



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
INSTITUTO DE BIOLOGÍA
BIOLOGÍA EVOLUTIVA

**Inferencias Funcionales de la madera dimórfica mediante el análisis transcriptómico
en Cactaceae.**

TESIS

QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS BIOLÓGICAS

PRESENTA:

EMILIO PETRONE MENDOZA

TUTORA PRINCIPAL DE TESIS: DRA. TERESA TERRAZAS SALGADO, Instituto de Biología, UNAM

**COMITÉ TUTOR: DRA. ALEJANDRA VASCO GUTIÉRREZ, Instituto de Biología, UNAM.
DR. LUIS DAVID ALCARAZ PERAZA, Instituto de Ecología, UNAM.**

MÉXICO, CD. MX. FEBRERO, 2018



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.




Lic. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
Presente

Me permito informar a usted que en la reunión del Subcomité por Campo de Conocimiento de Biología Experimental y Biomedicina del Posgrado en Ciencias Biológicas, celebrada el día 6 de noviembre de 2017, se aprobó el siguiente jurado para el examen de grado de **MAESTRO EN CIENCIAS BIOLÓGICAS** del alumno **PETRONE MENDOZA EMILIO** con número de cuenta **410066119** con la tesis titulada **"INFERENCIAS FUNCIONALES DE LA MADERA DIMÓRFICA MEDIANTE EL ANÁLISIS TRANSCRIPTÓMICO EN CACTACEAE"**, realizada bajo la dirección de la **DRA. TERESA MARGARITA TERRAZAS SALGADO**:

Presidente: DR. JORGE NIETO SOTELO
Vocal: DRA. MARIANA BENITEZ KEINRAD
Secretario: DR. LUIS DAVID ALCARÁZ PERAZA
Suplente: DRA. SVETLANA SHISHKOVA
Suplente: DR. ULISES YUNUÉN ROSAS LÓPEZ

Sin otro particular, me es grato enviarle un cordial saludo.

A T E N T A M E N T E
"POR MI RAZA HABLARA EL ESPIRITU"
Cd. Universitaria, Cd. Mx., a 23 de enero de 2018.


DR. ADOLFO GERARDO NAVARRO SIGÜENZA
COORDINADOR DEL PROGRAMA



c.c.p. Expediente del (la) interesado (a).

AGRADECIMIENTOS

Al posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México (UNAM), por el apoyo brindado en los trámites y la gestión de mi proyecto de maestría.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por la beca no. 700638 que fue otorgada durante el periodo 2015-2017.

Al financiamiento otorgado por el programa de Apoyos a Proyectos de Investigación e Innovación Tecnológica (PAPIIT IN209012 y 210115) a Dra. Teresa Terrazas Salgado.

A los miembros de mi comité tutor conformado por:

DRA. TERESA TERRAZAS SALGADO

DRA. ALEJANDRA VASCO GUTIÉRREZ

DR. LUIS DAVID ALCARAZ PERAZA

AGRADECIMIENTOS PERSONALES

A la Dra. Terrazas, por darme la confianza de tomar las riendas de un proyecto en el que los dos fuimos aprendiendo durante el camino. Por pensar en mi futuro académico y por las enseñanzas.

A la Dra. Ale Vasco por los consejos académicos y de vida. Por ser tan buena maestra de los helechos. Gracias

A Dr. Luis Alcaraz por sus oportunos comentarios en los tutorales.

A los miembros de mi jurado de examen conformado por: Dr. Jorge Nieto, Dra. Svetlana Shishkova, Dr. Ulises Rosas y Dra. Mariana Benítez. Por sus valiosas críticas y comentarios del manuscrito. En especial agradezco a la Dra. Svetlana por su cercana colaboración durante el 2017.

Al Dr. Jorge Reyes, por abrir las puertas a la investigación del presente proyecto y por ser buen compañero.

Al Mstro. Gustavo Rodríguez por los comentarios y sugerencias en los análisis transcriptómicos.

A toda la gente que está detrás del software libre. Por creer que la ciencia debe ser de todos y para todos.

A mi familia. Por estar presente y apoyarme siempre.

DEDICATORIA

Este trabajo está dedicado a la UNAM. Por ser una institución pública que brinda una oportunidad de formación educativa única en el mundo.

Todo el trabajo se desarrollo con software de acceso gratuito. El manuscrito se escribió en LibreOffice Writer. Las figuras se realizaron con R y se editaron con Inkscape o LibreOffice Impress. Por lo tanto este trabajo está dedicado directa e indirectamente a toda la gente detrás del software libre.

Índice

Abstract.....	8
Introducción.....	10
Método.....	17
Material vegetal, germinación y condiciones de crecimiento.....	17
Colecta, fijación de material y microtecnia.....	18
Extracciones de RNA y secuenciación.....	20
Análisis de calidad, limpieza de secuencias y ensamblado <i>de novo</i>	22
Anotación funcional.....	23
Cuantificación de los transcritos.....	23
Análisis comparativo de los transcriptomas de las zonas cambiales.....	24
Análisis de expresión diferencial.....	25
Análisis de categorías funcionales de GO enriquecidas.....	25
Resultados y discusión.....	27
Anatomía de la madera de las especies colectadas.....	27
Ensamblado <i>de novo</i> y análisis de calidad del transcriptoma de 5 especies de cactus.....	32
Cuantificación de transcritos y comparación de transcriptomas.....	42
Análisis de expresión diferencial en <i>E. platyacanthus</i> y <i>F. pilosus</i>	48
Análisis de categorías funcionales del GO enriquecidas.....	51
Conclusiones.....	57
Perspectivas.....	59
Referencias.....	60
Información suplementaria.....	65

Resumen

La variación anatómica del sistema vascular en las plantas es asombrosa. Esta variación se observa a nivel celular, a nivel de tejido y temporalmente durante distintos estadios del desarrollo de las plantas. A pesar de los grandes avances en el entendimiento de las bases genéticas del desarrollo del sistema vascular en las plantas, poco se sabe de los mecanismos genéticos que han dado lugar a la vasta diversidad anatómica. Las plantas de la familia Cactaceae tienen una gran diversidad en la anatomía del sistema vascular secundario o madera. Las maderas de los cactus se pueden clasificar en tres grandes grupos anatómicos: i) maderas con una matriz rica en fibras, ii) maderas con una matriz rica en traqueidas de banda ancha (WBT por sus siglas en inglés) y iii) las maderas dimórfas porque producen madera rica en WBT durante los primeros estadios de desarrollo y posteriormente forman madera fibrosa en función del aumento en la talla de los individuos. Por su diversidad de especies y diversidad anatómica, la familia Cactaceae permite estudiar las bases genéticas que han dado origen a la diversidad anatómica del sistema vascular. En este trabajo se reporta el transcriptoma de la zona cambial (cambium vascular y sus tejidos derivados) de cinco especies de cactus con diferencias anatómicas de la madera: *Ariocarpus retusus* (madera rica en WBT), *Cylindropuntia imbricata* (madera rica en fibras), *Echinocactus platyacanthus* (madera dimorfa), *Ferocactus pilosus* (madera dimorfa) y *Pereskia lychnidiflora* (madera rica en fibras). Después de evaluar la calidad y de eliminar los contigs duplicados se definió un set de unigenes para cada especie (*A. retusus*: 25 194, *C. imbricata*: 17 279, *E. platyacanthus*: 40 890, *F. pilosus*: 32 226 y *P. lychnidiflora*: 24 430). Se compararon los transcriptomas de los cactus con los transcriptomas del xilema de dos especies de uso forestal (*Populus trichocarpa* y *Eucalyptus grandis*), mediante las anotaciones funcionales del *Gene Ontology*. Las especies analizadas comparten muchas

categorías relacionadas con el sistema vascular secundario; sin embargo, los transcriptomas de las cuatro especies de cactus tienen más categorías de genes en común a pesar de sus diferencias anatómicas. Con el análisis de expresión diferencial en *F. pilosus* y *E. platyacanthus* se identificaron funciones relacionadas con actividad de transposones y metabolismo de ácidos nucleicos así como factores transcripcionales con unión a DNA en la zona cambial. Se encontró una duplicación de los posibles ortólogos de *KNAT7* en todos los cactus estudiados. En *E. platyacanthus* el posible ortólogo de *KNAT2-KNAT6*, perteneciente a la Clase I de la familia KNOX, se encuentra en mayor abundancia en la zona cambial con madera fibrosa que con respecto a los estadios que tienen WBT; mientras que uno de los posible homólogos de *KNAT7* (Clase II KNOX) está más representado en plantas con WBT que con respecto a la zona cambial. Se encontraron otros miembros de la familia KNAT pertenecientes a la clase I y II con expresión en la zona cambial en las otras especies. Este trabajo abre la posibilidad de hacer estudios de genética del desarrollo del sistema vascular secundario enfocados a genes relevantes con el fin de determinar las redes de regulación genética que subyacen la diferenciación de los distintos tipos de madera en la familia Cactaceae.

Abstract

The anatomic variation in the plant vascular system is amazing. This variation is observed at the cellular and tissue level and temporarily during different developmental stages. Despite advances in the understanding of the genetic basis of development of the vascular system in plant model species, little is known about the great anatomic diversity observed in nature. Plants of the Cactaceae family have great variation at the level of the secondary vascular system or wood. Cacti wood can be classified in three groups: i) woods rich in fiber cell types, ii) woods rich in wide band tracheids (WBT), iii) and dimorphic woods that produce WBT wood during the first developmental stages and later produce fiber wood. The change from one kind of wood to the other is related to the size of the plant. Cacti anatomy and species diversity makes them suitable for studying the genetic basis of vascular system diversity in plants. In this Master project we report the cambial zone (vascular cambium and the first derivative cells) transcriptomes of five cacti species with wood anatomic differences: Ariocarpus retusus (WBT wood), Cylindropuntia imbricata (fibrous wood), Echinocactus platyacanthus (dimorphic wood), Ferocactus pilosus (dimorphic wood), and Pereskia lychnidiflora (fibrous wood). After evaluating the quality of the transcriptomes and reducing the duplicated contigs, a set of unigenes was defined for each species (A. retusus: 25194, C. imbricata: 17279, E. platyacanthus: 4089, F. pilosus: 32226 and P. lychnidiflora: 24430). A comparison of the Cactaceae cambial zone transcriptomes and the xylem transcriptomes of two woody species (Populus trichocarpa and Eucalyptus grandis) was made using the functional categories of Gene Ontology. All analyzed species share Gene Ontology functional categories related to the secondary vascular system but cacti transcriptomes share more categories among them despite their different wood types.

A comparative expression analysis was performed for E. pilosus and E. platyacanthus transcriptomes in order to identify differentially expressed genes between young plants and cambial zones. The cambial zone of both species has an enrichment of functions related to transposon activity, nucleic acid functions, and transcription factors that bind to DNA. Importantly, in E. platyacanthus one of these transcripts encodes a Class I KNOX transcription factor (KNAT2-KNAT6) and has a higher expression in the cambial zone than the plants that have WBT. The possible ortholog of KNAT7 (Class II KNOX) is highly expressed in seedlings and plants with WBT compared to the cambial zone. Additional members of the KNOX family were also found in the cambial zone of cacti species and there is a possible duplication for the KNAT7 clade. This work opens the possibility to study the developmental genetics of wood development in cacti focusing in relevant genes. This knowledge, could allow the finding of the gene regulatory networks that underlie the wood differentiation in the Cactaceae family.

Introducción

La variación del sistema vascular en las plantas es muy extensa y refleja una gama amplia de adaptaciones e innovaciones evolutivas a la conducción de agua y nutrientes, así como al soporte mecánico (Lucas *et al.* 2013). Esta variación comprende distintos niveles de organización, observándose a nivel celular o a nivel de tejidos, dependiendo del estadio de desarrollo y del grupo filogenético (Wunderling *et al.* 2016). Por ejemplo, las innovaciones en las traqueofitas como la pared secundaria rígida en las células conductoras del xilema permiten transportar el agua en presiones negativas y además dan el soporte necesario para que las plantas aumenten su altura mejorando la captura de luz (Etchells y Turner 2017). Otro ejemplo de variación a nivel de tejidos y grupos filogenéticos son las diferencias entre las células del xilema de las gimnospermas y las angiospermas derivadas del cambium vascular. En las gimnospermas las traqueidas se encargan tanto del transporte del agua como del soporte mecánico; mientras que en las angiospermas los vasos transportan el agua y las fibras dan soporte mecánico (revisado en Ohtani *et al.* 2016).

El cambium vascular es el meristemo que se encarga del crecimiento secundario (engrosamiento radial) y de la formación del tejido vascular secundario en las Eudicotiledóneas y en las Gimnospermas. A diferencia de los meristemos apicales el cambium vascular es un meristemo lateral porque se encuentra en posición paralela a la superficie del tallo y raíz de la planta (Evert 2006). Las divisiones asimétricas de este meristemo dan origen a dos grupos de células: un grupo mantiene la identidad de células no diferenciadas y el otro grupo se vuelve a dividir para eventualmente diferenciarse en xilema secundario o en floema secundario (revisado en Bhalerao y Fischer 2006). En la Figura 1 se muestra un esquema de la orientación de las divisiones del cambium vascular y los tejidos

derivados en un tallo con crecimiento secundario. Las divisiones periclinales (paralelas a la superficie del órgano) del cambium vascular forman a las células que eventualmente se diferencian en el xilema secundario o madera al interior del tallo y floema secundario o corteza interna al exterior del tallo. Por otra parte, las divisiones anticlinales (perpendiculares a la superficie del órgano) generan un aumento en la circunferencia del tallo. El cambium vascular tiende a formar una circunferencia alrededor del xilema secundario, sin embargo si este se encuentra en hileras discretas, se restringe a formar bandas en la posición del xilema.

El cambium vascular, en conjunto con sus células derivadas tiene la capacidad de comunicarse entre sí, integrar señales internas y externas y al mismo tiempo mantener cierta independencia que permite la evolución modular (revisado en Aloni 2015; Olson y Rosell 2006). Para el clado de las angiospermas la actividad cambial se ha perdido de manera independiente tres veces y en varios grupos la alteración en las tasas de división entre xilema y floema tiene claros impactos en la aparición de anatomías novedosas (revisado en Spicer y Groover 2010; Pace, Lohmann, y Angyalossy 2011).

Los avances del conocimiento realizados en las causas de la variación anatómica del sistema vascular y su distribución filogenética, así como en el impacto de los factores ecológicos en la evolución del sistema vascular, sobrepasan los avances en el entendimiento de los mecanismos genéticos del desarrollo. Por ejemplo, los trabajos de anatomía ecológica inferencial a nivel de grupo taxonómico permiten determinar cuáles características anatómicas son relevantes en un contexto ecológico y tienen un valor adaptativo (Carlquist 1980; Lens *et al.* 2016). Con estos trabajos se ha determinado qué tanto de la variación observada se debe a factores ecológicos o a la inercia filogenética o adaptabilidad del grupo en estudio (Lens *et al.* 2004). Sin embargo, estos trabajos no contestan cuál es el origen de

la diversidad anatómica del sistema vascular entre linajes. Por otra parte, los sistemas que se han utilizado para entender las bases genéticas del desarrollo vascular se limitan a especies modelo como *Arabidopsis thaliana* (Miyashima *et al.* 2013) y especies del género *Populus* (Ohtani *et al.* 2011), cuya anatomía no refleja la diversidad de anatomías existentes. Por ello, un componente crucial para entender las diferentes anatomías vasculares son los estudios comparativos y evolutivos que estudien la función de los genes reguladores del sistema vascular en taxones faltantes.

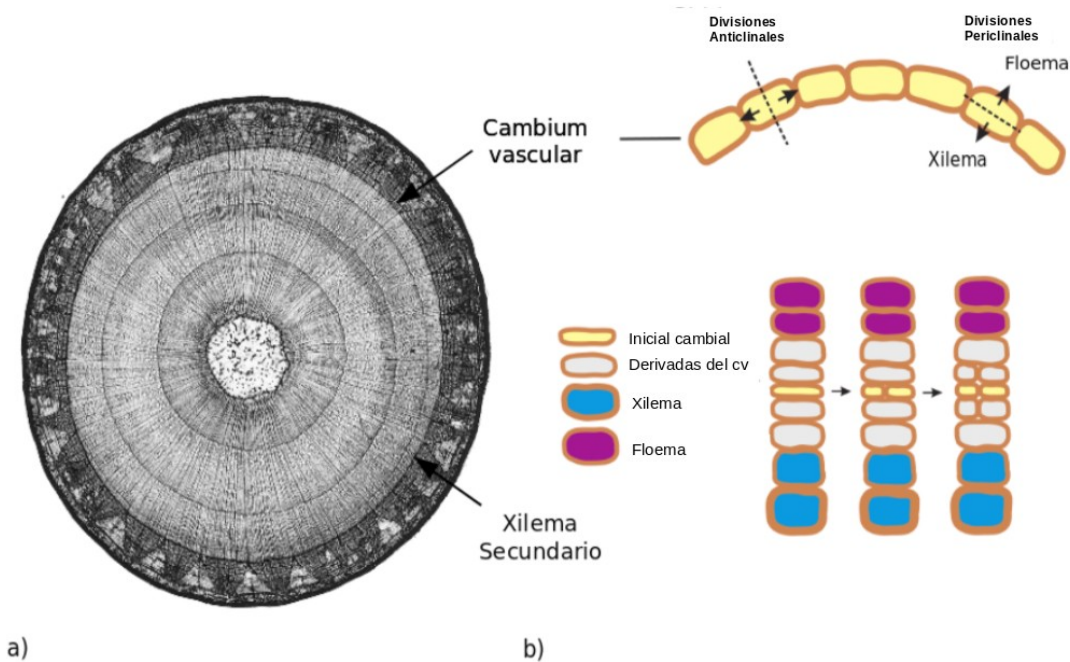


Figura 1. a) Sección transversal de un tallo del género *Tilia* con crecimiento secundario. b) Las iniciales cambiales se dividen anticlinalmente aumentando la circunferencia del tallo y periclinalmente para producir xilema hacia adentro y floema hacia afuera. Figura modificada de Evert (2006) y de Spicer y Groover (2010).

Los estudios de la evolución del desarrollo (*Evo-devo* por sus siglas en inglés) han comenzado a explicar cómo surgen las innovaciones evolutivas o cuáles son las bases genéticas que originan la diversidad morfológica (revisado en (Tickle y Urrutia 2017). Por ejemplo, la evolución de los mecanismos de desarrollo depende en gran medida de: i) la

duplicación de patrones o procesos (desde genes hasta órganos), ii) la modificación de dominios de expresión de genes o iii) la cooptación de distintos niveles de organización biológica (Moczek *et al.* 2015). Para estudiar la evolución del sistema vascular en las plantas se requiere de un sistema de estudio en el que i) las relaciones taxonómicas estén bien definidas, ii) la variación anatómica y ontogenética esté bien identificada y iii) se cuente con recursos genéticos para la identificación de genes relevantes (Spicer y Groover 2010). Con los avances en las tecnologías de secuenciación, la disminución en los costos y el mejoramiento de los algoritmos de ensamblaje de secuencias nucleotídicas (RNA-seq) sin tener un genoma de referencia (ensamblado *de novo*), el número de estudios con énfasis en la expresión génica en órganos, estadios de desarrollo y especies ha aumentado considerablemente (revisado en Oszolak y Milos 2011; Necsulea y Kaessmann 2014).

Un sistema útil para el estudio de la evolución del sistema vascular secundario en plantas es la familia Cactaceae. Los cactus exhiben una gran variación morfológica en el tallo asociada a una gran diversidad anatómica (Gibson 1973; revisado en Mauseth, 2006). Los estudios sobre la anatomía (Gibson 1973, 1977; Mauseth 2004); Terrazas y Arias 2002) y ontogenia del sistema vascular (Mauseth y Plemons-Rodriguez 1998; (Loza-Cornejo y Terrazas 2011) de los cactus son extensos. Se sabe que algunas especies retienen características juveniles en el xilema secundario de plantas adultas (Altesor, Silva, y Ezcurra 1994; Carlquist 2009). Además, un tipo celular derivado del cambium de algunas especies de cactus son las llamadas traqueidas de banda ancha (WBT, por sus siglas en inglés). Estas células son distintas de las traqueidas o fibro-traqueidas ya que tienen paredes primarias delgadas y paredes secundarias anulares o helicoidales que se proyectan profundamente en el lumen de la traqueida, no presentan perforaciones (Figura 2) y se les ha catalogado como una novedad evolutiva (Mauseth, 2006).

Las maderas en la familia Cactaceae pueden ser fibrosas, no fibrosas (ricas en WBT) o dimórficas, esto último significa que el cambium vascular produce madera no fibrosa durante los primeros estadios de desarrollo y madera fibrosa cuando la planta alcanza tallas más grandes (Reyes-Rivera *et al.* 2015). En general, los cactus con crecimiento de tipo arbóreo que alcanzan tallas grandes tienen madera fibrosa (Figura 2). Los cactus con formas globosas y discoidales que mantienen tallas pequeñas tienen una madera no fibrosa (Terrazas y Arias 2002), mientras que los cactus con formas columnares y cilíndricas pueden tener madera dimorfa. Estudios recientes indican que la composición de las maderas ricas en WBT y las maderas ricas en fibras tienen un bajo y alto porcentaje de lignina, respectivamente (Reyes-Rivera *et al.* 2015), de tal forma que el proceso de lignificación ocurre cuando se requiere más soporte mecánico, cambiando a una madera fibrosa. Una pregunta interesante sobre el desarrollo vascular de los cactus es si los genes implicados en el proceso de formación del xilema primario fueron cooptados en el desarrollo de las maderas ricas en WBT. En otras palabras, si las especies que tienen un proceso de retención de características del xilema primario en el xilema secundario mantienen los mismos patrones de expresión en los genes implicados en el proceso de formación del xilema primario en el tiempo.

Considerando que i) la transcripción es un nivel importante en la regulación y expresión genética de diversos procesos de desarrollo, incluyendo el crecimiento vascular secundario (Spicer y Groover 2010), ii) existe una mayor accesibilidad al análisis transcriptómico en especies no modelo y iii) la variación anatómica del sistema vascular secundario observada en la familia Cactaceae permiten estudiar las bases genéticas del desarrollo de la madera y su papel en el origen de la diversidad anatómica en linajes cercanos. Este trabajo tuvo como objetivo ensamblar *de novo* los transcriptomas de distintos representantes de la familia

Cactaceae, que comprenden especies con madera fibrosa, madera rica en WBT y madera dimórfica, con el fin de identificar genes implicados en el proceso de diferenciación anatómica y lignificación del xilema. De cada tejido colectado de cada especie se realizaron cortes anatómicos para entender a detalle la composición celular de cada estadio de desarrollo. Adicionalmente se hizo una comparación global de los transcriptomas en términos funcionales con otras especies maderables. También se llevó a cabo un análisis de expresión diferencial con dos especies y se identificaron factores transcripcionales con potencial para realizar estudios funcionales en el futuro.

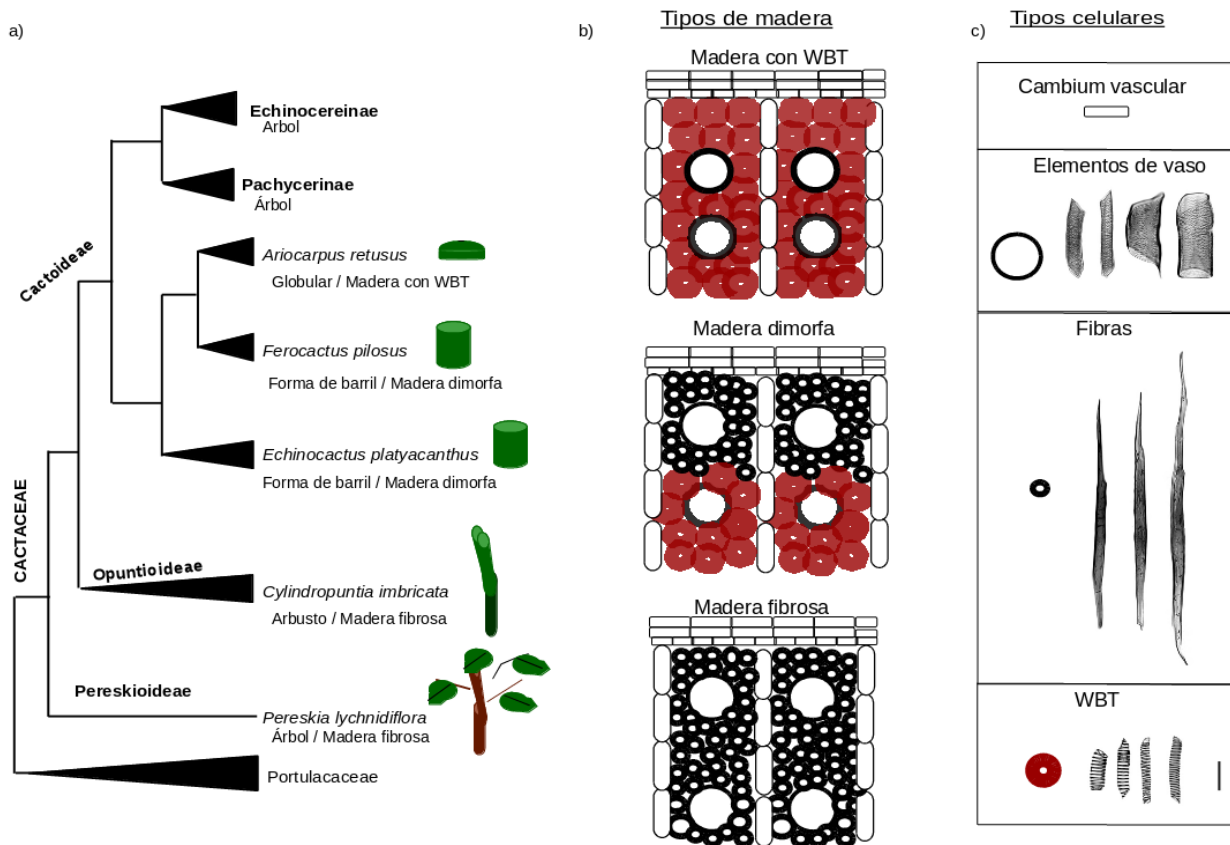


Figura 2. Filogenia de la familia Cactaceae y su tipo de madera. a) Filogenia con al menos un representante de cada subfamilia. b) Principales tipos de madera en el grupo y c) los tipos celulares asociados a cada madera. WBT= traqueidas de banda ancha. Topología basada en (Hernández-Hernández *et al.* 2011).

Objetivo general

Ensamblar *de novo* y analizar los transcriptomas de la zona cambial de distintas especies de la familia Cactaceae con diferentes tipos de madera.

Objetivos particulares

- Describir la anatomía de al menos una especie de Cactaceae con madera fibrosa, madera no fibrosa y madera dimorfa.
- Ensamblar el transcriptoma *de novo* de al menos una especie con madera fibrosa, madera no fibrosa y madera dimorfa.
- Realizar un análisis transcriptómico comparativo del sistema vascular entre las distintas especies de cactus.

Método

Material vegetal, germinación y condiciones de crecimiento

En este estudio se utilizaron cinco especies de la familia Cactaceae con diferencias en la anatomía de la madera: *Pereskia lychnidiflora* DC. (subfamilia: Pereskoideae); *Cylindropuntia imbricata* (Haworth) F. M. Knuth (Opuntioideae, tribu: Cylindropuntieae) con madera fibrosa; *Ariocarpus retusus* Scheidw (Cactoideae, tribu: Cactaeae) con madera no fibrosa; *Ferocactus pilosus* (Galeotti ex Salm-Dyck) Werderm (Cactoideae, Cactaeae) y *Echinocactus platyacanthus* (Link & Otto) (Cactoideae, Cactaeae) con madera dimorfa (Figura 3). Para identificar los transcritos presentes durante la ontogenia del sistema vascular, se colectaron frutos (el Cuadro 1 muestra los sitios de colecta) y se germinaron semillas de las especies anteriormente descritas con excepción de *A. retusus* debido a la dificultad para obtener y a su baja producción de semillas por fruto. Se extrajeron las semillas de los frutos y se desinfectaron por inmersión en una solución de hipoclorito de sodio comercial al 30% en agitación constante durante 15 minutos. Se realizaron cuatro lavados con agua destilada, dejándose reposar después del último por un periodo de 24 horas. Para la germinación, las semillas se sembraron sobre papel filtro saturado con agua destilada en cajas Petri estériles. Las cajas se colocaron en una cámara de crecimiento *Thermo Scientific™ Precision™ Plant Growth Chamber* a una temperatura constante de 25 °C con periodos de 12 horas luz y 12 horas oscuridad. Las plántulas obtenidas se trasplantaron a charolas con una mezcla de agrolita-tezontle-tierra para maceta (1:1:1) y se mantuvieron en la cámara de crecimiento bajo las mismas condiciones con períodos de riego de 15 días.

