

UNIVERSIDAD NACIONAL DEL LITORAL

Facultad de Bioquímica y Ciencias Biológicas



Tesis para la obtención del Grado Académico de Doctor en Ciencias Biológicas

“EVOLUCIÓN DE LA FOTOSÍNTESIS: DINÁMICA DEL DESARROLLO FOLIAR DE GRAMÍNEAS”

Lic. Santiago Prochetto

Directora de Tesis: Dr. Renata Reinheimer

Co-director de Tesis: Dr. Anthony J. Studer

Laboratorio de Evolución del Desarrollo (LED)

Instituto de Agrobiotecnología del Litoral (UNL-CONICET)

2021

AGRADECIMIENTOS

En primer lugar, quiero expresar mi agradecimiento a mis directores, que confiaron en mí y me acompañaron en este proyecto. En especial le agradezco a Renata por todo el apoyo y los consejos recibidos durante estos cinco años.

En segundo lugar, a CONICET, AGENCIA, La Fundación Fulbright y el Ministerio de Educación que a través de su financiamiento hicieron posible este trabajo.

A las personas que colaboraron directamente con el trabajo realizado. A Juan Acosta por brindarnos las semillas de las especies utilizadas, a Marcela Dotto por su asesoramiento y discusión sobre los resultados de transcriptómica, a Facundo Romani por lo propio con los resultados de filogenética, y a Virginia Sánchez Puerta por sus comentarios y acompañamiento como tutora de esta tesis.

A mis compañeros de laboratorio y a la comunidad del Instituto de Agrobiotecnología del Litoral por el acompañamiento diario. En especial a Federico Ariel y sus generosas cartas de recomendación.

A mis amigos y mi familia que compartieron estos años conmigo, especialmente a Luciano que me sostuvo en los momentos en que mis fuerzas flaquearon.

Finalmente, quiero agradecer a todas las personas que luchan por una educación y ciencia al servicio del desarrollo humano y la liberación de los pueblos. En especial a les docentes, estudiantes y no docentes que entienden a la educación pública como un derecho para todas las personas y, por lo tanto, siguen luchando para que la misma sea laica y gratuita en todos sus niveles.

PUBLICACIONES

Prochetto, S., Reinheimer, R. (2020) Step by step evolution of Indeterminate Domain (IDD) transcriptional regulators: from algae to angiosperms. *Annals of Botany*. doi:10.1093/aob/mcaa052.

ÍNDICE

UNIVERSIDAD NACIONAL DEL LITORAL	1
AGRADECIMIENTOS	2
PUBLICACIONES	3
ÍNDICE	2
ABREVIATURAS Y SÍMBOLOS	6
RESUMEN	11
ABSTRACT	13
INTRODUCCIÓN	15
Fotosíntesis C ₃ , C ₂ y C ₄	15
Reacciones dependientes de la luz	15
Reacciones independientes de la luz: RUBISCO y el ciclo de Calvin-Benson-Bassham	16
La fotosíntesis C ₄	18
Anatomía Kranz	20
Subtipos de fotosíntesis C ₄ : NADP-ME, NAD-ME y PEPCK	21
Origen y evolución de la fotosíntesis C ₄	23
Cambio climático y origen de la fotosíntesis C ₄	23
Modelo de evolución	24
Especies C ₃ cercanas a linajes C ₄	25
Especies C ₃ Proto-Kranz (PK)	26
Especies C ₂	26
Especies C ₄ -like	27
OBJETIVOS	29
Objetivo general	29
Objetivos específicos	29
CAPÍTULO 1	30
Evolución molecular de las subfamilias de reguladores transcripcionales SHR, SCR e IDD	30
INTRODUCCIÓN	31
Familia GRAS	31
Subfamilias SHR y SCR	32
Subfamilia IDD	33
OBJETIVOS ESPECÍFICOS	35

MATERIALES Y MÉTODOS	35
Obtención y filtrado de secuencias	35
Reconstrucción filogenética	36
Análisis de motivos conservados	38
RESULTADOS	38
Evolución de la familia GRAS	38
Evolución de la subfamilia SHR	38
Evolución de la subfamilia SCR	41
Evolución de la subfamilia IDD	44
Evolución en las angiospermas	44
Análisis en la división Streptophyta	45
Origen de la subfamilia IDD	47
Diversificación de motivos conservados	49
Estudio de los linajes	49
DISCUSIÓN	54
Evolución de las subfamilias SHR y SCR	54
Evolución de la subfamilia IDD	55
CAPÍTULO 2	59
Estudio de la anatomía foliar y la fisiología en Otachyriinae	59
INTRODUCCIÓN	60
La modificación de la anatomía foliar como uno de los primeros pasos en la evolución de la fotosíntesis C ₄	60
La subtribu Otachyriinae como modelo de estudio de la evolución de la fotosíntesis C ₄	61
OBJETIVOS ESPECÍFICOS	63
MATERIALES Y MÉTODOS	64
Selección de especies	64
Estudio fisiológico	64
Condiciones de cultivo y toma de muestra	64
Preparación y análisis de las muestras para microscopía óptica	65
Análisis estadístico	65
RESULTADOS	65
Aspectos cualitativos	65
Estudio de rasgos fenotípicos asociados al desarrollo	68

Parámetros del desarrollo foliar asociados a los haces vasculares	68
Parámetros del desarrollo foliar asociados a las células de la vaina y del mesófilo	69
Correlación entre parámetros del desarrollo foliar	71
Comparación de rasgos fenotípico entre especies	72
Fisiología de las especies estudiadas	75
DISCUSIÓN	77
Existen diferencias estructurales en la anatomía foliar desde el inicio del gradiente de desarrollo	77
La comunicación entre CV y M como indicador del grado “Kranz” de una especie	78
CAPÍTULO 3	81
Análisis de la transcriptómica del gradiente foliar en Otachyriinae	81
INTRODUCCIÓN	82
Ensamblado de transcriptomas de novo y el gradiente de maduración	82
OBJETIVOS ESPECÍFICOS	83
MATERIALES Y MÉTODOS	84
Condiciones de cultivo y toma de muestras	84
Extracción de ARN y secuenciación	84
Ensamblaje de novo de transcriptomas	84
Anotación de genes	85
Análisis de expresión diferencial	85
Análisis de enriquecimiento de términos de Gene Ontology (GO)	86
Análisis de redes de co-expresión de genes ponderado (WGCNA)	86
RESULTADOS	87
Obtención de los transcriptomas	87
Filtrado y control de calidad	87
Anotación funcional	89
Expresión diferencial de genes	91
La dinámica transcripcional durante el desarrollo	91
Las diferencias entre subtipos fotosintéticos	92
Análisis de GO para los OED entre las especies	94
Ortogrupos ortogrupos involucrados en procesos fotosintéticos	95
Ortogrupos involucrados en la formación de la vasculatura y expansión de la pared celular	97
Patrones de expresión de genes	98

Marcadores de zonas fuente/sumidero	98
Expresión de enzimas involucradas en el ciclo C ₄	99
Expresión de genes involucrados en la fotorrespiración	101
Expresión de reguladores transcripcionales SHR y SCR	102
Expresión de reguladores transcripcionales IDD	104
Redes de co-expresión (WGCNA)	105
Identificación de módulos de ortogrupos co-expresantes	105
Co-expresión de los reguladores transcripcionales IDD, SHR y SCR	108
DISCUSIÓN	111
La calidad de los transcriptomas	111
Diferentes dinámicas entre las especies	112
Conservación de patrones de expresión	113
CONCLUSIONES GENERALES	117
BIBLIOGRAFÍA	119
Anexo 1: Tablas, figuras e información suplementaria	135

