance*. The Belknap Press of Harvard University Press. Cambridge, Inglaterra. 974p.
- McCarthy R, Zhong R, Y ZH. 2011. Secondary wall NAC binding elements (SNBE), a key *cis*-acting element required for target gene activation by secondary wall NAC master switches. *Plant Signal Behav* 6: 1282-1285.
- McGinnis W, Garber RL, Wirz J, Kuroiwa A, Gehring WJ. 1984. A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. *Cell* 37: 403-408.

- McGregor AP, Orgogozo V, Delon I, Zanet J, Srinivasan DG, Payre F, Stern DL. 2007. Morphological evolution through multiple *cis*-regulatory mutations at a single gene. *Nature* 448: 587-590.
- McMahon AP, Ingham PW, Tabin CJ. 2003. Developmental roles and clinical significance of hedgehog signaling. *Curr Top Dev Biol* 53: 1-114.
- Melo A, Alves M. 2012. The discovery of *Lacandonia* (Triuridaceae) in Brazil. *Phytotaxa* 40: 21-25.
- Melzer R, Theissen G. 2011. MADS and more: transcription factors that shape the plant. En: Yuan L, Perry SE (eds.) *Plant Transcription Factors*. Methods in Molecular Biology, vol. 754.
- Messenguy F, Dubois E. 2003. Role of MADS box proteins and their cofactors in combinatorial control of gene expression and cell development. *Gene* 316: 1-21.
- Meyerowitz E. 2002. Plants compared to animals: the broadest comparative study of development. *Science* 295: 1482-1485.
- Morgenstern B, Frech K, Dress A, Werner T. 1998. DIALIGN: finding local similarities by multiple sequence alignment. *Bioinformatics* 14: 290-294.
- Müller GB, Wagner GP. 1991. Novelty in evolution: restructuring the concept. *Annu Rev Ecol Syst* 22: 229-256.
- Müller GB. 2007. Evo-devo: extending the evolutionary synthesis. *Nature Reviews Genetics* 8: 943-949.
- Müller GB. 2010. Epigenetic innovation. En: Pigliucci M, Müller GB (eds.) *Evolution – The Extended Synthesis*. MIT Press.
- Murray MG, Thompson WF. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* 8: 4321-4325.
- Nadeau JH, Sankoff D. 1997. Comparable rates of gene loss and functional divergence after genome duplications early in vertebrate evolution. *Genetics* 147: 1259-1266.
- Nakamura M, Tsunoda T, Obokata J. 2002. Photosynthesis nuclear genes generally lack TATA-boxes: a tobacco photosystem I gene responds to light through an initiator. *Plant J* 29: 1-10.
- Nakano T, Suzuki K, Fujimura T, Shinshi T. 2006. Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiol* 140: 411-432.

- Nagasawa N, Miyoshi M, Sano Y, Satoh H, Hirano H, Sakai H, Nagato Y. 2003. *SUPERWOMAN1* and *DROOPING LEAF* genes control floral organ identity in rice. *Development* 130: 705-718.
- Nam J, dePamphilis CW, Ma H, Nei M. 2003. Antiquity and evolution of the MADS-box gene family controlling flower development in plants. *Mol Biol Evol* 20: 1435-1447.
- Ng M, Yanofsky MF. 2001. Function and evolution of the plant MADS-box gene family. *Nat Rev Genet* 2: 186-195.
- Odom DT, Dowell RD, Jacobsen ES, Gordon W, Danford TW, MacIsaac KD, Rolfe PA, Conboy CM, Gifford DK, Fraenkel E. 2007. Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat Genet* 39: 730-732.
- Olds LC, Sibley E. 2003. Lactase persistence DNA variant enhances lactase promoter activity *in vitro*: functional role as a *cis*-regulatory element. *Human Molecular Genetics* 12: 2333-2340.
- Olsen AN, Ernst HA, Leggio LL, Skriver K. 2005. NAC transcription factors: structurally distinct, functionally diverse. *TRENDS in Plant Science* 10: 1360-1385.
- Page RB, Boley MA, Smith JJ, Putta S, Voss SR. 2010. Microarray analysis of a salamander hopeful monster reveals transcriptional signatures of paedomorphic brain development. *BMC Evolutionary Biology* 10: 199.
- Parcy F, Nilsson O, Busch MA, Lee I, Weigel D. 1998. A genetic framework for floral patterning. *Nature* 395: 561-566.
- Pařenicová L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, Cook HE, Ingram RE, Kater MM, Davies B, Angenent GC, Colombo L. 2003. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: new openings to the MADS world. *The Plant Cell* 15: 1538-1551.
- Peichel CL, Nereng KS, Ohgi KA, Cole BL, Colosimo PF, Buerkle CA, Schluter D, Kingsley DM. 2001. The genetic architecture of divergence between threespine stickleback species. *Nature* 414: 901-905.
- Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF. 2000. B and C floral organ identity functions require *SEPALLATA* MADS-box genes. *Nature* 405: 200-203.

- Pelaz S, Tapia-López R, Álvarez-Buylla ER, Yanofsky MF. 2001. Conversion of leaves into petals in *Arabidopsis*. *Current Biology* 11: 182-184.
- Pellegrini L, Tan S, Richmond TJ. 1995. Structure of serum response factor core bound to DNA. *Nature* 376: 490-498.
- Pigliucci M. 2008. What, if anything, is an evolutionary novelty? *Philosophy of Science* 75: 887-898.
- Piñeyro Nelson AA. 2013. De los mecanismos moleculares a la morfogénesis, el desarrollo y la evolución: el caso de la flor homeótica de *Lacandonia schismatica*. Tesis de doctorado. Universidad Nacional Autónoma de México.
- Piñeyro-Nelson A, Flores-Sandoval E, Garay-Arroyo A, García-Ponce B, Álvarez-Buylla ER. 2010. Development and evolution of the unique floral organ arrangement of *Lacandonia schismatica*. *Int J Plant Dev Bio* 4: 86-97.
- Pitzschke A, Djamei A, Teige M, Hirt H. 2009. VIP1 response elements mediate mitogen-activated protein kinase 3-induced stress expression. *Proc Nat Acad Sci USA* 106: 18414-18419.
- Prager EM, Wilson AC. 1975. Slow evolutionary loss of the potential for interspecific hybridization in birds: a manifestation of slow regulatory evolution. *Proc Nat Acad Sci USA* 72: 200-204.
- Puruggannan MD, Rounsley SD, Schmidt RJ, Yanofsky MF. 1995. Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. *Genetics* 140: 345-356.
- Pyo H, Demura T, Kufuda H. 2007. TERE, a novel *cis*-element responsible for a coordinated expression of genes related to programmed cell death and secondary wall formation during differentiation of tracheary elements. *Plant J* 51: 955-965.
- Ramlov KB, Laursen NB, Stougaard J, Marcker KA. 1993. Site-directed mutagenesis of the organ-specific element in the soybean leghemoglobin *lbc3* gene promoter. *Plant J* 4: 577-580.
- Reydon TAC. 2011. *The arrival of the fittest what?* En: Dieks D, Gonzalez WJ, Hartmann S, Uebel T, Weber M (eds.) *Explanation, Prediction, and Confirmation* (pp. 223-237). Springer, New York.

- Riechmann JL, Wang M, Meyerowitz EM. 1996. DNA-binding properties of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. *Nucleic Acids Research* 24: 3134-3141.
- Rieppel O. 2001. Turtles as hopeful monsters. *Bioessays* 23: 987-991.
- Rijkema AS, Vandenbussche M, Koes R, Heijmans K, Gerats T. 2010. Variations on a theme: changes in the floral ABCs of angiosperms. *Sem Cell Dev Biol* 21: 100-107.
- Robert JS, Hall BK, Olson WM. 2001. Bridging the gap between developmental systems theory and evolutionary developmental biology. *BioEssays* 23: 954-962.
- Rockman MV, Wray GA. 2002. Abundant raw material for *cis*-regulatory evolution in humans. *Mol Biol Evol* 19: 1991-2004.
- Rockman MV, Hahn MW, Soranzo N, Zimprich F, Goldstein DB, Wray GA. 2005. Ancient and recent positive selection transformed opioid *cis*-regulation in humans. *PLOS Biology* 3: e387.
- Rodríguez Mega E, Hernández Marroquín VR. 2013. Relatos breves sobre evolución y regulación genética. *Oikos* 9: 15-18.
- Rodríguez-Mega E, Piñeyro-Nelson A, Gutierrez C, García-Ponce B, Sánchez MP, Zluhan-Martínez E, Álvarez-Buylla E, Garay-Arroyo A. 2015. The role of transcriptional regulation in the evolution of plant phenotype. Manuscrito enviado para su revisión (ver **Apéndice 1**).
- Rogozin IB, Sverdlov AV, Babenko VN, Koonin EV. 2005. Analysis of evolution of exon-intron structure of eukaryotic genes. *Briefings in Bioinformatics* 6: 118-134.
- Ronse de Craene PS. 2003. The evolutionary significance of homeosis in flowers: a morphological perspective. *International Journal of Plant Sciences* 164: 225-235.
- Rose A, Meier I, Wienand U. 1999. The tomato I-box factor LeMYBI is a member of a novel class of myb-like proteins. *Plant J* 20: 641-652.
- Rounsley SD, Ditta GS, Yanofsky MF. 1995. Diverse roles for MADS-box genes in *Arabidopsis* development. *The Plant Cell* 7: 1259-1269.
- Rudall PJ. 2003. Monocot pseudanthia revisited: floral structure of the mycoheterotrophic family Triuridaceae. *International Journal of Plant Sciences* 164: S307-S320.
- Rui Z, Zhen Z. 2010. Cloning and sequence analysis of apetala3-like protein promoter from grapevine (*Vitis lambrusca* x *Vitis vinifera*). No publicado.

- Saghai-Marroof MA, Soliman KM, Jorgensen SA, Allard RW. 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc Nat Acad Sci USA* 81: 8014-8018.
- Sánchez Lara KC. 2013. Clonación, caracterización molecular y análisis comparativo *in silico* de los ortólogos de *LEAFY* y *UNUSUAL FLORAL ORGANS* de *Lacandonia schismatica*. Tesis de licenciatura. Universidad Nacional Autónoma de México.
- Schmidt RJ, Ketudat M, Aukerman MJ, Hoschek G. 1992. *Opaque-2* is a transcriptional activator that recognizes a specific target site in 22-kD zein genes. *Plant Cell* 4: 689-700.
- Schmidt D, Wilson M, Ballester B, Schwalie PC, Brown GD, Marshall A, Kutter C, Watt S, Martinez-Jimenez CP, Mackay S, Talianidis I, Flicek P, Odom DT. 2010. Five vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science* 328: 1036-1040.
- Schrack K, Nguyen D, Karlowski WM, Mayer KFX. 2004. START lipid/sterol-binding domains are amplified in plants and are predominantly associated with homeodomain transcription factors. *Genome Biol* 5: R41.
- Schwartz JH. 1999. Homeobox genes, fossils, and the origin of species. *The Anatomical Record* 257: 15-31.
- Schwarz-Sommer Z, Hue I, Huijser P, Flor PJ, Hansen R, Tetens F, Lönning WE, Saedler H, Sommer H. 1992. Characterization of the *Antirrhinum* floral homeotic MADS-box gene *deficiens*: evidence for DNA binding and autoregulation of its persistent expression through flower development. *EMBO J* 11: 251-263.
- Scott MP, Weiner AJ, Hazelrigg TI, Polisky BA, Pirrotta V, Scalenghe F, Kaufman TC. 1983. The molecular organization of the *Antennapedia* locus of *Drosophila*. *Cell* 35: 763-776.
- Sessa G, Morelli G, Ruberti I. 1997. DNA-binding specificity of the homeodomain-leucine zipper domain. *J Mol Biol* 274: 303-309.
- Shapiro MD, Bell MA, Kingsley DM. 2006. Parallel genetic origin of pelvic reduction in vertebrates. *Proceedings of the Natural Academy of Sciences* 103: 13753-13758.
- Sheppard L, Brunner AM, Krutovskii KV, Rottmann WH, Skinner JS, Vollmer SS, Strauss SH. 2000. A *DEFICIENS* homolog from the dioecious tree black

- cottonwood is expressed in female and male floral meristems of the two-whorled, unisexual flowers. *Plant Physiol* 124: 627-640.
- Shore P, Sharrocks AD. 1995. The MADS-box family of transcription factors. *European Journal of Biochemistry* 229: 1-13.
- Siebert PD, Chenchik A, Kellogg DE, Lukyanov KA, Lukyanov SA. 1995. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Research* 23: 1087-1088.
- Smaczniak C, Immink RGH, Angenent GC, Kaufmann K. 2012. Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. *Development* 139: 3081-3098.
- Soltis PS, Brockington SF, Yoo MJ, Piedrahita A, Latvis M, Moore MJ, Chanderball AS, Soltis DE. 2009. Floral variation and floral genetics in basal angiosperms. *American Journal of Botany* 96: 110-128.
- Sommer RJ. 2009. The future of evo-devo: model systems and evolutionary theory. *Nat Rev Genet* 10: 416-422.
- Spitz F, Furlong EM. 2012. Transcription factors: from enhancer binding to developmental control. *Nature Reviews Genetics* 13: 613-626.
- Steeves TA, Sussex IM. 1989. *Patterns in Plant Development*. Cambridge University Press, Cambridge, 388 p.
- Stern DL, Orgogozo V. 2008. The loci of evolution: how predictable is genetic evolution? *Evolution* 62: 2155-2177.
- Stevens PF. 1991. *Lacandonia schismatica* – a challenge to some recent theories of floral morphogenesis? *Flower Newsletter* 12: 32-33.
- Sucena E, Stern DL. 2000. Divergence of larval morphology between *Drosophila sechellia* and its sibling species caused by *cis*-regulatory evolution of *ovo/shaven-baby*. *Proc Nat Acad Sci* 97: 4530-4534.
- Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* 56: 564-577.
- Taylor M, Treisman R, Garrett N, Mohun T. 1989. Muscle-specific (CArG) and serum-responsive (SRE) promoter elements are functionally interchangeable in *Xenopus* embryos and mouse fibroblasts. *Development* 106: 67-78.

- Taylor S, Hofer J, Murfet II. 2001. *Stamina pistilloida*, the pea ortholog of *Fim* and *UFO*, is required for normal development of flowers, inflorescences and leaves. *Plant Cell* 13: 31-46.
- The Plant List. 2013. Versión 1.1. Publicada en la página de internet theplantlist.org (consultada el 7 de agosto de 2014).
- Theissen G. 2006. The proper place of hopeful monsters in evolutionary biology. *Theory in Biosciences* 124: 349-369.
- Theissen G. 2009a. Homeosis of the angiosperm flower: studies on three candidate cases of saltational evolution. *Paleodiversity* 3: 131-139.
- Theissen G. 2009b. Saltational evolution: hopeful monsters are here to stay. *Theory in Biosciences* 128: 43-51.
- Theissen G, Saedler H. 2001. Floral quartets. *Science* 409: 469-471.
- Theissen G, Melzer R. 2007. Molecular mechanisms underlying origin and diversification of the angiosperm flower. *Annals of Botany* 100: 603-619.
- Tilly JJ, Allen DW, Jack T. 1998. The CArG boxes in the promoter of the *Arabidopsis* floral organ identity gene *APETALA3* mediate diverse regulatory effects. *Development* 125: 1647-1657.
- Tournamille C, Colin Y, Cartron JP, Le Van Kim C. 1995. Disruption of a GATA motif in the *Duffy* gene promoter abolishes erythroid gene expression in *Duffy*-negative individuals. *Nature Genetics* 10: 224-228.
- Tsaftaris AS, Polidoros AN, Pasentsis K, Kalivas A. 2006. Tepal formation and expression pattern of B-class paleo*AP3*-like MADS-box genes in crocus (*Crocus sativus* L.). *Plant Science* 170: 238-246.
- Tsai WC, Kuoh CS, Chuang MS, Chen WH, Chen HH. 2004. Four *DEF-like* MADS-box genes displayed distinct floral morphogenetic roles in *Phalaenopsis* orchid. *Plant and Cell Physiology* 45: 831-844.
- Tsuchiya T, Nakamura I, Kanno A. Deduced promoter sequence of MADS-box gene (*LRDEF*) in *Lilium regale*. No publicado.
- Tuch BB, Galgoczy DJ, Hernday AD, Li H, Johnson AD. 2008. The evolution of combinatorial gene regulation in fungi. *PLoS Biology* 6: e38.

- Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J. 2006. A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* 24: 1298-1301.
- Ulmasov T, Hagen G, Guilfoyle TJ. 1999. Dimerization and DNA binding of auxin response factors. *Plant J* 19: 309-319.
- Valadez Moctezuma E, Kahl G. 2000. *Huellas de ADN en genomas de plantas (teoría y protocolos de laboratorio)*. Quinta edición. Ediciones Mundi-Prensa.
- Velasco Mosquera R. 2005. Marcadores moleculares y la extracción de ADN. *Biotecnología en el sector agropecuario y agroindustrial* 3: 14-18.
- Vergara-Silva F, Martínez-Castilla L, Álvarez-Buylla ER. 2000. MADS-box genes: development and evolution of plant body plans. *Journal of Phycology* 36: 803-812.
- Vergara-Silva F, Espinosa-Matías S, Ambrose BA, Vázquez-Santana S, Martínez-Mena A, Márquez-Guzmán J, Álvarez-Buylla ER. 2003. Inside-out flowers characteristics of *Lacandonia schismatica* evolved at least before its divergence from a closely related taxon, *Triuris brevistylis*. *International Journal of Plant Sciences* 164: 345-357.
- Vergara Silva F. 2009. *La homeosis y la naturaleza de la macroevolución: una perspectiva crítica*. En: Morrone JJ, Magaña P (eds.) *Evolución biológica: una visión actualizada desde la revista Ciencias* (pp. 421-435). Facultad de Ciencias, Universidad Nacional Autónoma de México.
- Wagner GP, Lynch VJ. 2010. Evolutionary novelties. *Current Biology* 20: 48-50.
- Wang ZY, Kenigsbuch D, Sun L, Harel E, Ong MS, Tobin EM. 1997. A Myb-related transcription factor is involved in the phytochrome regulation of an *Arabidopsis* *Lhcb* gene. *The Plant Cell* 9: 491-507.
- Weigel D, Meyerowitz EM. 1993. Activation of floral homeotic genes in *Arabidopsis*. *Science* 261: 1723-1726.
- Wellmer F, Graciet E, Riechmann JL. 2014. Specification of floral organs in *Arabidopsis*. *Journal of Experimental Botany* 65: 1-9.
- West-Eberhard MJ. 2003. *Developmental plasticity and evolution*. Oxford: Oxford University Press.
- West-Eberhard MJ. 2005. Developmental plasticity and the origin of species differences. *PNAS* 102: 6543-6549.