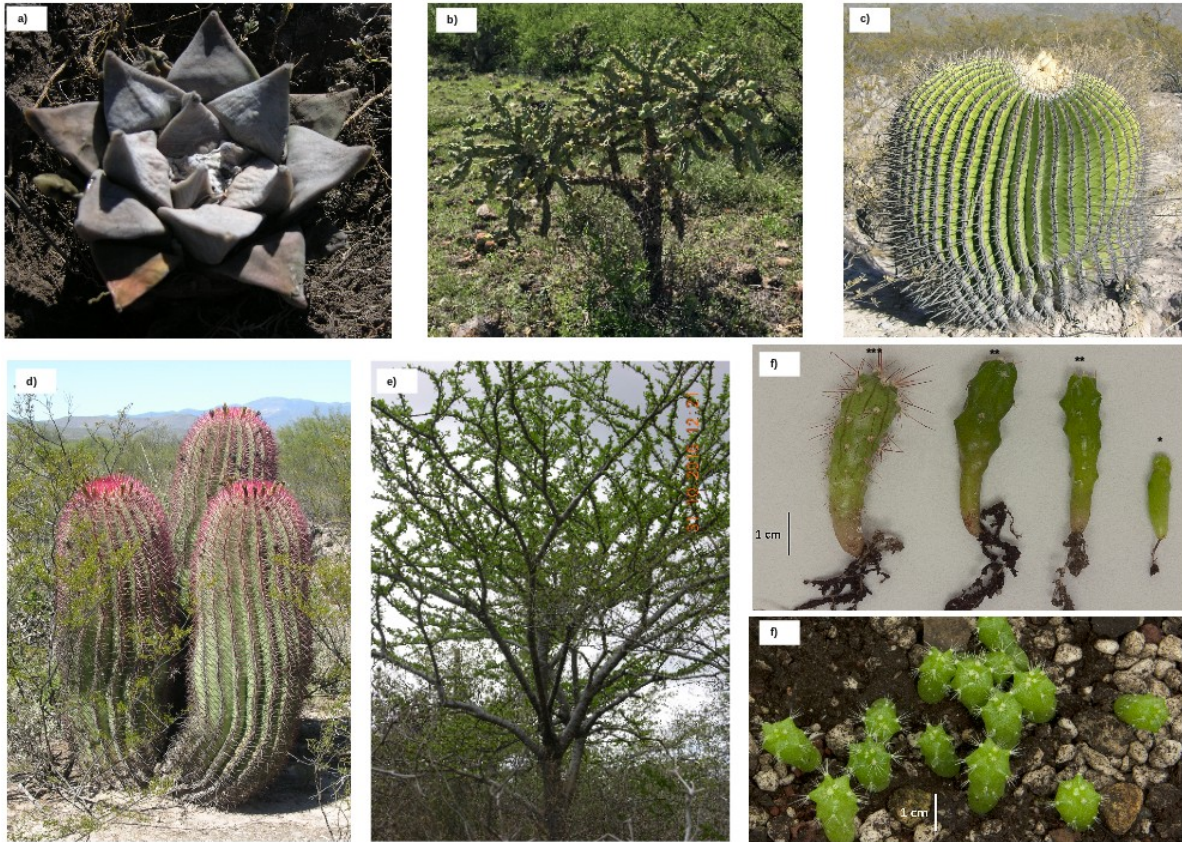


Figura 3. Especies de la familia Cactaceae utilizadas en el estudio. Plantas adultas de a) *Ariocarpus retusus*, b) *C. imbricata*, c) *E. platyacanthus*, d) *F. pilosus* y e) *P. lychnidiflora*. Plantas de f) *E. platyacanthus* de un año siete meses(***), ocho meses (**) y un mes con 15 días (*). Plantas de tres meses de g) *F. pilosus*.

Colecta, fijación de material y microtecnia

Dependiendo del estadio de desarrollo de los individuos de cada especie, se colectaron plantas completas, únicamente zona cambial o el cilindro vascular (Cuadro 2). La formación de tejido vascular en los cactus es escasa (Mauseth 2006b) y aislar el tejido vascular durante los primeros estadios de desarrollo resulta complicado; por ello se colectaron plantas completas de un mes con 15 días de *E. platyacanthus* y de tres y cinco meses de *F. pilosus*. Además, se colectó el cilindro vascular de plantas de *E. platyacanthus* con madera rica en WBT (ocho meses y un año con cinco meses). Para ello, se retiró la epidermis y el córtex dejando la zona cambial y la médula de la planta (Figura 4). Se fijaron en nitrógeno líquido inmediatamente después de haber sido removidas de la tierra y posteriormente se

resguardaron a -80°C . De los individuos de mayor edad que acumularon mayor cantidad de tejido vascular secundario se colectó la zona cambial. Se removieron los tejidos circundantes incluido el córtex, según lo describe Liu *et al.*, (2014). En las plantas adultas de tallas grandes de *F. pilosus*, *P. lychnidiflora* y *C. imbricata* se aisló la zona cambial de un individuo en su hábitat natural. De *A. retusus* se colectó un individuo de su hábitat natural junto con el suelo nativo y se colocó en maceta para su posterior procesamiento en el laboratorio. El espécimen adulto de *E. platyacanthus* fue donado por el Jardín botánico de la UNAM y colectado de igual forma que *A. retusus*. La colecta de la zona cambial en todas las especies se realizó durante la temporada de lluvias para asegurar la actividad cambial. Los tejidos de todas las especies y estadios de desarrollo se fijaron y procesaron por microtecnica de inclusión en parafina (Loza Cornejo y Terrazas 1996) para obtener cortes transversales y longitudinales y hacer preparaciones permanentes.

Cuadro 1. Lista de especies y sus sitios de colecta tanto de la zona cambial como de las semillas usadas para germinación.

Espece	Localidad	Tejido colectado
<i>Ariocarpus retusus</i>	Altiplano mexicano, estado de Nuevo León, 23 ° 23 39 N; 100 ° 21 57 W, voucher SA1976 MEXU	Zona cambial
<i>Echinocactus platyacanthus</i>	Valle de Tehuacán, estado de Oaxaca 18° 12 11 N; 97° 39 36 W	Semillas
	Jardín Botánico UNAM, Ciudad de México N 19° 19 8 N; 99° 11 43 W	Zona cambial
<i>Cylindropuntia imbricata</i>	Altiplano mexicano, estado de Guanajuato, 20°54´4"N;100°24´37".	Semillas y zona cambial
<i>Ferocactus pilosus</i>	Altiplano mexicano, estado de San Luis Potosí, 23° 19 31 N;100° 33 30 W, voucher TT890 MEXU	Semillas y zona cambial
<i>Pereskia lychnidiflora</i>	Istmo de Tehuantepec, estado de Oaxaca, 16°22 52 N;95° 19 03 W, voucher TT966 MEXU	Semillas y zona cambial

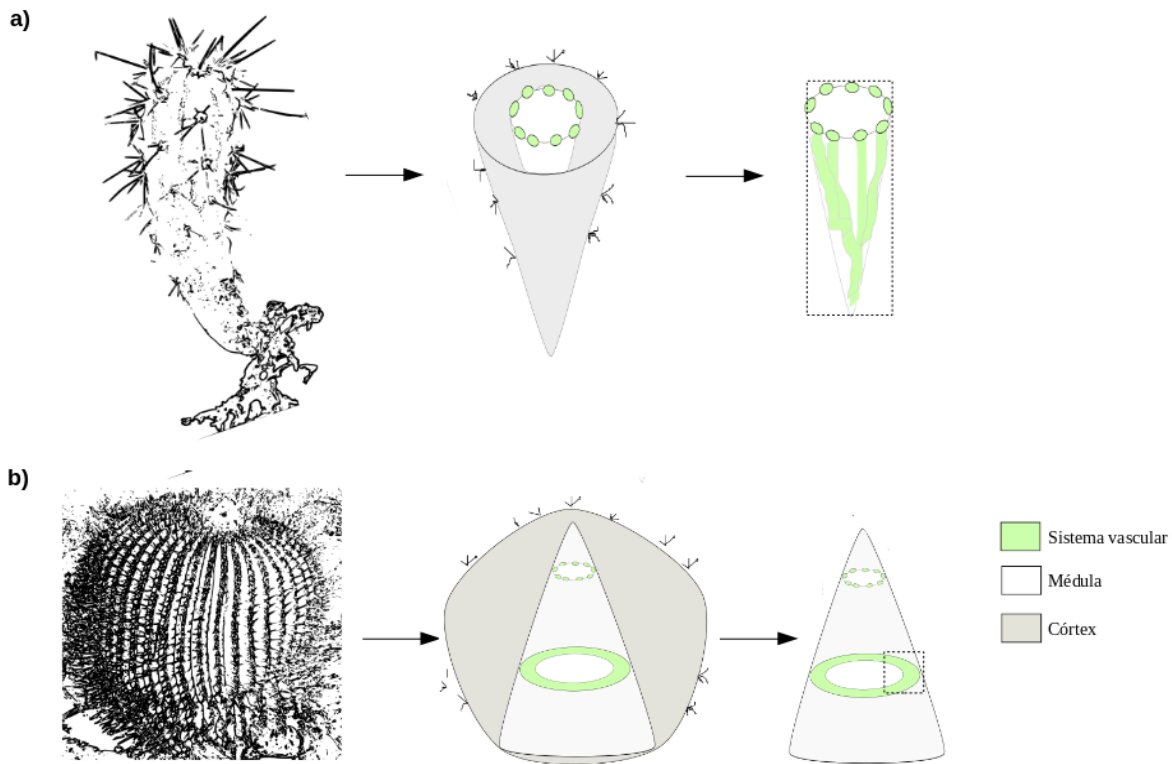


Figura 4. Aislamiento del cilindro vascular más la médula en plantas de a) un año cinco meses u ocho meses y b) de la zona cambial en planta adulta de *E. platyacanthus*. El cuadro punteado denota la zona aislada para posteriormente aislar en tejido vascular.

Extracciones de RNA y secuenciación

Las extracciones de RNA total se llevaron a cabo con el reactivo TRIzol (*Invitrogen*, CA, USA), de acuerdo con las instrucciones del fabricante (incluido el paso adicional de centrifugación previo a la separación de fases). Posteriormente, las muestras se procesaron de dos formas distintas: por método convencional de separación de fases fenol/cloroformo y precipitación por isopropanol o por purificación de columna con el *kit Direct-zol RNA MiniPrep* (*Zymo Research*, CA, USA). La calidad y cantidad del RNA se evaluaron por medio del *Agilent 2100 Bioanalyzer* (*Agilent Genomics*, CA, USA) y *Nanodrop 1000* (*Thermo Scientific*, MA, USA) respectivamente. Las muestras que se seleccionaron para secuenciar tienen un intervalo del número de integridad del RNA (RIN, por sus siglas en inglés) de 6.4 a 8.7 y un intervalo de concentración de 50-400 ng/μl (Cuadro S1).

Cuadro 2. Tejidos colectados por especie y edad para extracciones de RNA.

Especie	Edad de la planta	Tejidos colectados
<i>Ariocarpus retusus</i>	Planta adulta (edad desconocida)	Zona cambial
<i>Echinocactus platyacanthus</i>	1 mes con 15 días	Planta completa
	8 meses	Cilindro vascular
	1 año con 5 meses	Cilindro vascular
	Planta adulta (edad desconocida)	Zona cambial
<i>Ferocactus pilosus</i>	3 meses	Planta completa
	5 meses	Planta completa
	Planta adulta (edad desconocida)	Zona cambial
<i>Cylindropuntia imbricata</i>	Planta adulta (edad desconocida)	Zona cambial
<i>Pereskia lychnidiflora</i>	Planta adulta (edad desconocida)	Zona cambial

La preparación de las bibliotecas de cDNA y la secuenciación se realizó en dos sitios: en la Unidad de Secuenciación Masiva y Bioinformática del IBt de la UNAM y en el BGI-Hong Kong NGS Lab. Se llevaron a cabo 3 corridas de secuenciación. En la primera se prepararon dos muestras de *C. imbricata* y *A. retusus* y una muestra de *P. lychnidiflora* y *F. pilosus* con el *TruSeq Stranded RNA Library Preparation Kit (Illumina, USA)* y se secuenciaron lecturas pareadas de 75 bases en la plataforma *NextSeq 500 (Illumina, USA)* con insertos de 300 bases. En la segunda corrida se preparó una muestra de *A. retusus*, *F. pilosus* y *P. lychnidiflora* con el *TruSeq RNA Library Preparation Kit (Illumina, USA)* y se secuenciaron lecturas pareadas de 100 bases en la plataforma *HiSeq 2000*. En la tercera corrida se

prepararon 9 muestras de *E. platyacanthus* y tres muestras de *F. pilosus* con el *TruSeq Stranded RNA Library Preparation Kit* y se secuenciaron lecturas pareadas de 100 bases en la plataforma HiSeq 2000 con insertos de 300 bases (Cuadro 3 y Cuadro S2).

Análisis de calidad, limpieza de secuencias y ensamblado *de novo*

El análisis de calidad de las secuencias se realizó con el programa *Fastqc* (v.0.11.3). Para descartar las secuencias de baja calidad se utilizó el programa *Trimmoamtic* (v.0.25). Con las secuencias provenientes de la plataforma HiSeq 2000 se eliminaron las primeras 10 bases de todas las secuencias y se eliminaron las bases con una calidad phred promedio menor a 28 en ventanas de cuatro bases. Se descartaron las secuencias menores a 38 bases de longitud. Para minimizar la pérdida de secuencias provenientes de la plataforma NextSeq el set óptimo de parámetros cambió dependiendo de la biblioteca. En el cuadro S3 se presenta la lista de los parámetros utilizados. Las secuencias procesadas se combinaron para cada especie por separado y se ensamblaron con el programa *Trinity* (v.2.0.4) (Haas *et al.* 2013). La longitud de los k-meros para ensamblar las secuencias fue la preestablecida por el programa (25-mer). El parámetro de lecturas pareadas (--paired end) se utilizó para todas las especies; mientras que el de biblioteca hebra específica (--strand specific library) se utilizó con *C. imbricata* y *E. platyacanthus* ya que fueron las únicas especies cuyas bibliotecas de cDNA son enteramente hebra específica.

Para evaluar la calidad de los ensamblados se midió la proporción de lecturas pareadas alineadas correctamente con el programa *bowtie2* (Langmead y Salzberg 2012). Con el programa BUSCO (V.3) (Simão *et al.* 2015) se analizó la completitud de los transcriptomas. Para eliminar transcritos duplicados, fragmentados y quimeras formadas durante el ensamblado se utilizó el programa *tr2aacds* de *Evigene* (Gilbert y Gilbert 2016). Por último se

definió un set de transcritos que codifican para proteínas con *TransDecoder* (v.3.0.1) (TransDecoder source [2015] 2018).

Anotación funcional

La anotación funcional putativa de los transcritos se realizó siguiendo el flujo de instrucciones del programa *Trinotate* (V. 3.0.1). La similitud entre proteínas se detectó por medio de búsquedas de BLASTX ($E \leq 1 e^{-5}$) de una base de datos formada por SwissProt (Boeckmann *et al.* 2005) y UniRef90 (Camacho *et al.* 2009) descargada de UniProt (https://data.broadinstitute.org/Trinity/Trinotate_v3_RESOURCES/). Con los productos resultantes se realizaron búsquedas de similitud BLASTP ($E \leq 1 e^{-5}$), dominios de proteína conservados con Hmmer (Johnson, Eddy, y Portugaly 2010) y PFam (Finn *et al.* 2016) contra la misma base de datos anteriormente descrita. Se buscaron señales de péptido con SignalP (Petersen *et al.* 2011) y regiones transmembranales con TmHMM (Krogh *et al.* 2001). Los resultados se integraron en una base de datos relacional SQLite y se reportaron en un archivo delimitado por pestañas. Se asignaron los identificadores del *Gene Ontology* (Ashburner *et al.* 2000); (The Gene Ontology Consortium 2004) a las entradas de Swiss-Prot con los mejores alineamientos.

Cuantificación de los transcritos

La estimación de las abundancias de los transcritos por cada réplica biológica (cada muestra secuenciada es una réplica biológica) se realizó con el programa RSEM (V1.3.0), con los que se obtienen los conteos estimados de fragmentos de RNA-seq a nivel de genes e isoformas. Se evaluó la calidad entre réplicas biológicas por especie con el programa PtR de *Trinity*. PtR lleva a cabo una transformación de los conteos estimados por RSEM a conteos por millón (CPM, por sus siglas en inglés) seguida de una transformación logarítmica base dos.

Posteriormente, se grafica la dispersión de los conteos transformados entre réplicas y se hacen correlaciones de Pearson (Figura S1). En *E. platyacanthus* y *F. pilosus*, para los cuales se secuenciaron las muestras de al menos dos estadios de desarrollo con réplicas biológicas, el programa genera una matriz de correlaciones entre todas las muestras y se grafica un mapa de calor. Para preparar las matrices de comparación entre transcriptomas se realizó lo siguiente: primero se utilizó el programa EdgeR para estimar los CPM, después se filtraron los transcritos con menos de dos CPM por condición. Posteriormente, se normalizaron los conteos por medio del método de media ponderada recortada (TMM, por sus siglas en inglés) y se estimaron los transcritos por millón (TPM, por sus siglas en inglés). Para ello se dividió el número de conteos por transcrito entre su longitud, y a su vez este valor se dividió entre el total de conteos de la réplica y ese valor se multiplicó por 1×10^{-6} . Por último se eliminaron los transcritos con menos de un TPM. Las matrices resultantes se utilizaron para comparar transcritos de interés entre todas las especies.

Análisis comparativo de los transcriptomas de las zonas cambiales

Para comparar los transcriptomas de las zonas cambiales se extrajeron las categorías funcionales del *Gene Ontology* (GO, por sus siglas en inglés) del set de transcritos con al menos un TPM por condición. También se utilizaron las categorías del cambium vascular de dos especies maderables (*Eucalyptus* y *Populus*) (Zinkgraf, Gerttula, y Groover 2017) (<http://onlinelibrary.wiley.com/doi/10.1111/jipb.12538/supinfo>) con el fin de comparar con otros linajes de angiospermas. Se construyó una matriz de presencia/ausencia y se graficaron diagramas de *Venn* usando gplot en R.

Análisis de expresión diferencial

En *E. platyacanthus* y *F. pilosus* se llevaron a cabo análisis de expresión diferencial. Se utilizaron los programas edgeR (Robinson, McCarthy, y Smyth 2010) y DESeq2 (Love, Huber, y Anders 2014). Previo a la identificación de transcritos con expresión diferencial se eliminaron aquellos cuya suma fuera menor a dos CPM en alguna de las condiciones. Se verificó que la distribución de los conteos fuera similar en todas las réplicas (Figuras S2, S3) y se utilizó la función *mdsplot* de edgeR para graficar las distancias entre las tasas logarítmicas de cambio (logFC, por sus siglas en inglés) absolutas entre cada par de muestras para comprobar que las réplicas por condición se parecieran más entre sí (Figura S4). En *F. pilosus* se realizó una comparación entre las plantas de tres y cinco meses y la zona cambial, mientras que en *E. platyacanthus* se realizaron tres comparaciones distintas: plantas de un mes con 15 días (plantas completas) contra plantas de un año con cinco meses (cilindro vascular), plantas de un mes con 15 días contra zona cambial y plantas de un año con cinco meses contra la zona cambial. Los transcritos identificados por los métodos utilizados con un valor $p < 0.0001^*$ se consideraron como expresados diferencialmente.

Análisis de categorías funcionales de GO enriquecidas

Se utilizó el programa Goseq (Young *et al.* 2010) para determinar las categorías del GO enriquecidas. Goseq toma en cuenta el sesgo de la longitud de los transcritos diferencialmente expresados para determinar cuáles categorías se encuentran más representadas asociando un valor de probabilidad. Con base en la longitud de los transcritos genera una distribución nula de las categorías de cada transcrito. Se utilizó un valor $p < 0.005$ para extraer las categorías enriquecidas y con una tasa de descubrimiento falso (FDR) de 0.5.

De los factores transcripcionales pertenecientes a la familia KNOX con expresión diferencial se extrajeron las secuencias y se realizó un alineamiento múltiple contra las proteínas KNAT1,KNAT2, KNAT3, KNAT4, KNAT5, KNAT6, KNAT7 y STM de *A. thaliana* con el programa MAFFT (Kato *et al.* 2002). Posteriormente se realizó una filogenia de máxima verosimilitud con el programa PhyML bajo el modelo evolutivo de JTT para determinar su ortología (se utilizó un valor de 100 de bootstrap para probar la topología del árbol; Figura 16) (Guindon *et al.* 2010).

Resultados y discusión

Anatomía de la madera de las especies colectadas

Las plantas adultas de *P. lychnidiflora* y *C. imbricata* presentan una madera fibrosa con porosidad difusa y radios multiseriados lignificados. Los vasos en *C. imbricata* se agrupan levemente en bandas que van de tangenciales a diagonales; mientras que los vasos en *P. lychnidiflora* son solitarios. La madera de las plantas adultas de *A. retusus* consta de una matriz de WBT, vasos y algunas células de parénquima axial (Figura 3).

Las plantas de un mes 15 días de *E. platyacanthus* y de tres meses de *F. pilosus* forman una cantidad muy escasa de WBT inmediatamente después de establecer su xilema primario. La madera de plantas adultas de *F. pilosus* tiene una matriz de fibras con vasos agrupados y radios multiseriados no lignificados. Hacia la médula se observa una transición gradual de vasos y fibras por WBT en la que la parte intermedia de la transición se compone de una mezcla de WBT, vasos y fibras (Figura 4). En las plantas de un mes 15 días de *E. platyacanthus* la porción cercana al meristemo apical tiene numerosos haces vasculares con crecimiento vascular secundario y los espacios interfasciculares reducidos (Figura 5 a,b y e). Por otra parte, en las plantas de un año cinco meses de *E. platyacanthus* se observa una madera rica en WBT con escasos vasos y parénquima axial. En la parte del epicótilo se observan pocos espacios interfasciculares con numerosos haces vasculares casi formando una círculo de crecimiento vascular secundario (Figura 5c). En el hipocótilo se observan cuatro haces vasculares con mayor acumulación de xilema secundario y con espacios interfasciculares pequeños (Figura 5d). Mientras que la madera de la planta adulta tiene una matriz de fibras y vasos solitarios con radios multiseriados no lignificados. Los elementos de vaso presentan punteaduras escalariformes (Figura 5i).

En resumen, la zona cambial de plantas adultas de *C. imbricata*, *E. platyacanthus*, *F. pilosus* y *P. lychnidiflora* produce una madera fibrosa. La zona cambial de *A. retusus* y las plantas de ocho meses (no mostrada en la figura 5) y un año cinco meses de *E. platyacanthus* producen madera rica en WBT. Las plantas de tres y cinco meses (no mostrada en la figura 4) de *F. pilosus* y las de un mes con 15 días de *E. platyacanthus* tienen una cantidad mínima de WBT. Es importante resaltar que los estadios de tres y cinco meses (de aquí en adelante 3-5 meses) en *F. pilosus* tanto como los estadios de ocho meses y un año cinco meses en *E. platyanacthus* son equivalentes en términos anatómicos (todas producen madera rica en WBT) y por lo tanto se consideraron como iguales para los análisis transcriptómicos. En los estadios de desarrollo anteriormente descritos se encuentran los distintos tipos de madera que producen los cactus y por lo tanto se utilizaron para hacer los transcriptomas.

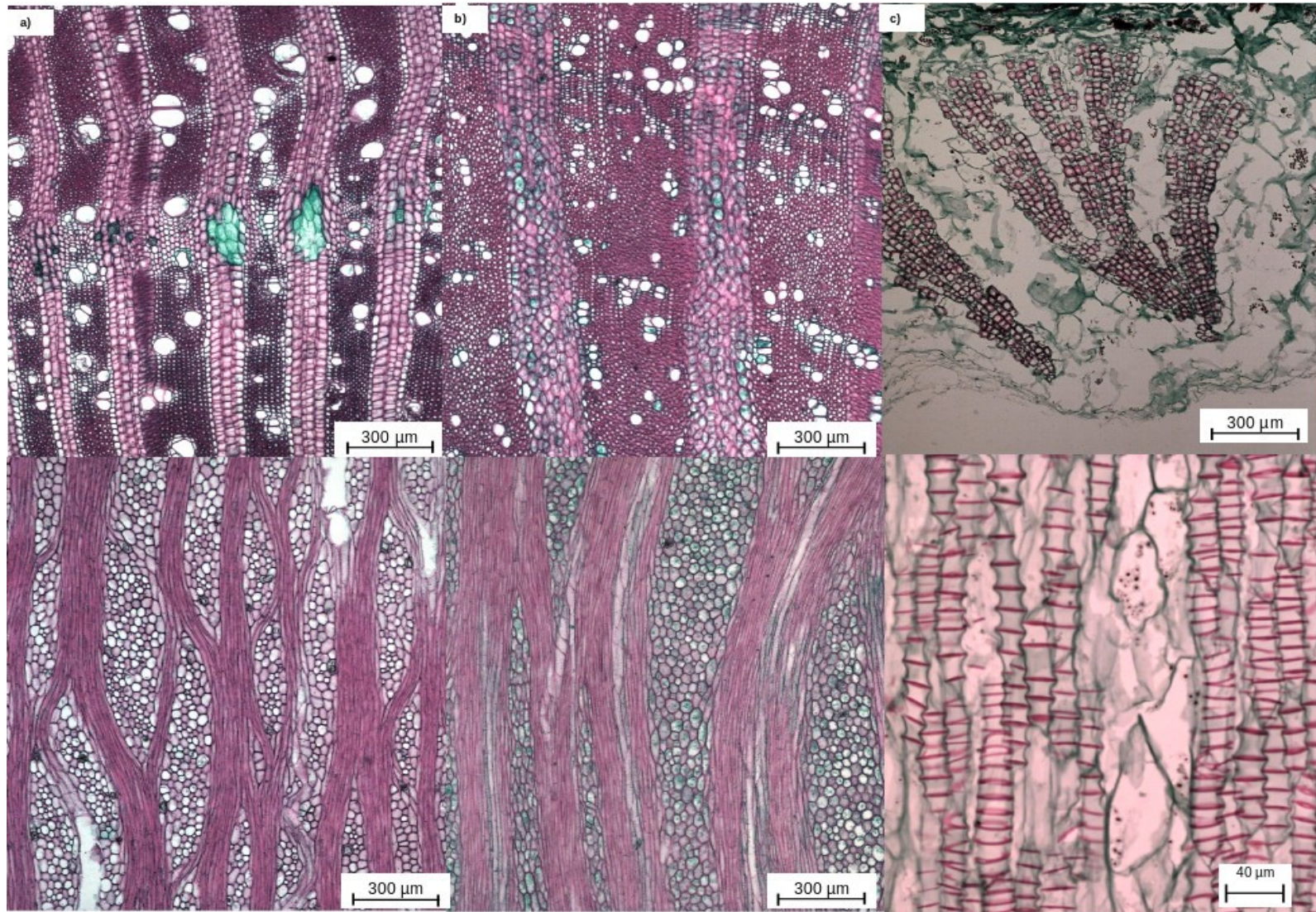


Figura 3. Anatomía de la madera de plantas adultas de a) *P. lychnidiflora*, b) *C. imbricata* y c) *A. retusus*. La parte de arriba muestra el plano transversal mientras que la parte de abajo muestra el plano tangencial.

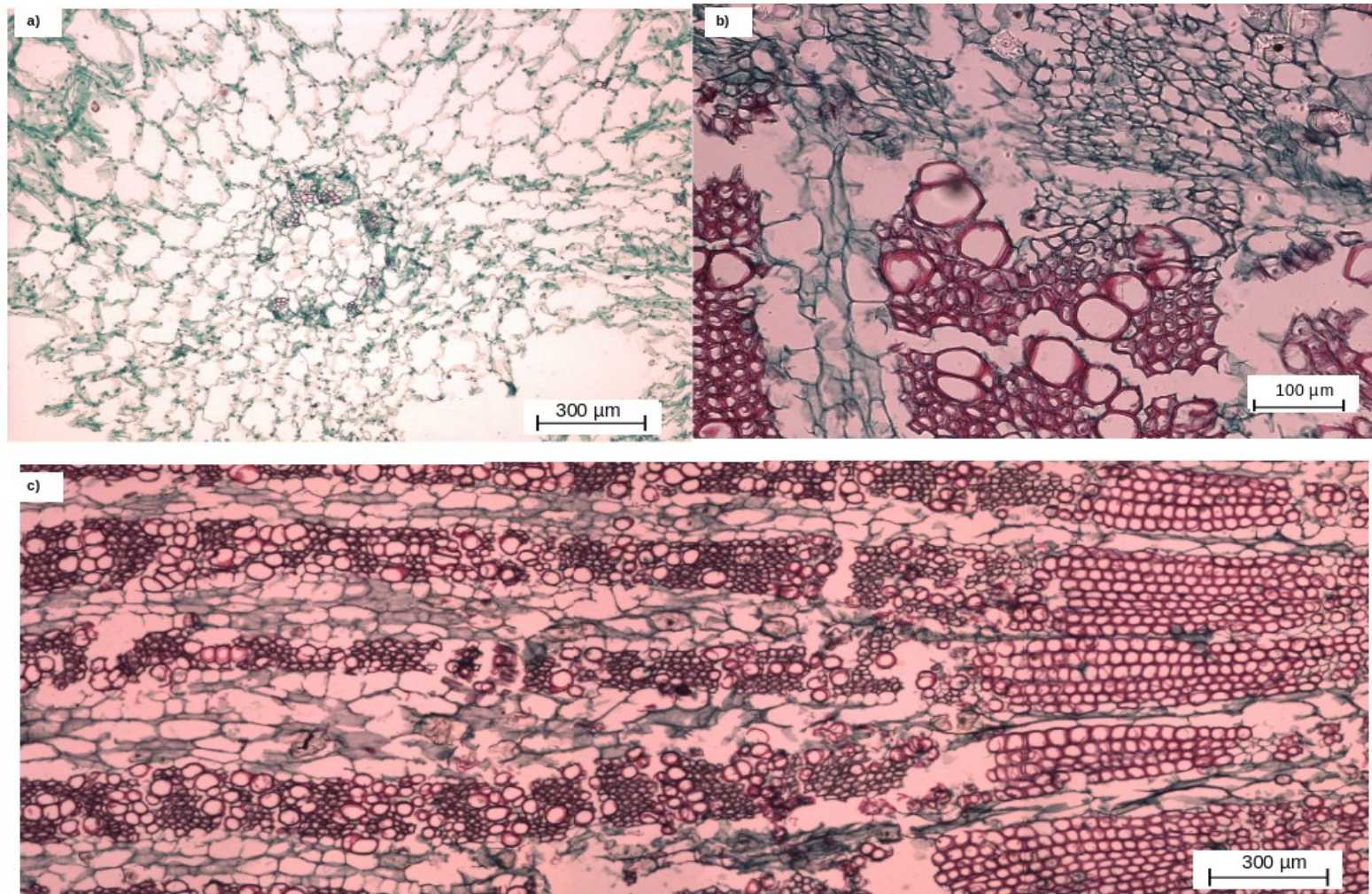


Figura 4. Anatomía de la madera de *F. pilosus* a) Plantas de tres meses con haces vasculares y los primeros elementos del protoxilema. b) Zona cambial de una planta adulta donde se observan los vasos y fibras y en c) la transición de madera rica en WBT a madera fibrosa. En la parte de la transición se observan algunas fibras, WBT y vasos.