ABREVIATURAS Y SÍMBOLOS

A: parámetro de actividad fotosintética o asimilación de CO₂

aa: Aminoácidos

ADN: Ácido desoxirribonucleico

ALA-AT: Alanina aminotransferasa

APX: Ascorbato peroxidasa

ASP-AT5: Aspartato aminotransferasa

ARN: Ácido ribonucleico

ARNm: Ácido ribonucleico mensajero

ATP: Adenosina trifosfato

CA: Anhidrasa carbónica

CAM: Metabolismo ácido de las crasuláceas

CAT1-2: catalasa

CCBB: Ciclo de Calvin-Benson-Bassham

Ci: Concentración de CO₂ en el espacio intercelular de la hoja

CycD4: Ciclina D4

CycD6: Ciclina D6

CNDa: Condor a

CNDb: Condor b

cpm: lecturas por millón (del inglés: *Counts per million*)

CV: Células de la vaina

CVE: Células de la vaina externa

CVI: Células de la vaina interna

C2H2: Motivo *zinc finger* de tipo Cys2-His2

EM: *Eigengen* del módulo

EMSA: Ensayo de cambio en la corrida electroforética (del inglés: *Electrophoretic Mobility-Shift Assay*)

Fd: Ferredoxina

GAF1: GAI Interacting Factor 1

GAI: GIBBERELLIN-INSENSITIVE

GDC: Complejo glicina descarboxilasa

GED: Genes expresados diferencialmente

GGT1: Glutamato-glioxilato amino transferasa

GOX: Glicolato oxidasa

GRN: Redes de regulación génicas (del inglés: *Gene Regulatory Networks*)

G3P: Gliceraldehído 3-fosfato

HAM: Hairy Meristem

HGC: Motivos de cobertura global alta (del inglés: *High Global Coverage Motifs*)

HPR1: Hidroxipiruvato reductasa

IDD: Dominio Indeterminado

IDD-i: Ortogrupo IDD indeterminado

ID1: Indeterminate 1

JKD: Jackdaw

k_{im} : Conectividad intramodular

k_{Total} : Conectividad total

LGC: Motivos de cobertura baja (del inglés: *Low Global Coverage Motifs*)

LISCL: Liliun Scarecrow-like

LogFC: Tasa de cambio logarítmica (del inglés: *logarithmic fold change*)

LS: Lateral Supresor

M: Células del mesófilo

MDH: Malato deshidrogenasa

ME: Enzima málica

MGC: Motivos de cobertura global media (del inglés: *Medium Global Coverage Motifs*)

MGP: Magpie

MM: Membresía al módulo (del inglés: *Module membership*)

MoRF: Caracteres de reconocimiento molecular (del inglés: *molecular recognition features*)

NAD: Nicotinamida adenina dinucleótido

NADH: Forma reducida de la nicotinamida adenina dinucleótido

NAD-ME: Enzima Málica dependiente de Nicotinamida Adenina Dinucleótido

NADP: Nicotinamida Adenina Dinucleótido Fosfato

NADPH: Forma reducida de la nicotinamida adenina dinucleótido fosfato

NADP-ME: Enzima Málica dependiente de Nicotinamida Adenina Dinucleótido Fosfato

NDH: Complejo de tipo NADH deshidrogenasa

NKD: Naked endosperm

NLS: Señal de localización nuclear (del inglés: *nuclear localization signal*)

NR1: Nitrato reductasa 1

NSP1: Nodulation Signaling Pathway

OAA: Ácido oxaloacético

OED: Ortogrupos expresados diferencialmente

OG: Ortogrupo

PAT1: Phytochrome A signal transduction 1

pb: Pares de bases

PC de CO₂: Punto de compensación de CO₂

PCA: Análisis de componentes principales (del inglés: *Principal Component Annalysis*)

PE: Lecturas de extremo emparejado (del inglés: *Pair end*)

PEPC: Fosfoenolpiruvato carboxilasa

PEPCK: Fosfoenolpiruvato carboxiquinasa

PGLP1: Fosfoglicolato fosfatasa

PK: Proto-Kranz

PPDK: Piruvato fosfato diquinasa

ppm: Partes por millón

PQ: Plastoquinona

pseudoTLE: vía de transporte de electrones pseudo cíclica o pseudo transporte lineal de electrones.

PSI: Complejo fotosistema I

PSII: Complejo fotosistema II

RGA: Repressor of ga1–3

RBCS1A: Subunidad pequeña de RUBISCO

RNAseq: Secuenciación del ARN a escala global

ROS: Especies reactivas del oxígeno

RUBISCO: Ribulosa-1,5-bisfosfato carboxilasa/oxigenasa

SCL3: Scarecrow-Like 3

SCR: Scarecrow

SG5: Shoot Gravitropism 5

SGAT: Serina-glioxilato aminotransferasa

SHBM: Motivo de union a SHR (del inglés: *SHR binding motif*)

SHM4: Serina hidroximetiltransferasa

SHM6-7: Serina hidroximetiltransferasa

SHMT1: Serina hidroximetiltransferasa

SHR: Shortroot

SOD: Superóxido dismutasa

STP1: transportador de sacarosa STP1

SUS: Sacarosa sintasa

SWEET: transportador de sacarosa SWEET

TCE: Transporte cíclico de electrones o fosforilación oxigénica

TLE: Transporte lineal de electrones o fosforilación anoxigénica

TPE: Transcritos poco expresados

TPM: Transcritos por millón de kilo bases (del inglés: *Transcripts Per Kilobase Million*)

WDP: Woodpecker

WGD: Duplicaciones completa del genoma (del inglés: *Whole Genome Duplications*)

WRB: Warbler

ZF: *Zinc fingers*

ZFP: Proteína *zinc finger*

RESUMEN

La fotosíntesis es un proceso complejo mediante el cual las plantas utilizan el CO_2 y la luz del sol para obtener energía. En la fotosíntesis C_3 , la enzima principal encargada de incorporar el CO_2 (llamada RUBISCO) también cataliza una reacción de oxigenación. Esta reacción da origen a un proceso denominado fotorrespiración, que implica una disminución en la eficiencia fotosintética. La reacción de oxigenación de RUBISCO se intensifica bajo ciertas condiciones climáticas como ambientes cálidos, áridos y salinos. Esta desventaja de la fotosíntesis C_3 fue resuelta por algunos grupos de plantas a través de la incorporación de enzimas más eficientes en la fijación de carbono y la relocalización de la maquinaria catalítica entre las células de mesófilo y las de la vaina. Dicha compartimentalización celular muestra una anatomía foliar diferencial denominada anatomía Kranz. Estas novedades evolutivas establecieron un nuevo tipo de fotosíntesis denominada fotosíntesis C_4 .