- West-Eberhard MJ. 2008. Toward a revival of Darwin's theory of evolutionary theory. *Philosophy of Science* 75: 899-908.
- Wilkinson MD, Haughn GW. 1995. *UNUSUAL FLORAL ORGANS* controls meristem identity and organ primordial fate in *Arabidopsis*. *The Plant Cell* 7: 1485-1499.
- Williams TM, Selegue JE, Werner T, Gompel N, Kopp A, Carroll SB. 2008. The regulation and evolution of a genetic switch controlling sexually dimorphic traits in *Drosophila*. *Cell* 134: 610-623.
- Willmore KE. 2012. An introduction to evolutionary developmental biology. *Evolution: Education & Outreach* 5: 181-183.
- Wilson AC, Sarich VM, Maxson LR. 1974a. The importance of gene rearrangement in evolution: evidence from studies on rates of chromosomal, protein, and anatomical evolution. *Proc Nat Acad Sci USA* 71: 3028-3030.
- Wilson AC, Maxson LR, Sarich VM. 1974b. Two types of molecular evolution. Evidence from studies of interspecific hybridization. *Proc Nat Acad Sci USA* 71: 2843-2847.
- Wilson MD, Barbosa-Morais NL, Schmidt D, Conboy CM, Vanes L, Tybulewicz VL, Fisher EM, Tavaré S, Odom DT. 2008. Species-specific transcription in mice carrying human chromosome 21. *Science* 322: 434-438.
- Wittkopp PJ, Kalay G. 2011. *Cis*-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics* 13: 59-69.
- Wray GA, Lowe CJ. 2000. Developmental regulatory genes and echinoderm evolution. *Syst Biol* 49: 28-51.
- Wray GA. 2002. Resolving the Hox paradox. *Science* 292: 2256-2257.
- Wray GA, Hahn MW, Abouheif E, Balhoff JB, Pizer M, Rockman MV, Romano LA. 2003. The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol* 20: 1377-1419.
- Wray GA. 2003. Transcriptional regulation and the evolution of development. *International Journal of Developmental Biology* 47: 675-684.
- Wray GA. 2007. The evolutionary significance of *cis*-regulatory mutations. *Nature Reviews Genetics* 8: 206-216.
- Xiong L, Schumaker KS, Zhu JK. 2002. Cell signaling during cold, drought, and salt stress. *The Plant Cell* 14 (suppl): s165-s183.

- Xiong Y, Liu T, Tian C, Sun S, Li J, Chen M. 2005. Transcription factors in rice: a genome-wide comparative analysis between monocots and eudicots. *Plant Mol Biol* 59: 191-203.
- Xu F, Park MR, Kitazumi A, Herath V, Mohanty B, Yu SJ, de los Reyes BG. 2012. *Cis*-regulatory signatures of orthologous stress-associated bZIP transcription factors from rice, sorghum and *Arabidopsis* based on phylogenetic footprints. *BMC Genomics* 13: 497.
- Xu Y, Teo LL, Zhou J, Kumar PP, Yu H. 2006. Floral organ identity genes in the orchid *Dendrobium crumenatum*. *The Plant Journal* 46: 54-68.
- Yamamoto YY, Ichida H, Abe T, Suzuki Y, Sugano S, Obokata J. 2007a. Differentiation of core promoter architecture between plants and animals revealed by LDSS analysis. *Nucleic Acids Research* 35: 6219-6226.
- Yamamoto YY, Ichida H, Matsui M, Obokata J, Sakurai T, Satou M, Seki M, Shinozaki K, Abe T. 2007b. Identification of plant promoter constituents by analysis of local distribution of short sequences. *BMC Genomics* 8: 67.
- Yanagisawa S. 2002. The Dof family of plant transcription factors. *Trends Plant Sci* 7: 555-560.
- Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, Zhaofeng G, Zhiqiang L, Yunfei Z, Xiaoxiao W, Xiaoming Q, Yunping S, Li Z, Xiaohui D, Jingchu L, Xing-Wang D, Zhangliang C, Hongya G, Li-Jia Q. 2006. The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Mol Bio* 60: 107-124.
- Zhang S, Sandal N, Polowick PL, Stiller J, Stougaard J, Fobert PR. 2003. *Proliferating floral organs (Pfo)*, a *Lotus japonicus* gene required for specifying floral meristem determinacy and organ identity, encodes an F-box protein. *Plant J* 33: 607-619.
- Zheng W, Zhao H, Mancera E, Steinmetz LM, Snyder M. 2010. Genetic analysis of variation in transcription factor binding in yeast. *Nature* 464: 1187-1191.
- Zhu JK. 2002. Salt and drought stress signal transduction in plants. *Annu Rev Plant Biol* 53: 247-273.

APÉNDICE 1

The role of transcriptional regulation in the evolution of plant phenotype

Emiliano Rodríguez-Mega¹, Alma Piñeyro-Nelson², Crisanto Gutierrez³, Berenice García-Ponce¹, María de la Paz Sánchez¹, Estephania Zluhan-Martínez¹, Elena R. Álvarez-Buylla¹
and Adriana Garay-Arroyo^{1,3,14}

¹Laboratorio de Genética Molecular, Desarrollo, Evolución y Epigenética de Plantas, Universidad Nacional Autónoma de México, 3^{er} Circuito Exterior junto al Jardín Botánico, Ciudad Universitaria, México, D.F. 04510, Mexico; ²Department of Plant and Microbial Biology, University of California, Berkeley, CA; ³Centro de Biología Molecular Severo Ochoa, CSIC-UAM, Nicolás Cabrera 1, Cantoblanco, 28049, Madrid, Spain.

Keywords: phenotypic evolution, gene regulation, epigenesis, evo-devo, gene regulatory
networks

Grant sponsor: CONACyT 180098, 81433, 180380, 167705, 152649, 147675 and DGAPA UNAM: IN204011-3; IN203214-3; EIN203113-3; IN203814-3. UC Mexus: ECO-IE415, ECO-IE416. BFU2012-34821 (MINECO) to CG and an institutional grant from Fundación Ramón Aceres to CBMSO. ERAB was supported by the Miller Institute for Basic Research in Science, University of California, Berkeley, USA. AGA was supported by a sabbatical fellowship by CONACYT (204001) and by PASPA, UNAM. APN was supported by a Postdoctoral Fellowship by UC-Mexus-CONACYT.

¹⁴ Corresponding author: garay.adriana@gmail.com. Telephone: 56229013

Abstract A growing body of evidence suggests that alterations in transcriptional regulation of key factors are at the basis of an important part of phenotypic evolution documented among species and within populations. While the effects of differential transcriptional regulation in organismal development have been preferentially studied in animal systems, this phenomenon has also been addressed in plants. In this review, we summarize scientific evidence of *cis*-regulatory mutations, *trans*-regulatory changes and epigenetic modifications as molecular events underlying important phenotypic alterations, and thus shaping the evolution of plant development and phenotype. We postulate that a mechanistic understanding of why such molecular alterations have a key role in development, morphology and evolution will have to rely on dynamic models of complex regulatory networks that consider the concerted action of genetic and non-genetic components, and that also incorporate the restrictions underlying the genotype to phenotype mapping process.

Keywords: phenotypic evolution, gene regulation, epigenetics, evo-devo, gene regulatory networks

Introduction

In the early 20th century, the modern evolutionary synthesis came to conciliate the mechanisms of inheritance studied through mendelian genetics with inquiries on the role of natural selection and gradual transition in an attempt to explain the evolution and diversification of life forms. The synthesis generated a strong conceptual framework but failed to satisfactorily address one of the oldest biological problems: how phenotypic variation is generated and modulated within and among species (Stern 2000; Müller 2007; Carroll 2008; Gilbert and Epel 2009; Willmore 2012).

In the modern synthesis, the findings and conceptual contributions of several disciplines such as embryology, paleontology, physiology and developmental biology in general were left aside. These disciplines continued pursuing their research lines, many addressing evolutionary problems, but outside of the realm of the conceptual framework set forth by the modern synthesis (Love and Raff 2003; Willmore 2012). The discovery of DNA as the material basis of inheritance, and the availability of different techniques to manipulate and test the phenotypic response to genetic manipulation marked the start of developmental genetics, first deployed in animal research (Love and Raff 2003; Müller 2007). Developmental genetics tried to attain mechanistic explanations based on molecular phenomena for the morphological and physiological observations, in an effort to bridge the gap between genotype and phenotype. Evolutionary developmental biology (also called evo-devo) was thus born in the beginning of the 1990s as a collective effort to assign development its proper place inside the modern evolutionary theory.

From genes to phenotype and back: the long and winding road

As a result of its recent creation, the core and shared tenets of evo-devo, and the practices that are considered part of its conceptual framework, have remained in flux. Nevertheless, it is of general consensus that one of its central objectives is the understanding of not just which genetic alterations during development influence organismal phenotypic change but how they do so (Raff 2000; Wilkins 2002; West-Eberhard 2003; Carroll 2008; Stern 2010; Wray 2013). In this regard, one of its most important and widely accepted contributions to evolutionary theory is the idea that

alterations in gene regulation represent the most significant causal basis of phenotypic change and evolution, at least in animals (Hoekstra and Coyne 2007).

A significant amount of research effort in evo-devo has focused on studying the regulatory logic of developmental networks that underlie evolutionary novelties (Müller 2007). While it is true that “development is larger than just developmental genetics”, since it is now well known that the phenotype of an organism arises from the interaction between its genotype, the environment and other non-genetic aspects, such as physical forces or chemical fields (Gilbert 2010), gene regulation and intracellular networks of gene interactions are key for cellular differentiation, pattern formation and morphogenesis that give rise to the phenotypic characteristics of multicellular organisms (Carroll 2008; Wray et al. 2003; Wray 2007; Wittkopp and Kalay 2012).

This view is not novel at all. Several cornerstone manuscripts published from the 1960s onward provided strong theoretical support to the observations that changes in regulatory sequences should be considered as an important element in the genetic component of the evolutionary process (Jacob and Monod 1961; Wallace 1963; Britten and Davidson 1969, 1971; King and Wilson 1975; Jacob 1977). Since then, molecular changes in regulatory sequences that can modify gene expression patterns started to be seen as more likely events involved in creating phenotypic variation than coding mutations (Carroll 2005; 2008). This was based in two main observations: 1) proteins from different species exhibited neutral mutations over time without any apparent effect on function and 2) proteins involved in gene regulation –such as transcription factors– exhibited a high structural and functional conservation between distantly related species (Zuckermandl 1968; King and Wilson 1975). Today, a growing body of evidence suggests that alterations in spatial and temporal patterns of gene expression can be associated with phenotypic change.

Nevertheless, the phenotype not only evolves by means of regulatory changes in the spatio-temporal regulation of transcription, but also due to changes in the mRNA, the miRNA or protein sequences, and both types of alterations have demonstrated to contribute in different levels to the generation of new phenotypes (Wittkopp et al. 2004; Rieseberg and Blackman 2010; Gruber et al. 2012; Meiklejohn et al. 2014).

Studying multicellular organisms poses a conundrum: if all the constituent cells share the same genome, why do they only synthesize the genetic products related with the

tissue type they belong to? A partial answer to this issue is provided by the observation that diverse cell lineages have differential gene expression profiles: each cell type expresses only a certain percentage of its genome in a specific space and time and such expression is inherited by subsequent generations of cells from a particular tissue, delimiting different body parts and changing in response to environmental stimuli (Latchman 2008). But the latter observation does not provide a mechanism, only a description and correlation of patterns.

More recently, a view that integrates different components and levels of regulation into formal and computational models of complex networks is emerging. Such approach aims at understanding the mechanistic basis of development and evolutionary change, by further comprehending how the genotype maps to, and feedbacks from the phenotype, and the environment (West-Eberhard 2003; Álvarez-Buylla et al. 2007; Benítez et al. 2013). Previous theoretical studies had already proposed that the emergent stable configurations of multi-stable complex regulatory networks underlie the process of cell differentiation (Kauffman 1969), thus solving the apparent paradox of multicellularity. But this model was dismissed because it was based on randomly generated theoretical networks; it was not until 1998 that the first regulatory network grounded on experimental data was proposed (Mendoza and Álvarez-Buylla 1998).

This area of inquiry has exploded since then, providing a repertoire of regulatory modules that map unto different, also modular, aspects of the phenotype (see for example Álvarez-Buylla et al. 2010a). Within such modules, some transcription factors have key positions, so that alterations in their expression levels or patterns of expression, or in those of their downstream genes, cause the multi-gene configurations or phenotypic state that emerges as a result of such complex regulatory networks to switch into another state. Undertaking mechanistic Gene Regulatory Networks (GRNs) approaches, can help contribute to understanding of systemic aspects of development such as robustness and evolvability of particular genetic modules involved in morphogenesis, and the role that epigenetic phenomena can have in explaining the developmental plasticity observed in studies documenting natural variation of a particular trait.

A complete algorithm for the complex series of events that occur during development to fully understand how the genotype maps unto the phenotype lies still ahead. As a complementary integrative task to understand such mapping, nonetheless, we review here

the cases in which transcriptional regulatory alterations due to different altered regulatory mechanisms have been associated to important phenotypic changes in plants.

This article aspires to set the bases to establish multiple hypotheses concerning how the position of the implied factors within regulatory modules may explain why their alterations are related to a modified phenotype. It is important for us to make clear that the approach we propose, then, does not imply a one-to-one mapping from the genetic to the phenotypic alteration focusing on candidate genes, but rather a network-based analysis in where the position of the altered component with respect to other components in the network explains the emergence of a novel trait. This approach will also help mechanistically and formally frame most of the ongoing discussions concerning the relative importance of different types of molecular alterations during the evolution of plant phenotype.

Complexity of regulatory pathways in plants

Presumably, the different regulatory processes affecting gene activity in eukaryotic cells allowed for an increase in genomic complexity (Beckerman 2005). This is particularly true for organisms like plants. Plant genomes may have one of the most sophisticated genetic regulatory logics of all, a likely consequence of being sessile, nonmotile, organisms whose development plastically responds to environmental conditions. Genome-wide comparative analyses suggest that transcriptional regulation may play a larger role in plants than in animals, since plants use a higher percentage of their genome to codify transcriptional regulators than animals or fungi (Riechmann et al. 2000; Gong et al. 2004; Xiong et al. 2005). An example that illustrates the contrasting strategies deployed by plants and animals to fine-tune gene expression and protein turnover is the case of the MADS-box family evolution in plants. Studies addressing the role of genomic duplications in plant evolution have highlighted that several whole genome duplications predate the diversification of plant lineages such as the angiosperms (Soltis et al. 2009). Whole genome duplications are now recognized as common phenomena that could have impelled lineage diversification through the inauguration of novel molecular interactions among homolog genes; differential selection of some duplicates involved in novel traits and, importantly, modifications of the underlying GRNs. Additionally, gene duplication events are also common, particularly in genes that control developmental processes

throughout the plant's life cycle, as is the case for the MADS-box genes (Martínez and Álvarez-Buylla 2003). The MADS-box gene family is composed by transcription factors that have a canonical DNA-binding domain of approximately 58 amino acids that binds to target DNA in consensus sequences called CArG boxes (Álvarez-Buylla et al. 2000; Parenicová et al. 2003). While MADS-box genes already existed in the common ancestor of animals and plants (Álvarez-Buylla et al. 2000), they have greatly diversified and expanded in plants, with over 100 type II MADS-box genes documented for *A. thaliana* (Martínez and Álvarez-Buylla 2003). Work in additional angiosperm species has shown that this is a general trend where the genes product of duplication events can either become redundant, be lost, subfunctionalize or neofunctionalize (Airoldi and Davies 2012); these molecular processes commonly have a direct impact on plant morphogenesis.

In the absence of any kind of migratory behavior, immune response or nervous system, plants display a range of phenotypes when faced with external biotic or abiotic agents and continuously adjust their development and morphology to cope with surrounding conditions (Schlichting 1986). This is known as phenotypic plasticity (West-Eberhard 2003; Fordyce 2006), and it is believed to drive –or at least facilitate– evolutionary change through its ecological effects (Price et al. 2003; Fordyce 2006; Pigliucci et al. 2006; Ghalambor et al. 2007; Whitman and Agrawal 2009).

How to link phenotypic plasticity with its underlying molecular mechanisms is a question that had been addressed since the early 1990s by Smith (1990), Schlichting and Pigliucci (1993; 1995), who proposed that differential control of gene expression represented a “simple means for generating altered, yet still functional, phenotypes”. Years before, Schlichting (1986) had implied that specific plastic responses in plants were determined by changes in the genetic bases controlling development.