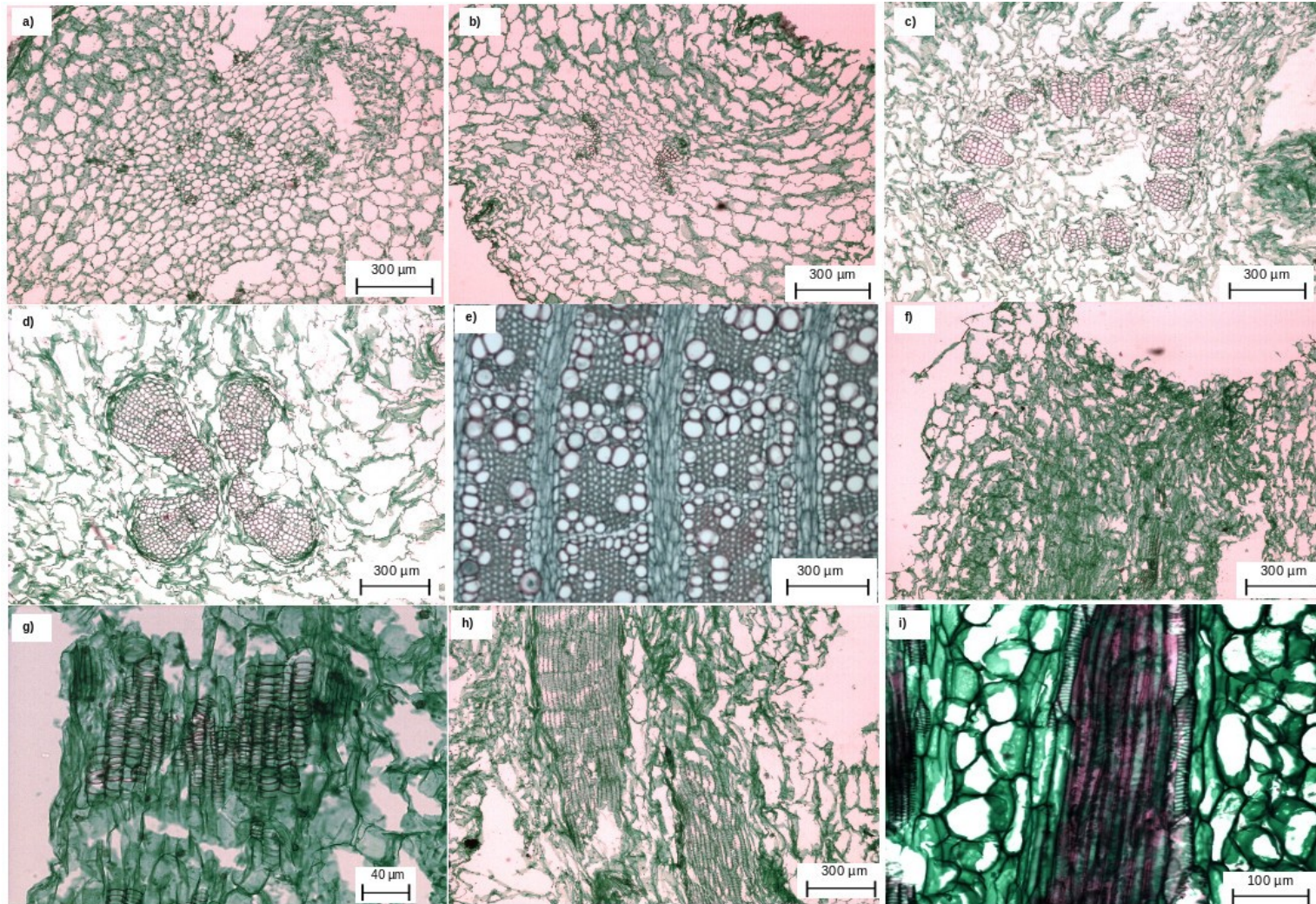


Figura 5. Anatomía de la madera de *E. platyacanthus*. Cortes transversales de plántulas de un mes y 15 días en a) epicótilo y b) hipocótilo. En c) epicótilo y d) hipocótilo de plantas de un año cinco meses y en e) planta adulta. Cortes longitudinales en f) epicótilo y g) hipocótilo de plántulas, h) plantas de año cinco meses y i) planta adulta.

Ensamblado de novo y análisis de calidad del transcriptoma de 5 especies de cactus

Después de limpiar las lecturas de baja calidad se obtuvieron 26 443 699, 5 221 561, 12 001 552, 65 790 652, 24 298 834 lecturas pareadas. En el cuadro 3 se muestran el número de lecturas pareadas antes y después de limpiar las lecturas de baja calidad. Se ensamblaron *de novo* las lecturas pareadas de cada especie con el programa Trinity (v.2.0.4). El porcentaje de lecturas pareadas alineadas correctamente fue mayor al 90% en *A. retusus*, *E. platyacanthus*, *F. pilosus* y *P. lychnidiflora*; mientras que en *C. imbricata* fue del 54.52% (Cuadro S4). Cabe destacar que la biblioteca de cDNA de *C. imbricata* se secuenció con lecturas de 75 bases lo cual influyó en el bajo porcentaje de lecturas ensambladas y en la reconstrucción de los transcritos. El cuadro 4 muestra el total de *contigs* y diversos estadísticos de los ensamblados por especie. *E. platyacanthus* fue la especie con más *contigs* ensamblados (300 798), seguida de *F. pilosus* (214 300), *A. retusus* (141 081) y *P. lychnidiflora* (119 980); mientras que de *C. imbricata* se ensambló el menor número de *contigs* (45 714) . El valor más alto de la longitud promedio de los *contigs* fue de 1 235 en *P. lychnidiflora* y el menor fue de 808 en *C. imbricata*. El *contig* N50 es una medida estadística de la longitud de los *contigs* considerando el 50% de las bases ensambladas. De los cinco ensamblados realizados, el N50 más alto fue de 2 023 bases en *P. lychnidiflora* mientras que el menor fue de 1 270 bases en *C. imbricata* (Cuadro 4).

Cuadro 3. Características de las bibliotecas de cDNA secuenciadas.

Especie	Tejido	Biblioteca de cDNA	Plataforma de Secuenciación	Total de lecturas	Total de lecturas procesadas
<i>Ariocarpus retusus</i>	Zona cambial (planta adulta)	TruSeq stranded mRNA	NextSeq 500	12 461 106	6 509 024
		Truseq mRNA	HiSeq 2000	22 274 178	19 934 675
<i>Cylindropuntia imbricata</i>	Zona cambial (planta adulta)	TruSeq stranded mRNA	NextSeq 500	9 910 468	5 221 561
<i>Echinocactus platyacanthus</i>	Plántas 1 mes y medio	TruSeq stranded mRNA	HiSeq 2000	42 976 733	37 176 687
	Zona cambial y médula de 8 meses y de 1 año 5 meses	TruSeq stranded mRNA	HiSeq 2000	46 697 662	40 276 213
	Zona cambial (planta adulta)	TruSeq stranded mRNA	HiSeq 2000	49 449 985	42 562 622
<i>Ferocactus pilosus</i>	Plantas 3 y 5 meses	TruSeq stranded mRNA	HiSeq 2000	44 636 673	38 480 860
	Zona cambial (planta adulta)	TruSeq stranded mRNA	NextSeq 500	12 125 407	7 110 986
		Truseq mRNA	HiSeq 2000	22 296 099	20 198 806
<i>Pereskia lychnidiflora</i>	Zona cambial (planta adulta)	TruSeq stranded mRNA	NextSeq 500	7 134 014	4 188 172
		Truseq mRNA	HiSeq 2000	22 394 026	20 110 662

Cuadro 4. Estadísticas generales de los transcriptomas ensamblados *de novo*.

<u>Especie</u>	<u>Bases ensambladas</u>	<u>Total de contigs</u>	<u>Porcentaje GC</u>	<u>Longitud promedio del Contig</u>	<u>Contig N50</u>	<u>Contig N10</u>	<u>Contig N90</u>	<u>Contig más grande</u>	<u>Contig Ex90N50</u>
<i>Ariocarpus retusus</i>	151 747 759	141 081	42.62	1 075	1 750	3 861	457	15 231	1 6945
<i>Cylindropuntia imbricata</i>	36 966 626	45 714	43.84	808	1 270	2 818	329	10350	1 294
<i>Echinocactus platyacanthus</i>	249 484 879	300 798	43.27	829	1 249	3 105	354	15 654	1 486
<i>Ferocactus pilosus</i>	204 522 982	214 300	42.62	954	1 537	3 563	396	13 805	1 683
<i>Pereskia lychnidiflora</i>	148 190 360	119 980	42.19	1 235	2 023	4 284	554	15 958	1 944

El número de *contigs* ensamblados, el N50 y la longitud promedio de los *contigs* es similar a la de otros transcriptomas ensamblados *de novo* de otras especies de plantas. El transcriptoma ensamblado *de novo* de *Spinacia oleraceae*, una especie perteneciente al mismo orden de los cactus tiene 72 151 *contigs*, una longitud promedio de los *contigs* de 644 bases y un N50 de 974 bases (Xu *et al.* 2015). En el transcriptoma de *Lophophora williamsii*, una especie de cactus de la subfamilia Cactoideae, reportan 63 704 *contigs* con longitudes promedio de 564.04 bases (Ibarra-Laclette *et al.* 2015). En el presente estudio, el transcriptoma de *C. imbricata* fue del que se obtuvo el menor número y longitud promedio de *contigs* y sin embargo la longitud promedio es mayor que en *Spinacia oleraceae* y *Lophophora williamsii*. En resumen, estos resultados nos indican que la longitud de los *contigs* obtenidos en las especies de cactus del presente estudio tienen valores ligeramente mayores a la de otros transcriptomas ensamblados *de novo*.

Los *contigs* ensamblados pueden representar transcritos fragmentados. Para revisar cuántos de los transcritos están completos o casi completos se realizaron alineamientos de blastx (BLAST+) contra la base de datos de proteínas del Uniprot de *A. thaliana* (<http://www.uniprot.org/uniprot/?query=taxonomy:3701>) reportando solamente los alineamientos más altos. La especie de la que se obtuvieron más transcritos con alineamientos en proteínas del 100% fue *E. platyacanthus* con 6443 (39.8%) y de la que se obtuvieron menos fue de *C. imbricata* con 3267 transcritos (29.9%) (Cuadro S4). En la Figura 6 se muestra el porcentaje de alineamiento entre los transcritos y las proteínas completas con respecto al total de alineamientos realizados. Como se puede ver, todos los transcriptomas tienen porcentajes de alineamiento muy similares. Cerca del 40% de los transcritos tienen alineamientos del 100% y menos del 30% de los transcritos tienen alineamientos con coberturas menores al 30%. La excepción es *C. imbricata* con solo el 25%

de proteínas alineadas al 100% y cerca del 35% de los alineamientos con coberturas menores al 30%. En resumen, estos resultados nos sugieren que cerca del 60% de los transcritos ensamblados están completos o casi completos con excepción de *C. imbricata*. Esto puede deberse al menor número y longitud de lecturas ensambladas en esta especie.

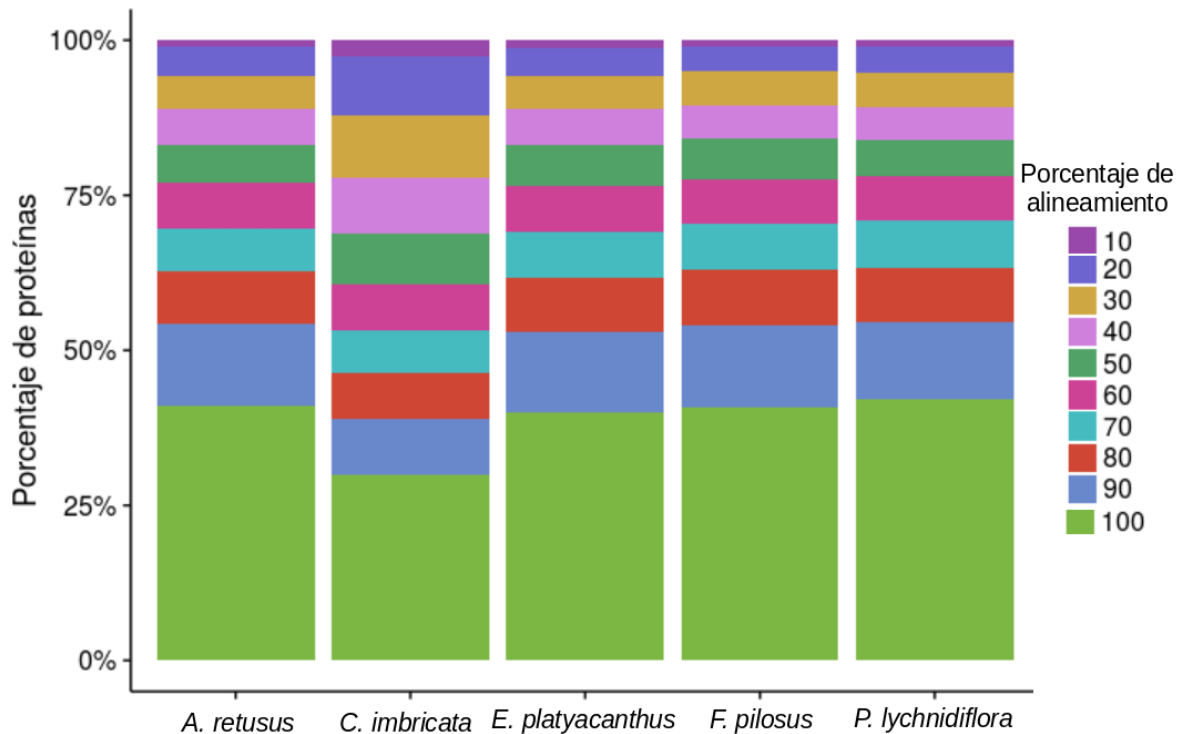


Figura 6. Porcentaje de transcritos ensamblados con distintos porcentajes de cobertura en los alineamientos contra las proteínas.

Un problema de los parámetros descritos anteriormente es que no toman en cuenta la variación en la abundancia de los transcritos y muchos de los que tienen bajos niveles de expresión tienden a ser cortos por la baja cobertura. Una alternativa que representa mejor la longitud de los contigs es el parámetro EXN50 cuyo principio se basa en los niveles de expresión del transcriptoma y el contig N50. Para obtener este valor, primero se estimaron las abundancias con RSEM y posteriormente se reportó en una tabla el N50 en función del porcentaje de datos de expresión (el número de transcritos que aportan el N% de la expresión total). Tomando el número de transcritos que aportan el 90% de los datos de

expresión y sacando el N50 el mayor contig fue de 1944 para *P. lychnidiflora* y el menor fue de 1294 en *C. imbricata*.

Además, graficar el N50 en función de los transcritos acomodados por su contribución en la expresión sirve como guía para saber si hace falta una mayor profundidad en la secuenciación y mejorar la calidad del ensamble. En la Figura 7A se puede ver cómo cambia el N50 en función de los transcritos acomodados por su contribución en la expresión del transcriptoma. Yendo de menos a más profundidad de secuenciación el N50 alcanza mayores valores cuando se considera el 90% de los datos de expresión. Cuando la profundidad en la secuenciación es mediana el N50 más alto se observa entre el 70 y 80% de los datos de expresión y cuando es muy baja el N50 tiende a ser muy homogéneo sin importar el porcentaje de datos de expresión. Como se puede ver en la Figura 7B el perfil del N50 en función de la expresión es muy similar en *A. retusus*, *E. platyacanthus*, *F. pilosus* y *P. lychnidiflora* en donde el N50 más alto se encontró cerca del 80% de los datos de expresión; mientras que en *C. imbricata* el N50 más alto se encontró cerca del 60% de los datos de expresión.

Otro problema de los ensamblados *de novo* es la imposibilidad de conocer *a priori* la composición de transcritos de una célula o tejido (Moreton, Izquierdo, y Emes 2016). El análisis BUSCO (Simão *et al.* 2015) permite medir de forma cuantitativa la completitud de un genoma o transcriptoma utilizando un set de genes ortólogos de copia simple. Se realizó dicho análisis con los transcriptomas ensamblados. Con base en el set de ortólogos de BUSCO, el transcriptoma más completo fue el de *E. platyacanthus* ya que sólo mostró el 7.9% de los genes del set de BUSCO como faltantes, mientras que *C. imbricata* fue el menos completo con 40% de genes faltantes y el 17% fragmentados. El set de genes para el resto de las especies se encuentra bastante bien representado con sólo un 8.5, 13 y 16% de

genes faltantes en *A. retusus*, *P. lychnidiflora* y *A. retusus* respectivamente (Figura 8). Un alto porcentaje de transcritos se encontró duplicado sugiriendo una alta similitud de secuencias entre contigs, una alta cantidad de variantes por *splicing* alternativo o la presencia de duplicaciones genómicas. Por lo anterior se utilizó el programa tr2aacds de evigene.

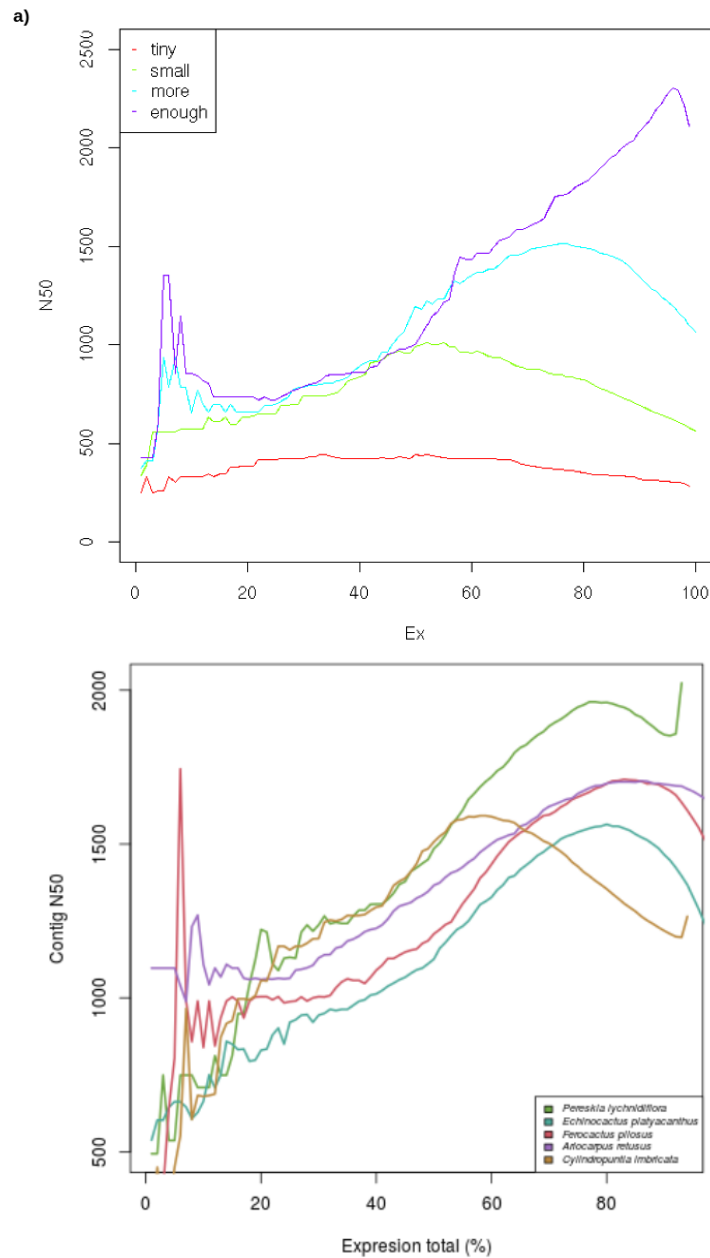
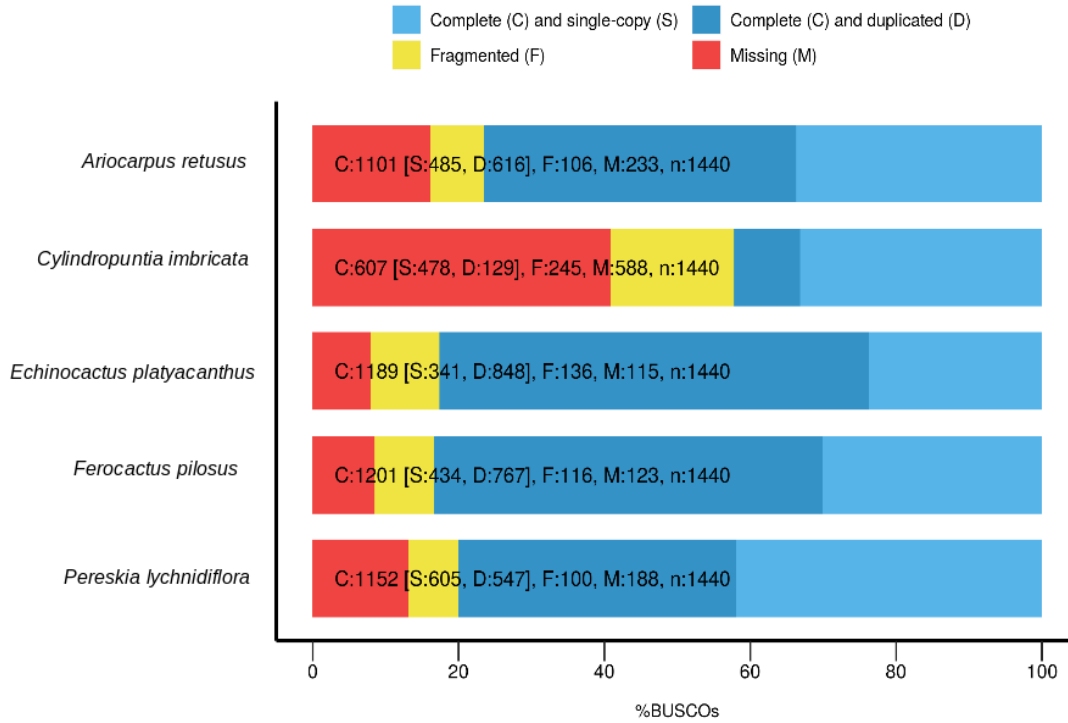


Figura 7. Contig N50 en función de los datos de expresión. El eje x representa el porcentaje de expresión del total de transcritos ensamblados; el eje y es el N50 de ese porcentaje de transcritos que contribuyen con el % de expresión. A) Perfiles del exN50 usando distintos números de lecturas (Tomada (trinityrnaseq: Trinity RNA-Seq de novo transcriptome assembly [2014] 2018). B) Perfil del exN50 para las especies ensambladas.

Evigene filtra transcritos parciales o fragmentados, transcritos de alta similitud de secuencia y transcritos con bajo potencial para codificar proteínas. En conjunto con *transdecoder* se definió un set de unigenes con un marco abierto de lectura (ORF por sus siglas en inglés). El mayor número de unigenes con ORF obtenidos fue 40 890 en *E. platyacanthus* y el menor número fue de 17 279 en *C. imbricata*. En *A. retusus*, *F. pilosus* y *P. lychnidiflora* el número de unigenes fue de 25 194, 32 226 y 24 430 respectivamente. Con este set de unigenes se volvió a llevar a cabo el análisis BUSCO obteniendo un set de unigenes con una composición muy similar en términos de completitud, pero esta vez el número de unigenes completos duplicados disminuyó considerablemente (Figura 8b). El porcentaje de genes faltantes fue de 14% en *A. retusus*, 42% *C. imbricata*, 9% *E. platyacanthus*, 9% *F. pilosus* y 13% *P. lychnidiflora*.

a)



b)

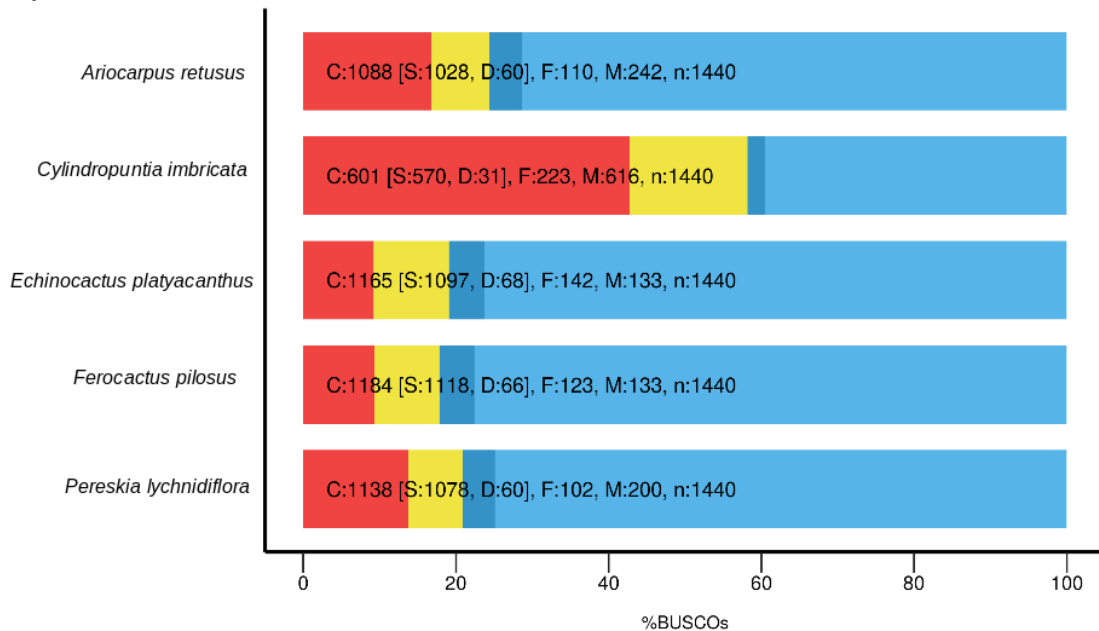


Figura8. Análisis de completitud BUSCO. El eje x representa el porcentaje de los distintos valores asignados por BUSCO en a) el set total de contigs obtenidos por Trinity, en b) con el set de contigs filtrados por Evigene. El set total de genes ortólogos de copia única es de n:1440.

Anotación de los transcritos

La anotación de los transcritos se realizó con el programa Trinotate (V.3.0). El número de transcritos homólogos (E-value < $1e^{-5}$) en proteínas de la base de datos de Uniprot fue mayor

al 60% en todas las especies (Cuadro 5). Para los transcritos en los que no se encontró ortología (tomando como criterio el E-value) en uniprot se volvieron a hacer búsquedas contra secuencias de proteínas de tres especies pertenecientes al orden Caryophyllales: *Portulaca oleracea* L., *Beta vulgaris* L. y *Spinacia oleracea* L. Las secuencias se descargaron de la base de datos UniProtKB/TrEMBL. Así se encontró similitud de entre un 15 y 18% más de los unigenes. De los que no se encontró similitud se realizaron búsquedas con blastp entre las mismas secuencias de cactus ensamblados en este trabajo. Al final, se encontraron 1 945, 789, 4 496, 2 926 y 2 217 unigenes específicos de *A. retusus*, *C. imbricata*, *E. platyacanthus*, *F. pilosus* y *P. lychnidiflora*, respectivamente. El total de transcritos con anotación funcional siguiendo el flujo de comandos de Trinotate fue de 16 291, 6 947, 22 371, 19 602 y 16 254 unigenes en *A. retusus*, *C. imbricata*, *E. platyacanthus*, *F. pilosus* y *P. lychnidiflora*, respectivamente.

Cuadro 5. Total de unigenes por especie y total de anotaciones de blastp contra distintas bases de datos.

<u>Especie</u>	<u>Total de unigenes</u>	<u>Transcritos Uniprot</u>	<u>Transcritos Caryophyllales</u>	<u>Transcritos cactus</u>	<u>Genes únicos</u>
<i>A. retusus</i>	25 194	17 468 (69%)	3 987(16%)	17 94(7%)	1 945
<i>C. imbricata</i>	17 279	12 842(74%)	3 146(18%)	502(3%)	789
<i>E. platyacanthus</i>	40 890	25 152(61%)	7 397(18%)	3 845(9%)	4 496
<i>F. pilosus</i>	32 226	20 909(65%)	5 511(17%)	2 880(9%)	2 926
<i>P. lychnidiflora</i>	24 430	17 335(70%)	3 820(15%)	1 058(4%)	2 217

Cuantificación de transcritos y comparación de transcriptomas

La correlación de las réplicas biológicas de cada especie es alta. En *P. lychnidiflora* y *C. imbricata* la correlación de Pearson entre las dos réplicas de zona cambial fue de 0.97 y 0.95 respectivamente. En *A. retusus* la correlación más baja fue de 0.85 (Figura S1). Por otra parte, en *F. pilosus* la correlación entre las réplicas de plantas de tres meses y de cinco meses fue de 0.95 mientras que en la zona cambial la correlación entre las dos muestras fue de 0.97. En *E. platyacanthus* la correlación más baja entre las réplicas de plantas de un mes 15 días fue de 0.97 y entre las réplicas de la zona cambial fue 0.98; mientras que en las plantas con WBT (cilindro vascular) la correlación de la muestra de 8 meses fue de 0.8 con respecto a las de un año cinco meses (Figura 9). Esta muestra es más parecida a las plantas de un mes 15 días por lo que en los análisis posteriores se eliminó.