La fotosíntesis C_4 les ha permitido a las plantas adaptarse a altas temperaturas e intensa radiación solar mediante un mejor uso del agua y del nitrógeno. A pesar de que solo un 3% de las angiospermas usa el circuito C_4 , las especies C_4 están entre los cultivos más importantes del planeta (por ejemplo, maíz, caña de azúcar, sorgo) y cubren cerca del 25% de la superficie de la tierra. Esta fotosíntesis es uno de los ejemplos más notables de homoplasia evolutiva, ya que se conocen al menos 70 orígenes independientes de la misma. Además, cerca de los linajes C_4 suelen encontrarse especies que presentan subtipos fotosintéticos con características intermedias. Algunas son especies C_3 con características anatómicas y fisiológicas que reducen los niveles de fotorrespiración, como las especies Proto-Kranz (PK) y C_2 de tipo I; y otras son especies donde además existe una incipiente vía C_4 , como en las C_2 de tipo II o las C_4 -like. A pesar de décadas de estudio, el mecanismo que promovió su evolución es un enigma biológico aún no resuelto. Los modelos actuales proponen que el ensamblado de la fotosíntesis C_4 implicó una serie de cambios evolutivos graduales en el cual la relocalización de las enzimas fotorrespiratorias, con el establecimiento de fotosíntesis intermedias, fue un paso clave para el surgimiento de la fotosíntesis C_4 .

Sorpresivamente, más de 1/3 de los orígenes C_4 ocurrieron dentro de la familia de las gramíneas (Poaceae). En particular, la subtribu Otachyriinae (tribu Paspaleae), constituida por 35 especies distribuidas en América, incluye a especies C_4 , C_2 , PK y C_3 . Esto la convierte en un linaje interesante para responder preguntas sobre la evolución y la optimización de la fotosíntesis.

Por estas razones, el objetivo general de este trabajo fue estudiar los mecanismos moleculares responsables del origen y diversificación de los sistemas fotosintéticos en gramíneas empleando como

grupo modelo a la subtribu Otachyriinae. Para ello se abordó esta problemática desde tres enfoques organizados en capítulos:

1) Evolución molecular de las subfamilias de reguladores transcripcionales IDD, SHR y SCR. Aunque poco se sabe respecto al control génico del establecimiento de la anatomía Kranz, se piensa que algunos factores de transcripción de las familias GRAS y C2H2 podrían jugar un rol importante. Las subfamilias GRAS SHORTROOT (SHR) y SCARECROW (SCR) fueron caracterizadas como participantes claves, junto a miembros de la subfamilia de *zinc finger* C2H2 INDETERMINATE DOMAIN (IDD), en el desarrollo del patrón radial del tejido basal, tanto en raíces como en tallos y hojas. Sólo un puñado de estudios se han enfocado en la filogenia y la identificación de ortólogos para los IDDs, SHRs y SCRs entre diferentes especies de plantas. Es por ello que se decidió abordar un estudio filogenético exhaustivo sobre la evolución molecular de estos genes en las plantas verdes.

2) Estudio de la anatomía foliar y la fisiología en Otachyriinae. Hasta el momento, ningún estudio se centró en el desarrollo foliar de la subtribu Otachyriinae y muy pocos reportan la caracterización de la anatomía foliar. Por este motivo, se propuso estudiar el desarrollo de la anatomía foliar para comprender los mecanismos responsables de la diversificación de los subtipos fotosintéticos en las gramíneas. Para ello se analizó, a través de microscopía óptica, la evolución de rasgos fenotípicos durante el desarrollo foliar.

3) Análisis de la transcriptómica del gradiente foliar en Otachyriinae. Para tratar de entender los mecanismos moleculares responsables de la evolución de la fotosíntesis C₄, se estudió en forma comparada la regulación transcripcional de genes a lo largo de la hoja en especies de la subtribu Otachyriinae. Para ello se secuenciaron y ensamblaron los transcriptomas de novo para tres especies con fotosíntesis C₃, PK y C₄. Luego, se investigó la dinámica transcripcional, la expresión de los genes a lo largo del gradiente foliar y se identificaron diferencias y similitudes entre las especies.

Los estudios filogenéticos en las subfamilias IDD, SHR y SCR permitieron identificar eventos de duplicación específicos y determinar las relaciones de ortología entre las especies. El estudio de la anatomía foliar, demostró la existencia de un gradiente de desarrollo con diferencias marcadas entre las especies con diferentes subtipos fotosintéticos. Finalmente, por medio de la secuenciación masiva del ARN, se lograron construir transcriptomas con un gran nivel de completitud, un bajo nivel de secuencias redundantes y una gran cantidad de transcriptos anotados. A través de un análisis de expresión diferencial se encontraron diferencias significativas entre las especies, con respecto a las dinámicas transcripcionales. Además, se encontró que la mayoría de los patrones de expresión de las enzimas asociadas al ciclo C₄ se encuentran conservados a lo largo del desarrollo entre la especie C₄ y las especies C₃.

ABSTRACT

Photosynthesis is a complex process by which plants use CO₂ and sunlight for obtaining energy. In C₃ photosynthesis, the main enzyme responsible for incorporating CO₂ (called RUBISCO) also catalyzes an oxygenation reaction. This reaction originates a process called photorespiration, which implies a decrease in photosynthetic efficiency. RUBISCO's oxygenation reaction intensifies under specific climatic conditions such as hot, arid and saline environments. This disadvantage of C₃ photosynthesis was solved by some groups of plants through the incorporation of more efficient enzymes in carbon fixation and the relocation of the catalytic machinery between mesophyll cells and bundle sheath cells. This cellular compartmentalization shows a distinctive foliar anatomy called Kranz anatomy. Altogether, these evolutionary novelties established a new type of photosynthesis called C₄ photosynthesis.

C₄ photosynthesis has allowed plants to adapt to high temperatures and intense solar radiation through better use of water and nitrogen. Despite the fact that only 3% of angiosperms use the C₄ cycle, C₄ species are among the most important crops on the planet (e.g. corn, sugar cane, sorghum) and cover about 25% of the surface of the Earth. This photosynthesis is one of the most remarkable examples of evolutionary homoplasy, since at least 70 independent origins are known. In addition, species with photosynthetic subtypes with intermediate characteristics are usually found near the C₄ lineages. Some are C₃ species with anatomical and physiological characteristics that reduce levels of photorespiration, such as Proto-Kranz (PK) and C₂ type I species; and others are species where there is also a weak C₄ pathway, as in C₂ type II or C₄-like. Despite decades of study, the mechanism that promoted its evolution is a biological enigma. Current models indicate that the assembly of C₄ photosynthesis involved a series of gradual evolutionary changes in which the relocation of photorespiratory enzymes, with the establishment of intermediate photosynthesis, was a key step to achieve C₄ photosynthesis.

Surprisingly, more than 1/3 of the C₄ origins occurred within the grass family (Poaceae). In particular, the Otachyriinae subtribe (Paspaleae tribe), with 35 species distributed in America, includes C₄, C₂, PK and C₃ species. This makes it an interesting lineage to answer questions about the evolution and optimization of photosynthesis.

For these reasons, the general objective of this work was to study the molecular mechanisms responsible for the origin and diversification of photosynthetic systems in grasses, using Otachyriinae subtribe as a model group. For this, the problem was addressed from three approaches organized in chapters:

1) Molecular evolution of the transcriptional regulator subfamilies IDD, SHR and SCR. Although little is known regarding the gene control of the establishment of the Kranz anatomy, it is thought that some transcription factors of the GRAS and C2H2 families may play an important role. The GRAS SHORTROOT (SHR) and SCARECROW (SCR) subfamilies were characterized as key participants, together with members of the *zinc finger* C2H2 INDETERMINATE DOMAIN (IDD) subfamily, in the development of the radial pattern of basal tissue, in both roots and stems and leaves. Only a handful of studies have focused on the phylogeny and identification of orthologues for IDDs, SHRs, and SCRs among different plant species. To complement the available information, in this work an exhaustive phylogenetic study on the molecular evolution of these genes was performed in green plants.