Indeed, developmental plasticity is mainly the result of paramount pluripotent activity in stem cell populations that will either differentiate into organ types or produce more stem cells within plant meristems (Kaufmann et al. 2010). Two main meristems are observed in plants: the Root Apical Meristem (RAM) localized in the root tip, that generates all underground structures, and the Shoot Apical Meristem (SAM), able to produce stems and leaves, and also inflorescences and flowers when it switches from a vegetative meristem to a reproductive one. Given that plants show an important postembryonic growth phase throughout their lifetime, these kind of developmental

switches govern most organismal traits through changes in complex GRN, where key transcriptional regulators orchestrate the expression of target genes involved in specific physiological processes, developmental pathways and cell type identities (Lemon and Tjian 2000; Kaufmann et al. 2010). Alterations in the interactions of a GRN can cause a switch from one stable state to another one, and thus underlie a plastic or permanent phenotypic alteration (Davidson and Erwin 2006; Espinosa-Soto et al. 2004; Álvarez-Buylla et al. 2007; Peter and Davidson 2011; Benítez et al. 2013; see Figure 2).

Much emphasis has been placed on studying changes in transcriptional regulation that account for shifts in gene expression patterns and, ultimately, phenotypic variation (Carroll 2005, 2008; Stern and Orgogozo 2008; Wagner and Lynch 2008). We would like to point out, nonetheless, that gene expression is altered at many other levels such that changes in additional regulatory mechanisms –alternative splicing, post-transcriptional regulation and translational regulation– (Cohen and Mayfield 1997; Alonso and Wilkins 2005; Simpson et al. 2010) require a specific review. In the forthcoming sections we will encompass empirical evidence documenting different instances where altered transcriptional regulation that underlie phenotypic evolution.

Different levels of gene transcriptional regulation in plants

An important conceptual contribution from the early proponents of evo-devo research is the notion that genetic networks likely underlying different developmental processes display a modular organization (Brakefield 2006). Such modularity occurs at two different levels: 1) genes encompassed in a particular genetic module involved in the development of a given structure/organ/trait can evolve in a semi-independent fashion from other modules; and 2) each gene has a modular structure (comprised of *cis*-regulatory sequences, exons, introns) that, in itself, could be modified in a modular fashion such that mutations occurring in the non-coding or coding sections can be independent from one another (Brakefield 2006).

From a broad genetic perspective, two basic units of transcriptional regulation can be identified: *trans*-acting factors and *cis*-acting elements (both *trans*- and *cis*- are latin prefixes which mean “on the other side of” and “on this side of”, respectively; see Figure 1). These units coordinate the spatiotemporal regulation of genetic expression, and work in a rather cooperative manner. *Trans*-acting factors include diffusible genetic products

like regulatory proteins (*i.e.* transcription factors, hormones and such) and noncoding regulatory RNAs acting on distant target genes through the recognition of small allele-specific non-diffusible DNA sequences called *cis*-acting elements which are, in turn, embedded in noncoding *cis*-regulatory regions: promoters, enhancers, silencers, UTRs, as well as in intronic sequences (Stern and Orgogozo 2008; Wagner and Lynch 2008; Emerson and Li 2010), although in rare cases they have also been discovered in exons (Neznanov 1997; Sandrelli et al. 2001; see Figure 1). According to their physical position in relation to the gene they regulate, they can be divided into local and distant elements (Emerson and Li 2010). These *cis*-acting elements, which work as binding sites, maintain transcriptional activity in a defined expression domain and are either involved in the basic process of transcription or in mediating the response to a particular stimulus (Wray et al. 2003; Latchman 2008).

Experimental studies have shown that sequence changes in binding sites affect the affinity to which transcription factors bind to DNA (Gompel et al. 2005; Jeong et al. 2006; Borneman et al. 2007; Odom et al. 2007; Tuch et al. 2008; Williams et al. 2008; Wilson et al. 2008; Schmidt et al. 2010; Zheng et al. 2010; Bradley et al. 2010; Kasowski et al. 2010). This is not trivial, since regulatory regions integrate information about the cell's developing state depending on which transcription factors bind to them (Latchman 2008). A eukaryotic genome, for example, could have as much as 5×10^4 genes that need to be differentially expressed in order to develop into a specific cell type during ontogeny and to respond to several environmental stimuli during its lifespan (e.g. White et al. 1999; Iyer et al. 2001; Kayo et al. 2001; Mody et al. 2001; Arbeitman et al. 2002).

Successful expression of a given gene is controlled by interplay among *trans*-acting factors, which recognize and bind specific *cis*-acting modules in regulatory regions, and subsequently recruit other protein intermediates in order to form molecular loops that bring closer all the elements of the transcriptional machinery necessary to activate gene expression. These protein interactions are complex, nonlinear and often strongly context dependent (Espinosa-Soto et al. 2004; Álvarez-Buylla et al. 2007; Benítez et al. 2013; Wray 2007).

Previous reviews have identified at least four main mechanisms capable of altering gene activity, and potentially transforming the architectural topology of GRNs that underlie phenotypic traits: 1) epigenesis or chromatin level regulation of transcriptional

expression without changes in the DNA sequence; 2) nonsynonymous coding mutations changing amino acid or mature mRNA sequences, although there are few studies showing that synonymous mutations may contribute to phenotypic evolution too (Stam and Laurie 1996; Nackley et al. 2006); 3) *cis*-regulatory mutations, where changes in the nucleotides of binding sites located on the same DNA strand of the gene they regulate can modify gene expression patterns; and 4) copy number mutations, such as gene loss and duplications, involving both coding and *cis*-regulatory changes (Stern and Orgogozo 2008; Peter and Davidson 2010; Gruber et al. 2012).

Altering GRNs by modifying the epigenetic landscape, mutating *cis*-regulatory elements or changing the structure of *trans*-regulatory factors may result in the generation of novel phenotypes in plants (Figure 1; Barrio et al. 2013). To what extent regulatory or coding mutations contribute to phenotypic evolution, however, is an issue that has not been satisfactorily answered yet, as stated above (Wray 2007; Hoekstra and Coyne 2007; Carroll 2008; Stern and Orgogozo 2008). In the case of plants, integrating the cases in which such types of alterations have been associated to a phenotypic change will contribute towards this aim.

Phenotypic changes associated to epigenetic variation

Epigenetics refers to stable heritable changes in gene function without altering the DNA sequence (Riggs and Porter 1996; Berger et al. 2009) through different mechanisms such as DNA methylation, synthesis of non-coding RNAs and histone modifications (e.g. methylation, acetylation, ubiquitination, phosphorylation and sumoylation). These epigenetic modifications constitute the epigenetic landscape that regulates chromatin states from a compact to a relaxed status and vice versa (Roudier et al. 2011; Sequeira-Mendes et al. 2014; see Figure 1) to adjust the interaction of transcriptional factors or DNA replication proteins to DNA. Recently, there has been an enormous amount of information regarding the fundamental role of epigenetics during the development and growth of living organisms; and *Arabidopsis thaliana* has been one of the most important model systems used (Desvoyes et al. 2014; Grimanelli and Roudier 2013; Law and Jacobsen 2010).

Plant DNA methylation occurs at any cytosine residue and is catalyzed by DNA methyltransferases such as MET1, CMT3, DRM1 and DRM2 (Bartee et al. 2001;

Lindroth et al. 2001; Cao and Jacobsen 2002; Kankel *et al.* 2003), which can promote DNA methylation after DNA replication or as *de novo* in small regions of the genome. Therefore, DNA methylation is crucial to genomic integrity maintenance and gene expression regulation (Furner and Matzke 2011). In this regard, it has been proposed that DNA methylation protects the host genome through silencing of transposable elements (TEs), maintaining centromeric and heterochromatin regions, and silencing imprinting genes (Hirochika et al. 2000; Lippman et al. 2004; Kohler and Grossniklaus 2005; Zhang et al. 2008). The silencing of a subset of transposons and genes by *de novo* DNA methylation can be mediated by the RNA-directed DNA methylation pathway, which requires small interfering RNAs (siRNAs) (Matzke and Mosher 2014). Unlike transposons, DNA methylation of promoter regions normally inhibits transcription, but methylation in coding regions does not generally affect gene expression, although some genes contain regulatory elements in their coding regions, which are also methylated when such genes are silenced (e.g. *SUPERMAN* and *AGAMOUS*).

DNA methylation is a stable although reversible epigenetic mark that regulates gene expression during plant development and is heritable across generations. Nonetheless, adaptation processes, external signals or intrinsic factors may induce erasure or modification of this mark at one or more points during plant development, and the inheritance acquired may be modified. This transgenerational epigenetic inheritance in plants usually occurs in response to environmental cues and adaptation processes, generating natural epialleles (Richards 2008).

Now we know that natural variation exists not only at the DNA sequence level but also at the epigenetic level (e.g. Vaughn et al. 2007; Herrera and Bazaga 2010). This may be particularly common in plants, and several studies suggest that natural epialleles can cause significant heritable variation that affects phenotypic traits. *Linaria vulgaris*, a plant described by Linneaus that could naturally produce two flower morphs: the most common; a bilaterally symmetrical flower but on occasions, flower with radial symmetry would develop. *L. vulgaris* was the first documented example of the existence of epialleles that produce morphological alterations depending on the DNA methylation state of the *Lcyc* gene (Figure 1A; Table 1). *Lcyc* is an homolog to the *CYCLOIDEA* gene from *Antirrhinum majus* and is involved in dorsoventral asymmetry, therefore, radial flowers are formed when *Lcyc* is silenced through DNA methylation (Cubas et al. 1999b; Luo et al.

1996). These morphological alterations can have ecological and evolutionary implications since changes in the shape and scent of a flower could alter its pollination syndrome, affecting its reproductive success.

The production of male or female or hermaphrodite flowers in melon (*Cucurbita melo*) is another example where sex determination depends on epigenetic variation. Melon plants are monoecious producing either unisexual flowers, or both unisexual and hermaphrodite flowers. This feature is correlated with the silencing of the gene *CmWIPI* through DNA methylation (Table 1). This gene is required to form stamens, therefore development of female flowers is induced when *CmWIPI* is silenced (Martin et al. 2009). Plants with unisexual flowers inhibit self-fertilization, generating genetic diversity (Turck and Coupland 2014).

Finally, experimental DNA demethylation experiments in *Arabidopsis thaliana* suggest that DNA methylation strongly affects ecologically important plant traits such as mortality, flowering time, plant height, biomass, and very importantly, phenotypic plasticity (Bossdorf et al. 2010). In non-model plants like the mangrove *Laguncularia racemosa*, natural morphological changes have been correlated with variation in methylation patterns (Lira-Medeiros et al. 2010). Populations of this plant grow in contrasting habitats where nutrients and salinity levels vary considerably, leading to significant differences in tree height, tree diameter, leaf width and leaf area. Lira-Medeiros and collaborators (2010) documented divergent epigenetic profiles between mangrove populations, showing that the genome from plants growing in salt marshes is hypomethylated in comparison with those growing in riversides. CpG-methylation changes, thus, may be associated with environmental heterogeneity and phenotypic plasticity between natural populations or ecotypes, as suggested by similar studies in other forest tree species, such as poplar and eucalypt (reviewed by Bräutigam et al. 2013).

Phenotypic alterations in flower development

A central theme in plant evo-devo has been the study of the development of flowers, since these structures are good model systems to investigate the genetic underpinnings and evolutionary consequences of phenomena such as: homeotic alteration of organ identity, changes in size and form of organs as well as modification of other characteristics like color, symmetry and scent. The effect of individual genes whose activity affects flower

organ development as well as flowering time have been studied in a subset of angiosperms, while their interactions in higher order gene regulatory networks have yielded important insights in floral evolution. In the next section we illustrate some of the salient findings in this area of research.

ABC model - Flowers are one of the defining structures of angiosperms and hermaphrodite flowers are organized in four different concentric whorls arranged from the outside to the inside in sepals (se), petals (pe), stamens (st) and carpels (ca). In the late 1980s, the knowledge attained through the study of floral homeotic mutants in the model plant species *Arabidopsis thaliana* and *Antirrhinum majus* enabled the proposal of the ABC model of floral organ specification (Bowman et al. 1991; Coen and Meyerowitz 1991). The ABC model integrated information from three classes of floral mutants, each one of them lacking adequate expression of a particular gene. This translated into homeotic transformations where two floral whorls were replaced: in “A” mutants, flowers were composed by ca, st, st and ca; in “B” mutants, flowers had se, se, ca and ca; and in “C” mutants, flowers had se, pe, pe and se, as well as an indeterminate floral meristem (Bowman et al. 1991). From previous observations, the functions of these genes were inferred and organized in the ABC model as follows: sepals are specified by the A function alone (genes *APETALA1* and *APETALA2*; *AP1* and *AP2*), petals require a combination of A and B functions (genes *APETALA3* and *PISTILLATA*; *AP3* and *PI*), stamens B and C functions (gene *AGAMOUS*; *AG*), and carpels C function alone. Additionally, A and C function genes mutually repress each other (Coen and Meyerowitz 1991).

The ABC model rendered a useful heuristic framework that helped establish a link between experimental genetics and evo-devo of flower formation, thus stimulating comparative studies of ABC orthologs and their gene expression patterns and correlations to floral morphology across angiosperms (Saedler et al. 2001; Melzer and Theißen 2011). A mechanistic explanation to the ABC combinatorial model came several years later when a robust GRN that integrated the ABC genes with other interacting components, was able to recover the observed ABC gene configurations documented in wild type and mutant flowers (Mendoza and Álvarez-Buylla 1998; Espinosa-Soto et al. 2004). The overall conservation of the ABC model, across angiosperms is still under debate, while analyses of recent developmental genetics studies in *A. thaliana* raise questions about some of the

main tenets of the ABC model, even in the realm of research conducted in this model species (Prunet and Jack 2014). In this context, GRN approaches will aid in generating testable hypotheses across flowering plants. In contrast, an overall widespread conservation of the ABC genes has been found across distantly related flowering plants; this is probably due to their molecular evolution being restricted by the GRN of which they are part. Such restrictions in a canonical GRN could have been formed during early stages of angiosperm evolution and thus explain the overall conservation of the ABC genes among many angiosperms (Davila et al. 2014).

MADS box genes - The ABC genes encode transcription factors and all but one (*AP2*) belong to the MADS-box gene family (Riechmann and Meyerowitz 1997): *AP1*, *PI*, *AP3* and *AG*. MADS-box genes participate in all major aspects of plant development although they have been mainly studied in specifying floral organ identity (Smaczniak 2012). MADS-box type II proteins possess four domains (MIKC), so they are referred to as MIKC proteins: the M represents an amino terminal domain that binds DNA, I is an intervening region involved in DNA binding and protein-protein interaction, K domain is involved in protein-protein interactions (Yang and Jack 2004), and the C terminal domain is implicated in transactivation and higher order MADS-box protein interactions (Egea-Cortines et al. 1999; Honma and Goto 2001). Since the 1990s, hybrid proteins interchanging MIKC domains were generated between *AP1*, *AP3*, *PI* and *AG*, showing that the MADS domain does not specify for their biological activity while the carboxyl region can contribute to DNA binding specificity (Krizek and Meyerowitz 1996; Riechmann and Meyerowitz 1997). In order to recognize and bind to the CA_nG boxes present on target DNA, MADS proteins form dimers –either homo or heterodimers– and tetrameric protein associations to carry out their function as transcriptional regulators. Furthermore, this gene family has undergone lineage-specific changes, some related with extensive duplication and subfunctionalization of genes through mutations in their coding or non-coding sequences (Saedler et al. 2001). Such mutations have been linked to the origin of morphological novelties, particularly in flowers (Khan et al. 2012; Gramzow and Theißen 2013).

It has been shown that ABC genes are required, but are not sufficient, to specify floral organ identity (Álvarez-Buylla et al. 2010a). They need, at least, the activities of another

four closely related MADS-box genes, *SEPALLATA1/2/3/4* (*SEP1/2/3/4*; dubbed “E” function genes), to develop floral organs; the inclusion of these genes as necessary for floral development resulted in what is known as the “floral quartet” model (Honma and Goto, 2001; Pelaz et al. 2001; Theissen and Saedler 2001). According to the ABCE modified model, in *Arabidopsis*, sepals are specified by the combination of the expression of *AP1/AP2* and *SEP* genes; petals by the expression of *AP1/AP2*, *AP3*, *PI*, and *SEP*; stamens by the expression of *AP3*, *PI*, *AG* and *SEP* and carpels by the expression of *AG* and *SEP* (Pelaz et al. 2000; Figure 2A).

The B class genes (*AP3* and *PI*) have received special attention, as petals are considered important structures for pollinator attraction and, as such, their diverse architecture has been correlated with the adaptive radiations documented in many angiosperm lineages (Irish 2009). Moreover, several duplications have taken place within the *AP3* and *PI* genes throughout the angiosperms (Kim et al. 2004; Hernández-Hernández et al. 2007; Mondragón-Palomino et al. 2009). Functional studies in different angiosperms encompassing basal angiosperms, non-model eudicot and monocot species indicate variations in the expression sites of B genes with implications on flower architecture (Kanno et al. 2007; Mondragón-Palomino and Theissen 2009), as it has been assumed that MADS-box genes exert their function where they are expressed (nonetheless, Urbanus et al. 2010 demonstrated that *AG* is capable of moving through cell layers). In this section we review various examples where changes in the spatial expression of a particular transcription factor (in these cases, MADS-box genes) have a direct impact on flower morphology.

In *Arabidopsis*, *AP3* and *PI* exert their function as obligate heterodimers that bind to DNA (Bowman et al. 1989; Goto and Meyerowitz 1994; Jack et al. 1992). Protein-protein binding occurs via the K domain, which does not show conservation at the primary structure level; nonetheless, the secondary structure is preserved in three amphipathic α -helices referred to as K1, K2, and K3 (Álvarez-Buylla et al. 2000). It has been shown that in the 80 amino acid region covered by the K domain, there are several heptad repeats (abcdefg)_n in which the bold positions are occupied by hydrophobic amino acids necessary to form the α -helices (Gramzow and Theissen 2010; Yang et al. 2003). Yang and collaborators (2003) demonstrated that the ability to form heterodimers between *AP3* and *PI* is fundamental to perform their function during plant development. Specifically, they

found that the critical amino acids for this protein-protein interaction are located in the hydrophobic regions of K1 and K2 domains and the hydrophobic region between them. When these amino acids are mutated, they cause defects in AP3/PI heterodimerization in yeast two-hybrid experiments and, correspondingly, in the phenotype of wild-type and mutant plants overexpressing these proteins in *Arabidopsis*.