Es interesante que si hay una fuerte correspondencia entre el transcriptoma y la anatomía de las plantas de tres y de cinco meses en *F. pilosus* (madera rica en WBT) mientras que *E. platyacanthus* las plantas de ocho meses (cilindro vascular) y las plantas de un mes con 15 días (planta completa) se parecen más entre sí a nivel transcriptómico que con las plantas de un año con 5 meses (cilindro vascular).

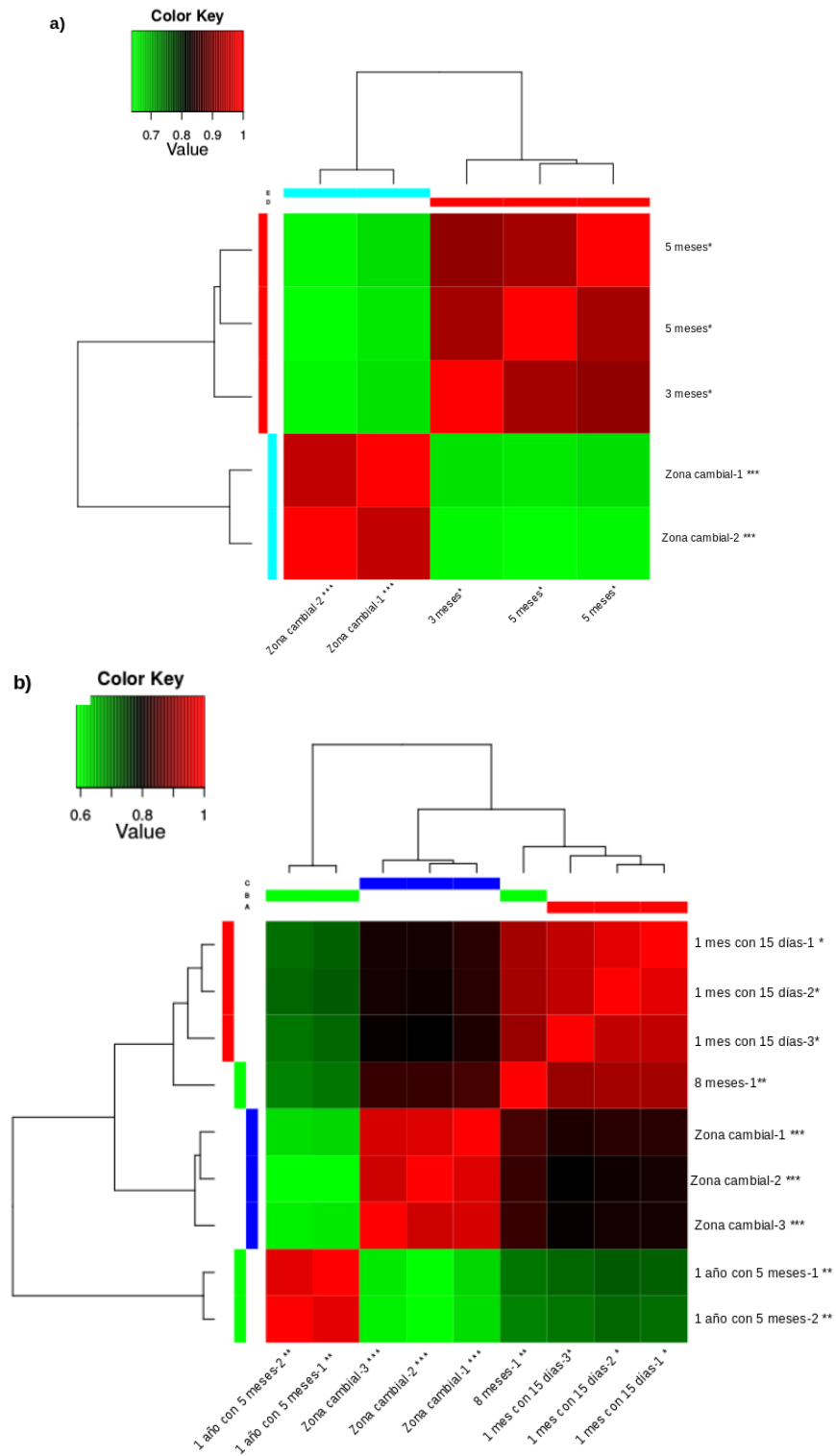


Figura 9. Mapa de calor con las correlaciones de Pearson entre réplicas biológicas de a) *F. pilosus* y de b) *E. platyacanthus*. * plantas completas, ** cilindro vascular, *** zona cambial.

Comparación del transcriptoma de zona cambial en cactus y dos especies maderables

La zona cambial de los cactus mostró muchas categorías funcionales de genes con el xilema secundario de dos especies maderables. Se extrajeron las categorías funcionales del Gene Ontology de los transcritos con al menos un TPM de la zona cambial de *A. retusus*, *C. imbricata*, *E. platyacanthus* (solo zona cambial con madera fibrosa), *F. pilosus* (solo zona cambial con madera fibrosa) y *P. lychnidiflora* y se compararon con las categorías funcionales de los transcriptomas del xilema secundario de *Eucalyptus* y *Populus*. De un total de 12 347 categorías se compartieron 1 288 (10.4%) entre todas, 2 149 (17.4%) se compartieron entre las especies de cactus y 862 de ellas fueron específicas a la familia Cactaceae. El número de categorías específicas por especie fue de 116, 1, 232, 67, 254, 191 y 44 para *A. retusus*, *C. imbricata*, *E. platyacanthus*, *Eucalyptus*, *F. pilosus*, *P. lychnidiflora* y *Populus* respectivamente. *C. imbricata* es la especie que tiene menos categorías específicas y esto se puede deber a su baja cobertura de transcritos. De hecho, el número de categorías compartidas para todas las especies, sin tomar en cuenta a *C. imbricata*, es de 3630 (29%) y de los cactus sin tomar en cuenta a la misma es de 3 468 (28%) (Figura 10). Dentro de las categorías que comparten las especies, sin tomar en cuenta a *C. imbricata*, cabe destacar que 68 de ellas tienen que ver con el mantenimiento y la diferenciación de células meristemáticas y 90 están relacionadas con el crecimiento secundario de las plantas (formación del xilema, componentes de la pared celular) (https://drive.google.com/open?id=1nhW94TykdeRTQ-P_Vn6mCKSV18CAkqtx). Esto nos indica que el xilema secundario de los cactus estudiados tiene una composición de transcritos similar a la de otros xilemas secundarios con crecimiento arbóreo. Las categorías funcionales no compartidas se pueden deber al hecho de que los transcriptomas de *Eucalyptus* y *Populus* no tienen la región del cambium vascular y del floema secundario. Haciendo una correlación de Pearson con el set

de categorías analizadas la zona cambial de *C. imbricata* es la menos parecido al resto de los cactus. *A. retusus* y *F. pilosus* son los que tienen más categorías compartidas (Figura S2).

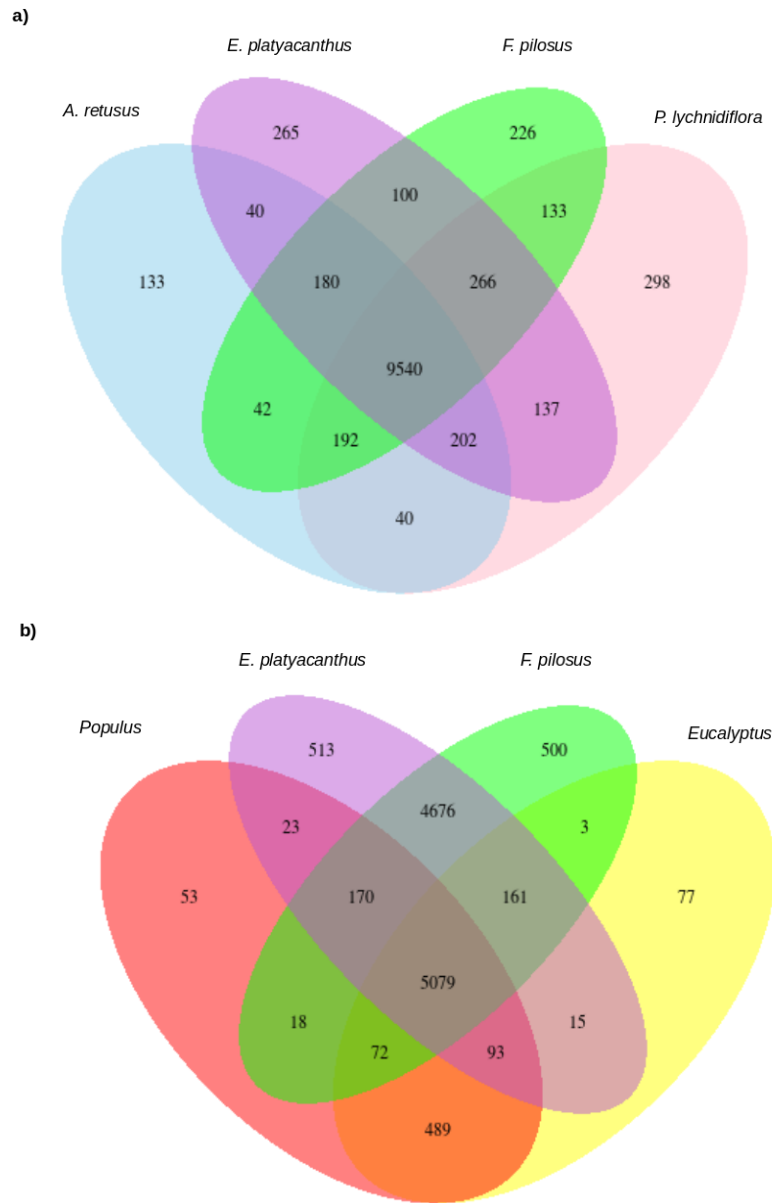


Figura 10. Diagramas de Venn con el total de categorías funcionales compartidas. a) Todas las posibles comparaciones de categorías compartidas entre *A. retusus*, *E. platyacanthus*, *F. pilosus* y *P. lychnidiflora*. b) Todas las posibles comparaciones de categorías compartidas entre *Populus*, *Eucalyptus*, *E. platyacanthus* y *F. pilosus*.

Una desventaja de solamente comparar el número de categorías funcionales del Gene Ontology es que no toma en cuenta los niveles de expresión de los transcritos. Para

comparar los transcriptomas tomando en cuenta los niveles de expresión se determinó la contribución de cada transcrito al total de expresión (se suma el total de CPM por réplica biológica y se divide entre los CPM de cada transcrito). Se extrajeron los transcritos que contribuyen con el 80% y con el 20% de la expresión total y se compararon las categorías funcionales. Para dicho análisis no se consideraron los datos de *Populus* y *Eucalyptus* pues no se cuenta con los valores de expresión y sí se consideraron los datos de plantas de 3-5 meses de *F. pilosus* y de plantas de un mes con 15 días, plantas de un año con cinco meses (cilindro vascular) y zona cambial de *E. platyacanthus*.

Cuando se consideró a los transcritos que contribuyen con el 80% de los datos de expresión y se extrajo su función, se formaron tres grupos de similitud (Figura 11). Un grupo se formó entre las plantas de 3-5 meses y la zona cambial de plantas adultas de *F. pilosus* junto con la zona cambial de *P. lychnidiflora*. Otro grupo se conformó por las plantas de un mes con 15 días y plantas de un año 5 meses (cilindro vascular) de *E. platyacanthus* junto con la zona cambial de plantas adultas de la misma especie y la de *A. retusus*; mientras que la zona cambial de *C. imbricata* se parece más al grupo de *E. platyacanthus* con *A. retusus*. Esto podría deberse a la baja representación de transcritos en *C. imbricata*. Cuando se consideró el 20% de los datos de expresión se formaron dos grupos. Uno formado por las plantas de un mes 15 días y las plantas de un año cinco meses (cilindro vascular) de *E. platyacanthus* junto con las plantas de 3-5 meses y la zona cambial de *F. pilosus*. El otro grupo, a su vez tiene dos subgrupos: la zona cambial de *A. retusus* y *E. platyacanthus* y el otro sub-grupo formado por la zona cambial de *C. imbricata* y *P. lychnidiflora*. Tomando en cuenta el 20% o el 80% de los datos de expresión, las plantas de un mes 15 días y las de un año cinco meses fueron más parecidas entre sí en *E. platyacanthus*. Además si se considera solo el 20% de los datos de expresión las plantas de un mes 15 días y de un año cinco

meses de *E. platyacanthus* junto con las plantas de 3-5 meses de *F. pilosus* formaron un grupo. Estos resultados, apoyan la idea de la retención de caracteres del xilema primario en el xilema secundario con WBT y la conservación de un perfil transcriptómico. Por otra parte, aunque no están todos los transcriptomas de especies con madera fibrosas agrupados, cuando se consideró el 20% de los datos de expresión, *C. imbricata* y *P. lychnidiflora* quedaron juntas. Estas dos especies tienen madera fibrosa, con radios lignificados. La zona cambial de *A. retusus* (con madera WBT) y de *E. platyacanthus* (con madera fibrosa) se agrupan considerando el 20% y el 80% por lo que tienen un transcriptoma muy similar a pesar de tener maderas distintas en la madurez.

La comparación de niveles de expresión por experimentos de RNA-seq entre especies tiene muchas dificultades (Dunn, Howison, y Zapata 2013). Para empezar cada tejido colectado está formado por una composición diferente de células (a pesar de ser fibrosas o no fibrosas los tipos celulares no se encuentran en las mismas proporciones, por ejemplo, la presencia de parénquima paratraqueal es escaso en la madera fibrosa de *F. pilosus* y nula en la madera fibrosa de *E. platyacanthus*). Por otra parte, la anotación de los transcritos puede sesgar las comparaciones entre las especies (Roux, Rosikiewicz, y Robinson-Rechavi 2015). Otra forma de analizar qué cambios en la expresión de los transcritos son importantes en las distintas anatomías observadas es haciendo comparaciones entre tejidos de la misma especie para identificar transcritos diferencialmente expresados y posteriormente poder analizar estos transcritos en las otras especies.

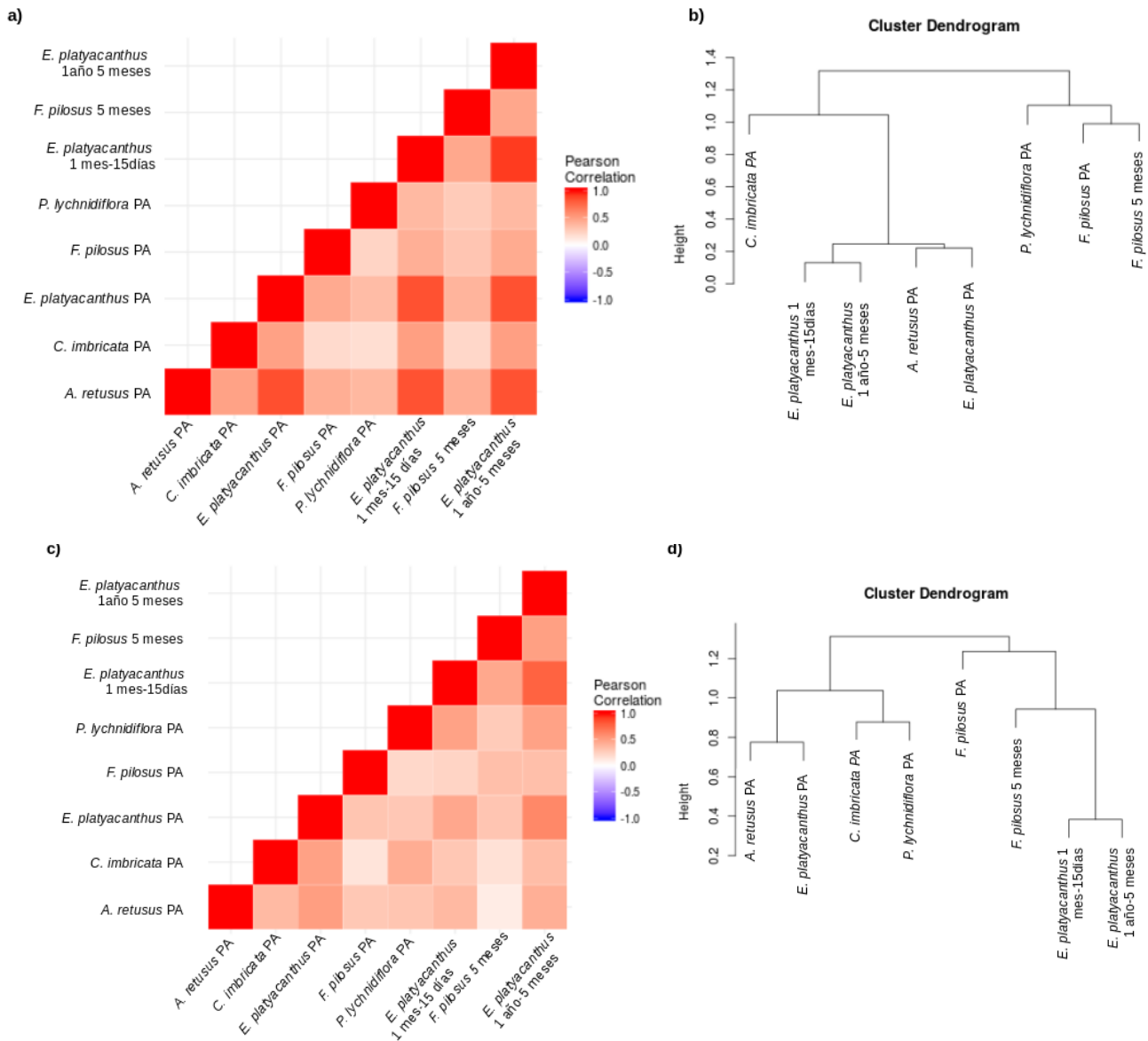


Figura 11. Correlaciones de Pearson considerando las categorías funcionales de los transcritos que contribuyen con el 80% (a y b) y el 20% (c y d) de los datos de expresión. a),c) Mapa de calor de las correlaciones de Pearson y b),d) dendrograma de las mismas correlaciones.

Análisis de expresión diferencial en *E. platyacanthus* y *F. pilosus*

Para el análisis de expresión diferencial con *E. platyacanthus* y *F. pilosus* se filtraron los transcritos con menos de dos CPM (ver Método). El análisis se realizó con 16 577 y 15 479 transcritos de *E. platyacanthus* y *F. pilosus* respectivamente. De *F. pilosus* se identificaron

3174 transcritos con expresión diferencial, 1871 se expresan mayormente en plantas de 3-5 meses y 1303 mayormente en zona cambial (Figura 12). En *E. platyacanthus* se identificaron 551 transcritos mayormente expresados en plantas de un mes 15 días, 446 en plantas de un año 5 meses (madera con WBT) y 891 en zona cambial (madera fibrosa). Además, se identificaron 1 888 transcritos mayormente expresados en plantas de un mes 15 días y un año 5 meses, 553 en plantas de un mes 15 días y zona cambial (madera fibrosa), y solo 261 en zona cambial (madera fibrosa) y plantas de un año 5 meses (madera con WBT). Esto destaca el número de transcritos con expresión diferencial comunes en el estadio de plántula y en ZC con WBT.

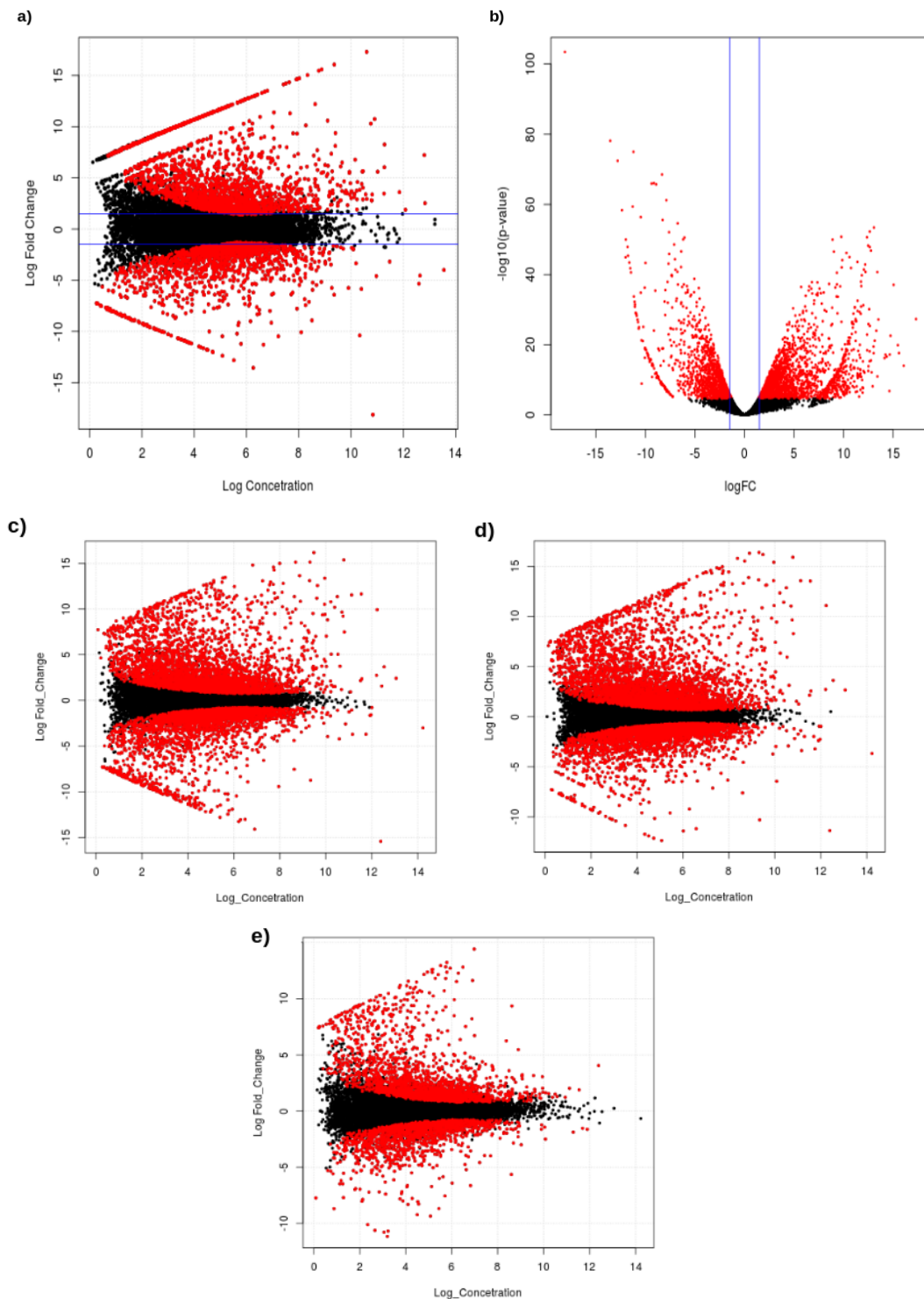


Figura 12. Transcritos con expresión diferencial en *F. pilosus* y en *E. platyacanthus*. a) Tasa de cambio logarítmica en función de los concentración por transcritos entre plantas de 3-5 meses y zona cambial y b) volcano-plot del p -value en función de la tasa de cambio logarítmica de *F. pilosus*; c) Tasa de cambio logarítmica en función de los concentración de transcritos entre plantas de un mes con 15 días y zona cambial, d) plantas de un año con cinco meses contra zona cambial y e) entre plantas de un mes con 15 días y un año con cinco meses de *E. platyacanthus*.

Análisis de categorías funcionales del GO enriquecidas*

Se utilizó goseq (Young *et al.*, 2010) para determinar las categorías funcionales enriquecidas en los distintos estadios. De los 15 479 transcritos analizados en *F. pilosus* solo 284 (1.8%) no tienen anotación funcional. En las plantas de 3-5 meses hubo 169 categorías enriquecidas mientras que en la zona cambial (madera fibrosa) hubo 104 (Cuadro S6). De los 16 577 transcritos analizados en *E. platyacanthus* solo 477 (2.9%) no tienen anotación funcional. En las plantas de un mes 15 días hubo 61 categorías enriquecidas; en las plantas de un año 5 meses (cilindro vascular con WBT) hubo 11 categorías enriquecidas mientras que en la zona cambial (madera fibrosa) hubo 113. La Figura 13 a y b muestra la frecuencia de categorías enriquecidas divididas en sus tres principales aspectos: proceso biológico (BP), componente celular (CC) y función molecular (MF). Las zonas cambiales de las dos especies son las que compartieron más categorías y la mayoría pertenecen a BP, seguidas de MF y después las de CC. Dentro de las categorías compartidas destacan las relacionadas con el metabolismo de ácidos nucleicos como transcripción génica, factores transcripcionales con unión a DNA, integración y síntesis de DNA mediada por RNA y la nucleocápside de retrotransposon; estas últimas sugieren una alta actividad transposónica por lo que será interesante revisar a qué clase de retrotransposones pertenecen las secuencias enriquecidas en bases de datos como *Gypsi* (Llorens *et al.* 2010). Las plantas de 3-5 meses en *F. pilosus* y de un mes 15 días en *E. platyacanthus* tuvieron una mayor proporción de categorías relacionadas con CC en comparación con los otros estadios y compartieron 35 categorías que se relacionan con componentes del cloroplasto, fotosíntesis u organización de la pared celular y metabolismo de ceras. Las plantas de un año con cinco meses de *E. platyacanthus* compartió un mayor número de categorías con las plantas de un mes 15 días y con las de 3-5 meses en *F. pilosus* que en comparación a las zonas cambiales. Destacan las funciones de

oxidoreductasa, unión a tetrapirrol, actividad de mono-oxigenasa, región extracelular y unión a iones de hierro. La categoría de síntesis de fenilpropanoides está enriquecida en las plantas de un año cinco meses de *E. platycanthus*. Dentro de los transcritos pertenecientes a la categoría de fenilpropanoides con una mayor expresión en el estadio de un año cinco meses se encuentra el posible ortólogo de la lacasa *LAC3* de *A. thaliana*. Mutantes del ortólogo de *lac3* en *Populus* afectan la integridad de las fibras del xilema y la composición de la pared secundaria (Ranocha *et al.* 2002). La acción de esta familia de enzimas es tejido específico, por ejemplo *LAC4* se expresa en fibras interfasciculares, *LAC7* en pelos radiculares mientras que *LAC8* se expresa en el floema (Turlapati *et al.* 2011). Será importante establecer las relaciones de ortología de los transcritos pertenecientes a las lacasas presentes en los transcriptomas del presente estudio y relacionar su expresión con los tejidos muestreados. Posiblemente la expresión de las lacasas que se expresan en las maderas ricas en WBT son distintas de las maderas fibrosas.

Otro de los transcritos de la vía de fenilpropanoides que tiene mayor expresión en las plantas de un año cinco meses (madera WBT) es el posible ortólogo de la peroxidasa clase III *PER72* de *A. thaliana*. Las peroxidasas Clase III son importantes en la formación de la lignina. Mutantes del gen *PER72* tienen paredes secundarias más delgadas y parece que su papel es más relevante durante la lignificación que durante la elongación celular (Fernández-Pérez *et al.* 2015). Las células derivadas (WBT) del xilema secundario de las plantas de un año cinco meses no tienen deposición de lignina a lo largo de la pared de estas células. Sin embargo, las WBT no se elongan como las fibras y sí depositan pared secundaria en engrosamientos helicoidales o anulares. Será interesante determinar si la acción de esta enzima es preponderante en las WBT sobre los vasos y fibras.

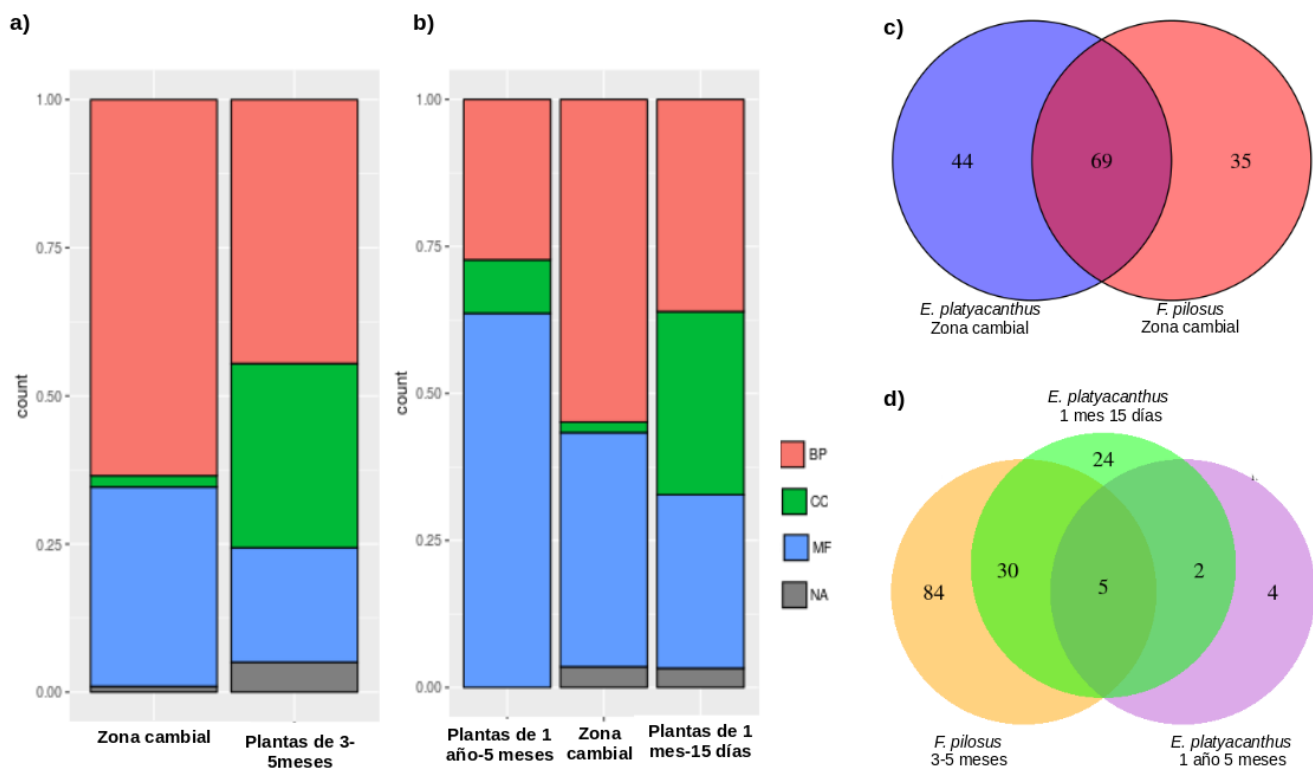


Figura 13. Categorías del GO enriquecidas y compartidas por especie. a) En *F. pilosus* se reportan las categorías enriquecidas en la zona cambial y en plantas de 3-5 meses. b) En *E. platyacanthus* se reportan las categorías enriquecidas en plantas de un año con cinco meses, zona cambial, y en plantas de un mes con 15 días. Diagrama de Venn con el total de categorías compartidas entre las c) zonas cambiales de ambas especies y entre d) las plantas de 3-5 meses de *F. pilosus*, y de un mes 15 días y un año cinco meses de *E. platyacanthus*. BP = Proceso biológico, CC = Componente celular, MF = Función molecular, NA = Función desconocida.