2) Study of foliar anatomy and physiology in Otachyriinae. So far, no studies have focused on the foliar development of the Otachyriinae subtribe and very few report the characterization of the foliar anatomy. Therefore, it was proposed to study the development of foliar anatomy in order to understand the mechanisms responsible for the diversification of photosynthetic subtypes in grasses. For this, the evolution of phenotypic traits during foliar development was analyzed through optical microscopy.

3) Analysis of foliar transcriptomics in Otachyriinae. To explore the molecular mechanisms responsible for the evolution of C₄ photosynthesis, the transcriptional regulation of genes throughout the leaf was studied comparatively among species of the Otachyriinae subtribe. For this, transcriptomes were sequenced and de novo assembled for three species with C₃, PK and C₄ photosynthesis. Then, the transcriptional dynamics and the expression of the genes along the foliar gradient were investigated and differences and similarities between the species were identified.

Phylogenetic studies in the IDD, SHR and SCR subfamilies identified specific duplication events and determined the orthological relationships between the species. The study of foliar anatomy demonstrated the existence of a development gradient with marked differences between species with different photosynthetic subtypes. Finally, by means of massive RNA sequencing, de novo transcriptomes were assembled with a high level of completeness, a low level of redundant sequences, and a large number of annotated transcripts. Through a differential expression analysis, significant differences were found between species on transcriptional dynamics. In addition, it was found that most of the expression patterns of the enzymes associated with the C₄ cycle are conserved throughout development between the C₄ and C₃ species.

INTRODUCCIÓN

Fotosíntesis C₃, C₂ y C₄

La fotosíntesis es un proceso complejo mediante el cual las plantas utilizan el CO₂ y la luz del sol para obtener energía. El proceso fotosintético consta de dos etapas: una dependiente de la luz, donde la energía lumínica del Sol es transformada en energía química y poder reductor en forma de ATP y NADPH respectivamente; y otra independiente de la luz, donde esas moléculas se utilizan para sintetizar carbohidratos a partir de la incorporación de CO₂ mediante el Ciclo de Calvin-Benson-Bassham (CCBB).

Reacciones dependientes de la luz

Las reacciones dependientes de la luz suceden en la membrana tilacoide de los cloroplastos. En ella, complejos multiproteicos llamados fotosistemas (PSI y PSII) captan la luz del sol a través de la excitación de pigmentos fotosintéticos, como la clorofila, generando un flujo de electrones que lleva en última instancia a la reducción de NADP a NADPH (Figura 1). Al mismo tiempo, el movimiento de electrones crea un gradiente de H⁺ a ambos lados de la membrana tilacoide que es usado por el complejo ATP sintetasa para la síntesis de ATP. La molécula de clorofila que perdió un electrón por la excitación de la luz solar, lo recupera cuando una molécula de H₂O es descompuesta liberando O₂ como producto de desecho ([1], [2]). A este circuito se lo conoce como fotofosforilación acíclica u oxigénica.

En el movimiento acíclico de electrones durante la fotofosforilación oxigénica, la relación de producción de ATP/NADP es aproximadamente 1,29, lo cual no satisface la relación 1,5 requerida por el CCBB ([1], [2]). Para compensar esto, dos vías de transportes de electrones alternativas contribuyen a la síntesis de ATP: la fotofosforilación cíclica o anoxigénica y la fotofosforilación pseudo-cíclica ([2]). En la fotofosforilación cíclica los electrones se desplazan alrededor del PSI desde una molécula de ferredoxina (Fd) a una de plastoquinona (PQ), contribuyendo al gradiente de H⁺ sin producción de NADPH. En el ciclo agua-agua, los electrones de PSI son transferidos a una molécula de O₂, generando un ion superóxido. Las especies reactivas al oxígeno (ROS) resultantes son eliminadas por la superóxido dismutasa (SOD) y la ascorbato peroxidasa (APX). Este sistema de eliminación de ROS consume moléculas de NADPH para la regeneración de ascorbato (Figura 1).

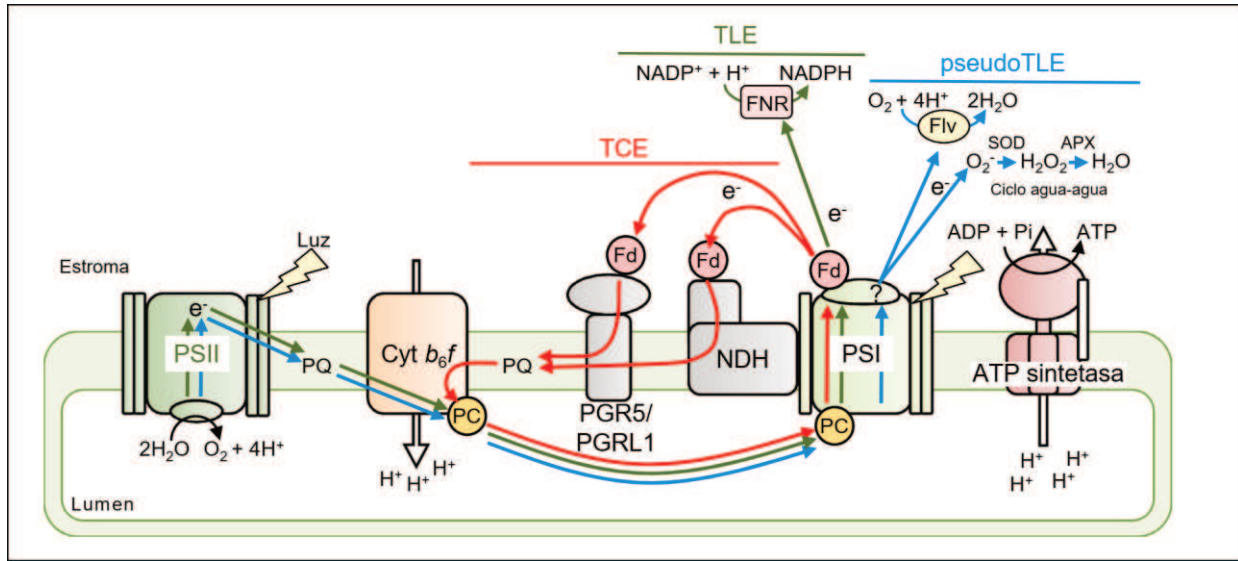


Figura 1. Representación de las reacciones dependientes de la luz en la membrana tilacoide. El transporte lineal de electrones (TLE, fotosíntesis oxigénica) desde el agua hasta NADP⁺ es indicado por flechas verdes. Dos vías de transporte cíclico de electrones (TCE, fotosíntesis anoxigénica) a través de los complejos NDH y PGR5/PGRL1, están indicadas por flechas rojas. Las flechas azules indican la vía de transporte de electrones pseudo-cíclica (pseudoTLE). El ciclo de agua-agua depende de la reducción de O₂ en las cercanías del PSI. Las especies reactivas del oxígeno (ROS) generadas son capturadas por la superóxido dismutasa (SOD) y la ascorbato peroxidasa (APX). Abreviaturas: fotosistema I (PSI), fotosistema II (PSII), plastoquinona (PQ), plastocianina (PC), citocromo b₆f (Cyt b₆f), ferredoxina (Fd), ferredoxina NADP⁺ reductasa (FNR) (Figura modificada de Shikanai y Yamamoto, 2017).