In *Papaver somniferum*, a basal eudicot, *AP3* is duplicated into *PapsAP3-1* and *PapsAP3-2*. Each paralog has a different expression pattern. *PapsAP3-1* is highly expressed in petals and stamens throughout floral development and weakly expressed in sepals and carpels during later stages. In contrast, *PapsAP3-2* is initially expressed in stamen primordia (stage 4) and petal primordia (stage 5); but from stage 8 onward, this gene is expressed in all floral organs (Drea et al. 2007). Silencing of *PapsAP3-1* results in homeotic transformation of petals to sepaloid organs, while inhibiting the expression of *PapsAP3-2* causes stamens to be transformed into carpeloid structures; stamen and petal homeotic transformations are observed when both genes are silenced simultaneously (Drea et al. 2007).

In tomato there are two paralogs for *AP3* (*TM6* and *TAP3*) and for *PI* (*TPI* and *TPIB*). De Martino et al. (2006) found that only the *TAP3* loss-of-function mutant shows a complete conversion of petals into sepals and stamens into carpels; the other B loss-of-function mutants (*TM6*, *TPI*, and *TPIB*) only affect stamen development (de Martino et al. 2006; Geuten and Irish 2010). The differences in function of the *AP3* orthologs could be partially attributed to the level of expression of the genes, as overexpression lines confer a similar, but not identical, degree of *tap3* mutant rescue (de Martino et al. 2006).

In *Antirrhinum majus*, there are two paralogs that fulfill the C function (*FARINELLI* and *PLENA*; *FAR* and *PLE*). Overexpression of *FAR* or *PLE* yields different flower morphologies in *Arabidopsis*. *PLE* has the same phenotypic effect as *AG* (a flower with whorls transformed into ca, st, st, ca) and *FAR* converts petals into stamens but does not affect the fate of sepals (Airoldi et al. 2010). Airoldi and collaborators (2010) showed that a single amino acid deletion in the K3 domain (important for protein-protein interaction) of *FAR* (Q173) or its introduction in the K3 domain of *AG* is sufficient to define the reproductive organ fate. The presence of Q173 correlates with male organ development, and its absence with the development of male and female organs (Airoldi et al. 2010). Moreover, these authors also demonstrate the importance of considering protein-protein

interactions in order to explain biological activity. In this case, AG interacts with the MADS-box proteins SEP1, SEP2 and SEP3 in stamen and carpel development while FAR only interacts with SEP3. SEP3 is not expressed at early developmental stages in sepals and the authors showed that the inability of FAR to develop carpels in the first whorl not only depends on its amino acid content but also on the expression pattern of *SEP3*, highlighting the importance of considering the underlying GRNs to understand how such molecular alterations map unto the phenotype.

Lacandonia – *Lacandonia schismatica* (Triuridaceae) is a monocotyledonous species endemic to the Mexican Lacandon rainforest that, along with *Lacandonia brasiliiana*, are the only examples of an angiosperm with bisexual flowers where the stamens are in the center of the flower surrounded by carpels (Martínez and Ramos 1989; Melo and Alves 2012). In *L. schismatica*, this phenotype is fixed in natural populations (Vergara-Silva et al. 2003). Developmental genetics studies have shown that while this plant has only one copy each of the *AP3* and *PI* orthologs, the unique arrangement of the reproductive whorls in *L. schismatica* depends on a the shift of the spatio-temporal expression patterns of *LsAP3* (*AP3* orthologs) to the center of the developing flower primordium (Álvarez-Buylla et al. 2010b). In contrast, the overall expression pattern of the other *L. schismatica* floral homeotic genes that participate in stamen and carpel development (*LsPI* and *LsAG*) is very similar to the expression patterns that have been reported for *Arabidopsis* orthologs (Álvarez-Buylla et al. 2010b). Thus, *L. schismatica* provides an extreme example of how a shift in the physical domain where a gene is expressed can yield a unique phenotype, without obvious modifications in its interacting partners (*LsPI* and *LsAG*). It is still not known if the displacement of *LsAP3* expression is due to changes in the *cis*-regulatory sequences of the gene, while the rest of the GRN is conserved, or it is due to alterations in the *trans*-acting factors that regulate *LsAP3* expression (probably, the transcription factor *LEAFY* (*LFY*) and its co-activator *UNUSUAL FLORAL ORGANS* (*UFO*), as documented in *Arabidopsis*; Chae et al. 2008; Percy et al., 1998; Weigel and Meyerowitz, 1993).

Flowering time - Flowering time is one of the most important events in the life cycle of plants, as it ensures reproductive success. There are winter- and summer-annual

Arabidopsis flowering habit ecotypes that depend on the variation in the expression of *FRI* (*FRIGIDA*) and *FLC* (*FLOWERING LOCUS C*). *FRI* is a positive regulator of *FLC*, a MADS-box gene that represses the expression of floral activators. It has been shown that *FLC* expression is low in summer-annual strains and high in winter-annual *Arabidopsis* plants (Michaels and Amasino 1999; Sheldon et al. 2000). Michaels and collaborators (2003) introduced a functional allele of *FLC* in summer-annual strains (Da (1)-12 and Shakh dara) and were able to restore winter-annual habit. Moreover, they also demonstrated that the low level of expression in the Ler ecotype is caused by an insertion in the first intron of *FLC*.

Novel floral morphologies - Structural change in the gynoecium of flowers, more specifically in the stigma –i.e. the pollen-receiving tip of carpels–, has also been linked to *cis*-regulatory mutations. The tomato genus *Lycopersicon* (tomato group) is composed of nine species that exhibit a wide range of mating systems: from the ones that are exclusively allogamous, forced to cross-pollinate, to the ones that are obligate autogamous and self-pollinated species, including other intermediate states (Rick 1988). Several reports have suggested that stigma length correlates with the evolution of either of these mating systems: allogamous species have exerted stigmas, while inserted stigmas are characteristic of autogamous species. It has been observed that *se2.1*, a major QTL and a putative transcription factor, is responsible for the phenotypic variation of this trait, and some studies have suggested that mutations on this gene may account for the evolution of mating systems in the *Lycopersicon* genus (Bernacchi and Tanksley 1997; Fulton et al. 1997; Tanksley and Loaizafigueroa 1985). Chen and collaborators (2007) reported that the transition from allogamy to autogamy is accompanied by a mutation in the *se2.1* promoter region, which explains its down-regulation during flower development in autogamous species.

Finally, an additional example of evolutionary studies in flower development, comes from the Solanaceae family, where the calyx (i.e. the sepals) exhibits many different morphologies ranging from a small sepal ring in *Tubocapsicum anomalum* to the inflated calyx syndrome (ICS) observed in 11 different species of *Withania* (Khan et al. 2012). In the ICS, sepals resume growth after pollination, encapsulate the mature fruit and give rise to a balloon-like structure. Previous studies have shown that the expression of a MADS-

box transcription factor (*MPF2*, a paralog of the MADS box gene *AGL24* in *Arabidopsis*; Khan et al. 2012) in the calyx is a prerequisite for ICS formation in the species *Physalis pubescens* (He and Saedler 2005, 2007). All *Withania* species have two orthologs of *MPF2*: one that is expressed in vegetative and flower tissues (*MPF2*-like-A) and another one that is only expressed in vegetative tissues (*MPF2*-like-B) (Khan et al. 2009). The expression divergence between paralogs (in this case, *MPF2*-like-A and *MPF2*-like-B) had been correlated with modifications in *cis*-regulatory regions (Chaudhary et al. 2009; Goode et al. 2011). By analyzing the *cis*-acting elements of *Withania* genes, it was found that *MPF2*-like-A regulatory regions lacks an ARE (Auxin Response Element, binding site for Auxin Response Factors) binding site in the 1st intron necessary to suppress the expression in the calyx (Khan et al. 2012). Furthermore, the *MPF2*-like-B promoter is devoid of a CARG-box (binding site for MADS proteins) near the transcriptional start site required for its expression in sepals (Khan et al. 2012). This reciprocal loss of *cis*-regulatory elements is characteristic of divergence events after duplication (Khan et al. 2012).

In contrast to *Withania*, *Tubocapsicum* does not have an *MPF2*-like-A gene and only features an *MPF2*-like-B. This gene is expressed in all floral tissues and has a CARG box element (Khan et al. 2009). The lack of *MPF2*-like-A in *Tubocapsicum* suggests a correlation between ICS formation and *MPF2*-like-A expression in *Withania*. Moreover, overexpression lines of *Withania* *MPF2*-like-A and *MPF2*-like-B genes and *Tubocapsicum* *MPF2*-like-B gene shows that only the overexpression of *MPF2*-like-A produces flowers with a large-sepal phenotype in *Arabidopsis* (Khan et al. 2009). These studies demonstrate that the variation of gene expression in *MPF*-like genes influence ICS formation and underlie the phenotypic variation in sepal development found within the Solanaceae family. The function of the ICS is not clear; however, could facilitate wind dispersal of the fruits (Knapp, 2002), provide a humid microclimate (Khan et al. 2009) or have a protective role in early developmental stages of some organs (He and Saedler, 2007).

LEAFY expression in plant development – *LFY* is a plant specific transcription factor that has been found in mosses, ferns, gymnosperms and angiosperms; in later it regulates flower development (Moyroud et al., 2010). Surprisingly, *LFY* orthologs have two conserved regions in the N-terminal and C-terminal domains in all land plants surveyed to date (Maizel et al. 2005; Moyroud et al. 2009). To test functional conservation, Maizel

and collaborators (2005) isolated *LFY* orthologs from mosses different land plants and introduced them into a *lfy* mutant from *Arabidopsis*, finding a direct correlation between the degree of complementation of the *LFY* phenotype and the phylogenetic distance from angiosperms. The *LFY* ortholog from the moss *Physcomitrella patens* was the only gene unable to rescue the *lfy* mutant (Maizel et al. 2005). In contrast, one amino acid change in the DNA binding domain of the moss gene was sufficient to partially complement a *lfy* *Arabidopsis* mutant phenotype, indicating that the moss ortholog has different DNA binding specificities.

Anthocyanin pathway – Anthocyanins are plant pigments (orange, pink, red, blue and purple) that are present in all tissues of higher plants; these pigments are synthesized via the flavonoid pathway that is highly conserved across angiosperms (Holton & Cornish, 1995). There are three precursors that require at least six different enzymes to produce all anthocyanin pigments. Interestingly, there is a coordinated gene regulation of the different pathways at the level of transcription in all the species studied (Sobel and Streisfeld 2013). The transcriptional complex that regulates gene expression is made up of proteins from three large transcription factor gene families: R2R3-MYB, basic helix-loop-helix (bHLH) and WD40 repeat (Sobel and Streisfeld 2013). It has been shown that the specificity of the complex is determined by the proteins from the R2R3-MYB protein while these transcription factors tend to be tissue specific in such a way that they affect anthocyanin production in a localized manner.

In the following section we will show examples of changes in flower and fruit color that rely either in mutation in regulatory (*cis*) or coding (*trans*) regions in different genes; many of them are in R2R3-MYB transcription factor genes.

Variation in flower color – Reproductive isolation results from prezygotic and postzygotic barriers to prevent gene flow between populations and one of the most important prezygotic barriers is pollinator isolation (Hopkins and Rausher 2012). In recent years, it has been strongly suggested that variation in flower color is somehow related with changes in pollinator behavior (Rieseberg and Blackman 2010; Hopkins 2013).

In an attempt to identify what is the relative importance of regulatory versus coding

mutations in the evolution of flower color, it has been found that regulatory changes play a significant role not only in generating phenotypic variation but also in underlying certain reproductive barriers in plants (Durbin et al. 2003; Rieseberg and Blackman 2010). Des Marais and Rausher (2010), for instance, performed gene expression assays and plant transformations experiments to determine the genetic basis of color shift in the hummingbird-pollinated flowers of two morning glory species: *Ipomoea quamocitl* and *Ipomoea coccinea*. Both of them belong to a monophyletic group called the *Mina* clade, which comprises 16 species, all of which are red or orange-flowered. Nevertheless, it is believed that blue/purple flowers, most often pollinated by bees, represent the ancestral state for the *Ipomoea* genus (Streisfeld and Rausher 2009). Thus it has been proposed that a flower color transition –from blue or purple to red– has occurred, bringing ecological consequences along the way since pigments are visual signals for animals.

Flower pigmentation in morning glories appears to be controlled by alternate precursors of the same anthocyanin biosynthetic pathway: blue-flowered species produce cyanidin whereas some red-flowered species synthesize pelargonidin (Zufall and Rausher 2004; Streisfeld and Rausher 2009). After revealing that gene expression of *F3'h* –an enzyme that blocks the cyanidin branch of the anthocyanin pathway– is modified in red flowers, Des Marais and Rausher (2010) examined whether *cis* or *trans* changes were responsible for the phenotypic shift from blue to red flowers in the *Mina* clade, and discovered that it is caused at least in part by a *cis*-regulatory mutation in the promoter of *F3'h*.

The Texas wildflower (*Phlox drummondii*) has also been the focus of similar discoveries, remarkably because it has been demonstrated that divergence of floral color increases prezygotic isolation among its populations (Levin 1985). Hopkins and Rausher (2011) found that the shift from light-blue to dark-red flowers in *P. drummondii* is related to variation in the expression of two genes from the anthocyanin biosynthetic pathway: *F3'5'h* and the R2R3-MYB related transcription factor. Based on analyses of allele-specific expression in heterozygous individuals, they were able to identify that both genetic changes contributing to flower color change involve *cis*-regulatory mutations.

A different study by Yuan et al. (2013) regarding *cis*-regulatory evolution as the origin of flower color variation between closely related species has been conducted in the hummingbird-pollinated *Mimulus cardinalis* and the bee-pollinated *Mimulus lewisii*. In

contrast to the red flowers of *M. cardinalis*, rich in carotenoids and anthocyanins, the pink flowers of *M. cardinalis* have small amounts of anthocyanins while carotenoids are absent in their petals. In a previous experiment, it had been identified that a higher anthocyanin concentration in *M. cardinalis* was due to a single QTL (Bradshaw et al. 1995). Through genetic mapping, mutagenesis essays, and transgenic research, Yuan et al. (2013) managed to relate the anthocyanin QTL to a R3-MYB repressor called *ROI1* and discovered that *cis*-regulatory mutations in this gene are responsible for pigment variation between *M. cardinalis* and *M. lewisii*.

Finally, there is an example where changes in *trans* acting factors underlie floral morphological variation. Bumblebee-pollinated *Petunia integrifolia* and hawkmoth-pollinated *Petunia axillaris* are closely related species with purple and white flowers, respectively (Figure 1C). Quatrocchio et al. (1999) and Hoballah et al. (2007) identified a key R2R3-MYB regulator that controls flower color differences between both plants called *AN2*. The authors from the latter article noticed that white-colored flowers of *P. axillaris* evolved as a consequence of multiple nonsense and frameshift mutations in the coding region of *AN2*, most likely generated by a transposable element (Hoballah et al. 2007). Furthermore, to investigate the role of *AN2* as a major determinant of pollinator attraction, they introgressed the functional *P. integrifolia AN2* variant into a *P. axillaris* background, transforming white flowers into purple ones, and tested pollinator preference. Surprisingly, restoring *AN2* function through purple transgenic flowers of *P. axillaris* resulted in a major pollination syndrome shift, as hawkmoths showed a preference for the wild-type white flowers over the transformed flowers, whereas bumblebees acted in the opposite way (Hoballah et al. 2007).

Variation in fruit color – Fruits are the structures derived from flowers that contain the fertilized ovules that develop into seeds. Fruit morphology is very diverse, ranging from dry and hard capsules that expel the seeds when mature to succulent tissues that are eaten by vertebrates who disperse the seeds, as well as indehiscent fruits, among many variations.

The underlying causes of shifts in flesh and skin color of fruits have also been researched. Red-fleshed apple varieties from Central Asia with high anthocyanin content contrast with the white colorflesh observed in most apples. In apples, the *MYB10* (R2R3-

MYB) gene regulates flesh color, and it has been observed that high expression levels of *MYB10* transcript correlate with a higher anthocyanin content in the flesh (Espley et al. 2007). Differences in transcriptional regulation of *MYB10* appear to be caused by regulatory changes: Espley et al. (2009) identified a 23-bp repeat motif in the promoter of *MYB10* specific to red-fleshed apples, which leads to an increase in *MYB10* expression levels.

In a similar investigation, Butelli et al. (2012) explored the genetic and molecular basis of color shift in oranges. Orange varieties and hybrids present a wide range of phenotypic differences. The blood orange (*Citrus sinensis*), for instance, has crimson colored flesh and presumed health-promoting properties due to its high anthocyanin content. As in flowers, R2R3-MYB transcription factors regulate the anthocyanin biosynthetic pathway and govern pigment intensity in fruits. Butelli et al. (2012) managed to isolate an R2R3-MYB gene named *Ruby* from blood orange tissue and, through expression analysis, found a clear correlation between levels of *Ruby* transcript and levels of anthocyanin in fruit flesh among 15 varieties of blond and blood oranges –higher levels of expression were observed in blood oranges whereas in blond oranges *Ruby* expression was not detected (Figure 1B). Afterwards, they cloned, sequenced, and isolated the *Ruby* gene from the genomic DNA of blood and blond varieties, discovering that in blood oranges there was a large DNA insertion in the upstream region that was responsible for controlling the expression of *Ruby* (Butelli et al. 2012).