Cabe destacar la presencia de las funciones de vías de señalización para la activación de etileno en las zonas cambiales de ambas especies. Se ha visto que ésta vía de señalización induce la proliferación celular en el cambium vascular (Etchells, Provost, y Turner 2012). También la vía de activación de citoquininas está enriquecida en *E. platyacanthus* (vía de activación del citoquininas). La mayor concentración de citoquininas se encuentra en el xilema secundario (Fischer y Teichmann 2017).

Las categorías de factores transcripcionales están más representadas en la zona cambial de las dos especies (122 transcritos en *E. platyacanthus* y 97 transcritos en *F. pilosus* de los

cuales 86 son posibles ortólogos). Destacan los factores de respuesta a etileno (familia ERF/AP2), proteínas de las familias bHLH, bZIP, GRAS, MADS-box, MYB y proteínas con dedos de ZINC (Cuadros S7 y S8). Dentro de los genes MADS-box que están mayormente expresados en la zona cambial de ambas especies se encuentran los posibles ortólogos de *AGL24*, *SOC1* y *AGL82*. En *A. thaliana*, *AGL24* y *SOC1* interactúan y en conjunto participan en una de las vías de floración de la planta. Su expresión se localiza en la parte apical del tallo vegetativo (Liu *et al.* 2008) Corroborar la expresión de estos factores transcripcionales con otros experimentos será determinante para saber si el dominio de expresión de estos MADS es más amplio en la familia Cactaceae. Con respecto a los factores transcripcionales de la familia GRAS, dos son posibles ortólogos de las proteínas inducibles por quitina y que responden a la giberelina (*chitin-inducible gibberellin-responsive protein en inglés (CIGR)*). Es posible que estos factores transcripcionales participen en la talla de los individuos ya que cuando se aumenta la expresión de estos genes se aumenta el tamaño de las plantas del arroz (Kovi *et al.* 2011). Por ello, estos genes son buenos candidatos para estudiar el aumento de las tallas en las especies de la familia Cactaceae que alcanzan tallas grandes y contrastar con plantas que se mantienen pequeñas.

En *E. platyacanthus* dos genes pertenecientes a la familia KNOX tienen expresión diferencial. Para identificar a qué clase de *KNOX* pertenecen se realizaron búsquedas de blast en los transcriptomas de los cactus con las secuencias de proteína de *KNAT1*, *KNAT2*, *KNAT3*, *KNAT4*, *KNAT5*, *KNAT6*, *KNAT7*, *STM* de *A. thaliana* y posteriormente se construyó un árbol filogenético para establecer las relaciones de homología.

Se encontró una duplicación del factor transcripcional clase I *STM* y del clase II *KNAT7* en los cactus. De *F. pilosus* no se recuperó la secuencia del clado *KNAT3-KNAT4-KNAT5*. Será importante determinar si se debe a una pérdida genética o solamente no se expresa en los

estadios de 3-5 meses o en la zona cambial de plantas adultas. Los transcritos de la familia *KNOX*, pertenecientes tanto a la clase I como a la clase II, se expresan en todas los estadios muestreados de los cactus pero son distintos dependiendo de la anatomía. Los transcritos diferencialmente expresados en *E. platyacanthus* son el posible ortólogo de *KNAT2-KNAT6* (mayor expresión en zona cambial) y uno de los posibles ortólogos de *KNAT7* con una mayor expresión en plantas de un mes 15 días y un año 5 meses (madera rica en WBT) (Figura 14 y 15). En la Figura 15 se puede ver un mapa de calor con la expresión logarítmica de los factores transcripcionales *KNOX* que se obtuvieron en el árbol filogenético. Aunque no hay un patrón claro, la expresión del ortólogo de *KNAT1* y uno de los posibles ortólogos de *KNAT7* es muy alta en todos los estadios y las especies. También lo es el posible ortólogo de *KNAT3-4-5*; sin embargo, en *F. pilosus* no se encontró ortólogo. Uno de los posibles ortólogos de *STM1* (*STM.1*) tiene una mayor expresión en la zona cambial de *E. platyacanthus* y *F. pilosus* que con respecto a los otros estadios de estas especies. Los dos posibles ortólogos de *KNAT2-6* tienen una menor expresión en las plantas de un mes 15 días, un año cinco meses de *E. platyacanthus* y de 3-5 meses de *F. pilosus* que con respecto a los otros estadios y especies. Por esta razón un posible candidato que participe en la formación de fibras son los *KNAT2-6*. Por otra parte, la co-expresión de genes *KNOX* de las clases I y II en la zona cambial no se contrapone con la idea de una acción antagonista entre las dos clases (Ohtani *et al.* 2011). Al tener las células de cambium vascular y las primeras células diferenciadas del xilema y floema no podemos saber a detalle el lugar donde su función es preponderante. Se necesitará hacer disecciones de tejido más finas para determinar esas expresiones.

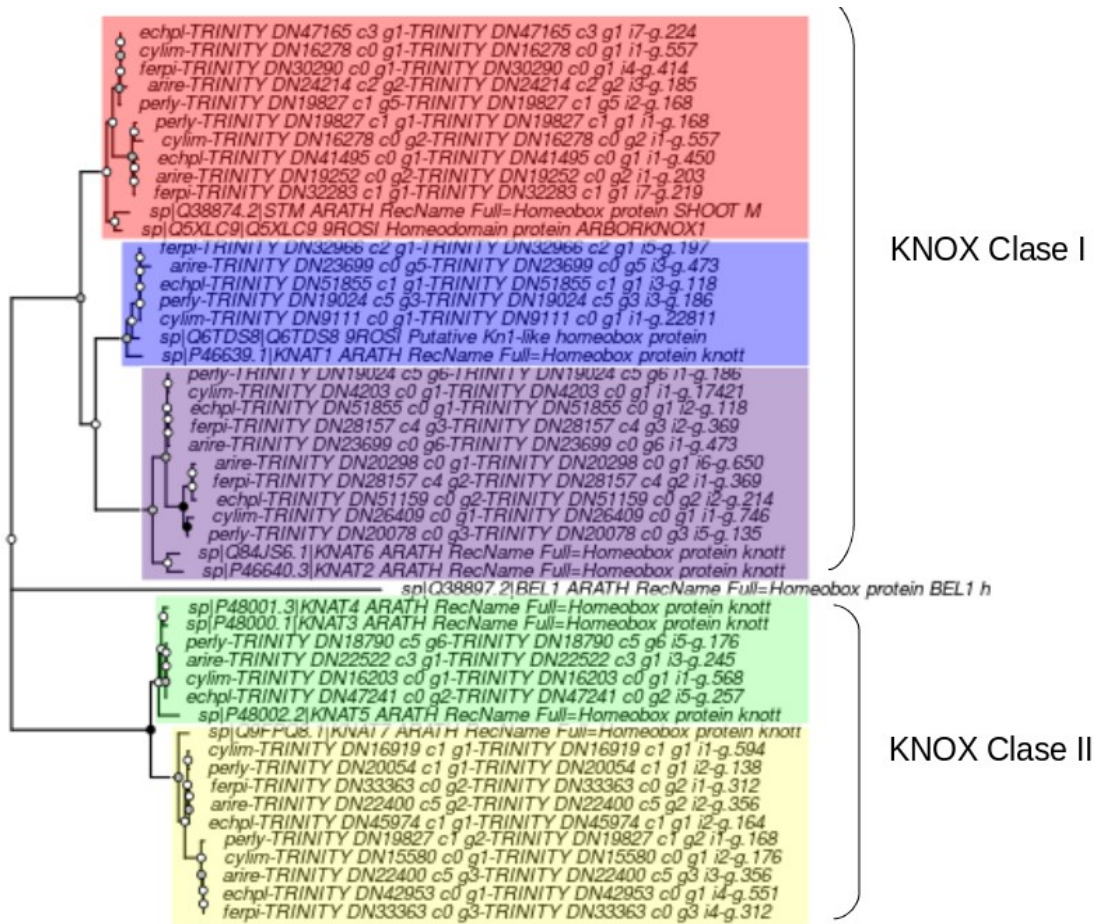


Figura 14. Árbol filogenético de las secuencias de proteínas KNOX seleccionadas de *A. thaliana*, *Populus tremula* y las especies de cactus estudiadas. Las proteínas de la clase KNOX I pertenecen a los clados KNAT1 (azul), KNAT2-KNAT6 (morado) y STM (rojo). Las proteínas KNOX de la clase II pertenecen los clados KNAT3-KNAT4, KNAT5 (amarillo) y KNAT7 (verde). Los círculos en la base de las ramas representan los valores de bootstrap (blanco menor a 75, gris mayor a 75 y menor a 95 y negro mayor a 95).

cambiales, abriendo la posibilidad de estudiar los FT importantes en el desarrollo secundario mediado por el CV. Dentro de los FT se encontraron genes pertenecientes a las familias, *bHLH*, *bZIP*, *GRAS*, *MADS-box*, *MYB*, *ERF* y *KNOX*. De los *KNOX* se encontró una duplicación del posible ortólogo de *STM* y del posible ortólogo de *KNAT7* para los cactus estudiados. En *E. platyacanthus* un posible ortólogo de *KNAT7* se expresa mayormente en plantas de un mes 15 días y de un año cinco meses (madera con WBT). Además, en el estadio de desarrollo de WBT se encontraron enriquecidas las funciones de síntesis de fenilpropanoides y de ligninas abriendo la posibilidad de estudiar la vía de síntesis de lignina en maderas con WBT. El posible ortólogo de la peroxidasa clase III *PER72* y de la lacasa *LAC3* tienen una mayor expresión en las maderas con WBT. Será importante determinar por otros experimentos, como PCR's tiempo real, la expresión de estos transcritos. Con este trabajo se identificaron genes candidatos para continuar estudiando la diferenciación vascular en la familia Cactaceae.

Perspectivas

Ahora que se cuenta con secuencias de transcriptomas de la zona cambial de cinco especies de la familia Cactaceae será importante hacer un mapa de expresión más detallado para alguna de las especies. Después de observar que la germinación y el crecimiento durante los primeros meses es fácil de seguir en *E. platyacanthus* y además de que tiene una transición de madera rica en WBT a madera fibrosa directa será conveniente centrar un análisis transcriptómico detallado con esta especie. Ampliando el número de estadios y tejidos por coleccionar. Por ejemplo en el estudio de (Sundell *et al.* 2017) disectan 25 secciones de la zona cambial para identificar grupos de transcritos con expresión diferencial a lo largo de esta región meristemática. Esto permitirá conocer más a detalle los cambios en la expresión de la zona cambial en especies que pueden producir dos tipos de madera durante su fase vegetativa. Posteriormente, se puede sacar ventaja de los transcritos ya ensamblados para las otras especies de cactus y ver si los patrones de expresión están conservados por medio de experimentos de PCR-tiempo real o hibridaciones *in situ*. Otro factor a considerar es escoger una especie para generar información a nivel genómico y a nivel proteómico.

Otra perspectiva importante del estudio, será curar las bases de datos de los transcritos para lograr una base de datos común de genes ortólogos. Esta base de datos permitirá hacer análisis evolutivos más profundos. Por ejemplo se podrán poner a prueba modelos de evolución de los transcriptomas y su relación con la anatomía (Harrison, Wright, y Mank 2012).

Referencias

- Aloni, Roni. 2015. "Ecophysiological implications of vascular differentiation and plant evolution". *Trees* 29 (1):1–16.
- Altesor, A., C. Silva, y E. Ezcurra. 1994. "Allometric neoteny and the evolution of succulence in cacti". *Botanical Journal of the Linnean Society* 114 (3):283–92.
- Ashburner, Michael, Catherine A. Ball, Judith A. Blake, David Botstein, Heather Butler, J. Michael Cherry, Allan P. Davis, Kara Dolinski, Selina S. Dwight, y Janan T. Eppig. 2000. "Gene Ontology: tool for the unification of biology". *Nature Genetics* 25 (1):25–29.
- Bhalerao, Rishikesh P., y Urs Fischer. 2006. "Environmental and hormonal control of cambial stem cell dynamics". *Journal of Experimental Botany* 68 (1):79–87.
- Boeckmann, Brigitte, Marie-Claude Blatter, Livia Famiglietti, Ursula Hinz, Lydie Lane, Bernd Roechert, y Amos Bairoch. 2005. "Protein variety and functional diversity: Swiss-Prot annotation in its biological context". *Comptes Rendus Biologies* 328 (10):882–99. <https://doi.org/10.1016/j.crv.2005.06.001>.
- Camacho, Christiam, George Coulouris, Vahram Avagyan, Ning Ma, Jason Papadopoulos, Kevin Bealer, y Thomas L. Madden. 2009. "BLAST+: architecture and applications". *BMC bioinformatics* 10 (1):421.
- Carlquist, Sherwin. 1980. "Further concepts in ecological wood anatomy, with comments on recent work in wood anatomy and evolution". *Aliso: A Journal of Systematic and Evolutionary Botany* 9 (4):499–553.
- . 2009. "Xylem heterochrony: an unappreciated key to angiosperm origin and diversifications". *Botanical Journal of the Linnean Society* 161 (1):26–65.
- Dunn, Casey, Mark Howison, y Felipe Zapata. s/f. "Agalma: an automated phylogenomics workflow | BMC Bioinformatics | Full Text". Consultado el 29 de enero de 2018. <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-14-330>.
- Etchells, J. Peter, Claire M. Provost, y Simon R. Turner. 2012. "Plant vascular cell division is maintained by an interaction between PXY and ethylene signalling". *PLoS Genetics* 8 (11):e1002997.
- Etchells, J. Peter, y Simon R. Turner. 2017. "Realizing pipe dreams – a detailed picture of vascular development". *Journal of Experimental Botany* 68 (1):1–4. <https://doi.org/10.1093/jxb/erw482>.
- Evert, Ray F. 2006. *Esau's Plant Anatomy: Meristems, Cells, and Tissues of the Plant Body: Their Structure, Function, and Development*. John Wiley & Sons.
- Fernández-Pérez, Francisco, Federico Pomar, María A. Pedreño, y Esther Novo-Uzal. 2015. "Suppression of Arabidopsis peroxidase 72 alters cell wall and phenylpropanoid metabolism". *Plant Science* 239 (octubre):192–99. <https://doi.org/10.1016/j.plantsci.2015.08.001>.
- Finn, Robert D., Teresa K. Attwood, Patricia C. Babbitt, Alex Bateman, Peer Bork, Alan J. Bridge, Hsin-Yu Chang, Zsuzsanna Dosztányi, Sara El-Gebali, y Matthew Fraser.

2016. "InterPro in 2017—beyond protein family and domain annotations". *Nucleic Acids Research* 45 (D1):D190–99.
- Fischer, Urs, y Thomas Teichmann. 2017. "The ERECTA and ERECTA-like genes control a developmental shift during xylem formation in Arabidopsis". *New Phytologist* 213 (4):1562–63.
- Gibson, Arthur C. 1973. "Comparative anatomy of secondary xylem in Cactoideae (Cactaceae)". *Biotropica*, 29–65.
- . 1977. "Wood anatomy of opuntias with cylindrical to globular stems". *Botanical Gazette* 138 (3):334–51.
- Gilbert, Don, y Don Gilbert. 2016. "Gene-omes built from mRNA seq not genome DNA". *F1000Research* 5 (julio). <https://doi.org/10.7490/f1000research.1112594.1>.
- Guindon, Stéphane, Jean-François Dufayard, Vincent Lefort, Maria Anisimova, Wim Hordijk, y Olivier Gascuel. 2010. "New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0". *Systematic Biology* 59 (3):307–21.
- Haas, Brian J., Alexie Papanicolaou, Moran Yassour, Manfred Grabherr, Philip D. Blood, Joshua Bowden, Matthew Brian Couger, David Eccles, Bo Li, y Matthias Lieber. 2013. "De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis". *Nature Protocols* 8 (8):1494–1512.
- Harrison, Peter W., Alison E. Wright, y Judith E. Mank. 2012. "The evolution of gene expression and the transcriptome–phenotype relationship". *Seminars in Cell & Developmental Biology*, Compartmentalised Cell Signalling & Next Generation Sequencing in Health and Disease, 23 (2):222–29. <https://doi.org/10.1016/j.semcdb.2011.12.004>.
- Hernández-Hernández, Tania, Héctor M. Hernández, J. Arturo De-Nova, Raul Puente, Luis E. Eguiarte, y Susana Magallón. 2011. "Phylogenetic relationships and evolution of growth form in Cactaceae (Caryophyllales, Eudicotyledoneae)". *American Journal of Botany* 98 (1):44–61.
- Ibarra-Laclette, Enrique, Flor Zamudio-Hernández, Claudia Anahí Pérez-Torres, Victor A. Albert, Enrique Ramírez-Chávez, Jorge Molina-Torres, Araceli Fernández-Cortes, *et al.* 2015. "De novo sequencing and analysis of Lophophora williamsii transcriptome, and searching for putative genes involved in mescaline biosynthesis". *BMC Genomics* 16 (septiembre):657. <https://doi.org/10.1186/s12864-015-1821-9>.
- Johnson, L. Steven, Sean R. Eddy, y Elon Portugaly. 2010. "Hidden Markov model speed heuristic and iterative HMM search procedure". *BMC Bioinformatics* 11 (agosto):431. <https://doi.org/10.1186/1471-2105-11-431>.
- Katoh, Kazutaka, Kazuharu Misawa, Kei-ichi Kuma, y Takashi Miyata. 2002. "MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform". *Nucleic Acids Research* 30 (14):3059–66. <https://doi.org/10.1093/nar/gkf436>.
- Kovi, Mallikarjuna Rao, Yushan Zhang, Sibin Yu, Gaiyu Yang, Wenhao Yan, y Yongzhong Xing. 2011. "Candidacy of a Chitin-Inducible Gibberellin-Responsive Gene for a Major

- Locus Affecting Plant Height in Rice That Is Closely Linked to Green Revolution Gene *sd1*. *Theoretical and Applied Genetics* 123 (5):705. <https://doi.org/10.1007/s00122-011-1620-x>.
- Krogh, Anders, BjoÈrn Larsson, Gunnar Von Heijne, y Erik LL Sonnhammer. 2001. "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes". *Journal of Molecular Biology* 305 (3):567–80.
- Langmead, Ben, y Steven L. Salzberg. 2012. "Fast gapped-read alignment with Bowtie 2". *Nature methods* 9 (4):357–59.
- Lens, Frederic, James L. Luteyn, Erik Smets, y Steven Jansen. 2004. "Ecological trends in the wood anatomy of Vaccinioideae (Ericaceae sl)". *Flora-Morphology, Distribution, Functional Ecology of Plants* 199 (4):309–19.
- Lens, Frederic, Rutger A. Vos, Guillaume Charrier, Timo van der Niet, Vincent Merckx, Pieter Baas, Jesus Aguirre Gutierrez, Bart Jacobs, Larissa Chacon Dória, y Erik Smets. 2016. "Scalariform-to-simple transition in vessel perforation plates triggered by differences in climate during the evolution of Adoxaceae". *Annals of Botany* 118 (5):1043–56.
- Liu, Chang, Hongyan Chen, Hong Ling Er, Hui Meng Soo, Prakash P. Kumar, Jin-Hua Han, Yih Chheng Liou, y Hao Yu. 2008. "Direct Interaction of AGL24 and SOC1 Integrates Flowering Signals in Arabidopsis". *Development* 135 (8):1481–91. <https://doi.org/10.1242/dev.020255>.
- Liu, Lijun, Vladimir Filkov, y Andrew Groover. 2014. "Modeling transcriptional networks regulating secondary growth and wood formation in forest trees". *Physiologia Plantarum* 151 (2):156–63.
- Llorens, Carlos, Ricardo Futami, Laura Covelli, Laura Domínguez-Escribá, Jose M. Viu, Daniel Tamarit, Jose Aguilar-Rodríguez, Miguel Vicente-Ripolles, Gonzalo Fuster, y Guillermo P. Bernet. 2010. "The Gypsy Database (GyDB) of mobile genetic elements: release 2.0". *Nucleic Acids Research* 39 (suppl_1):D70–74.
- Love, Michael I., Wolfgang Huber, y Simon Anders. 2014. "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2". *Genome biology* 15 (12):550.
- Loza Cornejo, S., y T. Terrazas. 1996. "Anatomía del tallo y de la raíz de dos especies de Wilcoxia Britton et Rose (Cactaceae) del noreste de México". *Stem and root anatomy of two species of Wilcoxia Britton et Rose (Cactaceae) of northeast Mexico.* *Bol. Soc. Bot. Mex* 59:13–23.
- Loza-Cornejo, Sofía, y Teresa Terrazas. 2011. "Morfo-anatomía de plántulas en especies de Pachycereeae: ¿ hasta cuándo son plántulas?" *Boletín de la Sociedad Botánica de México*, núm. 88:1–13.
- Lucas, William J., Andrew Groover, Raffael Lichtenberger, Kaori Furuta, Shri-Ram Yadav, Ykä Helariutta, Xin-Qiang He, Hiroo Fukuda, Julie Kang, y Siobhan M. Brady. 2013. "The plant vascular system: evolution, development and functions". *Journal of Integrative Plant Biology* 55 (4):294–388.

- Mauseth, James D. 2004. "Wide-Band Tracheids Are Present in Almost All Species of Cactaceae". *Journal of Plant Research* 117 (1):69–76. <https://doi.org/10.1007/s10265-003-0131-5>.
- . 2006a. "Structure–function relationships in highly modified shoots of Cactaceae". *Annals of Botany* 98 (5):901–26.
- . 2006b. "Structure–function relationships in highly modified shoots of Cactaceae". *Annals of Botany* 98 (5):901–26.
- Mauseth, James D., y Brandon J. Plemons-Rodriguez. 1998. "Evolution of extreme xeromorphic characters in wood: a study of nine evolutionary lines in Cactaceae". *American Journal of Botany*, 209–18.
- Miyashima, Shunsuke, Jose Sebastian, Ji-Young Lee, y Yka Helariutta. 2013. "Stem cell function during plant vascular development". *The EMBO Journal* 32 (2):178–93.
- Moczek, Armin P., Karen E. Sears, Angelika Stollewerk, Patricia J. Wittkopp, Pamela Diggle, Ian Dworkin, Cristina Ledon-Rettig, David Q. Matus, Siegfried Roth, y Ehab Abouheif. 2015. "The significance and scope of evolutionary developmental biology: a vision for the 21st century". *Evolution & Development* 17 (3):198–219.
- Moreton, Joanna, Abril Izquierdo, y Richard D. Emes. 2016. "Assembly, assessment, and availability of de novo generated eukaryotic transcriptomes". *Frontiers in Genetics* 6:361.
- Necsulea, Anamaria, y Henrik Kaessmann. 2014. "Evolutionary dynamics of coding and non-coding transcriptomes". *Nature Reviews Genetics* 15 (11):734–48.
- Ohtani, Misato, Nobuhiro Akiyoshi, Yuto Takenaka, Ryosuke Sano, y Taku Demura. 2016. "Evolution of plant conducting cells: perspectives from key regulators of vascular cell differentiation". *Journal of Experimental Botany* 68 (1):17–26.
- Ohtani, Misato, Nobuyuki Nishikubo, Bo Xu, Masatoshi Yamaguchi, Nobutaka Mitsuda, Nadia Goué, Fusun Shi, Masaru Ohme-Takagi, y Taku Demura. 2011. "A NAC domain protein family contributing to the regulation of wood formation in poplar". *The Plant Journal* 67 (3):499–512.
- Olson, Mark E., y Julieta A. Rosell. 2006. "Using heterochrony to detect modularity in the evolution of stem diversity in the plant family Moringaceae". *Evolution* 60 (4):724–34.
- Ozsolak, Fatih, y Patrice M. Milos. 2011. "RNA sequencing: advances, challenges and opportunities". *Nature Reviews Genetics* 12 (2):87–98.
- Pace, Marcelo R., Lúcia G. Lohmann, y Veronica Angyalossy. 2011. "Evolution of disparity between the regular and variant phloem in Bignonieae (Bignoniaceae)". *American Journal of Botany* 98 (4):602–18.
- Petersen, Thomas Nordahl, Søren Brunak, Gunnar von Heijne, y Henrik Nielsen. 2011. "SignalP 4.0: discriminating signal peptides from transmembrane regions". *Nature Methods* 8 (10):785–86.
- Ranocha, Philippe, Matthieu Chabannes, Simon Chamayou, Saida Danoun, Alain Jauneau, Alain-M. Boudet, y Deborah Goffner. 2002. "Laccase Down-Regulation Causes

- Alterations in Phenolic Metabolism and Cell Wall Structure in Poplar". *Plant Physiology* 129 (1):145–55. <https://doi.org/10.1104/pp.010988>.
- Reyes-Rivera, Jorge, Gonzalo Canche-Escamilla, Marcos Soto-Hernández, y Teresa Terrazas. 2015. "Wood chemical composition in species of Cactaceae: the relationship between lignification and stem morphology". *PloS One* 10 (4):e0123919.
- Robinson, Mark D., Davis J. McCarthy, y Gordon K. Smyth. 2010. "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data". *Bioinformatics* 26 (1):139–40.
- Roux, Julien, Marta Rosikiewicz, y Marc Robinson-Rechavi. 2015. "What to compare and how: Comparative transcriptomics for Evo-Devo". *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 324 (4):372–82.
- Simão, Felipe A., Robert M. Waterhouse, Panagiotis Ioannidis, Evgenia V. Kriventseva, y Evgeny M. Zdobnov. 2015. "BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs". *Bioinformatics* 31 (19):3210–12.
- Spicer, Rachel, y Andrew Groover. 2010. "Evolution of development of vascular cambia and secondary growth". *New Phytologist* 186 (3):577–92.
- Sundell, David, Nathaniel R. Street, Manoj Kumar, Ewa J. Mellerowicz, Melis Kucukoglu, Christoffer Johnsson, Vikash Kumar, et al. 2017. "AspWood: High-Spatial-Resolution Transcriptome Profiles Reveal Uncharacterized Modularity of Wood Formation in *Populus Tremula*". *The Plant Cell*, enero, tpc.00153.2017. <https://doi.org/10.1105/tpc.17.00153>.
- Terrazas, Teresa, y Salvador Arias. 2002. "Comparative stem anatomy in the subfamily Cactoideae". *The Botanical Review* 68 (4):444–73.
- The Gene Ontology Consortium. 2004. "The Gene Ontology (GO) Database and Informatics Resource". *Nucleic Acids Research* 32 (suppl_1):D258–61. <https://doi.org/10.1093/nar/gkh036>.
- Tickle, Cheryll, y Araxi O. Urrutia. 2017. *Perspectives on the history of evo-devo and the contemporary research landscape in the genomics era*. The Royal Society.
- TransDecoder source*. (2015) 2018. Perl. TransDecoder. <https://github.com/TransDecoder/TransDecoder>.
- trinityrnaseq: Trinity RNA-Seq de novo transcriptome assembly*. (2014) 2018. Perl. trinityrnaseq. <https://github.com/trinityrnaseq/trinityrnaseq>.
- Turlapati, Phanikanth V., Kye-Won Kim, Laurence B. Davin, y Norman G. Lewis. 2011. "The Laccase Multigene Family in *Arabidopsis Thaliana*: Towards Addressing the Mystery of Their Gene Function(S)". *Planta* 233 (3):439–70. <https://doi.org/10.1007/s00425-010-1298-3>.
- Wunderling, Anna, Mehdi Ben Targem, Pierre Barbier de Reuille, y Laura Ragni. 2016. "Novel tools for quantifying secondary growth". *Journal of Experimental Botany* 68 (1):89–95.
- Xu, Chenxi, Chen Jiao, Yi Zheng, Honghe Sun, Wenli Liu, Xiaofeng Cai, Xiaoli Wang, et al. 2015. "De novo and comparative transcriptome analysis of cultivated and wild spinach". *Scientific Reports* 5 (diciembre). <https://doi.org/10.1038/srep17706>.

- Young, Matthew D., Matthew J. Wakefield, Gordon K. Smyth, y Alicia Oshlack. 2010. "Gene ontology analysis for RNA-seq: accounting for selection bias". *Genome Biology* 11 (2):R14.
- Zinkgraf, Matthew, Suzanne Gerttula, y Andrew Groover. 2017. "Transcript profiling of a novel plant meristem, the monocot cambium". *Journal of Integrative Plant Biology* 59 (6):436–49.