Reacciones independientes de la luz: RUBISCO y el ciclo de Calvin-Benson-Bassham

Las moléculas de ATP y NADPH generadas en las reacciones dependientes de la luz, se consumen para fijar CO₂ en el ciclo de reducción de carbono fotosintético, más conocido como Ciclo de Calvin-Benson-Bassham (CCBB) o Ciclo C₃ (Figura 2). Este ciclo, que ocurre en el estroma de los cloroplastos, comienza cuando una molécula de CO₂ se combina con un azúcar de 5 carbonos, la ribulosa-1,5-bisfosfato, en una reacción catalizada por la enzima ribulosa-1,5-bisfosfato carboxilasa/oxigenasa (RUBISCO). La reacción forma un compuesto intermediario inestable de 6 carbonos que se rompe en dos moléculas de 3-fosfoglicerato. Este compuesto es primero fosforilado por la 3-bisfosfoglicerato quinasa usando ATP para formar 1,3-bisfosfoglicerato, que a su vez es reducido por la gliceraldehído 3-fosfato deshidrogenasa usando NADPH para formar gliceraldehído 3-fosfato (G3P), un azúcar de 3 carbonos. Por cada 3 moléculas de CO₂ inicialmente combinadas con la ribulosa-

1,5-bifosfato, se producen 6 moléculas de G3P. Una de ellas es considerada el producto del CCBB mientras que las otras 5 son utilizadas para regenerar a la ribulosa-1,5-bifosfato en una compleja serie de reacciones que también requieren de ATP. La molécula de G3P producida puede ser rápidamente convertida por una variedad de vías metabólicas en aminoácidos, lípidos o carbohidratos. En total, la síntesis de una molécula de G3P requiere de 9 moléculas de ATP y 6 de NADPH (relación ATP/NADPH = 1,5) [1],[3].

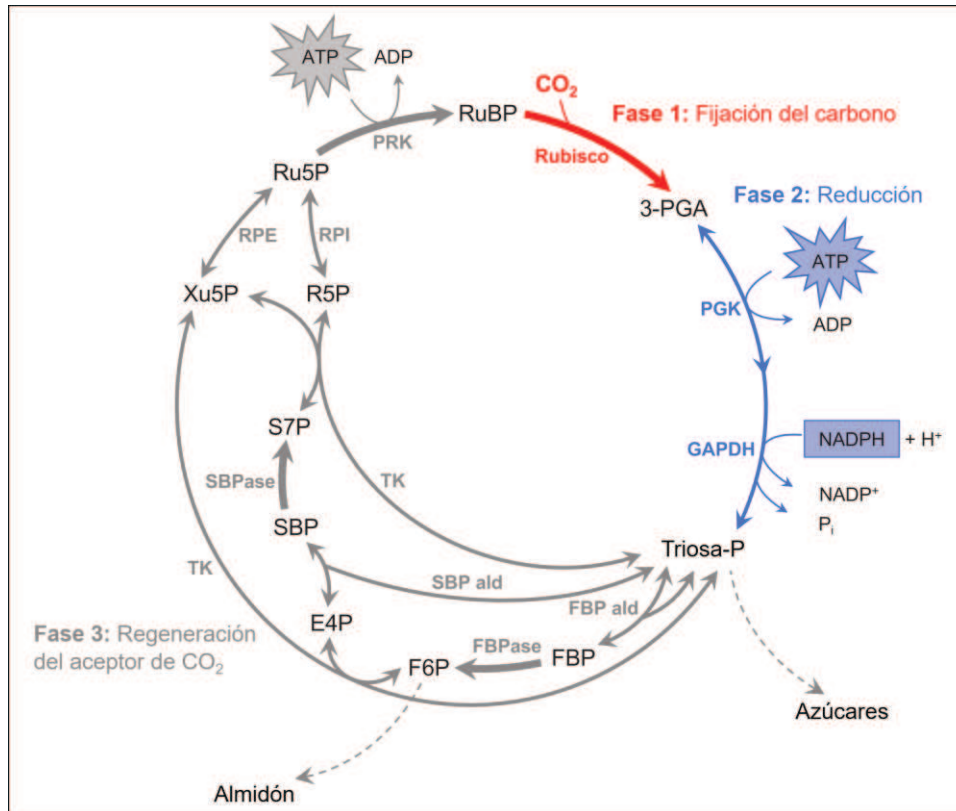


Figura 2. Ciclo de Calvin-Benson-Basham con sus tres fases: (1) fijación del carbono(rojo), (2) reducción (azul) y (3) regeneración del aceptor de CO₂ (gris). La carboxilación es realizada por RUBISCO, que fija CO₂ usando ribulosa-1,5bifosfato (RuBP) como aceptor, produciendo dos moléculas de 3-fosfoglicerato (3-PGA). El 3-PGA es luego fosforilado por la fosfoglicerato kinasa (PGK) y reducido a triosa fosfato (triosa-P) por la gliceraldehído-3-fosfato deshidrogenasa (GAPDH) en la fase reductiva. El ciclo usa 3 ATP y 2 NADPH por cada molécula de CO₂ fijada. La triosa-P puede ser transportada al citosol para producir sacarosa. La fructosa 1,6-bifosfato aldolasa (FBP ald) puede convertir a la triosa-P en fructosa-6 fosfato (F6P), como intermediario para producir almidón. Además, la triosa-P puede ser convertida a RuBP en una serie de reacciones de regeneración. Abreviaturas: fructosa-1,6 bifosfato (FBP), fructosa-1,6-bifosfatasa (FBPase), eritrosa-4-fosfato (E4P), sedoheptulosa 1,7-bifosfato aldolasa (SBP ald), sedoheptulosa-1,7-bifosfato (SBP), sedoheptulosa-1,7- bifosfatasa (SBPase), sedoheptulosa-7-fosfato (S7P), transketolasa (TK), ribosa-5-fosfato (R5P), xilulosa-5-fosfato (Xu5P), ribosa-5-fosfato isomerasa (RPI), ribulosa-5-fosfato epimerasa (RPE), ribulosa-5-fosfato (Ru5P), fosforibulokinasa (PRK).

Las enzimas que catalizan reacciones irreversibles están resaltadas por flechas más anchas (Figura modificada de Schreier y Hibberd, 2019).

A pesar de la complejidad del CCBB, el paso de fijación del carbono es realizado por la enzima RUBISCO. Este complejo proteico soluble se localiza en el estroma de los cloroplastos y consiste de 8 subunidades grandes (56 kDa), que contienen los dominios catalíticos y regulatorios, y 8 unidades pequeñas (14 kDa), que potencian las funciones catalíticas de las subunidades grandes ([1]).