Fruit skin color change has also been documented within species. Perhaps one of the best well-known examples is the domesticated grapevine (*Vitis vinifera* L.), where the color of grape skin is determined by the accumulation of anthocyanins in several cell layers. Since wild vines have dark-colored berries, it is believed that white-colored varieties are the result of mutations in the anthocyanin biosynthetic pathway, specifically in MYB regulators. Grapevine has two paralogous copies of R2R3-MYB transcriptional regulators called *VvMYBA1* and *VvMYBA2* (Kobayashi et al. 2002). In an early approach, Kobayashi et al. (2004) examined the molecular basis of color variation in grapevines and found a retrotransposon-induced mutation in the promoter of *VvMYBA1* related with transcription interruption of this gene, which they associated to a loss of pigmentation in white cultivars.

Although Kobayashi et al. (2004) argue that loss of pigmentation in skin color of grapevines is due to changes in the promoter of the gene *VvMYBA1*, Walker et al. (2007) focused on the transcription factor *VvMYBA2* and discovered that the white berry allele is inactivated by two nonsynonymous changes. Therefore, white grapes may have arisen through regulatory and coding mutations in the MYB regulators *VvMYBA1* and *VvMYBA2*, rendering both genes inactive.

Fruit morphology in the Brassicaceae family - In *Arabidopsis*, two modified carpels or valves protect the developing seeds; these valves are joined by a structure called the replum. At maturity, the valves detach from the replum to open the fruit and disperse the seeds (Roeder and Yanofsky 2006; Roeder et al. 2003). Roeder and collaborators (2003) identified a homeodomain transcription factor called *REPLUMLESS* (*RPL*) important in the development of the valve/replum boundary (Marsch-Martínez et al. 2014). The replum morphology varies considerably in the Brassicaceae family: while in *Arabidopsis* the replum has ten cell files that separate the valves, in its closest genus (*Brassica* sp.) the replum is very compact and the valves are closer together. The expression level of *RPL* can explain the difference in valve phenotype between *Arabidopsis* and *Brassica*. Moreover, a single nucleotide change in the *cis*-acting elements of *RPL* gene is sufficient to convert the *Arabidopsis* fruit phenotype into a *Brassica* fruit phenotype, and vice versa (Arnaud et al. 2011).

Floral scent - Flowers of many plant species produce floral scent, a mixture of low molecular weight compounds that act as an attractant to pollinators. In some plant species, enzymatic activities and emission of some floral scent compounds are regulated at the transcriptional level (Dudareva et al. 1996). *Petunia axillaris* is an annual herbaceous plant native to South America with strong scented flowers pollinated by a hawkmoth. In contrast, the closely related *P. exerta* displays scentless bright red flowers pollinated by hummingbirds. In order to dissect the genetic basis of fragrance production differences between *P. axillaris* and *P. exerta*, Klahre et al. (2011) performed hybrid crosses and identified two QTLs. One of these loci is the *ODO1* gene, an R2R3-MYB transcription factor responsible of regulating benzenoid volatile production in *Petunia* flowers (Verdonk et al. 2005). Sequencing of the gene's promoter from several species and assaying

expression levels in the F1 hybrids suggested that the differences in gene expression are explained by a *cis*-acting polymorphism in *ODO1* (Klahre et al. 2011).

Leaf development - Leaves are specialized organs where photosynthesis takes place in many plant species. In angiosperms, they can have a simple (with an undivided blade) or compound (with subdivided blade) structure and both types of leaves can be found in related species or even in the same individual (heterophylly), suggesting the different types of leaf morphology could have evolved recently (Chen et al. 1997) and that they can develop contrasting morphologies based on environmental queues. Among the genes that have been shown to alter leaf development are the homeodomain transcription factors knotted-like (KNOX) genes. KNOX genes are classified in two groups, based on expression patterns and sequence similarity within the homeodomain (Hay and Tsiantis 2010). Here, we give examples of the function of class I genes *KNOX1*, *STM* (*SHOOTMERISTEMLESS*), *TKN1* (*TOMATEKN1*), in leaf development (Hay and Tsiantis 2006, 2010).

In model species with simple leaves, like maize and *Arabidopsis*, KNOX genes are expressed only in meristem cells of the SAM and are excluded from the leaves; in contrast, in many species that have compound leaves (tomato, among others), KNOX genes are expressed both in the SAM and in leaf primordia cells (Hareven et al. 1996; Piazza et al. 2010). It has been shown that transgenic expression of any member of Class 1 KNOX genes in the leaves of *A. thaliana* yield lobed leaves (Piazza et al. 2010; Shani et al. 2009). Moreover, are other *Arabidopsis* species, like *A. lyrata*, have compound leaves and express KNOX genes in the developing leaf primordia (Piazza et al. 2010). Additionally, other species of the Brassicaceae family, like *Cardamine hirsute*, have compound leaves and present *STM* expression in older leaf primordia; if KNOX expression is reduced using weak iRNA lines, the plants develop fewer leaflets per leaf (Hya and Tsiantis 2006).

Moreover, the aquatic Brassicaceae species, *Neobeckia aquatica*, bares both simple and compound leaves. Leaf development in this species depends on light intensity which in turn, regulates *KNOX1* expression: under low light intensity, plants develop compound leaves and *KNOX1* is expressed; in contrast, under high light intensity, plants develop simple leaves and *KNOX1* is repressed (Bharathan et al. 2002).

Finally, leaf margin dissection patterns, which range from smooth to serrated or lobed,

represent a well-studied trait that has evolved multiple times across land plants, and for some cases it has been correlated with the modification of certain loci involved in complex regulatory networks. One of such loci is the homeobox *RCO* gene. *RCO* is specifically expressed in leaves, where it is required for leaflet development. When deleted, the *RCO* mutants (*rc*) exhibit leaf simplification, from dissected into simple lobed leaves (Vlad et al. 2014). Within the Brassicaceae, some naturally evolved changes in leaf shape may be attributed to differences in the regulation of *RCO*, as suggested by a comparative analysis between the sister species *Capsella rubella* and *Capsella grandiflora* (Sicard et al. 2014). In their study, Sicard et al. (2014) show that an allelic variation at *RCO-A* is related to leaf shape variation in *Capsella*, and they later discover that *cis*-regulatory variation at this locus underlies naturally occurring leaf shape changes by responding differently to ambient temperature.

Leaf morphogenesis – Angiosperm mature leaves are typically flat with two anatomically distinct faces that correspond to the upper (adaxial) and lower (abaxial) side of the leaf. It has been proposed that the development of a flat lamina was a key event in leaf evolution as it enables this organ to capture light more efficiently and to generate complex organ morphologies, such as stamens (Yamaguchi et al. 2012). Arabidopsis leaf growth occurs in cells located in the intersection of abaxial and adaxial zones and partially depends on the participation of genes of two subfamilies of the **WUSCHEL** (*WUS*) homeobox (*WOX*) transcription factor family: *WOX1/STF* and *WOX3/PRS* (Fukushima and Hasebe, 2013; Yamaguchi et al. 2012). In Arabidopsis, the 15 members of the *WOX* family are involved in different aspects of plant developmental (Haecker et al. 2004; Mayer et al. 1998; Park et al., 2005; Sarkar et al. 2007). This family of genes has been classified in three clades based, mainly, on their phylogenetic relations but also on conserved motifs in the N and C terminal parts of the protein. Within each family, the expression pattern of the gene members is different although they have retained some similar functions since some of them are able to functionally complement the function of other members (Costanzo et al. 2014; Lin et al. 2013). For example *WOX5* and *WUS* of *Arabidopsis* who participate in maintenance of stem cells in root and shoot stem cells, respectively, are functionally equivalent and can complement each other's mutant (Mayer et al. 1998; Sarkar et al. 2007). Interestingly, the clades reflect the gene ancestry; genes like *WUS* and *WOX1-*

WOX7 are found in seed plants and belongs to a derived clade; another intermediate clade consists of the genes *WOX8*, *WOX9*, *WOX11* and *WOX12* found in vascular plants including lycophytes, and the basal clade comprised of *WOX10*, *WOX13* and *WOX14*, which are found in vascular and nonvascular plants, including mosses and green algae (Lin et al. 2013; Lian et al. 2014). Genes from the intermediate clade have acquired protein motifs (up to seven) after the divergence from the ancestor that seems to have had only one motif, the homeodomain, as found in many members of the ancient clade (Lian et al. 2014; Lin et al. 2013). Curiously, all members of the modern clade contained only two motifs: the homeodomain and the so-called WUS box domain that is essential for repressor activity (Lin et al. 2013).

In many plant species, *WOX1* and its orthologs are expressed in the adaxial-abaxial boundary of leaf primordia and many of these gene's mutants have severe leaf blade phenotypes; the most drastic is *lam1* of *Nicotiana sylvestris* that shows a bladeless leaf; this mutant has been used as a model system to study leaf morphogenesis. Lin and collaborators (2013) reported that *lam1* leaf mutant phenotype can be complemented with *stenofilia* (*STF*; *WOX1* ortholog of *Medicago truncatula*), and with all the genes of the modern WOX clade from *Arabidopsis* using *STF* as the promoter. Moreover, they showed that proteins of the intermediate and ancient clades cannot complement *lam1* leaf phenotype as they do not contain the WUS box or other repressor motif. The evolution of this gene family provides an example of *cis* (*WUS* and *WOX5*) and *trans* regulatory evolution (acquisition of different proteins motifs). The shared target of the WOX family members are still yet to be identified: these could shed light onto conserved molecular mechanisms involved in leaf morphogenesis.

Maize case - Maize was domesticated from a Mexican teosinte species (*Zea mays* ssp. *parviglumis* or spp. *mexicana*) approximately 9,000 years ago. One of the main differences between maize and teosinte is the inflorescence architecture: in teosinte, develops several long branches with many tassels (male inflorescences) at their tips and many small dehiscent ears (female inflorescences) at their nodes while maize has one compound tassel and one or two large ears (indehiscent and compacted) (Wang et al. 1999; Studer et al. 2011). As an additional outcome of domestication, the grains contained in each ear expanded from 10-12 in teosinte to around 300 in maize. One of the main genes involved

in such phenotypic differences is the transcription factor *teosinte branched1* (*tb1*), which is part of the TCP family (Cubas et al. 1999a).

During development, *tb1* acts as a repressor of the emergence of axillary branches; it is expressed in the axillary buds at different rates and it is more highly expressed in maize than in teosinte or other plant species like *Arabidopsis* (Finlayson 2007), bamboo (Peng et al. 2007), barley (Ramsay et al. 2011), rice (Takeda et al. 2003), sorghum (Kebrom et al. 2006), and wheat (Lewis et al. 2008). The genetic basis of this phenotype in maize is related with the insertion of two transposable elements (called *HOPSCOTCH* and *TOURIST*) in the regulatory region of this gene. These insertions are missing in the regulatory region of teosinte's *tb1*, and it has been shown that Hopscotch insertion functions as an enhancer of *tb1* expression in maize (Studer et al. 2011). Accordingly, changes in the levels of *tb1* expression underlie maize plant architecture evolution.

Fitness costs of natural regulatory variation

Whilst most of the phenotypic diversity related to regulatory mutations reviewed here occurs in the wild (see Table 1), we practically know nothing about the stability of these variations in natural populations. However, it has been observed in greenhouse experiments that changes in the regulatory context of plants may derive in altered fitness and in a dramatic impact on their ecological traits (Bossdorf et al. 2010).

Perhaps a case of a naturally occurring phenotypic change that has evolved ever since the last glaciation period is the inside-out flower of *Lacandonia schismatica*, discussed above. This has been observed by Vergara-Silva et al. (2003), who analyzed more than 1000 *L. schismatica* inflorescences and discovered that only 2% of the flowers were unisexual –the rest of them were all bisexual flowers with central stamens surrounded by peripheral carpels. As mentioned before, the unique floral arrangement of this Mexican species may be the result of either *cis*- or *trans*-regulator changes, although we do not rule out the possibility that both mechanisms may have had an important role during its evolution.

Another example is the peloric variety of *Linaria vulgaris*, also discussed above, which displays actinomorphic flowers as a consequence of an epimutation in the *CYCLOIDEA* gene (Cubas et al. 1999a). Although it has been suggested that the *Linaria* epimutant may

have a reduced fitness when compared with the wild-type flower, its populations have persisted for more than 200 years (Hintz et al. 2006).

A plant known as shepherd's purse, *Capsella bursa-pastoris*, also exhibits stable populations in the wild of a floral homeotic variation, the decandric mutant. In decandric flowers, the four petals of *C. bursa-pastoris* are converted into functional stamens (Hintz et al. 2006). Although not entirely researched, it has been suggested that decandric mutants may have evolved as a result of an ectopic expression of a class C gene in the second whorl. Despite the apetalous appearance of these flowers, Hintz et al. (2006) confirm they have persisted in wild habitats as stable populations, perhaps by shifting their pollination syndrome –from bees to beetles, to which the amount of pollen produced by extra stamens might be more appealing.

A more recent analysis performed by Wang et al. (2011) also documented that some double flowered varieties, *i.e.* those where reproductive organs transform into extra petals, such as those present in natural populations of the *Vinca minor flore pleno* species, may persist in the wild for centuries, potentially establishing new evolutionary lineages.

Understanding the phenotypic impact of altering complex GRNs

Contemporary biology is emphasizing the importance of a systems perspective to understanding how genetic and epigenetic alterations map unto phenotypical changes. In this paper we have reviewed the published examples of phenotypical alterations that have been related, mainly, to alterations of transcriptional regulatory components, but this does not mean that the associated phenotypes solely depend on the function of such individual factors in isolation.

In contrast, a systems approach considers that such phenotypes are the result of the concerted action of the identified components, in conjunction with other directly and indirectly related entities or factors (both genetic and non-genetic). Hence, the challenge becomes to understand the collective behavior that emerges from the interactions among identified regulators and other components, and not from the function of the individual components themselves. The latter must be considered as part of complex GRNs that we need to unravel and study as dynamical systems.

Multicellular development, that implies cell differentiation and morphogenesis, is an example of emergent behavior from collective and concerted action of many underlying

components, among which transcriptional regulators seem to be key. Such dynamical processes, as cell differentiation, proliferation and their spatio-temporal organization self-organize without a central controller or choreographer. Cell differentiation, for example, has been successfully studied in terms of underlying regulatory networks (Kaufmann et al. 2009) and characterized by the stable multi-genic configurations to which such GRN converge. These configurations are referred to as attractors (Davila and Álvarez-Buylla 2015; Figure 2).

The first GRN that was grounded on experimental data and formally analyzed was put forward for floral organ specification during early flower development (Mendoza and Álvarez-Buylla 1998) and since then, the same approach has been used for other cases (Mendoza et al. 1999, 2000; Benítez et al. 2013; Azpeitia et al. 2013; Alvarez-Buylla et al. 2010). Once such GRNs are established, uncovering the emergence and maintainance of the dynamical behavior from the concerted action of the molecular components being considered and their regulatory interactions, requires the use of mathematical/computational models (See Davila and Álvarez-Buylla 2015 for a description of methods used and examples). Such methods enable mechanistic understanding of the role of particular genes or other types of molecular components in the context of the GRN to which they belong. For example, it is possible to simulate loss and gain of function mutations and assess the role of the altered components in the emergent multigenic configurations that, in turn, alter cell fates. Thus a first approach to understanding why alterations in some genes are more important than others in underlying a particular phenotype (an altered cell-fate, in this case) is established.

For example, in Álvarez-Buylla and Mendoza (1998) and later in Espinosa-Soto et al. (2004) and further updatings of the floral organ specification GRN (Álvarez-Buylla et al. 2010), a mechanistic GRN-based model was put forward to explain the ABC model. The GRN model provided then a mechanistic explanation of why the A, B and C type genes are homeotic, and their mutations yield floral organ phenotypes in which one organ type is substituted by another one, as described above. Moreover, the same GRN model was able to explain the phenotypic impact in terms of altered cell fates for loss and gain of function mutations of all the components considered in the GRN model (see Espinosa-Soto et al. 2004).

In terms of such type of models, for example, we now understand that when a homeotic gene is turned off, the system switches from one attractor to another one (see Davila and Álvarez-Buylla 2015 for a complete methodological description). An attractor is a stable multigenic configuration that characterizes a cell type. Hence, for example, when *AP3*, a petal and stamen specific transcriptional regulator is turned off and the B function is lost, petals and stamens differentiate into sepals and carpels (Coen and Meyerowitz 1991). This is so because this gene occupies a position within a positive regulatory loop of the GRN and functions in a “switch-like” fashion, in such a way that its activation makes the system follow a particular trajectory until reaching a stable point or attractor (that of petals and stamens), and when it is turned off, it makes the system follow a different trajectory and reach another attractor (or cell-fate; that of sepals and carpels). Such contrasting emergent behaviors depend on the way the components are interconnected among them within the GRN module to which they belong (see Espinosa-Soto et al. 2004 for further details), and not on their individual function.

GRN models only consider the interactions that restrict the system to particular multigenic configurations that are characteristic of particular cell-fates. Physical and chemical fields, as well as cell proliferation dynamics, cell-cell communication, stochastic fluctuations, among others, have to be considered to understand how specific genetic and epigenetic alterations alter morphogenetic patterns during development and underlie phenotypical evolution (Álvarez-Buylla et al. 2007; Álvarez-Buylla et al. 2008; Barrio et al. 2010; Azpeitia et al. 2013).

The present review identifies several transcriptional regulatory nodes and alterations that seem to be key in underlying phenotypical changes. Thus the challenge now becomes to put together new GRN modules in which these nodes participate (see de Almeida et al. 2014 for an example of a critical node within the polarity GRN module) and, based on their dynamic analyses, provide a mechanistic explanation for their role in development and phenotypic shifts during plant evolution, as the one that was achieved for the ABC mutants (Espinosa-Soto et al. 2004 and references that followed), which we have summarized above.