Información suplementaria

Cuadro S.1. Calidad de las muestras de RNA seleccionadas para secuenciación

Especie	Tejido	Muestra	Concentración	RIN
<i>Ariocarpus retusus</i>	Cambium vascular	A2.C	124	6.8
		A2.K	107	6.4
		J14-A	81	6.8
<i>Cylindropuntia imbricata</i>	Cambium vascular	Co.C	246	7.5
		Co.K	62	6.9
<i>Echinocactus platyacanthus</i>	PC (2 individuos) 1.5 meses	Terrazas_15	87	6.9
	PC (3 individuos) 1.5 meses	Terrazas_25	142	7.7
	PC (3 individuos) 1.5 meses	Terrazas_26	93	7.5
	PC 1.3 años	Terrazas_23	127	8.3
	PC 1.3 años	Terrazas_24	156	8.3
	PC 8 meses	Terrazas_13	130	7.4
	ZC, planta adulta	Terrazas_39	203	7.2
	ZC, planta adulta	Terrazas_40	343.8	7
	ZC, planta adulta	Terrazas_41	189	7.5
<i>Ferocactus pilosus</i>	PC (2 Individuos) 5 meses	Terrazas_10	51	7.6
	PC (3 individuos) 3 meses	Terrazas_29	89	7.7
	PC (3 individuos) 5 meses	Terrazas_30	51	8.3
	Cambium vascular	890.C	376	6.9
	Cambium vascular	E2	114	7.8
<i>Pereskia lychnidiflora</i>	Cambium vascular	P3.C	496	7.3
		J11_P'	370	8.7

Cuadro S2. Parámetros para eliminar bases de baja calidad de limpieza utilizados con Trimmomatic de las bibliotecas de RNA-seq

Especie	Headcrop	Slidingwindow	Minlen	Tophred
<i>A. retusus</i>	9	4:25	31	Tophred33
<i>C. imbricata</i>	8	4:24	33	Tophred33
<i>E. platyacanthus</i>	10	4:28	38	Tophred33
<i>F. pilosus</i>	10	4:28	38	Tophred33
<i>P. lychnidiflora</i>	9	4:25	31	Tophred33

Cuadro S.3. Porcentaje de lecturas mapeadas correctamente

Especie	Pares correctos	Pares no alineados	Total de fragmentos alineados	Porcentaje de bases ensambladas
<i>B. retusus</i>	24,078,151	18,211,24	25,899,275	92.97
<i>C. imbricata</i>	2,665,793	2,223,261	4,889,054	54.52
<i>E. platyacanthus</i>	113,275,367	4,884,504	118,159,871	95.86
<i>F. pilosus</i>	59,377,721	524,581	64,623,002	91.88
<i>P. lychnidiflora</i>	22,056,876	1782,509	23,942,308	92.12

Cuadro S4. Porcentaje de cobertura de los alineamientos únicos más altos entre los transcritos ensamblados y de proteínas de la base de datos uniprot.

Especie	% de cobertura	Número de proteínas con alineamientos mayores	Acumulado de proteínas con alineaientos
<i>A. retusus</i>	100	5773	5773
	90	1865	7638
	80	1190	8828
	70	998	9826
	60	1029	10855
	50	865	11720
	40	808	12528
	30	745	13273
	20	667	13940
	10	147	14087
<i>C. imbricata</i>	100	3267	3267
	90	994	4261
	80	815	5076
	70	743	5819
	60	808	6627
	50	902	7529
	40	983	8512
	30	1084	9596
	20	1059	10655
	10	269	10924
<i>F. pilosus</i>	100	6211	6211
	90	2002	8213
	80	1356	9569
	70	1111	10680

Cuadro S4. Continuación

Especie	% de cobertura	Número de proteínas con alineamientos mayores	Acumulado de proteínas con alineamientos
<i>F. pilosus</i>	60	1106	11786
	50	983	12769
	40	827	13596
	30	822	14418
	20	620	15038
	10	148	15186
<i>P. lychnidiflora</i>	100	5836	5836
	90	1748	7584
	80	1199	8783
	70	1056	9839
	60	984	10823
	50	814	11637
	40	739	12376
	30	754	13130
	20	584	13714
10	155	13869	

Cuadro S5. Categorías funcionales enriquecidas en *E. platyacanthus*

Estadio	category	Valor p sobrerrepresentado	Término	Ontología
1 año 5 meses	GO:0020037	8.80677603071876E-10	heme binding	MF
1 año 5 meses	GO:0016491	1.06414744013642E-09	oxidoreductase activity	MF
1 año 5 meses	GO:0046906	1.53843112251189E-08	Tetrapyrrole binding	MF
1 año 5 meses	GO:0016705	2.47220315812157E-07	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	MF
1 año 5 meses	GO:0004497	1.17705570131191E-06	monooxygenase activity	MF
1 año 5 meses	GO:0009698	3.84157192619904E-06	phenylpropanoid metabolic process	BP
1 año 5 meses	GO:0005506	6.28844583907602E-06	iron ion binding	MF
1 año 5 meses	GO:0005576	9.80873630446952E-06	extracellular region	CC
1 año 5 meses	GO:0016115	2.77544025274776E-05	terpenoid catabolic process	BP
1 año 5 meses	GO:0046914	3.17950619567153E-05	transition metal ion binding	MF
1 año 5 meses	GO:0008300	5.0472948406896E-05	isoprenoid catabolic process	BP
Zona cambial	GO:0003964	1.0250496791956E-34	RNA-directed DNA polymerase activity	MF
Zona cambial	GO:0034061	1.66168930581653E-34	DNA polymerase activity	MF
Zona cambial	GO:0016779	5.09225501118835E-26	nucleotidyltransferase activity	MF
Zona cambial	GO:0090304	2.35329398512578E-23	nucleic acid metabolic process	BP
Zona cambial	GO:0015074	4.99403047356131E-22	DNA integration	BP
Zona cambial	GO:0016893	1.12210812060096E-21	endonuclease activity, active with either ribo- or deoxyribonucleic acids and producing 5'-phosphomonoesters	MF
Zona cambial	GO:0006139	9.68719474091459E-21	nucleobase-containing compound metabolic process	BP
Zona cambial	GO:0003700	1.44203287594121E-19	DNA binding transcription factor activity	MF
Zona cambial	GO:0001071	1.68900426198213E-19	NA	NA
Zona cambial	GO:0003677	2.99176522981353E-19	DNA binding	MF
Zona cambial	GO:0004190	8.63425051566494E-19	aspartic-type endopeptidase activity	MF
Zona cambial	GO:0070001	8.63425051566494E-19	aspartic-type peptidase activity	MF
Zona cambial	GO:0046483	5.28920847250726E-18	heterocycle metabolic process	BP
Zona cambial	GO:0004523	5.59414942348178E-18	RNA-DNA hybrid ribonuclease activity	MF
Zona cambial	GO:0034654	1.03488358447049E-17	nucleobase-containing compound biosynthetic process	BP
Zona cambial	GO:0006259	1.37728344973373E-17	DNA metabolic process	BP
Zona cambial	GO:0003676	3.50887392171341E-17	nucleic acid binding	MF
Zona cambial	GO:0006725	4.18434221321431E-17	cellular aromatic compound metabolic process	BP
Zona cambial	GO:0004519	9.21831238583373E-17	endonuclease activity	MF
Zona cambial	GO:0016891	2.0469756907139E-16	endoribonuclease activity, producing 5'-phosphomonoesters	MF
Zona cambial	GO:0006351	2.80489230486074E-16	transcription, DNA-templated	BP
Zona cambial	GO:0097659	2.80489230486074E-16	nucleic acid-templated transcription	BP
Zona cambial	GO:0032197	3.18693673960529E-16	transposition, RNA-mediated	BP

Continuación. Cuadro S5. Categorías funcionales enriquecidas en *E. platyacanthus*

Estadio	category	Valor p sobrerrepresentado	Término	Ontología
Zona cambial	GO:1901360	3.43042559809722E-16	organic cyclic compound metabolic process	BP
Zona cambial	GO:0004518	6.8996140789087E-16	nuclease activity	MF
Zona cambial	GO:0004521	7.08058719690801E-16	endoribonuclease activity	MF
Zona cambial	GO:0032196	2.12698620082711E-15	transposition	BP
Zona cambial	GO:0018130	3.2893083262627E-15	heterocycle biosynthetic process	BP
Zona cambial	GO:0032774	1.04356716546393E-14	RNA biosynthetic process	BP
Zona cambial	GO:0019438	2.90400957192059E-14	aromatic compound biosynthetic process	BP
Zona cambial	GO:0034641	3.81240959573338E-14	cellular nitrogen compound metabolic process	BP
Zona cambial	GO:0004540	6.63853567122602E-14	ribonuclease activity	MF
Zona cambial	GO:1901362	5.70927885905818E-13	organic cyclic compound biosynthetic process	BP
Zona cambial	GO:0043565	9.54920315969074E-13	sequence-specific DNA binding	MF
Zona cambial	GO:0034645	1.14894746046641E-12	cellular macromolecule biosynthetic process	BP
Zona cambial	GO:0006310	1.35459809447486E-11	DNA recombination	BP
Zona cambial	GO:0004175	1.63745271066175E-11	endopeptidase activity	MF
Zona cambial	GO:0005634	2.21871430927875E-11	nucleus	CC
Zona cambial	GO:0044271	1.11457903105072E-10	cellular nitrogen compound biosynthetic process	BP
Zona cambial	GO:0009059	1.21833582885179E-10	macromolecule biosynthetic process	BP
Zona cambial	GO:0044260	5.88073253665585E-10	cellular macromolecule metabolic process	BP
Zona cambial	GO:0003887	7.5971727207063E-10	DNA-directed DNA polymerase activity	MF
Zona cambial	GO:0016070	1.59520793433036E-09	RNA metabolic process	BP
Zona cambial	GO:0046872	1.81537904835688E-09	metal ion binding	MF
Zona cambial	GO:0043169	2.12347950474959E-09	cation binding	MF
Zona cambial	GO:0000943	5.54602760739935E-09	retrotransposon nucleocapsid	CC
Zona cambial	GO:1901363	7.50111480507386E-09	heterocyclic compound binding	MF
Zona cambial	GO:0097159	8.44438614361506E-09	organic cyclic compound binding	MF
Zona cambial	GO:0001171	9.64765079360317E-09	reverse transcription	BP
Zona cambial	GO:0009036	9.64765079360317E-09	Type II site-specific deoxyribonuclease activity	MF
Zona cambial	GO:0015666	9.64765079360317E-09	restriction endodeoxyribonuclease activity	MF
Zona cambial	GO:0032199	9.64765079360317E-09	reverse transcription involved in RNA-mediated transposition	BP
Zona cambial	GO:0016888	2.01872305847546E-08	endodeoxyribonuclease activity, producing 5'-phosphomonoesters	MF
Zona cambial	GO:0006355	2.71809660622124E-08	regulation of transcription, DNA-templated	BP
Zona cambial	GO:0000160	2.76098600627239E-08	phosphorelay signal transduction system	BP
Zona cambial	GO:1903506	2.81239511410165E-08	regulation of nucleic acid-templated transcription	BP
Zona cambial	GO:2001141	3.00707465912873E-08	regulation of RNA biosynthetic process	BP
Zona cambial	GO:0043170	3.89589968945779E-08	macromolecule metabolic process	BP
Zona cambial	GO:0008270	4.15481568597855E-08	zinc ion binding	MF

Continuación. Cuadro S5. Categorías funcionales enriquecidas en *E. platyacanthus*

Estadio	category	Valor p sobrerrepresentado	Término	Ontología
Zona cambial	GO:0009755	7.82167106624369E-08	hormone-mediated signaling pathway	BP
Zona cambial	GO:2000112	8.52814876827269E-08	regulation of cellular macromolecule biosynthetic process	BP
Zona cambial	GO:0010556	1.16665732638025E-07	regulation of macromolecule biosynthetic process	BP
Zona cambial	GO:0006278	1.24135022831546E-07	RNA-dependent DNA biosynthetic process	BP
Zona cambial	GO:0019219	1.31420127844767E-07	regulation of nucleobase-containing compound metabolic process	BP
Zona cambial	GO:0032501	1.42088044533838E-07	multicellular organismal process	BP
Zona cambial	GO:0071897	2.28299504626979E-07	DNA biosynthetic process	BP
Zona cambial	GO:0051252	2.30605463962795E-07	regulation of RNA metabolic process	BP
Zona cambial	GO:0044707	2.70171645401765E-07	NA	NA
Zona cambial	GO:0016772	8.26364471745039E-07	transferase activity, transferring phosphorus-containing groups	MF
Zona cambial	GO:0031326	9.53570113767697E-07	regulation of cellular biosynthetic process	BP
Zona cambial	GO:0009889	1.03486915128829E-06	regulation of biosynthetic process	BP
Zona cambial	GO:0010468	1.4618421348733E-06	regulation of gene expression	BP
Zona cambial	GO:0070011	1.79723299195105E-06	peptidase activity, acting on L-amino acid peptides	MF
Zona cambial	GO:0004520	1.90458759742033E-06	endodeoxyribonuclease activity	MF
Zona cambial	GO:0004536	2.17911784836647E-06	deoxyribonuclease activity	MF
Zona cambial	GO:0009299	2.18480963150096E-06	mRNA transcription	BP
Zona cambial	GO:0048856	2.46146710249825E-06	anatomical structure development	BP
Zona cambial	GO:0008233	2.98519747698763E-06	peptidase activity	MF
Zona cambial	GO:0051171	6.21870922483437E-06	regulation of nitrogen compound metabolic process	BP
Zona cambial	GO:0070042	6.224759893204E-06	rRNA (uridine-N3-)-methyltransferase activity	MF
Zona cambial	GO:0016788	6.39843015869648E-06	hydrolase activity, acting on ester bonds	MF
Zona cambial	GO:0000976	9.15990301501456E-06	transcription regulatory region sequence-specific DNA binding	MF
Zona cambial	GO:0009736	1.68824119557316E-05	cytokinin-activated signaling pathway	BP
Zona cambial	GO:0044212	2.65255358481137E-05	transcription regulatory region DNA binding	MF
Zona cambial	GO:0006807	2.79806526711845E-05	nitrogen compound metabolic process	BP
Zona cambial	GO:0000975	2.83665944501633E-05	regulatory region DNA binding	MF
Zona cambial	GO:0001067	2.83665944501633E-05	regulatory region nucleic acid binding	MF
Zona cambial	GO:0016436	2.92427717744621E-05	rRNA (uridine) methyltransferase activity	MF
Zona cambial	GO:0022414	3.50508956872905E-05	reproductive process	BP
Zona cambial	GO:1990837	3.8127687318078E-05	sequence-specific double-stranded DNA binding	MF
Zona cambial	GO:0031323	3.81421144195831E-05	regulation of cellular metabolic process	BP
Zona cambial	GO:0060255	4.06392572248359E-05	regulation of macromolecule metabolic process	BP
Zona cambial	GO:0080090	6.44025330557019E-05	regulation of primary metabolic process	BP
Zona cambial	GO:0016427	6.84845541039844E-05	tRNA (cytosine) methyltransferase activity	MF
Zona cambial	GO:0044767	9.25160813315676E-05	NA	NA

Continuación. Cuadro S5. Categorías funcionales enriquecidas en *E. platyacanthus*

Estadio	category	Valor p sobrerrepresentado	Término	Ontología
Zona cambial	GO:0035556	0.000115101073129	intracellular signal transduction	BP
Zona cambial	GO:0015691	0.0001187460789	cadmium ion transport	BP
Zona cambial	GO:0070574	0.0001187460789	cadmium ion transmembrane transport	BP
Zona cambial	GO:0009873	0.000128858005816	ethylene-activated signaling pathway	BP
Zona cambial	GO:0044702	0.000152684917952	NA	NA
Zona cambial	GO:0005488	0.000152847310624	binding	MF
Zona cambial	GO:0043531	0.000185028036493	ADP binding	MF
Zona cambial	GO:0046914	0.000202987684975	transition metal ion binding	MF
Zona cambial	GO:0048608	0.000218389278146	reproductive structure development	BP
Zona cambial	GO:0042991	0.000230649866147	transcription factor import into nucleus	BP
Zona cambial	GO:0019222	0.000235931414722	regulation of metabolic process	BP
Zona cambial	GO:0007165	0.000269432232923	signal transduction	BP
Zona cambial	GO:0001046	0.000298285149655	core promoter sequence-specific DNA binding	MF
Zona cambial	GO:0032502	0.000318143464542	developmental process	BP
Zona cambial	GO:0043167	0.000324673016085	ion binding	MF
Zona cambial	GO:0009791	0.000448790030236	post-embryonic development	BP
Zona cambial	GO:0090305	0.000488103338371	nucleic acid phosphodiester bond hydrolysis	BP
Zona cambial	GO:0050794	0.000509093988649	regulation of cellular process	BP
1 mes 15 días	GO:0005576	2.19909097985627E-15	extracellular region	CC
1 mes 15 días	GO:0016491	2.26029989940499E-15	oxidoreductase activity	MF
1 mes 15 días	GO:0046906	5.17021392082079E-12	tetrapyrrole binding	MF
1 mes 15 días	GO:1901568	4.68908763933092E-10	fatty acid derivative metabolic process	BP
1 mes 15 días	GO:0010025	2.71302370539851E-09	wax biosynthetic process	BP
1 mes 15 días	GO:0010166	2.71302370539851E-09	wax metabolic process	BP
1 mes 15 días	GO:1901570	6.85649545495049E-09	fatty acid derivative biosynthetic process	BP
1 mes 15 días	GO:0045229	1.7143960696811E-08	external encapsulating structure organization	BP
1 mes 15 días	GO:0005618	2.63186989810149E-08	cell wall	CC
1 mes 15 días	GO:0030312	2.63186989810149E-08	external encapsulating structure	CC
1 mes 15 días	GO:0044710	3.79676413834996E-08	NA	NA
1 mes 15 días	GO:0006629	4.12264945994314E-08	lipid metabolic process	BP
1 mes 15 días	GO:0016168	2.32734691790718E-07	chlorophyll binding	MF
1 mes 15 días	GO:0005506	4.6361217659651E-07	iron ion binding	MF
1 mes 15 días	GO:0020037	5.21596985240817E-07	heme binding	MF
1 mes 15 días	GO:0004497	6.85257251121112E-07	monoxygenase activity	MF
1 mes 15 días	GO:0009535	1.58004282457499E-06	chloroplast thylakoid membrane	CC
1 mes 15 días	GO:0055035	1.58004282457499E-06	plastid thylakoid membrane	CC

Continuación. Cuadro S5. Categorías funcionales enriquecidas en *E. platyacanthus*

Estadio	category	Valor p sobrerrepresentado	Término	Ontología
1 mes 15 días	GO:0044436	2.48666726943957E-06	thylakoid part	CC
1 mes 15 días	GO:0034357	2.76132223425254E-06	photosynthetic membrane	CC
1 mes 15 días	GO:0042651	2.76132223425254E-06	thylakoid membrane	CC
1 mes 15 días	GO:0102336	3.63106808441924E-06	3-oxo-arachidoyl-CoA synthase activity	MF
1 mes 15 días	GO:0102337	3.63106808441924E-06	3-oxo-cerotoyl-CoA synthase activity	MF
1 mes 15 días	GO:0102338	3.63106808441924E-06	3-oxo-lignoceronyl-CoA synthase activity	MF
			oxidoreductase activity, acting on paired donors, with incorporation or reduction of	
1 mes 15 días	GO:0016705	3.86954896357104E-06	molecular oxygen	MF
1 mes 15 días	GO:0009521	6.9048223187508E-06	photosystem	CC
1 mes 15 días	GO:0031224	1.08551207472464E-05	intrinsic component of membrane	CC
1 mes 15 días	GO:0071555	1.24147351299256E-05	cell wall organization	BP
1 mes 15 días	GO:0006631	1.31956484425556E-05	fatty acid metabolic process	BP
1 mes 15 días	GO:0009523	1.66323539882486E-05	photosystem II	CC
1 mes 15 días	GO:0010584	1.67563991722304E-05	pollen exine formation	BP
1 mes 15 días	GO:0009922	2.17165014689982E-05	fatty acid elongase activity	MF
1 mes 15 días	GO:0044711	2.20764446370396E-05	NA	NA
1 mes 15 días	GO:0006633	2.21926217184248E-05	fatty acid biosynthetic process	BP
1 mes 15 días	GO:0042335	2.41629291938034E-05	cuticle development	BP
1 mes 15 días	GO:0009534	2.73051844171231E-05	chloroplast thylakoid	CC
1 mes 15 días	GO:0031976	2.73051844171231E-05	plastid thylakoid	CC
1 mes 15 días	GO:0009522	3.26984829086827E-05	photosystem I	CC
1 mes 15 días	GO:0008610	3.49982853652017E-05	lipid biosynthetic process	BP
1 mes 15 días	GO:0009579	3.54236272124072E-05	thylakoid	CC
1 mes 15 días	GO:0072330	4.19771097464011E-05	monocarboxylic acid biosynthetic process	BP
1 mes 15 días	GO:0010927	4.33038132912416E-05	cellular component assembly involved in morphogenesis	BP
1 mes 15 días	GO:0080110	4.64148495991928E-05	sporopollenin biosynthetic process	BP
1 mes 15 días	GO:0004553	6.83642561936412E-05	hydrolase activity, hydrolyzing O-glycosyl compounds	MF
1 mes 15 días	GO:0016747	6.94213374614175E-05	transferase activity, transferring acyl groups other than amino-acyl groups	MF
1 mes 15 días	GO:0018298	7.90464354249924E-05	protein-chromophore linkage	BP
1 mes 15 días	GO:0071554	8.14798771265577E-05	cell wall organization or biogenesis	BP
1 mes 15 días	GO:0046658	0.000106577778718	anchored component of plasma membrane	CC
1 mes 15 días	GO:0008422	0.00012409331896	beta-glucosidase activity	MF
1 mes 15 días	GO:0010208	0.000133274213781	pollen wall assembly	BP
1 mes 15 días	GO:0085029	0.000133274213781	extracellular matrix assembly	BP
1 mes 15 días	GO:0016614	0.000139509257044	oxidoreductase activity, acting on CH-OH group of donors	MF
1 mes 15 días	GO:0016798	0.000142844553764	hydrolase activity, acting on glycosyl bonds	MF

Continuación. Cuadro S5. Categorías funcionales enriquecidas en *E. platyacanthus*

Estadio	category	Valor p sobrerrepresentado	Término	Ontología
1 mes 15 días	GO:0042973	0.000161165268667	glucan endo-1,3-beta-D-glucosidase activity	MF
1 mes 15 días	GO:0016746	0.000193481218582	transferase activity, transferring acyl groups	MF
1 mes 15 días	GO:0010143	0.000202910304327	cutin biosynthetic process	BP
1 mes 15 días	GO:0044425	0.000204272879804	membrane part	CC
1 mes 15 días	GO:0031225	0.000213240890151	anchored component of membrane	CC
1 mes 15 días	GO:0048046	0.000221521321568	apoplast	CC
1 mes 15 días	GO:0048235	0.000251610720925	pollen sperm cell differentiation	BP
1 mes 15 días	GO:0009765	0.000306662858844	photosynthesis, light harvesting	BP

*Las categorías de Zona cambial de plantas adultas están enriquecidas con respecto a plantas de un mes 15 días y un año cinco meses. Las categorías enriquecidas en plantas de 1 año 5 meses son contra zona cambial y plantas de un mes 15 días y las categorías enriquecidas en plantas de un mes 15 días son contra la zona cambial y las plantas de un año cinco meses.

Cuadro S6. Categorías funcionales enriquecidas en *F. pilosus*

Estadio	Categoría	Valor p sobrerrepresentado	Término	Ontología
Zona cambial	GO:0003964	6.18870544251215E-17	RNA-directed DNA polymerase activity	MF
Zona cambial	GO:0034061	2.39086861835623E-15	DNA polymerase activity	MF
Zona cambial	GO:0003700	1.77934049363643E-14	DNA binding transcription factor activity	MF
Zona cambial	GO:0001071	2.41898253861907E-14	NA	NA
Zona cambial	GO:0010200	1.24286557876963E-13	response to chitin	BP
Zona cambial	GO:0004190	4.56846583176321E-13	aspartic-type endopeptidase activity	MF
Zona cambial	GO:0070001	4.56846583176321E-13	aspartic-type peptidase activity	MF
Zona cambial	GO:0016779	3.89428618207611E-12	nucleotidyltransferase activity	MF
Zona cambial	GO:0006952	4.17058422718323E-12	defense response	BP
Zona cambial	GO:0043565	1.04142565822173E-11	sequence-specific DNA binding	MF
Zona cambial	GO:0010243	3.89632953926718E-11	response to organonitrogen compound	BP
Zona cambial	GO:0003677	3.93914115611282E-11	DNA binding	MF
Zona cambial	GO:0015074	4.29684901056537E-11	DNA integration	BP
Zona cambial	GO:0006351	2.50252066432455E-10	transcription, DNA-templated	BP
Zona cambial	GO:0097659	2.50252066432455E-10	nucleic acid-templated transcription	BP
Zona cambial	GO:0050896	5.91053687783494E-10	response to stimulus	BP
Zona cambial	GO:0009607	9.40948416249816E-10	response to biotic stimulus	BP
Zona cambial	GO:0016893	9.54434137741387E-10	endonuclease activity, active with either ribo- or deoxyribonucleic acids and producing 5'-phosphomonoesters	MF
Zona cambial	GO:1901700	1.19290512089313E-09	response to oxygen-containing compound	BP
Zona cambial	GO:1901698	1.3556932013055E-09	response to nitrogen compound	BP
Zona cambial	GO:0032774	2.2853066213899E-09	RNA biosynthetic process	BP
Zona cambial	GO:0043207	2.55489440236174E-09	response to external biotic stimulus	BP
Zona cambial	GO:0046872	3.85333469544657E-09	metal ion binding	MF
Zona cambial	GO:0043169	4.2364379708346E-09	cation binding	MF
Zona cambial	GO:0051707	7.43062120314271E-09	response to other organism	BP
Zona cambial	GO:0006950	7.43568880305888E-09	response to stress	BP
Zona cambial	GO:0016772	7.5282109390175E-09	transferase activity, transferring phosphorus-containing groups	MF
Zona cambial	GO:0005634	7.668274333555E-09	nucleus	CC
Zona cambial	GO:0098542	9.20724799985803E-09	defense response to other organism	BP
Zona cambial	GO:0032197	9.34610283013839E-09	transposition, RNA-mediated	BP
Zona cambial	GO:0042221	1.0574352031247E-08	response to chemical	BP
Zona cambial	GO:0010033	1.30818953041155E-08	response to organic substance	BP
Zona cambial	GO:0004175	1.3476405109443E-08	endopeptidase activity	MF
Zona cambial	GO:0006355	2.02862399253602E-08	regulation of transcription, DNA-templated	BP
Zona cambial	GO:1903506	2.02862399253602E-08	regulation of nucleic acid-templated transcription	BP

Cuadro S6. Categorías funcionales enriquecidas en *F. pilosus*

Estadio	Categoría	Valor <i>p</i> sobrerrepresentado	Término	Ontología
Zona cambial	GO:2001141	2.22838414989375E-08	regulation of RNA biosynthetic process	BP
Zona cambial	GO:0051704	3.64612450384183E-08	multi-organism process	BP
Zona cambial	GO:0051252	4.62784866377565E-08	regulation of RNA metabolic process	BP
Zona cambial	GO:0016891	9.25621276866684E-08	endoribonuclease activity, producing 5'-phosphomonoesters	MF
Zona cambial	GO:0009719	1.00411694605806E-07	response to endogenous stimulus	BP
Zona cambial	GO:0010556	1.03331283050061E-07	regulation of macromolecule biosynthetic process	BP
Zona cambial	GO:2000112	1.0547028496403E-07	regulation of cellular macromolecule biosynthetic process	BP
Zona cambial	GO:0051171	1.08476455371846E-07	regulation of nitrogen compound metabolic process	BP
Zona cambial	GO:0019219	1.10776042814991E-07	regulation of nucleobase-containing compound metabolic process	BP
Zona cambial	GO:0065007	1.46129111265866E-07	biological regulation	BP
Zona cambial	GO:0009889	1.95742674430484E-07	regulation of biosynthetic process	BP
Zona cambial	GO:0031326	1.99409090256958E-07	regulation of cellular biosynthetic process	BP
Zona cambial	GO:0070011	3.96371994430664E-07	peptidase activity, acting on L-amino acid peptides	MF
Zona cambial	GO:0032196	5.31593760251844E-07	transposition	BP
Zona cambial	GO:0034654	8.13168172708888E-07	nucleobase-containing compound biosynthetic process	BP
Zona cambial	GO:0046914	9.53340741032979E-07	transition metal ion binding	MF
Zona cambial	GO:0004540	1.3288277726239E-06	ribonuclease activity	MF
Zona cambial	GO:0004523	1.4469423331063E-06	RNA-DNA hybrid ribonuclease activity	MF
Zona cambial	GO:0000943	1.87929239034547E-06	retrotransposon nucleocapsid	CC
Zona cambial	GO:0050794	1.99502873959374E-06	regulation of cellular process	BP
Zona cambial	GO:0004521	2.16731697948853E-06	endoribonuclease activity	MF
Zona cambial	GO:0009617	2.29988550575111E-06	response to bacterium	BP
Zona cambial	GO:0010468	2.59423375953653E-06	regulation of gene expression	BP
Zona cambial	GO:0008233	4.0974333158372E-06	peptidase activity	MF
Zona cambial	GO:0044212	4.52546208209502E-06	transcription regulatory region DNA binding	MF
Zona cambial	GO:0008270	4.88899914976852E-06	zinc ion binding	MF
Zona cambial	GO:0000975	5.10686567902311E-06	regulatory region DNA binding	MF
Zona cambial	GO:0001067	5.10686567902311E-06	regulatory region nucleic acid binding	MF
Zona cambial	GO:0016740	5.36943298291745E-06	transferase activity	MF
Zona cambial	GO:0050789	7.58174630076087E-06	regulation of biological process	BP
Zona cambial	GO:0031323	8.44189214007502E-06	regulation of cellular metabolic process	BP
Zona cambial	GO:0009605	9.46870920251226E-06	response to external stimulus	BP
Zona cambial	GO:0018130	1.48321234077372E-05	heterocycle biosynthetic process	BP
Zona cambial	GO:0009266	1.59411274270293E-05	response to temperature stimulus	BP
Zona cambial	GO:0001101	1.59705188776251E-05	response to acid chemical	BP