Además de la actividad carboxilasa, RUBISCO cataliza una reacción de oxigenación al combinar ribulosa-1,5-bifosfato con O_2 en lugar de CO_2 . Esta reacción origina una molécula de 3-fosfoglicerato y una molécula de fosfoglucolato (un azúcar de 2 carbonos), que luego de una serie de reacciones regenera a una molécula de 3-fosfoglicerato y una molécula de CO_2 , consumiendo ATP en el proceso. Este proceso se conoce como fotorrespiración. Si bien la actividad oxigenasa de RUBISCO es mucho menos favorable que la carboxilasa, la alta concentración del O_2 ambiental (210.000 ppm) comparada con la de CO_2 (411 ppm) permite que la fotorrespiración ocurra habitualmente ([1], [4], [5]).

La fotosíntesis C_4

La vía metabólica C_4 concentra el CO_2 de los espacios de aire intercelulares de las hojas en un compartimento interno donde se localiza RUBISCO. De esta manera, se contrarrestan los efectos inhibitorios de los bajos valores de CO_2 atmosféricos en la fotosíntesis. Este mecanismo representa un rasgo evolutivo complejo que resultó de una sustancial reorganización de la anatomía foliar y del metabolismo [6].

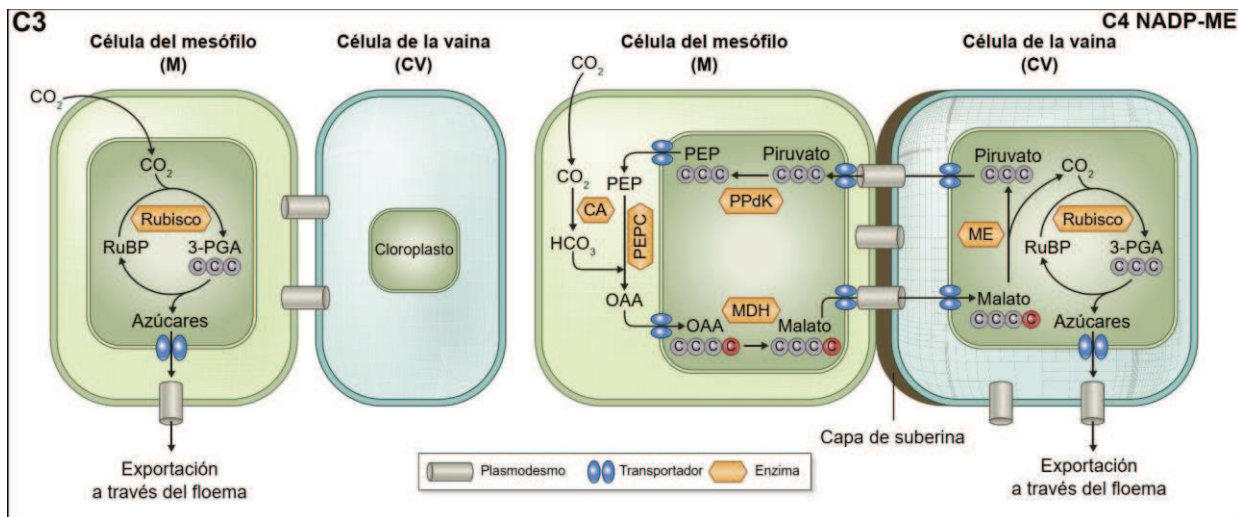


Figura 3. Representación esquemática de las vías fotosintéticas C_3 y C_4 . El ciclo C_3 ocurre en el M, donde RUBISCO fija el CO_2 en 3-PGA. En el ciclo C_4 , el CO_2 es reducido inicialmente por la CA en el citoplasma del M, y luego es fijado por la PEPC para producir OAA. El OAA es reducido a malato por la MDH, que es transportado a las CV por medio de plasmodesmos. En las CV, la ME descarboxila al malato liberando CO_2 , el cual es reducido por RUBISCO produciendo azúcares en el CCBB. El piruvato generado durante la descarboxilación es enviado al M donde se usa para regenerar PEP por medio de la PPdK y continuar el ciclo. Abreviaturas: ribulosa-1,5-bisfosfato (RuBP), 3-fosfoglicerato(3-PGA), anhidrasa carbónica (CA), PEP-carboxilasa (PEPC), oxaloacetato (OAA), malato deshidrogenasa (MDH), enzima málica dependiente de NADP (ME), PEPdikinasa (PPdK) (Figura modificada de Kumar y Kellogg, 2018).

El primer paso en la vía C_4 se produce en el citoplasma de las células del mesófilo (M), donde la anhidrasa carbónica (CA) incorpora al CO_2 en carbonato que es transformado luego por la fosfoenolpiruvato carboxilasa (PEPC) en un compuesto de 4 carbonos, llamado oxaloacetato (OAA) (Figura 3, [7], [8]). El OAA es convertido a malato en los cloroplastos y transportado al compartimento interno (las células de la vaina, CV), donde es descarboxilado por una enzima málica (ME) para producir piruvato y CO_2 . El piruvato es transportado al M donde sirve para regenerar el PEP por medio de la Piruvato fosfato dikinasa (PPdK) y el CO_2 es capturado por RUBISCO en los cloroplastos de las CV dando inicio al CCBB ([8], Figura 3).

En hojas C_4 el compartimento donde se concentra el CO_2 suele ser una capa celular que rodea a los haces vasculares (HV) llamada célula de la vaina externa (CVE). Sin embargo, las CVE representan sólo uno de los compartimentos modificados para tal fin a lo largo de la evolución. Las células del mestoma o células de la vaina interna (CVI), junto con células del parénquima entre células del M y tejido almacenador de agua, son ejemplos de sitios de concentración de CO_2 en algunos linajes de especies C_4 ([9], [10]).

Además de las diversas alternativas descritas como compartimentos de concentración de CO_2 entre especies C_4 , se han identificado variaciones en el tipo de enzimas reclutadas para llevar a cabo la vía metabólica C_4 . Según la principal enzima descarboxilante usada en la CV, las especies con fotosíntesis C_4 se clasifican en subtipos NADP-enzima málica (NADP-ME), NAD-enzima málica (NAD-ME) y PEP carboxiquinasa (PEPCK).

En la mayoría de los linajes C_4 , la combinación de rasgos característicos resulta en la concentración de CO_2 alrededor de RUBISCO dentro un compartimento interno. Sin embargo, en ciertas especies de la familia Amaranthaceae, se ha descrito que el proceso fotosintético C_4 puede ocurrir dentro de una única célula. En estos casos, el proceso de concentración de CO_2 ocurre en un lugar intracelular especializado; por ejemplo, en *Bienertia cycloptera* la concentración de CO_2 ocurre

en el centro celular, mientras que en *Suaeda aralocaspica* ocurre en el polo interno de una célula particularmente elongada [10].

Las diferentes estrategias de concentración de CO₂ alrededor de RUBISCO descritas en el reino vegetal y sus orígenes independientes convierten a la vía C₄ en un síndrome, al considerarse una combinación de rasgos que producen un resultado característico [11].

Anatomía Kranz

El término “*Kranz*”, proviene del alemán y significa “corona” o “guirnalda”. Fue usado por primera vez por el botánico austriaco Gottlieb Haberlandt para referirse a la “corona” de células del M organizadas radialmente alrededor de las CV en especies de gramíneas y juncos. Sin embargo, el término fue aplicado tanto a las CV agrandadas y ricas en cloroplastos (células Kranz), como a todo el conjunto de características estructurales distintivas (“anatomía Kranz”) [9].