Technological approaches to discover GRN components

A point of great concern for many evolutionary biologists has been the restrictions on experimental approaches that can be applied in order to unravel the functional effects of particular genes, as well as the lack of reference genomes/transcriptomes in non-model species; both approaches have been instrumental in the identification of interactions amongst genes embedded in a GRN. Nowadays, recent advances using Next Generation Sequencing (NGS) technologies and bioinformatics pipelines are generating a large amount of sequence information that can help study various aspects of genome evolution and architecture, as well as the interconnection among genes that are part of a GRN. NGS can be applied to DNA genome sequencing, transcriptomics (RNA-Seq), detection of Single Nucleotide Polymorphisms (SNPs), Chromatin Immunoprecipitation (ChIP) or ChIP-chip, identification of sRNA and epigenome characterization.

Specifically, the construction of transcriptoma libraries derived from organ, tissue, structure, treatment-specific cDNA assemblies can allow for the characterization of the transcriptional profile of an organism during different developmental or physiological states. This enables the isolation of genes that are likely involved in a particular trait, as well as the isolation of candidate genes that have presumably conserved functions –given what has been thoroughly characterized in model systems–. Furthermore, the study of their particular expression/repression patterns during specific times of development of distinct organ types –such as the floral organ whorls described above– can be used as a proxy to infer logical rules for all genes of interest across all tissue types assayed. Such inference can feed into formal biomathematical models of gene interaction such as the GRNs described.

Whole-genome and tandem duplications are ubiquitous phenomena in angiosperm evolution and have given rise to the presence of paralogs in many plant lineages; the role of such duplicates in modifying the dynamics and interactions among genes that are part of underlying GRN involved in different aspects of plant development could be formally and systematically tested, as has already been suggested for the case of the *Petunia* B-class gene paralogs by Espinosa-Soto et al. (2004; see Figure 2).

Concluding remarks and perspectives

Our brief overview of cases where the gene(s) that affect a particular trait have been characterized and the molecular processes that yield contrasting phenotypes have been

investigated, suggests that in plants as well as in animals, *cis*-regulatory evolution is an important driver in the diversification of form (Table 1). In plants, changes in *cis*-regulatory sequences in orthologous genes explain variations in phenotype between closely related species for phenomena such as anthocyanin synthesis (*Ipomea*, *Phlox*, *Mimulus*, *Petunia*, *Lycopersicon*, apple, oranges and grapes); repression of secondary branching structures (maize), leaf development (Brassicaceae family and *Capsella*) and vernalization requirements by ecotypes of *A. thaliana*. Gene duplication seems to also stimulate mutations in the regulatory sequence of one or both paralogs, which have been correlated with differential gene expression, as exemplified by the ICS in *Physalis* and relatives, as well as with the expression of genes involved in pigmentation in *Petunia* species. Nonetheless, gene duplication has also led to gene subfunctionalization via *trans* changes in the cDNAs of developmentally important genes, as has been documented for flower development in *Papaver*, *Solanum* and *Antirrhinum*, while the role of epigenetic regulation of plant development and its capacity to generate contrasting phenotypes based on changes in floral symmetry, sexual organ development or phenotypic plasticity have been documented for *Linaria*, melon and *Arabidopsis thaliana* and *Laguncularia racemosa* respectively.

The examples integrated here point to some key components of regulatory modules underlying the emergence of phenotypic shifts in plant vegetative and reproductive structures as a result of alterations in epigenetic regulation, mutations in regulatory regions of implicated genes or in the coding sequences of transcription factors themselves. Arguing about the relative importance of how different types of molecular alterations contribute to phenotypic evolution is a sterile debate (Hilscher et al. 2009). Rather, we need mechanistic and predictive frameworks, such as those being proposed by GRN and epigenetic landscape models, which enable us to understand why the alterations in the expression of particular components are more critical for phenotypic change than others.

Each one of the cases reviewed here opens the possibility to integrate or further explore the components of the relevant regulatory module in which the studied genes participate. Formal GRN models for these modules will then be useful to predict additional phenotypes and provide mechanistic and dynamic explanations to such phenotypic variation (Álvarez-Buylla et al. 2007, 2010). This approach will yield a

repertoire of regulatory modules in the search of general principles concerning the mechanisms by which phenotypic variation arises from changes at the molecular level.

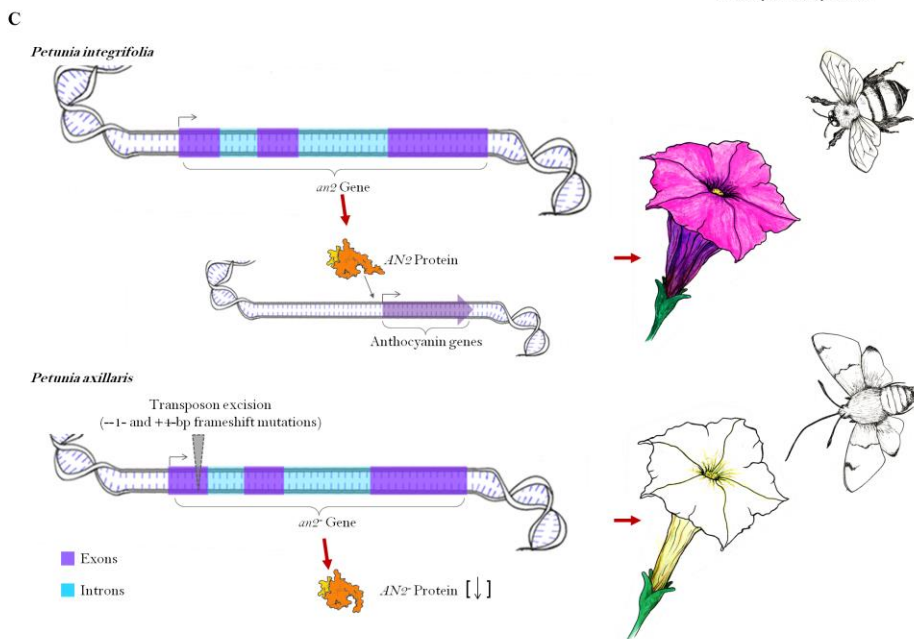
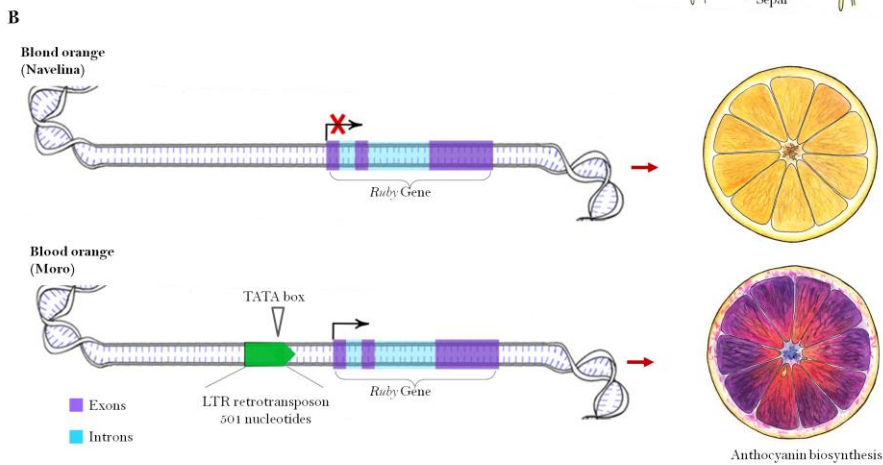
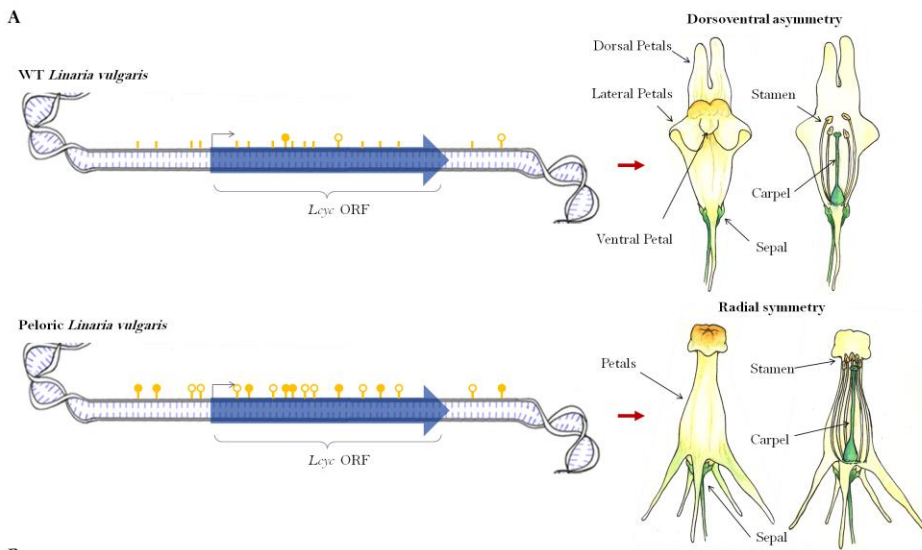
It is noteworthy that most of the cases in which a particular and relatively simple molecular alteration has been associated with a clear phenotypic change correspond to genes that code for transcription factors. These are probably the nodes with more interactions within their respective GRN. Hence, it is likely that alterations in the expression patterns of these genes are sufficient to cause an alteration in the GRN stable states, or in the resulting phenotype. More detailed explanations will benefit from further elaborations of developmental models grounded on experimental data for particular functional or morphological modules.

FIGURE LEGENDS

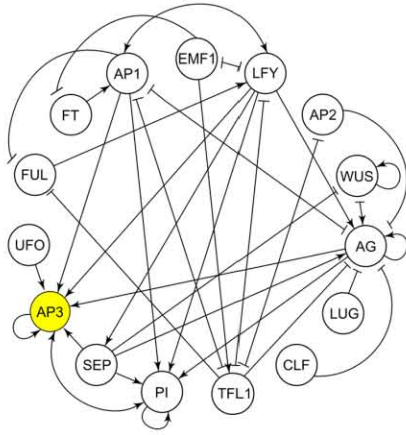
Figure 1. Changes in transcriptional regulation account for plant phenotypic variation. (A) DNA methylation patterns are related to phenotypic shifts among angiosperms, as in the case of *Linaria vulgaris*, where floral symmetry is changed from bilateral (top) to radial (bottom) as a result of extensive methylation in the *Lcyc* gene (Cubas et al. 1999b). In addition, the WT flowers of *L. vulgaris* fully develop only four stamens, while the peloric mutants develop five functional stamens. All methylation sites within the *Lcyc* locus are represented by filled orange circles; partially methylated sites are represented by empty orange circles; orange sticks represent non-methylated sites. (B) Modifying the regulatory regions of genes may alter the GRNs governing phenotypic traits, as observed with the *Ruby* gene: a single 501 bp retrotransposon insertion in its upstream sequence increases anthocyanin levels in orange fruit flesh (Butelli et al. 2012). A TATA box present in the retrotransposon facilitates *Ruby* transcription in blood oranges. (C) *Trans*-regulatory changes may also play a significant role in plant phenotypic evolution by altering protein structure of key transcription factors. Multiple mutations in the coding sequence of *AN2* account for the evolution of flower color between *Petunia integrifolia* and *Petunia axillaris*: while the *AN2* protein of *P. integrifolia* regulate the expression of genes that belong to the anthocyanin pathway, the frameshift mutations in the *P. axillaris* copy, derived from the excision of a transposable element, originate a

truncated product (Hoballah et al. 2007). It is believed that this regulatory mutation may have contributed to the evolution of the pollination syndrome of *P. axillaris*.

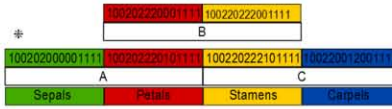
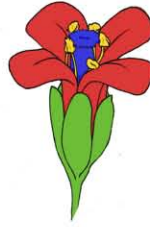
Figure 2. Corresponding GRN models to floral phenotypic changes. (A) The GRN for the *Arabidopsis* floral organ fate determination, according to Espinosa-Soto et al. (2004) and Álvarez-Buylla et al. (2007). The topology of the network shows the regulatory interactions between node pairs, with activations as arrowheads and repressions as blunt-end lines. The basins of attraction that lead to floral primordial cell types are shown in green (sepals), red (petals), yellow (stamens), and blue (carpels). Each attractor is defined for WT and *ap3* flowers as the genetic steady-state combination of all genes represented in the network –listed in the figure with their corresponding order–, established by three levels of expression: 0 (no expression), 1 (intermediate level of expression) and 2 (overexpression). (B) The GRN for the *Petunia* floral organ fate determination, as reconstructed by Espinosa-Soto et al. (2004); the topology and logical rules are the same as the *Arabidopsis* GRN. In *Petunia*, nonetheless, two B-class genes govern the development of petals and stamens: *PhDEF* and *PhTM6*. In *phdef*⁻ mutants, functional stamens are observed but in the double mutants *phtm6*⁻/*phdef*⁻ stamens do not develop. The yellow nodes of both networks indicate which genes are turned off in floral mutants.



A *Arabidopsis thaliana*



WT

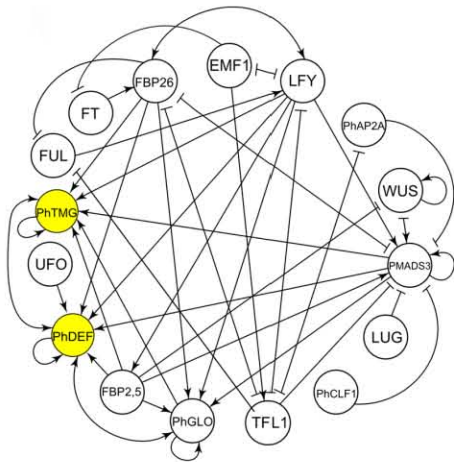


ap3

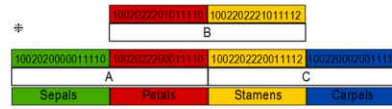


- * FT
- EMF1
- TFL1
- LFY
- FUL
- AP1
- AP3
- PI
- AG
- UFO
- WUS
- AP2
- SEP
- LUG
- CLF

B *Petunia*



WT



phdef



phtm6/phdef



- * FT
- EMF1
- TFL1
- LFY
- FUL
- FBP26
- PhDEF
- PhGL01,2
- pMADS3
- UFO
- WUS
- PhAP2A
- FBP2,5
- LUG
- PhCLF1,2
- PhTM6

Table 1. Transcriptional regulation shaping plant phenotype

Gene	Species	Epigenetic modification	<i>Cis</i> -regulatory change	<i>Trans</i> -regulatory change	Present in natural populations	Phenotypic change	Gene function	References
Floral morphology								
<i>Lyc</i>	<i>Linaria vulgaris</i>	☐	☐	☐	☐	Floral symmetry	TF (TCP)	Cubas et al. (1999b); Luo et al. (1996)
<i>CmWIPI</i>	<i>Cucurbita melo</i>	☐	☐	☐	☐	Male or female flowers	TF (WIP)	Martin et al. (2009); Turck and Coupland (2014)
unknown	<i>Laguncularia racemosa</i>	☐	☐	☐	☐	Mangrove development	?	Lira-Medeiros et al. (2010)
Floral development								
<i>AP3/PI</i>	<i>Arabidopsis thaliana</i>	☐	☐	☐	☐	Flower development	TF (MADS)	Yang et al. (2003)
<i>PapsAP3-1/ PapsAP3-2</i>	<i>Papaver somniferum</i>	☐	☐	☐	☐	Flower development	TF (MADS)	Drea et al. (2007)
<i>TM6/ TAP3</i>	<i>Solanum lycopersicum</i>	☐	☐	☐	☐	Flower development	TF (MADS)	de Martino et al. (2006)
<i>LsAP3</i>	<i>Lacandonia schismatica</i>	☐	?	?	☐	Flower development	TF (MADS)	Álvarez-Buylla et al. (2010b)
<i>FLC</i>	<i>Arabidopsis thaliana</i> (Ler ecotype)	☐	☐	☐	☐	Flower development	TF	Michaels et al. (2003)
<i>FAR/ PLE</i>	<i>Antirrhinum majus</i>	☐	☐	☐	☐	Flower development	TF (MADS)	Airoidi et al. (2010)
Flowering time								
<i>FRI/ FLC</i>	<i>Arabidopsis</i> ecotypes	☐	☐	☐	☐	Flowering time	FRI protein TF (MADS)	Michaels et al. (2003)
Novel floral morphologies								
<i>Se2.1</i>	<i>Lycopersicon</i> genus	☐	☐	☐	☐	Mating system	Putative TF	Chen et al. (2007)
<i>MPF2</i>	<i>Withania</i> genus	☐	☐	☐	☐	Inflated calyx syndrome	TF (MADS)	Khan et al. (2009; 2012)
LFY expression								
<i>LFY</i>	Land plants	☐	☐	☐	☐	Flower development	TF	Maizel et al. (2005)
Anthocyanin pathway								
<i>F3'h</i>	<i>Mina</i> clade	☐	☐	☐	☐	Flower color	Enzyme	Des Marais and Rausher (2010)
<i>F3'5'h</i>	<i>Phlox drumondii</i>	☐	☐	☐	☐	Flower color	Enzyme	Hopkins and Rausher (2011)