Cuadro S6. Categorías funcionales enriquecidas en *F. pilosus*

Estadio	Categoría	Valor p	Término	Ontología
Zona cambial	GO:0043167	1.72803993949766E-05	ion binding	MF
Zona cambial	GO:0019438	1.77086508786809E-05	aromatic compound biosynthetic process	BP
Zona cambial	GO:0080090	2.15965629254437E-05	regulation of primary metabolic process	BP
Zona cambial	GO:0090304	2.73733946640899E-05	nucleic acid metabolic process	BP
Zona cambial	GO:0000302	2.92680192727925E-05	response to reactive oxygen species	BP
Zona cambial	GO:0042742	2.96090340534301E-05	defense response to bacterium	BP
Zona cambial	GO:0005488	3.61797101972936E-05	binding	MF
Zona cambial	GO:0044271	3.83989672660546E-05	cellular nitrogen compound biosynthetic process	BP
Zona cambial	GO:0009816	4.77458640254467E-05	defense response to bacterium, incompatible interaction	BP
Zona cambial	GO:1901362	5.13400625734133E-05	organic cyclic compound biosynthetic process	BP
Zona cambial	GO:0003887	5.37157944663406E-05	DNA-directed DNA polymerase activity	MF
Zona cambial	GO:0060255	6.01665262453808E-05	regulation of macromolecule metabolic process	BP
Zona cambial	GO:0019222	6.26309917504074E-05	regulation of metabolic process	BP
Zona cambial	GO:0002376	9.28022858977363E-05	immune system process	BP
Zona cambial	GO:0009755	0.000116318671174	hormone-mediated signaling pathway	BP
Zona cambial	GO:0004519	0.000117699506888	endonuclease activity	MF
Zona cambial	GO:0003676	0.000163051043931	nucleic acid binding	MF
Zona cambial	GO:0009751	0.000168431564252	response to salicylic acid	BP
Zona cambial	GO:0009410	0.000170771325727	response to xenobiotic stimulus	BP
Zona cambial	GO:1901363	0.000187093838693	heterocyclic compound binding	MF
Zona cambial	GO:0010035	0.000199582329917	response to inorganic substance	BP
Zona cambial	GO:0097159	0.000208251987786	organic cyclic compound binding	MF
Zona cambial	GO:0007165	0.000209703141091	signal transduction	BP
Zona cambial	GO:0009873	0.000214473689532	ethylene-activated signaling pathway	BP
Zona cambial	GO:0010193	0.000216191946691	response to ozone	BP
Zona cambial	GO:0043531	0.000220465565914	ADP binding	MF
Zona cambial	GO:0046740	0.000253490158193	transport of virus in host, cell to cell	BP
Zona cambial	GO:1902586	0.000253490158193	multi-organism intercellular transport	BP
Zona cambial	GO:0009414	0.000310965520589	response to water deprivation	BP
Zona cambial	GO:0009415	0.000310965520589	response to water	BP
Zona cambial	GO:0006979	0.000386928394611	response to oxidative stress	BP
Zona cambial	GO:0000160	0.000397140378183	phosphorelay signal transduction system	BP
Zona cambial	GO:0016888	0.000406684772552	endodeoxyribonuclease activity, producing 5'-phosphomonoesters	MF
Zona cambial	GO:0004518	0.000496068079845	nuclease activity	MF
3-5meses	GO:0044435	3.99315824376753E-45	plastid part	CC

Cuadro S6. Categorías funcionales enriquecidas en *F. pilosus*

Estadio	Categoría	Valor p sobrerrepresentado	Término	Ontología
3-5meses	GO:0044434	5.33766835908759E-45	chloroplast part	CC
3-5meses	GO:0044436	1.63978682772406E-40	thylakoid part	CC
3-5meses	GO:0009507	2.11788566397613E-39	chloroplast	CC
3-5meses	GO:0034357	1.41856717156593E-37	photosynthetic membrane	CC
3-5meses	GO:0042651	1.41856717156593E-37	thylakoid membrane	CC
3-5meses	GO:0009536	3.13566659083835E-37	plastid	CC
3-5meses	GO:0009535	9.30398900374545E-36	chloroplast thylakoid membrane	CC
3-5meses	GO:0055035	9.30398900374545E-36	plastid thylakoid membrane	CC
3-5meses	GO:0015979	3.49179962010147E-29	photosynthesis	BP
3-5meses	GO:0009521	4.02110230538982E-26	photosystem	CC
3-5meses	GO:0009579	8.47350006260932E-23	thylakoid	CC
3-5meses	GO:0009534	1.18585340572316E-21	chloroplast thylakoid	CC
3-5meses	GO:0031976	1.18585340572316E-21	plastid thylakoid	CC
3-5meses	GO:0031984	3.59480739296899E-19	organelle subcompartment	CC
3-5meses	GO:0005576	2.23077515761738E-18	extracellular region	CC
3-5meses	GO:0009523	1.02392254882435E-17	photosystem II	CC
3-5meses	GO:0009570	5.52974206719975E-17	chloroplast stroma	CC
3-5meses	GO:0009532	9.28105292890285E-17	plastid stroma	CC
3-5meses	GO:0009941	9.40898704253026E-17	chloroplast envelope	CC
3-5meses	GO:0009526	1.06569415653878E-16	plastid envelope	CC
3-5meses	GO:0009522	1.38316509582989E-14	photosystem I	CC
3-5meses	GO:0044710	1.97870758785756E-14	NA	NA
3-5meses	GO:0031967	3.02684219146144E-13	organelle envelope	CC
3-5meses	GO:0031975	3.02684219146144E-13	envelope	CC
3-5meses	GO:0016168	2.55934491687616E-12	chlorophyll binding	MF
3-5meses	GO:0048046	5.24026219082347E-12	apoplast	CC
3-5meses	GO:0019752	6.11388309543664E-12	carboxylic acid metabolic process	BP
3-5meses	GO:0010287	6.73395630163846E-12	plastoglobule	CC
3-5meses	GO:0055114	7.51191196623413E-12	oxidation-reduction process	BP
3-5meses	GO:0044711	7.91925672930059E-12	NA	NA
3-5meses	GO:0043436	1.57691292806855E-11	oxoacid metabolic process	BP
3-5meses	GO:0006082	1.64508678970467E-11	organic acid metabolic process	BP
3-5meses	GO:0016491	9.35348545069504E-11	oxidoreductase activity	MF
3-5meses	GO:0008610	2.32285526569522E-10	lipid biosynthetic process	BP
3-5meses	GO:0044281	6.43450550408802E-10	small molecule metabolic process	BP

Cuadro S6. Categorías funcionales enriquecidas en *F. pilosus*

Estadio	Categoría	Valor p sobrerrepresentado	Término	Ontología
3-5meses	GO:0044283	8.02705984852199E-10	small molecule biosynthetic process	BP
3-5meses	GO:0006629	8.3939304638043E-10	lipid metabolic process	BP
3-5meses	GO:0032787	1.17822421889264E-09	monocarboxylic acid metabolic process	BP
3-5meses	GO:0016053	4.00355743685238E-09	organic acid biosynthetic process	BP
3-5meses	GO:0046394	4.00355743685238E-09	carboxylic acid biosynthetic process	BP
3-5meses	GO:0009768	8.8779475612611E-09	photosynthesis, light harvesting in photosystem I	BP
3-5meses	GO:0005975	1.17095874214519E-08	carbohydrate metabolic process	BP
3-5meses	GO:0044425	2.92506607001261E-08	membrane part	CC
3-5meses	GO:0031224	2.95599139230657E-08	intrinsic component of membrane	CC
3-5meses	GO:0072330	3.72762518705957E-08	monocarboxylic acid biosynthetic process	BP
3-5meses	GO:0044699	4.30381229525427E-08	NA	NA
3-5meses	GO:0006633	4.36100916258489E-08	fatty acid biosynthetic process	BP
3-5meses	GO:0046906	5.20206551028597E-08	tetrapyrrole binding	MF
3-5meses	GO:0019685	5.48283639248585E-08	photosynthesis, dark reaction	BP
3-5meses	GO:0031409	5.90997712378845E-08	pigment binding	MF
3-5meses	GO:0055085	6.64938967973481E-08	transmembrane transport	BP
3-5meses	GO:0022857	9.77074183173016E-08	transmembrane transporter activity	MF
3-5meses	GO:0009765	1.21406180416694E-07	photosynthesis, light harvesting	BP
3-5meses	GO:0016020	2.27370739728837E-07	membrane	CC
3-5meses	GO:0019253	2.84696120142668E-07	reductive pentose-phosphate cycle	BP
3-5meses	GO:0046148	4.14348060709771E-07	pigment biosynthetic process	BP
3-5meses	GO:0042440	4.50125977417213E-07	pigment metabolic process	BP
3-5meses	GO:0006006	6.53298574051438E-07	glucose metabolic process	BP
3-5meses	GO:0010025	6.58346213054569E-07	wax biosynthetic process	BP
3-5meses	GO:0010166	6.58346213054569E-07	wax metabolic process	BP
3-5meses	GO:0015994	8.53647259914858E-07	chlorophyll metabolic process	BP
3-5meses	GO:0006091	1.06459111996893E-06	generation of precursor metabolites and energy	BP
3-5meses	GO:0006546	1.98362789097146E-06	glycine catabolic process	BP
3-5meses	GO:0016052	2.02676602404253E-06	carbohydrate catabolic process	BP
3-5meses	GO:0004497	2.18937253994221E-06	monooxygenase activity	MF
3-5meses	GO:0008964	2.54410175253566E-06	phosphoenolpyruvate carboxylase activity	MF
3-5meses	GO:0030104	2.6655554648697E-06	water homeostasis	BP
3-5meses	GO:0005618	2.69414855236936E-06	cell wall	CC
3-5meses	GO:0030312	2.69414855236936E-06	external encapsulating structure	CC
3-5meses	GO:0015672	3.06518764167711E-06	monovalent inorganic cation transport	BP

Cuadro S6. Categorías funcionales enriquecidas en *F. pilosus*

Estadio	Categoría	Valor p	Término	Ontología
3-5meses	GO:0004375	3.21401098774033E-06	glycine dehydrogenase (decarboxylating) activity	MF
3-5meses	GO:0016642	3.21401098774033E-06	oxidoreductase activity, acting on the CH-NH2 group of donors, disulfide as acceptor	MF
3-5meses	GO:0010206	3.570218366528E-06	photosystem II repair	BP
3-5meses	GO:0015995	3.83756355669426E-06	chlorophyll biosynthetic process	BP
3-5meses	GO:0006778	4.0012059790136E-06	porphyrin-containing compound metabolic process	BP
3-5meses	GO:0031977	4.06441902140824E-06	thylakoid lumen	CC
3-5meses	GO:0022804	4.25268652380628E-06	active transmembrane transporter activity	MF
3-5meses	GO:0044712	4.27729792924087E-06	NA	NA
3-5meses	GO:0009538	4.34228788933239E-06	photosystem I reaction center	CC
3-5meses	GO:0015977	4.49398430086486E-06	carbon fixation	BP
3-5meses	GO:0004611	4.50943282069255E-06	phosphoenolpyruvate carboxykinase activity	MF
3-5meses	GO:0009071	6.6043894777065E-06	serine family amino acid catabolic process	BP
3-5meses	GO:0016021	6.62269349132046E-06	integral component of membrane	CC
3-5meses	GO:0006631	7.91678531728625E-06	fatty acid metabolic process	BP
3-5meses	GO:0033013	8.9697259933381E-06	tetrapyrrole metabolic process	BP
3-5meses	GO:0044255	9.0544789850254E-06	cellular lipid metabolic process	BP
3-5meses	GO:0016829	1.00779530973637E-05	lyase activity	MF
3-5meses	GO:0005372	1.13848361316294E-05	water transmembrane transporter activity	MF
3-5meses	GO:0015250	1.13848361316294E-05	water channel activity	MF
3-5meses	GO:0006090	1.16321255505515E-05	pyruvate metabolic process	BP
3-5meses	GO:0005215	1.35778301156472E-05	transporter activity	MF
3-5meses	GO:0015254	1.59660458571474E-05	glycerol channel activity	MF
3-5meses	GO:0031226	2.23939356116687E-05	intrinsic component of plasma membrane	CC
3-5meses	GO:0009416	2.8624758819789E-05	response to light stimulus	BP
3-5meses	GO:0102336	2.92690660389739E-05	3-oxo-arachidoyl-CoA synthase activity	MF
3-5meses	GO:0102337	2.92690660389739E-05	3-oxo-cerotoyl-CoA synthase activity	MF
3-5meses	GO:0102338	2.92690660389739E-05	3-oxo-lignoceronoyl-CoA synthase activity	MF
3-5meses	GO:0042170	3.46074561547992E-05	plastid membrane	CC
3-5meses	GO:0044723	3.91775389286743E-05	NA	NA
3-5meses	GO:0009992	4.12570795600188E-05	cellular water homeostasis	BP
3-5meses	GO:0006779	4.30806865421502E-05	porphyrin-containing compound biosynthetic process	BP
3-5meses	GO:0009767	4.58915304938478E-05	photosynthetic electron transport chain	BP
3-5meses	GO:0033014	4.98037559465033E-05	tetrapyrrole biosynthetic process	BP
3-5meses	GO:0005506	4.98645863045407E-05	iron ion binding	MF
3-5meses	GO:0006073	4.98934994074047E-05	cellular glucan metabolic process	BP

Cuadro S6. Categorías funcionales enriquecidas en *F. pilosus*

Estadio	Categoría	Valor <i>p</i> sobrerrepresentado	Término	Ontología
3-5meses	GO:0071555	5.58390142622432E-05	cell wall organization	BP
3-5meses	GO:0006818	6.41454352446267E-05	hydrogen transport	BP
3-5meses	GO:0015992	6.41454352446267E-05	proton transport	BP
3-5meses	GO:0006884	6.44533192419863E-05	cell volume homeostasis	BP
3-5meses	GO:0015168	6.48229198640965E-05	glycerol transmembrane transporter activity	MF
3-5meses	GO:0044042	6.50131787390336E-05	glucan metabolic process	BP
3-5meses	GO:0047274	7.29124626301653E-05	galactinol-sucrose galactosyltransferase activity	MF
3-5meses	GO:0016820	7.55112504629303E-05	hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances	MF
3-5meses	GO:0044262	7.67529266832551E-05	cellular carbohydrate metabolic process	BP
3-5meses	GO:0009654	7.92070079534407E-05	photosystem II oxygen evolving complex	CC
3-5meses	GO:0031969	8.33382606981833E-05	chloroplast membrane	CC
3-5meses	GO:0016138	8.69651205564936E-05	glycoside biosynthetic process	BP
3-5meses	GO:0044724	9.94339038925669E-05	NA	NA

*Las categorías de Zona cambial de plantas adultas están enriquecidas con respecto a plantas de 3-5 meses de *F. pilosus*. Las categorías enriquecidas en plantas de 3-5 meses son con respecto a la zona cambial de plantas adultas

Cuadro S7. Factores transcripcionales más expresados en los distintos estadios de *E. platyacanthus*

Tejido	Factor transcripcional
Zona Cambial	Full=Putative Myb family transcription factor At1g14600
Zona Cambial	Full=DNA-binding protein SMUBP-2
Zona Cambial	Full=Dehydration-responsive element-binding protein 3
Zona Cambial	Full=Homeobox-leucine zipper protein ATHB-20
Zona Cambial	Full=TATA box-binding protein-associated factor RNA polymerase I subunit B
Zona Cambial	Full=MADS-box protein SVP {ECO:0000303 PubMed:10758486}
Zona Cambial	Full=MADS-box protein SVP {ECO:0000303 PubMed:10758486}
Zona Cambial	Full=Squamosa promoter-binding protein 1
Zona Cambial	Full=Agamous-like MADS-box protein AGL61
Zona Cambial	Full=Transcription factor bHLH51
Zona Cambial	Full=Protein SHORT-ROOT
Zona Cambial	Full=Cyclic dof factor 2
Zona Cambial	Full=Transcription factor bHLH25
Zona Cambial	Full=Heat stress transcription factor C-1
Zona Cambial	Full=Ethylene-responsive transcription factor ERF035
Zona Cambial	Full=Zinc finger BED domain-containing protein DAYSLEEPER
Zona Cambial	Full=Zinc finger BED domain-containing protein DAYSLEEPER
Zona Cambial	Full=Homeobox-leucine zipper protein ATHB-16
Zona Cambial	Full=Zinc finger protein 1 {ECO:0000305}
Zona Cambial	Full=MADS-box protein SOC1
Zona Cambial	Full=MADS-box protein SOC1
Zona Cambial	Full=Homeobox-leucine zipper protein ATHB-12
Zona Cambial	Full=Two-component response regulator ORR26 {ECO:0000305}
Zona Cambial	Full=Transcription factor bHLH96
Zona Cambial	Full=Zinc finger protein 3 {ECO:0000305}
Zona Cambial	Full=Zinc finger protein ZAT12
Zona Cambial	Full=Probable WRKY transcription factor 13
Zona Cambial	Full=Protein FAR1-RELATED SEQUENCE 7
Zona Cambial	Full=Dehydration-responsive element-binding protein 3
Zona Cambial	Full=Ethylene-responsive transcription factor 9
Zona Cambial	Full=Probable WRKY transcription factor 74
Zona Cambial	Full=MADS-box protein JOINTLESS
Zona Cambial	Full=Zinc finger protein ZAT10
Zona Cambial	Full=bZIP transcription factor 44 {ECO:0000305}
Zona Cambial	Full=Zinc finger protein CONSTANS-LIKE 2
Zona Cambial	Full=Transcription factor MYB86
Zona Cambial	Full=Zinc finger protein WIP2
Zona Cambial	Full=Zinc finger protein WIP3
Zona Cambial	Full=Zinc finger protein WIP2
Zona Cambial	Full=Ethylene-responsive transcription factor ERF114 {ECO:0000303 PubMed:16407444}
Zona Cambial	Full=NAC transcription factor 32 {ECO:0000305}
Zona Cambial	Full=NAC domain-containing protein 2
Zona Cambial	Full=AP2-like ethylene-responsive transcription factor ANT
Zona Cambial	Full=Transcription factor TGA9 {ECO:0000303 PubMed:20805327}
Zona Cambial	Full=Transcription factor bHLH68
Zona Cambial	Full=Transcription factor bHLH68
Zona Cambial	Full=Zinc finger protein 1 {ECO:0000305}
Zona Cambial	Full=Uncharacterized protein At4g00950

Continuación Cuadro S7. Factores transcripcionales con expresión diferencial en <i>E. platyacanthus</i>	
Zona Cambial	Full=Myb family transcription factor PHL5 {ECO:0000305}
Zona Cambial	Full=Cyclic dof factor 2
Zona Cambial	Full=Zinc finger protein JACKDAW
Zona Cambial	Full=Auxin response factor 17
Zona Cambial	Full=Probable WRKY transcription factor 57
Zona Cambial	Full=Chitin-inducible gibberellin-responsive protein 1
Zona Cambial	Full=Chitin-inducible gibberellin-responsive protein 1
Zona Cambial	Full=Zinc finger CCCH domain-containing protein 47
Zona Cambial	Full=Squamosa promoter-binding protein 1
Zona Cambial	Full=Ethylene-responsive transcription factor ERF008
Zona Cambial	Full=Trihelix transcription factor PTL
Zona Cambial	Full=Transcription factor MYBS3 {ECO:0000303 PubMed:12172034}
Zona Cambial	Full=Transcription factor bHLH68
Zona Cambial	Full=Transcription factor bHLH68
Zona Cambial	Full=Protein REVEILLE 1
Zona Cambial	Full=Ethylene-responsive transcription factor-like protein At4g13040
Zona Cambial	Full=Protein ENHANCED DOWNY MILDEW 2 {ECO:0000303 PubMed:23940361}
Zona Cambial	Full=Auxin response factor 7
Zona Cambial	Full=Nuclear transcription factor Y subunit B-7
Zona Cambial	Full=Transcription factor HHO5 {ECO:0000305}
Zona Cambial	Full=bZIP transcription factor 44 {ECO:0000305}
Zona Cambial	Full=Ethylene-responsive transcription factor CRF4
Zona Cambial	Full=Trihelix transcription factor ASIL2 {ECO:0000305}
Zona Cambial	Full=Myb-related protein 1 {ECO:0000303 PubMed:12008900}
Zona Cambial	Full=Myb-related protein 2 {ECO:0000303 PubMed:21255164}
Zona Cambial	Full=Heat shock factor protein HSF30
Zona Cambial	Full=Ethylene-responsive transcription factor CRF2
Zona Cambial	Full=Transcription factor MYB124 {ECO:0000303 PubMed:11597504}
Zona Cambial	Full=Transcription factor MYB124 {ECO:0000303 PubMed:11597504}
Zona Cambial	Full=MADS-box transcription factor 27
Zona Cambial	Full=Zinc finger protein CONSTANS-LIKE 14
Zona Cambial	Full=Transcription factor bHLH130
Zona Cambial	Full=Myb family transcription factor EFM {ECO:0000303 PubMed:25132385}
Zona Cambial	Full=AP2-like ethylene-responsive transcription factor At2g41710
Zona Cambial	Full=AP2-like ethylene-responsive transcription factor At2g41710
Zona Cambial	Full=Transcription factor bHLH121
Zona Cambial	Full=Heat stress transcription factor B-4
Zona Cambial	Full=Scarecrow-like protein 14
Zona Cambial	Full=Heat shock factor protein HSF30
Zona Cambial	Full=Transcription factor HHO5 {ECO:0000305}
Zona Cambial	Full=Two-component response regulator ORR26 {ECO:0000305}
Zona Cambial	Full=Common plant regulatory factor 1
Zona Cambial	Full=Heat stress transcription factor B-2a
Zona Cambial	Full=Transcription factor MYC2
Zona Cambial	Full=Myb family transcription factor PHL7 {ECO:0000305}
Zona Cambial	Full=Ethylene-responsive transcription factor CRF6
Zona Cambial	Full=Two-component response regulator-like PRR1
Zona Cambial	Full=Auxin response factor 8
Zona Cambial	Full=CRC domain-containing protein TSO1

Continuación Cuadro S7. Factores transcripcionales con expresión diferencial en <i>E. platyacanthus</i>	
Zona Cambial	Full=Homeobox-leucine zipper protein HAT22
Zona Cambial	Full=Protein ENHANCED DOWNY MILDEW 2 {ECO:0000303 PubMed:23940361}
Zona Cambial	Full=Nucleolar and coiled-body phosphoprotein 1 {ECO:0000305}
Zona Cambial	Full=Homeobox protein knotted-1-like 6
Zona Cambial	Full=Probable WRKY transcription factor 19
Zona Cambial	Full=Probable WRKY transcription factor 19
Zona Cambial	Full=Probable WRKY transcription factor 19
Zona Cambial	Full=Zinc finger CCCH domain-containing protein 20
Zona Cambial	Full=Transcription factor MYC2
Zona Cambial	Full=Zinc finger CCCH domain-containing protein 66
Zona Cambial	Full=Zinc finger BED domain-containing protein RICESLEEPER 2
Zona Cambial	Full=Zinc finger BED domain-containing protein RICESLEEPER 2
Zona Cambial	Full=Ethylene-responsive transcription factor ERF098
Zona Cambial	Full=Zinc finger BED domain-containing protein DAYSLEEPER
Zona Cambial	Full=Auxin response factor 7
Zona Cambial	Full=Protein ETHYLENE INSENSITIVE 3
Zona Cambial	Full=Ethylene-responsive transcription factor ERF017
Zona Cambial	Full=Ethylene-responsive transcription factor ERF109
1 año 5 meses	Full=MADS-box protein SVP {ECO:0000303 PubMed:10758486}
1 año 5 meses	Full=bZIP transcription factor 2 {ECO:0000305}
1 año 5 meses	Full=Zinc finger protein 1 {ECO:0000305}
1 año 5 meses	Full=MADS-box protein SOC1
1 año 5 meses	Full=Probable transcription factor At5g61620 {ECO:0000305}
1 año 5 meses	Full=Homeobox-leucine zipper protein ATHB-12
1 año 5 meses	Full=Zinc finger protein WIP3
1 año 5 meses	Full=NAC transcription factor 25
1 año 5 meses	Full=Nuclear transcription factor Y subunit A-7
1 año 5 meses	Full=Auxin response factor 9
1 año 5 meses	Full=Zinc finger BED domain-containing protein RICESLEEPER 2
1 año 5 meses	Full=Auxin response factor 7
1 año 5 meses	Full=bZIP transcription factor 44 {ECO:0000305}
1 año 5 meses	Full=Chromatin remodeling protein SHL {ECO:0000305}
1 año 5 meses	Full=Protein NLP2
1 año 5 meses	Full=Homeobox-leucine zipper protein HDG11
1 año 5 meses	Full=Zinc finger BED domain-containing protein DAYSLEEPER
1 año 5 meses	Full=Auxin response factor 7
1 año 5 meses	Full=Transcription factor TCP15
1 mes 15 días	Full=Ethylene-responsive transcription factor SHINE 2
1 mes 15 días	Full=WRKY transcription factor 22
1 mes 15 días	Full=Ethylene-responsive transcription factor ERF023
1 mes 15 días	Full=Protein FAR-RED IMPAIRED RESPONSE 1
1 mes 15 días	Full=Auxin-responsive protein IAA7
1 mes 15 días	Full=Putative GATA transcription factor 22
1 mes 15 días	Full=Homeobox-leucine zipper protein PROTODERMAL FACTOR 2
1 mes 15 días	Full=Transcription factor MYB30 {ECO:0000303 PubMed:10929106}
1 mes 15 días	Full=Basic leucine zipper 34 {ECO:0000303 PubMed:11906833}
1 mes 15 días	Full=Floral homeotic protein DEFICIENS
1 mes 15 días	Full=Probable WRKY transcription factor 41
1 mes 15 días	Full=Zinc finger protein CONSTANS-LIKE 16

Continuación Cuadro S7. Factores transcripcionales con expresión diferencial en *E. platyacanthus*

1 mes 15 días	Full=Homeobox-leucine zipper protein PROTODERMAL FACTOR 2
1 mes 15 días	Full=Homeobox-leucine zipper protein ANTHOCYANINLESS 2
1 mes 15 días	Full=Homeobox-leucine zipper protein ANTHOCYANINLESS 2
1 mes 15 días	Full=Calmodulin-binding transcription activator 6
1 mes 15 días	Full=Transcription factor UNE10
1 mes 15 días	Full=Heat shock factor protein HSF30
1 mes 15 días	Full=Transcription factor bHLH60
1 mes 15 días	Full=Transcription factor bHLH148
1 mes 15 días	Full=Light-inducible protein CPRF2
1 mes 15 días	Full=Calmodulin-binding transcription activator 5
1 mes 15 días	Full=RNA polymerase sigma factor sigB
1 mes 15 días	Full=Transcription factor RF2b
1 mes 15 días	Full=Protein indeterminate-domain 6, chloroplastic {ECO:0000303 PubMed:16784536}

Cuadro S8. Factores transcripcionales con expresión diferencial en *F. pilosus*

Tejido	Factor transcripcional
Zona Cambial	Full=Common plant regulatory factor 1
Zona Cambial	Full=Basic leucine zipper 43 {ECO:0000303 PubMed:11906833}
Zona Cambial	Full=G-box-binding factor 4
Zona Cambial	Full=Ethylene-responsive transcription factor ERF114 {ECO:0000303 PubMed:16407444}
Zona Cambial	Full=Trihelix transcription factor ASR3 {ECO:0000303 PubMed:25770109}
Zona Cambial	Full=Trihelix transcription factor ASIL2 {ECO:0000305}
Zona Cambial	Full=Probable WRKY transcription factor 50
Zona Cambial	Full=Myb family transcription factor PHL11 {ECO:0000305}
Zona Cambial	Full=Transcription factor ILR3
Zona Cambial	Full=Transcription factor bHLH47
Zona Cambial	Full=NAC domain-containing protein 41 {ECO:0000303 PubMed:15029955}
Zona Cambial	Full=Transcription factor bHLH147
Zona Cambial	Full=Transcription factor bHLH13
Zona Cambial	Full=Zinc finger protein 3 {ECO:0000305}
Zona Cambial	Full=Transcription factor bHLH80
Zona Cambial	Full=Transcription factor TGA4
Zona Cambial	Full=Transcription factor TGA4
Zona Cambial	Full=B-box zinc finger protein 21 {ECO:0000303 PubMed:19920209}
Zona Cambial	Full=B-box zinc finger protein 21 {ECO:0000303 PubMed:19920209}
Zona Cambial	Full=Ethylene-responsive transcription factor 5
Zona Cambial	Full=B-box zinc finger protein 32 {ECO:0000303 PubMed:19920209}
Zona Cambial	Full=Transcription factor bHLH66
Zona Cambial	Full=Protein REVEILLE 6
Zona Cambial	Full=Transcription factor MYBC1 {ECO:0000305}
Zona Cambial	Full=Probable WRKY transcription factor 70
Zona Cambial	Full=GATA transcription factor 5
Zona Cambial	Full=Transcription factor bHLH68
Zona Cambial	Full=Zinc finger protein CONSTANS-LIKE 15
Zona Cambial	Full=Zinc finger protein CONSTANS-LIKE 15
Zona Cambial	Full=Transcription factor HHO5 {ECO:0000305}
Zona Cambial	Full=Transcription factor HHO5 {ECO:0000305}
Zona Cambial	Full=GATA transcription factor 1
Zona Cambial	Full=Uncharacterized protein At4g00950
Zona Cambial	Full=Homeobox-leucine zipper protein HAT5
Zona Cambial	Full=Cyclic dof factor 3
Zona Cambial	Full=Cyclic dof factor 3
Zona Cambial	Full=Transcription factor HHO3 {ECO:0000305}
Zona Cambial	Full=Ethylene-responsive transcription factor ERF071
Zona Cambial	Full=Ethylene-responsive transcription factor ERF071
Zona Cambial	Full=Probable WRKY transcription factor 32
Zona Cambial	Full=Ethylene-responsive transcription factor CRF4
Zona Cambial	Full=NAC domain-containing protein 2
Zona Cambial	Full=Ethylene-responsive transcription factor ERF010
Zona Cambial	Full=Dof zinc finger protein PBF
Zona Cambial	Full=Transcription factor LUX
Zona Cambial	Full=Dehydration-responsive element-binding protein 2C
Zona Cambial	Full=Zinc finger protein CONSTANS-LIKE 13
Zona Cambial	Full=Zinc finger protein ZAT10