A pesar de la considerable variación en la arquitectura de los tejidos de las hojas entre las familias en las que ha evolucionado la fotosíntesis C₄, ciertas características anatómicas están asociadas invariablemente con la vía y se consideran esenciales para su funcionamiento. Ellas son: la especialización de dos tipos de células fotosintéticas; la configuración espacial de los tejidos de modo que las células del M se dispongan en el exterior de la hoja y en contacto con el espacio aéreo intercelular, mientras que las CV se ubican hacia el interior de la hoja, en estrecho contacto con los HV; una corta trayectoria de difusión para los metabolitos fotosintéticos entre las células de M y CV, reflejada en una alta densidad de HV y bajas proporciones de M sobre CV; y características que limiten la pérdida de CO₂ de las CV, incluidas la minimización de la exposición de la superficie de CV al espacio intercelular y la modificación química de la pared celular [9].

Las modificaciones en las CV de las especies C₄ en comparación con las especies C₃ incluyen: un incremento en el tamaño celular, la presencia de gránulos de almidón, numerosos cloroplastos de mayor tamaño que en las células del M y una organización asimétrica de las organelas. Con respecto a las paredes celulares, se observa generalmente la presencia de suberina que dificulta la difusión pasiva de gases, y una mayor presencia de plasmodesmos que facilita el intercambio de metabolitos entre CV y M de hojas C₄ ([9], [12], [13], [14]).

Así mismo, las células del M de las especies C₄ suelen presentar una organización más compacta que disminuye el espacio aéreo intercelular, una menor cantidad de cloroplastos y una organización diferencial de las mitocondrias, asociadas estrechamente a los cloroplastos ([9], [14], [15]).

En particular, en las gramíneas C₃ los HV longitudinales están rodeados por dos capas celulares: la interna (CVI), formada por células de esclerénquima sin cloroplastos; y la externa (CVE),

formadas por células de parénquima con cloroplastos. En las especies de gramíneas C_4 , estas capas están modificadas de diferentes formas según el subtipo bioquímico fotosintético ([9], [10]).

Subtipos de fotosíntesis C_4 : NADP-ME, NAD-ME y PEPCK

Como se mencionó anteriormente, dependiendo de la enzima responsable del mecanismo de descarboxilación, las especies C_4 se agrupan en diferentes subtipos bioquímicos: NADP-ME, NAD-ME y PEPCK. En algunas especies C_4 , especialmente aquellas que usan NADP-ME como descarboxilasa primaria, la PEPCK opera como descarboxilasa secundaria [16]. Por esta razón, hay cierto debate sobre si el subtipo PEPCK debería o no considerarse independiente. Sin embargo, la clase de descarboxilasa primaria se asocia generalmente con un conjunto de características anatómicas, bioquímicas y fisiológicas distintivas ([17] [10]).

En las gramíneas, cada subtipo bioquímico está fuertemente asociado a una anatomía Kranz particular (Figura 4). En la mayoría de las especies NADP-ME las CVE están ausentes y el CCBB se realiza en las CVI. Éstas poseen cloroplastos que presentan menos granas, debido probablemente a una menor presencia del PSII, y organelas organizadas centrífugamente. ([10], [18], Figura 4 a).

Las especies NAD-ME poseen las dos capas de CV desarrolladas, con la CVE funcionando como compartimento Kranz. En estos casos, las CVE carecen de la capa de suberina en su pared celular y presentan cloroplastos con numerosas granas, una gran cantidad de mitocondrias y una distribución centrípeta de las organelas (Figura 4 b).

Las especies PEPCK también poseen dos capas de CV, ambas con láminas de suberina en sus paredes. Las CVE presentan cloroplastos con granas, un gran número de mitocondrias, y las organelas están organizadas centrífugamente o dispersas en el citoplasma (Figura 4 c).

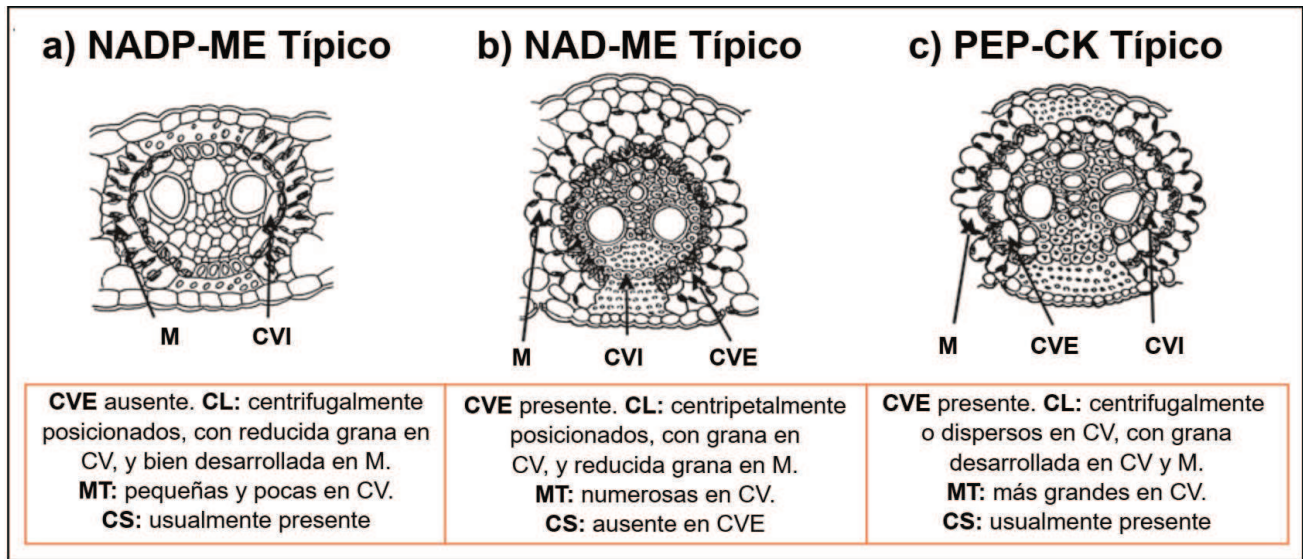


Figura 4. Ilustración de las formas de anatomía Kranz en las gramíneas. Características de las hojas del subtipo NADP-ME (a), características de las hojas del subtipo NAD-ME (b), y características de las hojas del subtipo PEP-CK (c). Abreviaturas: células del mesófilo (M), células de la vaina interna (CVI), células de la vaina externa (CVE), cloroplastos (CL), mitocondrias (MT), capa de suberina (CS) (Figura modificada de Edwards y col., 2010).

Si bien existe una clara compartimentación de las reacciones de carboxilación entre CV y M, la localización de las reacciones dependientes de la luz es menos clara en las especies C_4 . En las especies NADP-ME *Sorghum bicolor* y *Zea mays*, las células del M contienen una cadena de transporte de electrones lineal completa como en las especies C_3 . En cambio, las CV poseen cantidades reducidas del PSII, indicando una preferencia por un transporte de electrones cíclico. Además, en las CV de los tres subtipos se pueden identificar mega complejos formados por PSI y NDH a los cuales puede asociarse la mayor actividad en el flujo de transporte de electrones cíclicos. La existencia de este transporte de electrones más activo proveería los ATP adicionales necesarios para compensar los costos energéticos originados por regenerar el PEP en las células del M y el sobreciclaje de CO_2 en la CV. Por último, la fotólisis del agua por el PSII incrementa la relación O_2/CO_2 , lo que favorecería la fotorrespiración en las CV ([2], [18]–[20]).