<i>ROI1</i>	<i>Mimulus</i> species	□	□	□	□	Flower color	TF (R3-MYB)	Yuan et al. (2013)
<i>AN2</i>	<i>Petunia</i> species	□	□	□	□	Flower color	TF (R2R3-MYB)	Quatrocchio et al. (1999); Hoballah et al. (2007)
<i>MYB10</i>	<i>Mattiana</i> (apple)	□	□	□	□	Fruit color	TF (R2R3-MYB)	Espley et al. (2007; 2009)
<i>Ruby</i>	<i>Citrus</i> (orange)	□	□	□	□	Fruit color	TF (R2R3-MYB)	Butelli et al. (2012)
<i>VvMYBA1/ VvMYBA2</i>	<i>Vitis vinifera</i> L.	□	□	□	□	Skin fruit color	TF (R2R3-MYB)	Walker et al. (2007)
<i>VvMYBA1/ VvMYBA2</i>	<i>Vitis vinifera</i> L.	□	□	□	□	Skin fruit color	TF (R2R3-MYB)	Kobayashi et al. (2002; 2004)
Fruit morphology								
<i>RPL</i>	<i>Arabidopsis thaliana</i>	□	□	□	□	Valve phenotype	TF (Homeodomain)	Arnaud et al. (2011); Roeder et al. (2003); Roeder and Yanofsky (2006)
Floral scent								
<i>ODO1</i>	<i>Petunia</i> species	□	□	□	□	Floral scent	TF (R2R3-MYB)	Klahre et al. (2011)
Leaf development								
KNOX genes	Brassicaceae family	□	□	□	□	Leaf development	TF (Homeodomain)	Bharathan et al. (2012); Hya and Tsiantis (2006); Piazza et al. (2010); Shani et al. (2009)
<i>RCO-A</i>	<i>Capsella rubella</i>	□	□	□	□	Leaf shape	TF (Homeodomain)	Sicard et al. (2014)
<i>WOX</i>	<i>Arabidopsis thaliana</i>	□	□	□	□	Leaf development	TF (WOX)	Costanzo et al. (2014); Lian et al. (2014); Li et al. (2013); Yamaguchi et al. (2012)
Maize case								
<i>tb1</i>	<i>Zea mays</i>	□	□	□	□	Maize development	TF (TCP)	Cubas et al. (1999a); Studer et al. (2011)

APÉNDICE 2

Regiones promotoras de los ortólogos de *APETALA3*

MONOCOTILEDÓNEAS

>T_brevistylis TbAP3v1 (esta tesis)

```
AAATCAACATCTGAATGAGGTAGTATTATCCTTGAAAATTGAAACAAGGTGGTGTGAAGTTT
GAAGTAATAAATAAATATTTCTTTTTTGCTTGGATGGAAGTGATGATGAATACAATCCGAGA
AGATCATTCAGTGGTTTTGTCAGTGGGGTGGGGACGTGACGCATTCTTGGACTGGGGCTTT
AATCCCATCACATGCTTTCATGCTTTTCCATTCCCTTCTCTCTCTCTCGGCCGTCTTCTT
CTTGATCTTCTTCCTTCTTCCCCTCCTGTCACGAAAGCTGGAGGTCGCCATG
```

>T_brevistylis TbAP3v2 (esta tesis)

```
CTGGAGAGGTGGTGACCCCTTCCATAATCCGGCCAAGACACCACACGTTCCCCTCCTCCTC
CTCCCTCGTGACTTTCATCGATTTCGAAACCCACATCCAGAAAGCTTCCCCCATCCTTTCC
TCATTTAGAAACCCTGCGGCCACCACCCTTCTCTCCCTCCATCCCTCTATATGAAACCCTA
CCGTGGTCCCATAAGCTTGGAGGGAAACCCTAGTTGTGTCTCCATG
```

>L_schismatica LsAP3 (Piñeyro-Nelson 2013)

```
TTGGTTCTCAAGTATGAAGGCAAACCAACAAAAGCATATTATAATATTTGCATAAACATAAA
TGCATCTCATTTAACTGGAAACATAAAGCTGCCCTTCCCATCTCTCTGAAAGGATCCCAGTT
TTCAAAGTGTTTCCTCACACCTCTCAGTTGTTCAATCTAATTCCGATCTCCTGGAATCATGC
ATTTTATCCCTTAGGATTTAACAACAAGCAACTGCTTACTTAACAACCTGATTTGCATTTGTT
AACCTCTTTTAAACAAACCTTAATCCCAACTGGCTTAGCTAAAACCTGAGTCTAACTGACTCTA
ACAAATCTCCTCTCTAGAGTCCCTGCTGACTCCTCCCCTTTGAATCCAAGATTAGCCAACCG
CGGCTACTTAATTATTATTGCTAACAACGAAGTAGGCCACATGCAGACTAGGCCCTGGTTCT
AACTTAGCTCCATAACCTCAATCAAGTGATGAATCGATCTCTTTATCTGAAGGAACCTTTAT
TTGGTTCCCTAGACTAATAATCCCTCGTCTAATCTGGCACCACGGGCTCTGCGCTCTAAATTA
ACAATTTATATAAAGAGGGCGAGGCCCGCTTAGCGATTTCAAGGCAATTAAGAGTCTGAAGT
AACCATATATATTTCAATTTATTAATTTCCCAAGCATAGACAACATTTTCATCCTTGATATTG
CAAGCAACACCAAGTGAGCACATGCAGCCTATGACTTGGGTCCTAAATAACCTAACTAAATG
ACCAACATGCAAACATCTTTTCTTTTTAAAGATAAAGAAATACCAGCAAGACAGAATGCCTG
GAACGACCAGGAAAGGCCGTACAGTTTGAACAGCTCTTCCAGAAAGACAAGTTCTAAGCAA
ATTCCGTCCACACGATAATATCTGCAAGAAAGAGAGATCGATAGTCATTGGTTCTCTCTGTC
ATTAATTGTTGCCATCAAGGCATCTCCATCTCTAACTTCCCTCCACCGCCCGCTCTCCCTCTC
CTCCGGCCATG
```

>L_regale LRDEF (AB158291.1)

```
ATACTCAGCCTTAATTATCTGTACTTTGATTCATAAATAAAGAACCCTTGATCTGTACCGTT
AGACTAAATTTGGATGATAATTTTATTGCATTTTCAAAGAGGAGCTTGGCTTGTGGGGTGAG
TTTGCATATGTATCCAAGAGTTGGAGAGGATGGGGATTTTGGGAGGGAATGGAGGGAATTGT
GGTTAGTGGGAGATATTGTGTTTTTTCAGTGAAATAAGTAGTGAGGTATGATGGGTGAGGCTTG
```

TAGTTGGATGTTAGGGTATGTGAATGGGAGGCCATGATGAAAAGTGTGTTGGCAGAAGAACGA
TAAGGAGGAGGGTATTGGTGGATAATAGTGGAACGTATAATTGGGGGAATATGATTAGGGC
CACCACAATTTTGGAGAGTTCGGATTCAGAAGGGTAGTTTTTTTTTTAACATCCGATCTTTT
TATTTTTCTGAATAGTAGAAAAATGAAAAAGAAATGGAGGGTGTGATATGATGAGAGAGAAGAG
TGATAGAGACAATTATGGATATTGAAGGAATGGTGTGAGGTGTCGATGATAAATGGTGTAGGAT
GATTGTTAATGGCGGGGGAGTGAACAAAAAACTAAGGAAATAGGAAGCAGGGAGTATCTCAC
CAAATCTAGATACATGGGAGCCAACATATGGGATTTCTGGTCAGCCGTCCGCCTTATCCCAG
CCCCTACTCCCATCCCTATTCTCAACAGTCTTGCGGTAAATAAAGCCACAGAGGGTGGCGGG
CTTCATAGGAAAATGTGAAACAACACAAAGGAAAAGTGACCGATTCTTAAACAGGCAAGACA
CAGCACGTTGCCCTCGTACCCCCACTCTTTCATCAATTTGGCAAACCTTCCATTTATAGCA
ATTCTAGCTTTTGCAGCCCCCTCCCAAATTCAGACATTACCCTTCCTCTTGCTTTCCCCCTCA
GTCCAAACCTCCCACCCACCGCTCCTTTATATCCCACCACCGCTCCACCATCCCCTCAACA
GATCAGT**ATG**

>A_officinalis AODEF (AB180962.1)

AGCTTGATGGACTTGCTCTTAGCTGGCTTTAATTTTCTTCTTTCCTTGTTTTCTCGTGGGC
TCATCTTTCTCCTCTTCCCTTAATTTGCTTCTCGCTTGCATCTTTGAAAGTAAGTAAGGATA
CATTAAATCCAACATGTCGCTGTAAGCGTTGAAACCAAAAAATCAAACAAAACGAGTTGAT
TTTGCTTGGAAGTGAATAATTTGTCAAAATAAATAAACTTTGAAAAAAGAAAAAGCTTAAT
TAGGAGAAAGTTCTAATAAAAATAAAAATAAAAAGCTTGAAAAACAATAGAATCTAAACTAAA
TGAAGGGGAAGACAAAAACAACCTCAAGGGTAATAATAATGTGGGTACCATATGGATAAAA
TTAACAAGGTGAAGATTGATCAAGTGATCAGTTAGAGGTTGGTATAAGTGGTAGCAAAAAAA
ATAGCTAGGGGCAAGAGAAGAGAAGGAGAGGAAGAAATGGAGGGATTCTAATTAAAGTGGGA
AAACATGACACGTTGAACCGTACGCACAAAGCTTTCATGAGTTTGGCAAGCCTACCCATTTA
TAGAACCCCTACCTTCGCTCCCTCTCCCTCTCTCTTCTCTATTTAGAAATAATTAGCAAAG
CAAGAACCCTAGCTATCAGCTGGGTGGGGGGTGGGGGGGGGGGGCTTCTCTTTCTCTCTC
CTGTCCCCACATGTAGATAGCCTATAGCTACCTTTTTTCTCTTGTCCCTCCCTCTCTCAAAC
CCACCCACCACTCCGTAATACTAGCTTCTTTTTTTCTTTTTTTGGGTTTCGACCCTTCCC
CTTTTATAAGATAATAATATAATCAAGCCCATGCCCCCTTCTTCTTTCATCTTCAACT
CCTTGCTTGCTCTTCTTTCTTTCTTTCTTTCTTCTCGTGTGCTTAGGTATTTATTTACAGC
TCTGGGCTCTGGGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCGATCTCGAAAAGAA
AAGGAGAG**ATG**

>P_equestris MADS5 (AY496050.1)

ATTCAGACTCAAATTTGGATTGCCATGGTAGATTTGGTAGGGGGAAATTGATTGGGATGTGC
TAAGGGTCCCTCAGGGAAGCTCTCTGAAGAGTTTAGAAGGCCCTCAGTGAAAAATTTTTTTAC
CACCCCTCTCCATGGCCGTGAGGGGTAGATTTTATAGAAAATGGCTATAAGGATGCTTAGCT
GGCCAATTGAGAGCTTAGCCCCACTGCTTAGGTGTCAAACAGAGGGGTCCTCAACTCAATTC
GAGCTGAGGTGGCCTCTTGGCATGAATTTGGGTTCTAACCATGTGGGGGCTCACGGTTTTGA
GCTTTAGCTTGGGTTAAGGTGGTTAACCAAGACTTGGACCATGGTTAAGGTTTATGATGCTT
AGTAAAAATGACACCATAGAGTTTAGACATGTGGTATGACAAGTACATTGATTTAACCTCCT
TATTATCTCTTTAAAAGCTTTTTTCAAAAATTTACCCATGGGATAAATATCAAGCCTAACCA
TGGGTTGATGACTGACTTGGAACTCAACCCTAGGGTCTAAGGTGAGTCAAGGTAGAGCTCAT

TTGACCATAACCAAGCTTAATTGGTTTTGGATTTGGTCATAAGTCCAAAATGGTGTTCATGACC
AAATCTTGTGTTTTGGGCTAGGTTAGGCATGTCAATTGGCCCACATGTCTAAAGCACATGTTTC
CACTTCGCTTTCACAGAGAAGGAAATCACTTTCCTTAGGGAACTTTAGCTTTTGCAATCT
CTCCCTAAAGTGGAATTCCTTCTTCTTCTTTCATCTTCTTTCATCCGCTCCTCTCCACAGCATC
CAGTCTCCACTAACTGCCTCTTTATCACTTTAAAAATCCCTCCTTTTACACTTTCTACGTAT
GCGGCCTTCAACCACCCATTGAATTGTAGAACACGACTTTGTTGCCTTTCAGGGGGAGAGAGA
GGGAGAGGCAGAGAGAGAGAGAGAGAGAGAGAAGTGAATTTGTGCTGAAGGAAGCTAAGGGGGG
TGGAGGGA**ATG**

>Z_mays SILKY1 (CM000782.3)

CGTTTACAATTGGGCCATGTGCCGTGTAATAGAACGGCGTCGTTGCCATATAAGGAGAAAGT
CAGACCGGAAGAGGTAACGAAGAACATAAACCTGATTCCTTCGCCAACTCTTCCCTCTTCT
AGTTCTTCTTCTTCAATCCACCTTCTCTTACCCGCTAACTTAACAGTGGAGTACATCTAAA
CTAGCTTGTCACTTGTCACTTGGAGTAGAACCAAACAGACTTTGATCATGCATGTAATCATG
GCCCCGTAATTAATGCATGGTTGTGTAGAGCGACCGAGCGAGCGGTTTGATTTGATGCCGGT
GAGGTGACGTGACGACGCTAGGTAGCTAACAAAGGAGGTAGACGATAGCTGCTACCTGCTAGC
AGGAGGTGTGTGATCGAGAGTTGACAGCCGGTGCAAAGGAGGAGGCCAAGAGAAGAAGAGA
AAGCAAAGGAGGAGGGAGTGTGCCGGAGAAGAGGATCCGGAAGCCAAAACCGTGCTAACCGT
TGTGCCAAAAGCCGCCACCACGGCTGACCGACGGACGGCACGGTGGATTGATTGGAAACGCG
CCGAGATCGACCCAGCCGGCCGGCGCTTACACCCACGCCGCGTCAGATCACAGGGGCCGGGC
GCGGGCGGGCACTGCACGGCACCCACAATACGGCACGCGGGCGGGCGGGCGGGCACGGTGCC
CCACGCCTTTTCAGGATCGGGCAGCTGTCCCCGTCCACGCCGCGCACCGCGCCCCCGTCTCG
CCACCCCGAAATGCACACGCACACGCCTTGTCCTTGCTTGCTTCCCTTGGCACCACCGCCTCC
CCTCCTCTCCTTATTACCACCACCTAGCTAGCAGCGGCAGCGGCAGCACCGGCCTGTGTGC
TGCTCGCTCAGACAGCTCTGCTAGCTGCATCCTCCTAACTCTCCAGGTCTCTCTCTCCTCTC
CCAATCCCAAGTCCCATCCGGATCGAGACGCTGGAGGCGGAGCGCCCCCGGGACGGCGG
CGGCGAC**ATG**

>O_sativa SPW1/OsMADS16 (AP005463.3)

GGCCTGGGTTCAAACCTCACCCCTTCTAATTATTTGATATTATGTCATTTTCTAATATTCG
CGTCTGTTTTGTTGTCATCGTCTTTCTAAGCGTGCGTGCTTGCAAAGGAAAGGAGTAGAGAG
AGAGAGAGGCACATCAGCAGATCGAGAGAGAGAAGCTGGTATGACAAAAGAAAAGAAAAGAC
AGAAGATGACTCATGCACACTTGCGAAATCAAACCAAAGAGAATCAAAGCGAGGCAATATG
TACGCGCAGAAGCAATGATGAAGCGAAAGCCGAAATGGTGCTAACCGTTGGGCGGCCAGCCA
CGCATGCAGCGAATCGGTCGATCGATCATGCAGCTGCTAACACACACACAAGAGAAGACGAG
ACGAAACTCAGCCGGCCGGTGATTTCGATCGAGAGTAGGTGACACAGCCAAGCAGGATAGAGG
CTATTTTTAGGAGAGCTTCTAGCTGCTGCAACTTCTCTCAGAATTAGAAGCTCCCCCAAACA
GTATATCTTTTTGGTTTCATATTTCTAAGAAGTTGTAGTTGTAGAATCCAGAAAGTGAACCAGAA
GCTAGAAGCTGAAAAACCTAGCTTTTTCAATATTTCTCAGAAGCTGTAGTTGTAAAATCTAGAA
AATGAATCAGAAACCAAAAAAATGAAAAATCCAGCTTTTCTAGATTCTTATAAATTGTCTAC
AAATCAGCTGCTTCTCATAATTTTAAACTCCACAAATACTGCAGGATCCGCTCTCCGCGAACC
GAGAGACGTTACGGTATACGGAGTGGGCACGATTGCTTTCACGAGTTGGGAATCCGTTGCTG
TCGTACGTCTCGTCTATGGCGTGCTGTATGTCCTCTCCTTCTTTCTTCTTCCAAATAAGC

ACTCCTCTTCTTCCTCCTCCTCCTCCTCCGCCTCCTCTTCTCATTCCCAACTCATCGATCCA
TCAGTAGCTAGCTAGCTAGCTAGCTAGCTGCATTGTCCGGCGAGAGATAGCTGCTGCAGG
GGCGGCC**ATG**

DICOTILEDÓNEAS

>A_thaliana AP3 (U30729.1)