Cuadro S9. Factores transcripcionales con expresión diferencial en *F. pilosus*

Zona Cambial	Full=Heat stress transcription factor C-1
Zona Cambial	Full=Protein NLP1
Zona Cambial	Full=B-box domain protein 31 {ECO:0000303 PubMed:19920209}
Zona Cambial	Full=Scarecrow-like protein 4
Zona Cambial	Full=Transcription factor MYB102 {ECO:0000305}
Zona Cambial	Full=Scarecrow-like protein 15
Zona Cambial	Full=Probable WRKY transcription factor 7
Zona Cambial	Full=Probable WRKY transcription factor 63
Zona Cambial	Full=Scarecrow-like protein 30
Zona Cambial	Full=Scarecrow-like protein 14
Zona Cambial	Full=Trihelix transcription factor GT-3b
Zona Cambial	Full=Nuclear transcription factor Y subunit A-7
Zona Cambial	Full=bZIP transcription factor 44 {ECO:0000305}
Zona Cambial	Full=Calmodulin-binding transcription activator 4
Zona Cambial	Full=Transcription factor TGA3
Zona Cambial	Full=Ethylene-responsive transcription factor CRF2
Zona Cambial	Full=Ethylene-responsive transcription factor CRF2
Zona Cambial	Full=Probable WRKY transcription factor 3
Zona Cambial	Full=Transcription factor bHLH155
Zona Cambial	Full=Trihelix transcription factor GT-2
Zona Cambial	Full=Trihelix transcription factor GT-2
Zona Cambial	Full=Homeobox-leucine zipper protein HAT5
Zona Cambial	Full=Transcription factor bHLH123
Zona Cambial	Full=Protein REVEILLE 6
Zona Cambial	Full=Myb-like protein L
Zona Cambial	Full=Scarecrow-like protein 9
Zona Cambial	Full=Squamosa promoter-binding-like protein 12
Zona Cambial	Full=Protein REVEILLE 6
Zona Cambial	Full=Protein NLP2
Zona Cambial	Full=Protein NLP2
Zona Cambial	Full=Myb family transcription factor EFM {ECO:0000303 PubMed:25132385}
Zona Cambial	Full=Probable WRKY transcription factor 7
Zona Cambial	Full=MADS-box protein SVP {ECO:0000303 PubMed:10758486}
Zona Cambial	Full=Transcription factor PIF3
Zona Cambial	Full=Histone-lysine N-methyltransferase EZ3
Zona Cambial	Full=Histone-lysine N-methyltransferase EZ3
Zona Cambial	Full=Squamosa promoter-binding protein 1
Zona Cambial	Full=Transcription factor bHLH25
Zona Cambial	Full=Basic leucine zipper 43 {ECO:0000303 PubMed:11906833}
3-5 meses	Full=MADS-box protein CMB2
3-5 meses	Full=Transcription factor MYB61 {ECO:0000305}
3-5 meses	Full=Basic leucine zipper 34 {ECO:0000303 PubMed:11906833}
3-5 meses	Full=NAC domain-containing protein 100
3-5 meses	Full=NAC domain-containing protein 100
3-5 meses	Full=Homeobox protein ATH1
3-5 meses	Full=Homeobox-leucine zipper protein ANTHOCYANINLESS 2
3-5 meses	Full=Homeobox-leucine zipper protein ANTHOCYANINLESS 2
3-5 meses	Full=Transcription factor MYB124 {ECO:0000303 PubMed:11597504}
3-5 meses	Full=Zinc finger BED domain-containing protein DAYSLEEPER

Continuación. Cuadro S9. Factores transcripcionales con expresión diferencial en *F. pilosus*

3-5 meses	Full=Protein RADIALIS-like 6
3-5 meses	Full=Transcription factor bHLH148
3-5 meses	Full=Transcription factor MYB124 {ECO:0000303 PubMed:11597504}
3-5 meses	Full=Trihelix transcription factor ASIL2 {ECO:0000305}
3-5 meses	Full=Transcription factor bHLH61
3-5 meses	Full=CRC domain-containing protein TSO1
3-5 meses	Full=Putative AC transposase
3-5 meses	Full=Zinc finger protein MAGPIE {ECO:0000303 PubMed:17785527}
3-5 meses	Full=Ethylene-responsive transcription factor SHINE 3
3-5 meses	Full=Transcription factor CPC
3-5 meses	Full=AP2/ERF and B3 domain-containing transcription repressor RAV2
3-5 meses	Full=Transcription factor PIF7
3-5 meses	Full=RNA polymerase sigma factor sigB
3-5 meses	Full=Zinc finger protein CONSTANS-LIKE 6
3-5 meses	Full=bZIP transcription factor 50 {ECO:0000303 PubMed:18065552}
3-5 meses	Full=Scarecrow-like protein 3
3-5 meses	Full=High mobility group B protein 15
3-5 meses	Full=Protein RADIALIS-like 6
3-5 meses	Full=High mobility group B protein 7
3-5 meses	Full=Zinc finger BED domain-containing protein RICESLEEPER 2
3-5 meses	Full=Zinc finger BED domain-containing protein RICESLEEPER 3
3-5 meses	Full=Transcription factor bHLH67
3-5 meses	Full=Transcription factor bHLH147
3-5 meses	Full=Auxin-responsive protein IAA4
3-5 meses	Full=Homeobox-leucine zipper protein HAT5
3-5 meses	Full=Zinc finger BED domain-containing protein DAYSLEEPER
3-5 meses	Full=Scarecrow-like protein 28
3-5 meses	Full=Zinc finger BED domain-containing protein RICESLEEPER 2
3-5 meses	Full=Squamosa promoter-binding-like protein 2
3-5 meses	Full=Squamosa promoter-binding-like protein 6
3-5 meses	Full=Transcription factor MYBC1 {ECO:0000305}
3-5 meses	Full=Basic leucine zipper 34 {ECO:0000303 PubMed:11906833}
3-5 meses	Full=Transcription factor BPE
3-5 meses	Full=Putative GATA transcription factor 22
3-5 meses	Full=Transcription factor SCREAM2
3-5 meses	Full=Transcription factor bHLH93
3-5 meses	Full=Protein FAR1-RELATED SEQUENCE 12
3-5 meses	Full=Protein FAR1-RELATED SEQUENCE 12
3-5 meses	Full=Protein FAR1-RELATED SEQUENCE 12
3-5 meses	Full=Transcription factor BIM2
3-5 meses	Full=Light-harvesting complex-like protein 3 isotype 1, chloroplastic {ECO:0000305}
3-5 meses	Full=Putative AC transposase
3-5 meses	Full=Zinc finger protein CONSTANS-LIKE 16
3-5 meses	Full=Protein indeterminate-domain 5, chloroplastic {ECO:0000303 PubMed:16784536}
3-5 meses	Full=Protein indeterminate-domain 5, chloroplastic {ECO:0000303 PubMed:16784536}
3-5 meses	Full=MADS-box protein SOC1
3-5 meses	Full=CRC domain-containing protein TSO1
3-5 meses	Full=Zinc finger BED domain-containing protein RICESLEEPER 2
3-5 meses	Full=Nuclear transcription factor Y subunit A-2

Continuación. Cuadro S9. Factores transcripcionales con expresión diferencial en *F. pilosus*

3-5 meses	Full=Transcription factor UNE10
3-5 meses	Full=GATA transcription factor 15
3-5 meses	Full=Transcription factor bHLH62
3-5 meses	Full=Homeobox-leucine zipper protein ATHB-7
3-5 meses	Full=Transcription factor UNE10
3-5 meses	Full=Ethylene-responsive transcription factor ERF023
3-5 meses	Full=Myb family transcription factor PHL13 {ECO:0000305}
3-5 meses	Full=Transcription factor bHLH25
3-5 meses	Full=Homeobox-leucine zipper protein PROTODERMAL FACTOR 2
3-5 meses	Full=Homeobox-leucine zipper protein PROTODERMAL FACTOR 2
3-5 meses	Full=Transcription factor MYB30 {ECO:0000303 PubMed:10929106}
3-5 meses	Full=Auxin response factor 7
3-5 meses	Full=Auxin response factor 7
3-5 meses	Full=Common plant regulatory factor 1
3-5 meses	Full=Zinc finger protein CONSTANS-LIKE 2
3-5 meses	Full=Zinc finger protein CONSTANS-LIKE 2
3-5 meses	Full=RNA polymerase sigma factor sigE, chloroplastic/mitochondrial
3-5 meses	Full=Homeobox protein BEL1 homolog
3-5 meses	Full=Ethylene-responsive transcription factor RAP2-7
3-5 meses	Full=Protein NLP4
3-5 meses	Full=Calmodulin-binding transcription activator 5
3-5 meses	Full=BEL1-like homeodomain protein 1
3-5 meses	Full=Auxin response factor 19
3-5 meses	Full=Transcription factor TCP2 {ECO:0000303 PubMed:10363373}
3-5 meses	Full=Transcription factor bHLH144
3-5 meses	Full=Putative AC transposase

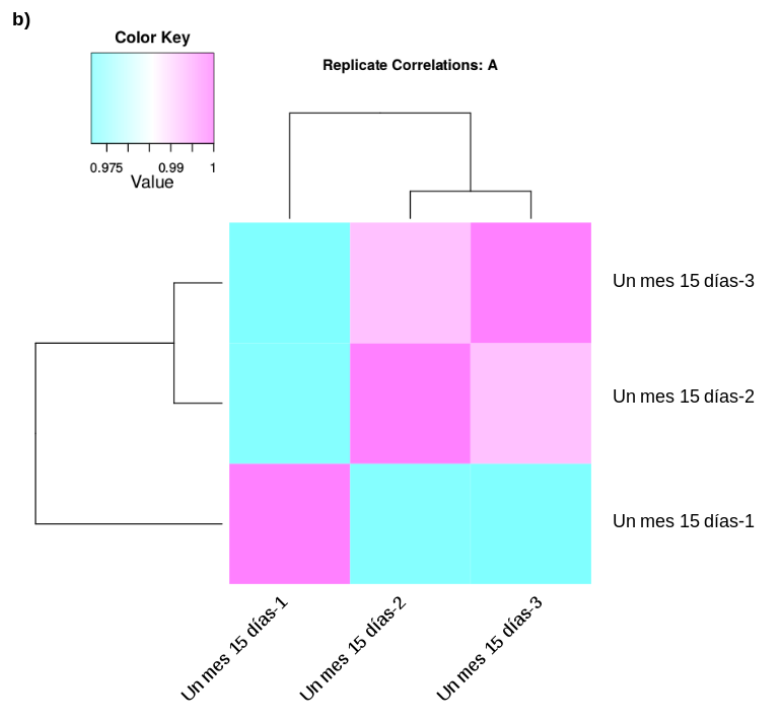
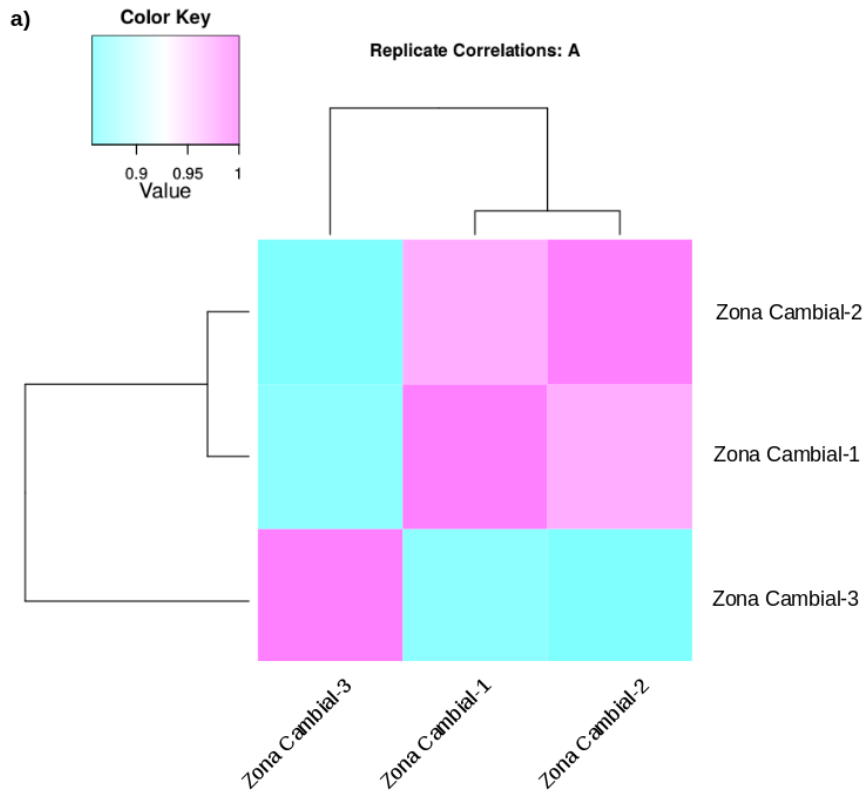


Figura S1. Correlaciones de Pearson entre réplicas biológicas de la misma condición. a) Zona cambial de *A. retusus*. b) Plántulas de *E. platyacanthus*.

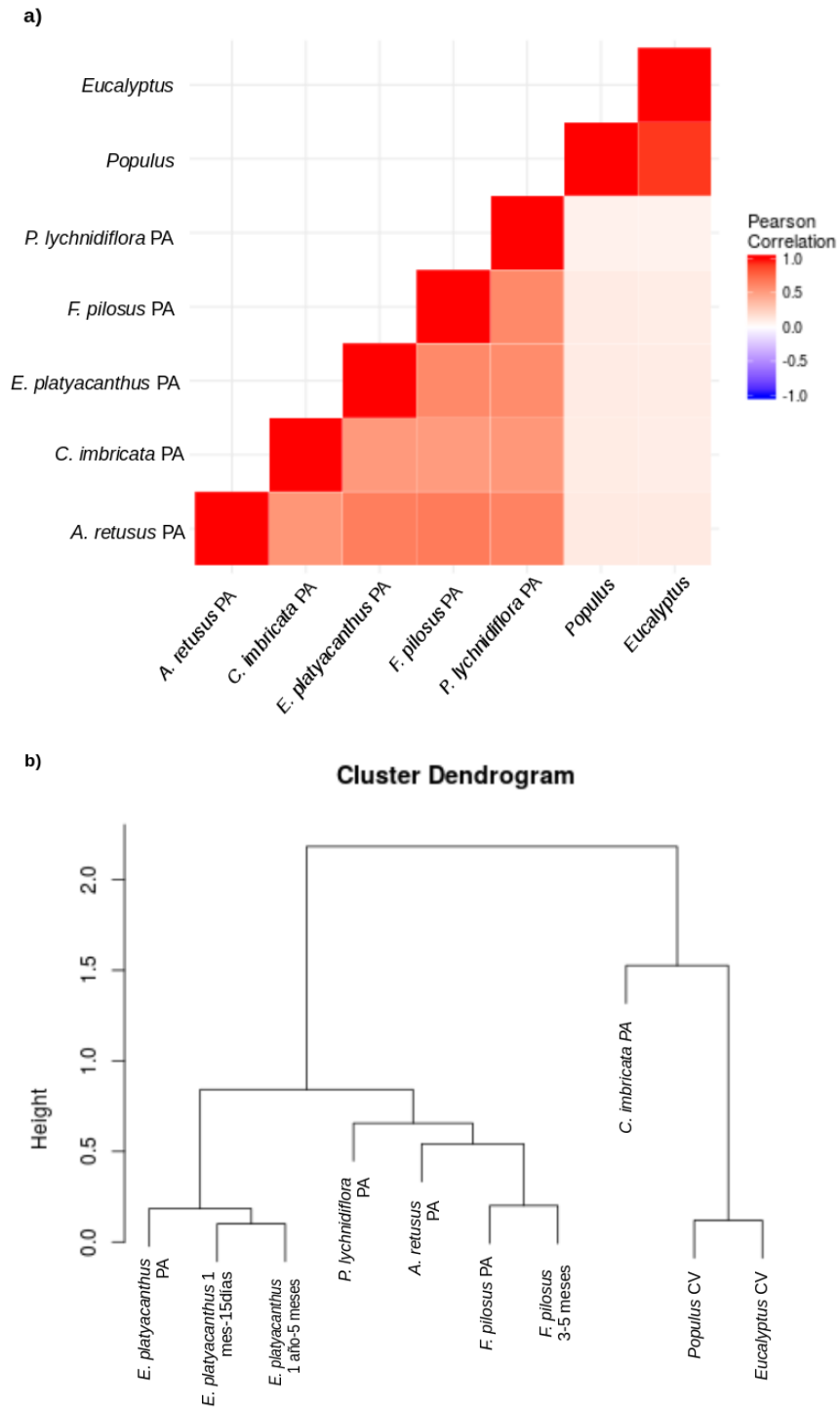


Figura S2. Correlaciones de Pearson considerando las categorías funcionales de los transcritos de Zona Cambial en las especies de cactus y en *Populus* y *Eucalyptus*. a) Mapa de calor de las correlaciones de Pearson y b) dendrograma de las mismas correlaciones.

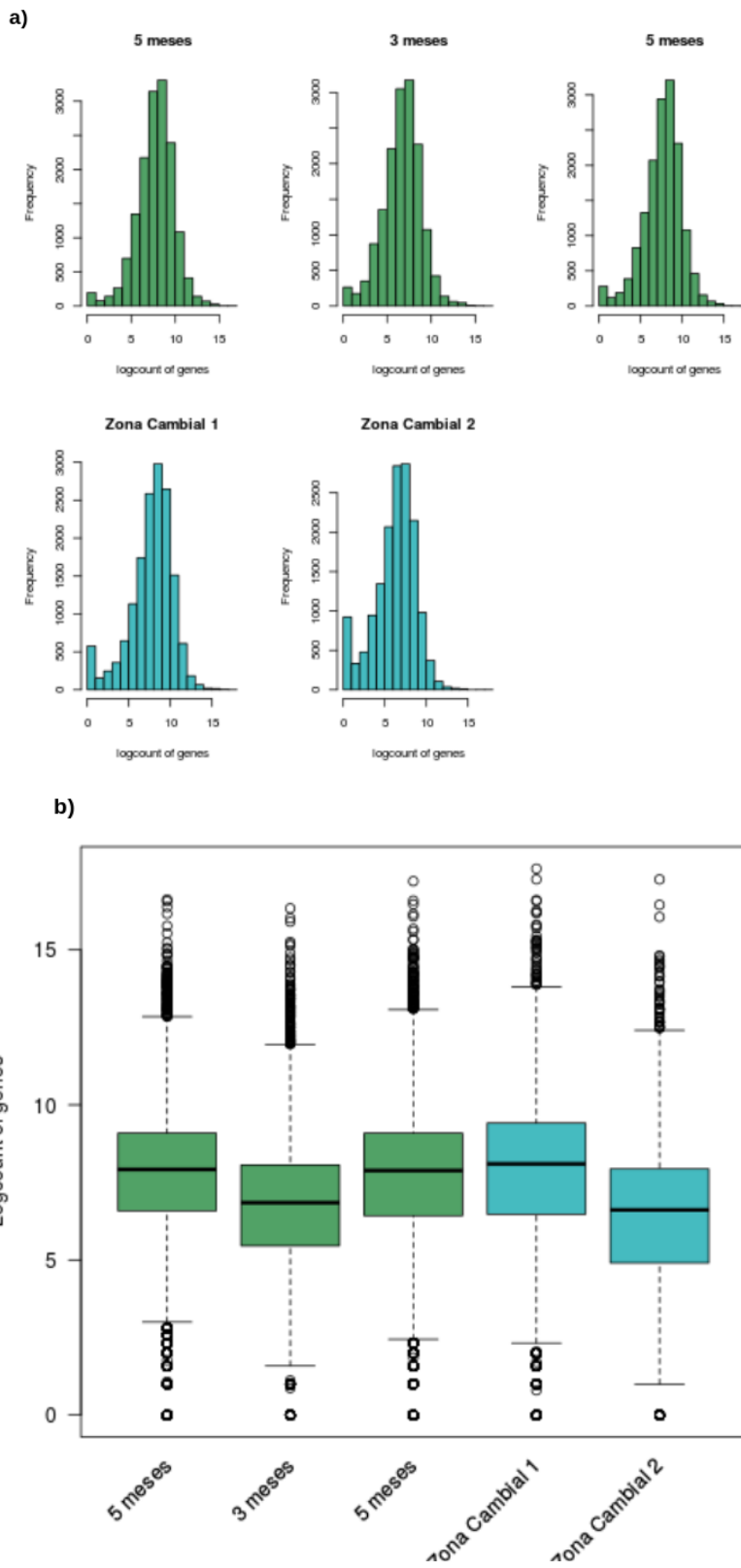


Figura S3. Distribución de los conteos filtrados por muestra en *F. pilosus*. a) histogramas y b) gráfica de cajas mostrando los conteos en escala logarítmica.

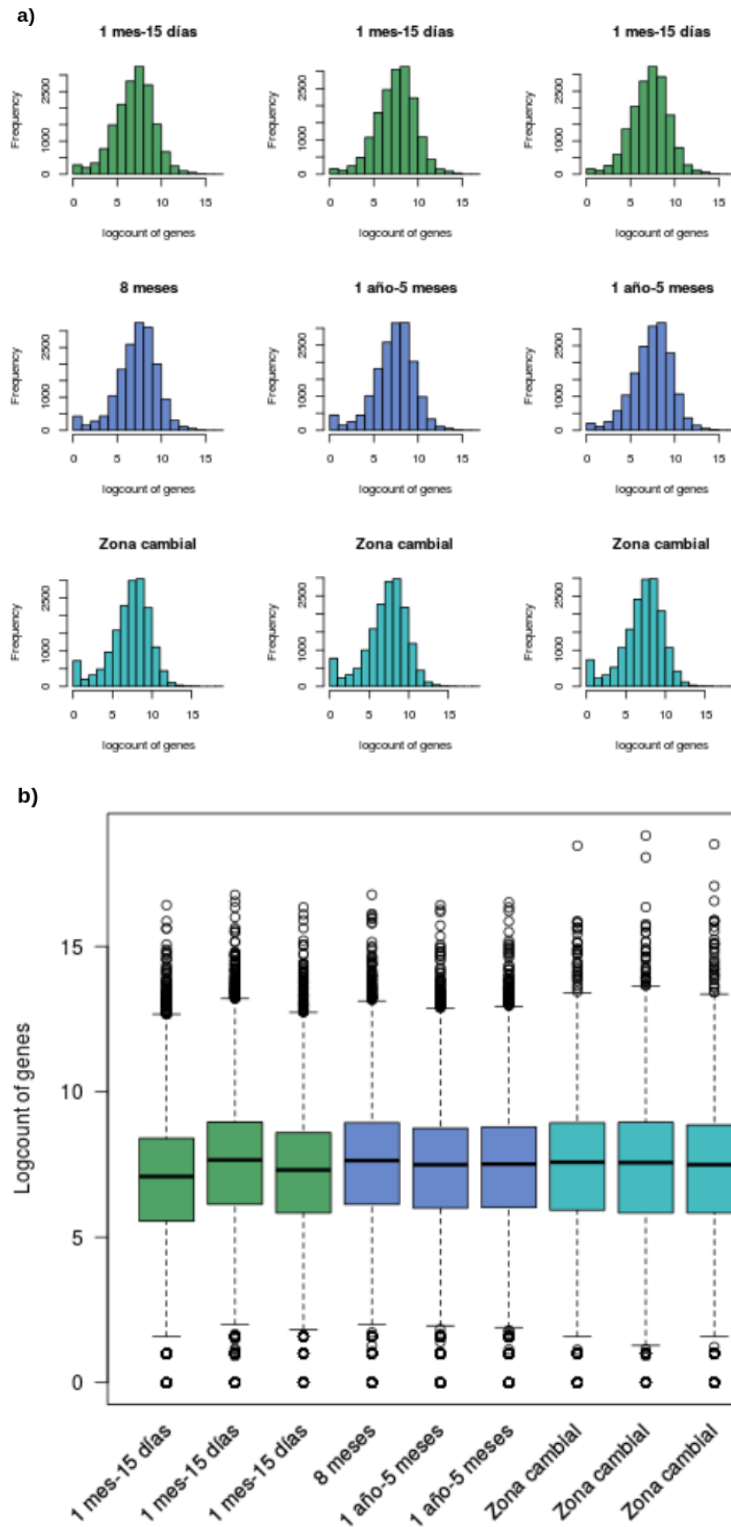


Figura S4. Distribución de los conteos filtrados por muestra en *E. platyacanthus* a) histogramas y b) gráfica de cajas mostrando los conteos en escala logarítmica.

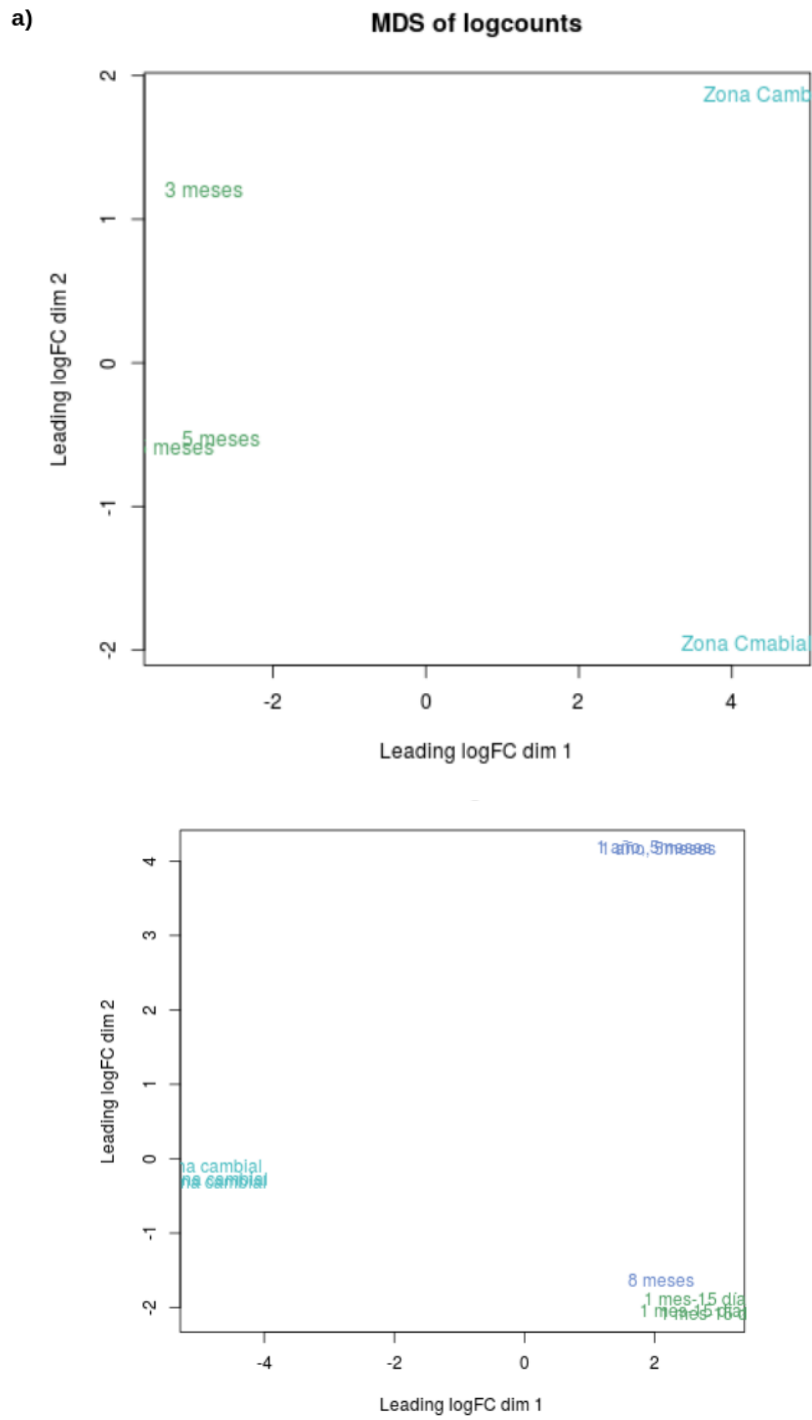


Figura S5. Gráfica MDS de los conteos logarítmicos en a) *F. pilosus* y en b) *E. platyacanthus*.