ATTGTTCTTGTGAAGTTGATACACAAGTTCTTTGGATATCCTATCAGTATAAAGGATAGGTT
TCCATTTTCGTGACTCACTCACTGATTTCCATTGCTGAAAATTGATGATGAACTAAGATCAA
TCCATGTAGTTCAAACAACAGTAACTGTGCCACTAGTTTGAACAACACTAACTGGTCGAGCA
AAAGAAAAAGAGTTCATCATATATCTGATTTGATGGACTGTTTGGAGTTAGGACCAAACATT
ATCTACAAACAAAGACTTTTCTCCTAACTTGTGATTCCTTCTTAAACCCTAGGGGTAATATT
CTATTTTCCAAGGATCTTTAGTTAAAGGCAAATCCGGGAAATTATTGTAATCATTTGGGGCC
ACATATAAAAGATTTGAGTTAGATGGAAGTGACGATTAATCCAAACATATATATCTCTTTCT
TCTTATTTCCCAAATTAACAGACAAAAGTAGAATATTGGCTTTTAAACCCAATATAAAAAC
TGCTTCACACCTAAACACTTTTGTTTACTTTAGGGTAAGTGCAAAAAGCCAACCAAATCCAC
CTGCACTGATTTGACGTTTACAAACGCCGTTAAGTTTGTCCACCGTCTAAACAAAAACAAAGT
AGAAGCTAACGGAGCTCCGTTAATAAATTGACGAAAAGCAAACCAAGTTTTTAGCTTTGGTC
CCCCTCTTTTACCAAGTGACAATTGATTTAAGCAGTGTCTTGTAATTATACAACCATCGATG
TCCGTTGATTTAAACAGTGTCTTGTAATTAATAAATCAGTTTACATAAATGGAAAATTTAT
CACTTAGTTTTTCATCAACTTCTGAACTTACCTTTTCATGGATTAGGCAATACTTTCCATTTTT
AGTAACTCAAGTGGACCCTTTACTTCTTCAACTCCATCTCTCTCTTTCTATTTCACTTCTTT
CTTCTCATTATATCTCTTGTCTCTCCACCAAATCTCTTCAACAAAAAGATTAAACAAAGAG
AGAAGAAT**ATG**

>A_majus DEFA (X62810.1)

GGGGCTTGTCCATAGAAATTGTTGGTCAGCAATACTAGTAACACGGAATGAAAATGTGAAATC
AAAAATATAAATATGTTAACAATGAAGAACTAAATCCACCACAATTGAAAGAAAACCCCG
AATTTTATAAAGGATATGAAAACGAGCTATATATAGTTTATTAGCTCTATCTCTAGCTATTT
GATGTAGTAGAACAAGATGCCCTAATATTAAGTTTGAATAATCTCGAAAGTAAACAAGCCG
CACGAAACAACCGTAACTTTGGACTCATTCTTTCAATTGAGGCTAATAATTTATTAAATTAG
ACGTATTTCTGTGATCCACGATTGATATTGAGGCCACTAGTTAGTCCCTAGCTAAAACCATC
GTTGAAGGACATAGTTTTCGGAGCGAAGTAGAAGGAAATCGAGTGCTAACTAGTAACAAATTT
TAAAGCTAAAATCTAATTTTACATGCGAGATAAACTACTTCGCTCCGAACTATATGTCCT
TGAACGATGGTTTTAGCTAGGGACTAACTAGTGGCCTCAATATCAATGGTGGATCACAGAAA
TACGTCTAATTTAATAAATTATTAGCCTCAATTGAAAGAATGAGTCAAAGTTACGGTTGTT
TCGTGCAGCTTGTTTACTTTTCGAGATTTTTTCAAACCTTAAATATTAGGGCATCTTGTTCTACT
ACATCAAATAGCTAGAGATAGAGCTAATAAACATTGAACTCACAACCTCGTGAACAAAACCTTT
GACGTGTGACAAATAAATAAATGTTACACATTAATAATTTTGCAACAGACCTTCTACTAAT
TAAGAATACTTTTCCCTTTTATCTTAATAGTACCGCAACGATACGTAACATAGATAGATATA
TATATATATCTATTCTTCTTACTTTCATATCTGTACTTGCAAATCTCAAATCTTATCAC
AGTTTTTAGGAAAGTAAGGAGAGAAAGGAAAGCTGGAAGAAGAAAACAAGAGCAGTAGTGTA

GTGGTTC**ATG**

>S_tuberosum DEF4 (X67511.1)

AGTTTGTGTGTGTGTGTGTGTACATGAGTGAGTATTGAAGAAATCTCACATCGGTCAATTATGG
AATGGGTACTAATTAACACTACTATCCATGTGTTTTAACTTAATTTCTATAGGCCATTTCTATA
ATTATCATATGTTTCATATAACATAATTTATAATTATCTTTTTAGGTGGCCTAATTATCTAGAT
ATTTTATCTACCGCTTGAGAGGGCTAACTCGTATAGTAGAGGCCATCAAACCTTTTCATAA
TATTCGTTGTCTGTTTTTTTTATTAACGGGATAATTCATTGAATACGTTGGTAAAGTACATAA
TTATGGCACTACAAACCAATAAACAGGAAAATAAAAACAGATTAAGGAATTGATTGATCAAA
AAGATTAGAAAAAATGTCTCAAGTCAAATCTAACTATATTGAGCATCAACGATTATTTCTAC
TCTAACCATTTTCAACATATATATTTTTGAACAATTACAATACTTCAACGCGGTCAATTACA
CACTTTAATATGATATTAATTATTCAAATCCATATCTCATGATCTCGACGTCATCCAGTG
AATAAAAAGAATTTGCATTAGAAATAAAGAACCAAATCTATATACGAGGAGACGTATAATA
TATAATCCCTCTTAGTATTAAGTATATCTTGAATTTTAAAAATAAGATGATCACACATTTTT
CCTCGTTTAATTCATAATAAAAATATAAATACTGGTGGGTGACAAGTGACTGATAACTGAGA
CGTTGTGATGGTAGGGACAACACTACAACAAGTGACAATTGATTGATTCTTAAATTACTTTTTAT
CAAATTCGCAATCACCTTTTCAATTGTTTTGGCATACTTTCCATTTTTAGTAACTCCATCTT
TCTAAGACTCTTCTCCTCCTCCAATATCTTATCACAAATCAAATAAAAAATAAAAAATTGC
ACAATAAAAAGTTAACTTGACCTTTTAGGGTTTGAGTATTCAAGATCTCAAAAAAAAAGAAA
AGAGAGTT**ATG**

>V_lambrusca AP3-like (HM192809.1)

TGACCTGCCTGTTCGTAGCGGGTTCTATCGCAAAGGAAAAACAAAAGTGTAATTAACTAA
ACCTGTCAAATCATCAAGATTATGTATTTTCAAATTTGAATACAAAAGAACCCAGCTTTT
TCTCGGCTCACCTACGTTCTTGCCTACCTCAAATTTGGACAAAGATACCTCAAATTTAAATCT
TGCTTGCATATAACCAGAAATCTACGTACCACCTTAGCCGCTAAACATTGTTTTAAAAATAAAT
AGATAAATCAGGATTCAGGACATGCAAATAAGGTTGACCATGGGCCAAGCCCTAAAATTGG
GTATATGGAGAGCCCAATGATGGCAGAAGAGAAGCCCATCTTTGTCCATGGTGCAACTGCGG
AGTCCGCCTCGTTTTTTCAGGAACATGATAATAATTTGGTGGAAAATGGGTAGGGTAAAAGCC
ACACGGTATCGTTTTTATTAGCATTTCAAAGGTATTATTGCTTTTTTAAAGCAAAAACCTGTTT
TCATTTTATTAACTATTAGAAGTAAGAAAATCACTTCTGAAATCATGCTTATCGGTTTCTAT
AACTTCACTCCTTCGAGGATCAATTTAAATCATGATTGGACACGTGAGAGAAAGGAAAAGGC
GTTACATTTTGGAGGCTCAGCAACGTAATCACTTTAGTACCTGCAAACCTTCGTATTTTCCCA
AGCGCAAGCCCAAGCCTATTCCCATGCTTTCTGCACATCTACACCGTCTGATGCTTGTAAT
ATTCCCATCTCGATCTGAGCCGTTGAGATAATAACTTAACGGGCGTAACGGAGAAGGACGGA
AAGCTGTTGTCCCATGACTGAAATGCGGTTAGCGACGTTATCGATCCCCTGCAAACCTTTA
CGAAATGAGCAAATATTTCCGTTAATAGTATGAAGTTGGAGTTGTTCTGCTTTGATCTTTCT
TCTTTGTTCGTGTAGTGGAATCTTATGAAACAGAAGCGACCTGGAAATCGGAGGGAGAAGAG
AGGGAGT**ATG**

>P_trichocarpa PTD (AF057708.2)

AGGAAAAAATTTTTTAAAAAAAACCTTTCAATCACTAATACATAGAAGGTGTGGGGAAAGC
CACAGTGATTTCCCGTACCTTTTAAAGTATTACTTAATATATAGGTGAATTTAATTGACCG

TCACGAAAAAGACTATTCTGGCTTCCTCTTACAATGGACGCTATCTAAATTCAAATACTTTG
AAAAAAGATTTAATCCTGTAACCTTCTTTTCGTTTTTTTATGCCTTCAATCCATCTATTTATT
GTTTTTATGATTTTTCTTAGATACAAAAGAGCATATTTTAAAGAAGAAAAAATAAGCTAAG
CACCTCAAGTTTTGATTTTTTTTTTATTTTGCAGCCAATTTTTTAAATATTTAAATTTTCAT
AATAGATCAAAGGATAATTCAAATTCATCCAATAACAACATTAGTAATGGAAGGACTTA
TGGTATGAATGGATCAATAATATAAGGGCTGAATTAACAACATTTTTTTTTATTTAGATCCTG
TTTATTTTTACGTTTTAAAAATATTTTTGAAATTATTTATTTTTTATTATAAATTAATATT
TTTAGATCATTTTAATACGTTAATATAAAAAATAATTTTTTTAAAAAATTTATTTTAATAT
ATTTTTTAAAAATAATTTTTAAAAACAATCATAACAATATCTCATTACCTAACACAGTC
ATGGAACAGGAATGAGAAAAGGTCTTATCAGTAAATTGCTTGCATGTCATGTCAAGGTGTAT
GAACCTCCAATACTTCTCACGCTACCCTTCAGAAATCCAATCTCAGAAGCCACAGACAATC
TAAGTTACGCTACAATCAACTTCCATCACCCCTTCCCTTATTTAGAACTCCACTTAATCAC
ATTTACCCTTTTTCATCATCTTCTCTTCCCTTCAAGAAGCCTAGGTACTGTGCAAGAAAC
CCTTATCTCTCCCCCTCAGTATTTACTTTTGTTTAGTGCTACAGCTTTCACAAAGAAGTAAG
GAAAAAAT**ATG**

APÉNDICE 3

Secuencias completas de *TbAP3v1* y *TbAP3v2*

>TbAP3v1 (1059 pb)

AAATCAACATCTGAATGAGGTAGTATTATCCTTGAAAATTGAAACAAGGTGGTGTGAAGTT
TGAAGTAATAAATAAATATTTCTTTTTTGGCTTGGATGGAAGTGATGATGAATACAATCCGA
GAAGATCATTCACTGGTTTTGTCACTGGGGGTGGGGACGTGACGCATTCTTGGACTGGGGC
TTTAATCCCATCACATGCTTTCATGCTTTTCCATTCCCTTCTCTCTCTCTCGGCCGTCT
TCTTCTTGATCTTCTTCCCTTCTTCCCTCCTGTACAGAAAGCTGGAGGTGCC**ATGGGCAG**
GGGAAGATCGAGATCAAGAAGATCGAGAACCCACCAACCGGCAGGTGACCTATTCCAAG
CGCCGGCTGGGGCTGATGAAGAAGGCCAAGGAGCTCACCGTCTCTGCGACGCCCAGGTCT
CCATCGTCATGGTCTCCTCCACCAACAAGTTCGCCGAGTACTGCAGCCCC**TCCACCGAGTA**
CCCATCCCTCCCTTCCCTCTTCTTTTTTTAACTAACACACCAAACCTAATTCCAACCT
CCACACAGTCTGAAGACCATCTACGACCGCTACCAAGAAGCCTCCGGGATCAACCTCTGGA
CGGCTCAATACGAGAAGATGCAGAACGCGCTGAGCCATCTGAAGGAGATCAACCGCAACCT
GAGGAGAGAGATCAGGCAGAGGATGGGGGAGGAGCTGGAGGAGATGGAGGCGCAGGAGCTG
CGCGGTCTTGAGCAAACCTTGGACGAGTCACTGAAGCTCGTCCGCCAGAGGAAATATCATG
TGATCACCACACAGACAGATACTTACAAGAAGAAGATAAAGAATTCTCATGAAGCCTACAA
GACTTTGCTGCGAGAGCTGGAAATCCGAGATGGGGGAGGCGATATGGGTTATGTTATCGAT
GAGGAGGCTGTCACCGGTTATGATGGAGGTACGCTGGCGATGGGGAACGGAGGTAGCCCTC
ACATGTTACATTCGCGGTTACGCCAGCCAGCCCAATCTCCACCTCATGGGCAATGGTTT
CCATGATCTTCGTCTCGCC**TAG**

>TbAP3v2 (971pb)

CTGGAGAGGTGGTGACCCCTTCCCTAAATCCGGCCAAGACACCACACGTTCCCACTCCTCCT
CCTCCCTCGTGACTTTCATCGATTTCGAAACCCACATCCAGAAAGCTTCCCCCATCCTTT
CCTCATTTAGAAACCCTGCGGCCACCACCCTTCTCTCCCTCCATCCCTCTATATGAAACC
CTACCGTGGTCCATAAGCTTGGAGGGAAACCCTAGTTGTGTCTCC**ATGGGCAGGGGGAAG**
ATCGAGATCAAGAAGATCGAGAACCCGACCAGCCGCGCAGGTGACCTACTCGAAGCGCCGCC
TCGGCCTCATGAAGAAAGCCAAGGAGCTCACCGTCTCTGCGACGCCCAGGTGGGCGTCAT
CATGATCTCCAGACCAAGAAGCTCGCGGAATACTGCAGCCC**TTCACCAAGTACCCAAA**
ACCCTAATCTCCCTCCCTCTACTTCTATAGTTGTATAATAAATCCTAAACCCTAAACCCTT
TACAGCCTCAAGTCGATGTACGACCGCTACCAGGAAGCAACGGGCATCTATCTCTGGGGCT
CCCAATACGAGAAGATGCAGAACGCGCTGAGCCACCTGAAGGAGAACAACATGAATCTCAG
CCATGAAATCAAGCTGAGGATGGGGGAGGAGCTCGAGGGGATGGACGTGCAGGAATTGCGT
GGTCTTGAGCAAACAATGGATGAGGCACTCAAGATCGTGCGCCAGAGGAAGTATCATGTGA
TCACTACACAGACAGAGACATAACAAGAAGAAGTTAAGAACTCCAACGAGGCGTACAAGAA
CCTCGTCCGGGAGCTGGAGATGCGCAACGAACATGTCGACATTGTCTACGCAGACGTGATT
GGTGCGGGCGGCCGCTATGCTGGGCCGGTGCCGGTCGAGATGCTCACCTTCTTGTGGCGC
CGCCGGGCCACCCTTCGACAATGAACTATGCTATTCACGACCTCCGCCTGGCC**TAA**

Dominio MADS

Dominio I

Dominio K

Dominio C-terminal

Promotor/1er intrón

APÉNDICE 4

Motivos detectados con FrameWorker

Nombre del sitio	Factor de transcripción	Procesos en los que participa	Referencias
AHBP	HB (HD-ZIP)	Desarrollo; estrés biótico y abiótico	Sessa et al. (1997)
AREF	ARF	Señalización de auxinas; desarrollo	Ulmasov et al. (1999)
BRRE	ARF	Desarrollo y crecimiento de la raíz	Vert et al. (2008)
CCAF	Myb	Ritmos circadianos	Harmer et al. (2000)
CE1F	N/A	Señalización de respuesta a ABA y azúcares	Busk y Pages (1998)
CNAC	NAC/NAM	Estrés biótico	Olsen et al. (2005)
DOFF	DOC-zinc finger	Activador o represor del desarrollo y el crecimiento	Yanagisawa (2002)
DREB	AP2/ERF	Estrés abiótico	Liu et al. (1998); Nakano et al. (2006)
EREF	AP2/ERF	Estrés biótico y abiótico	Nakano et al. (2006)
GAGA	HSF	Respuesta a luz; desarrollo	Shopland et al. (1995)
GBOX	bZIP	Señalización de luz; estrés biótico y abiótico	Kang et al. (2002)
GTBX	Myb	Señalización de cAMP; desarrollo	Liu et al. (1999)
IBOX	GATA-zinc finger; Myb-like	Respuesta a luz; desarrollo	Rose et al. (1999)
IDDF	C2H2/C3H-zinc finger	Desarrollo y crecimiento	Kozaki et al. (2004)
L1BX	Homeodominio	Especificación de órganos	Schrack et al. (2004)
LEGB	ABI3	Desarrollo de semillas	Bäumlein et al. (1992); Mönke et al. (2004)
MADS	MADS	Crecimiento y desarrollo	Shore y Sharrocks (1995); Álvarez-Buylla et al. (2000); Honma y Goto (2001)
MYBL	Myb-like	Desarrollo	Wang et al. (1997); Yanhui et al. (2006)
MYBS	Myb-R1	Desarrollo	Wang et al. (1997); Yanhui et al. (2006)

MYCL	bHLH (Myc)	Desarrollo; estrés biótico y abiótico	Cole y Cowling (2008)
NACF	NAC/NAM	Desarrollo; estrés biótico y abiótico	Uauy et al. (2006)
NCS1	N/A	Interacción planta-patógeno	Ramlov et al. (1993)
OCSE	bZIP	Estrés oxidativo, biótico y abiótico	Kobayashi et al. (2008)
OPAQ	bZIP	Desarrollo del endospermo; señalización de ABA	Lohmer et al. (1991); Schmidt et al. (1992)
PSRE	N/A	Expresión específica en polen	Eyal et al. (1995)
SALT	PHD	Estrés abiótico	Xiong et al. (2002); Zhu (2002)
SBPD	SBP	Desarrollo; desarrollo y crecimiento de raíz	Guo et al. (2008)
SNBE	SWN	Biosíntesis de pared secundarias	McCarthy et al. (2011)
TERE	N/A	Muerte celular programada; biosíntesis de pared secundaria	Pyo et al. (2007)
VRES	bZIP	Estrés abiótico	Pitzschke et al. (2009)