A diferencia de lo que ocurre en las especies NADP-ME, las CV de los otros dos subtipos mantienen una cantidad substancial de PSII, lo cual explicaría su menor rendimiento cuántico de la fotosíntesis [18].

Origen y evolución de la fotosíntesis C₄

Cambio climático y origen de la fotosíntesis C₄

En los últimos 40 millones de años el sistema climático de la Tierra pasó de un mundo cálido y húmedo con polos templados, a un planeta más seco, con casquetes polares, desiertos extremos y grandes praderas. Asociado a este cambio climático, se produjo una reducción en el contenido de CO₂ atmosférico desde 1.000 $\mu\text{mol CO}_2\text{mol}^{-1}$ de aire, hace 50 millones de años, hasta menos de 200 $\mu\text{mol mol}^{-1}$, hace 20 millones de años (Figura 5; [6], [21]).

Estos cambios en el clima y la atmósfera causaron drásticas respuestas evolutivas en los seres vivos del planeta que contribuyeron al surgimiento de la biósfera moderna. Muchas de estas respuestas ocurrieron en la fisiología de las plantas, reflejando el impacto directo del CO₂, la temperatura, la humedad y la disponibilidad de agua en la fotosíntesis ([22], [23]).

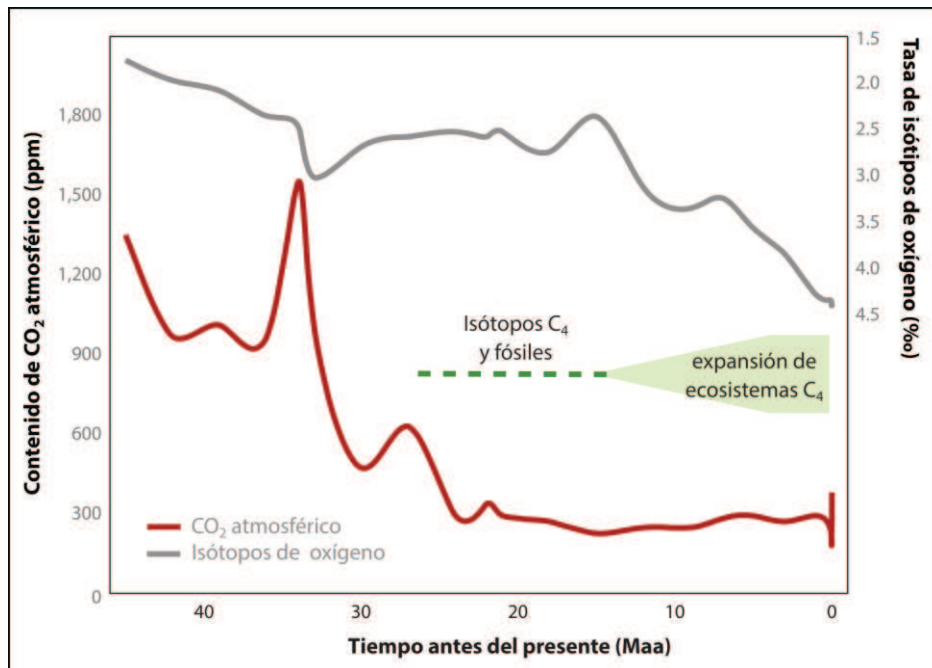


Figura 5. Concentraciones atmosféricas de CO₂ y tasas de isótopos de oxígeno ($\delta_{18}\text{O}$) durante los últimos 46 millones de años. En general, niveles bajos de $\delta_{18}\text{O}$ indican la existencia de climas cálidos y húmedos a lo largo del planeta, mientras que valores altos corresponden a climas más fríos y secos (Figura modificada de Sage y col., 2012).

Entre los cambios evolutivos originados por esta profunda modificación climática, se destaca el incremento en la diversidad de vías fotosintéticas como el metabolismo ácido de las crasuláceas (CAM) y la vía fotosintética C_4 . Estas nuevas vías fotosintéticas impactaron en la biósfera de manera significativa, contribuyendo al surgimiento de nuevos ecosistemas y formas de vida [6].

El linaje C_4 más antiguo conocido surgió hace unos 28,5 millones de años en el género *Chloridoideae* perteneciente a la familia de las gramíneas. La mayoría de los linajes C_4 surgieron en el intervalo 5-10 millones de años atrás, durante el mioceno tardío, momento que se correlaciona con una expansión de las praderas C_4 a escala global [24].

Modelo de evolución

La fotosíntesis C_4 es uno de los ejemplos más notables de homoplasia evolutiva, ya que se conocen al menos 70 orígenes independientes de la misma, que se agrupan en tres grandes grupos filogenéticos: la familia de las gramíneas o Poaceae (22-24 orígenes independientes), la familia Caryophyllaceae (23 orígenes independientes) y la familia Cyperaceae (6 orígenes independientes) ([6], [24], [25]). En total se conocen alrededor de 8.145 especies C_4 distribuidas a lo largo 418 géneros y 19 familias de angiospermas, con alrededor del 60% de éstas concentradas en la familia de las gramíneas. No existen especies que realicen fotosíntesis C_4 en gimnospermas, otras plantas vasculares o briofitas [24].

Además, cerca de los linajes C_4 suelen encontrarse especies que presentan subtipos fotosintéticos con características intermedias entre C_3 y C_4 . Algunas son especies C_3 con características anatómicas y fisiológicas que reducen los niveles de fotorrespiración, como las especies Proto-Kranz (PK) y C_2 de tipo I; y otras son especies donde además existe una incipiente vía C_4 , como en las C_2 de tipo II o las C_4 -like [26].

A partir de trabajos en especies con características intermedias C_3 - C_4 en los géneros *Alternanthera*, *Flaveria*, *Mollugo*, *Moricandia*, *Neurachne*, y *Steinchisma*, Monson, Rawsthorne y colaboradores (2000) fueron desarrollando un modelo conceptual para explicar la progresión evolutiva de especies C_4 a partir de especies C_3 [27]. Recientemente, Sage y colaboradores han contribuido a este modelo a partir del estudio de nuevos linajes C_4 (Figura 6. [6], [11], [24], [26]–[28]). A pesar de que este modelo supone un progreso secuencial donde cada cambio habilita el paso posterior, el orden secuencial de estos cambios pareciera ser una característica típica de cada linaje, pudiendo haber diferentes caminos hacia una fotosíntesis C_4 [29].

Especies C₃ cercanas a linajes C₄

La ausencia de fotosíntesis C₄ en la gran mayoría de las plantas sugieren que la mayoría de los linajes del reino vegetal tiene un potencial bajo de desarrollar la vía C₄. Por otro lado, la distribución en cúmulos de los linajes C₄ indica una predisposición por la evolución de la vía en estos grupos ([11], [26]). La adquisición de rasgos que incrementan el potencial para evolucionar a C₄ representa una fase hipotética de preconditionamiento (Figura 6). Entre ellos se encuentran: (1) una mayor densidad de HV, (2) una mayor proporción del tejido ocupado por las CV y (3) genomas de gran tamaño ([6], [26], [30], [31]).

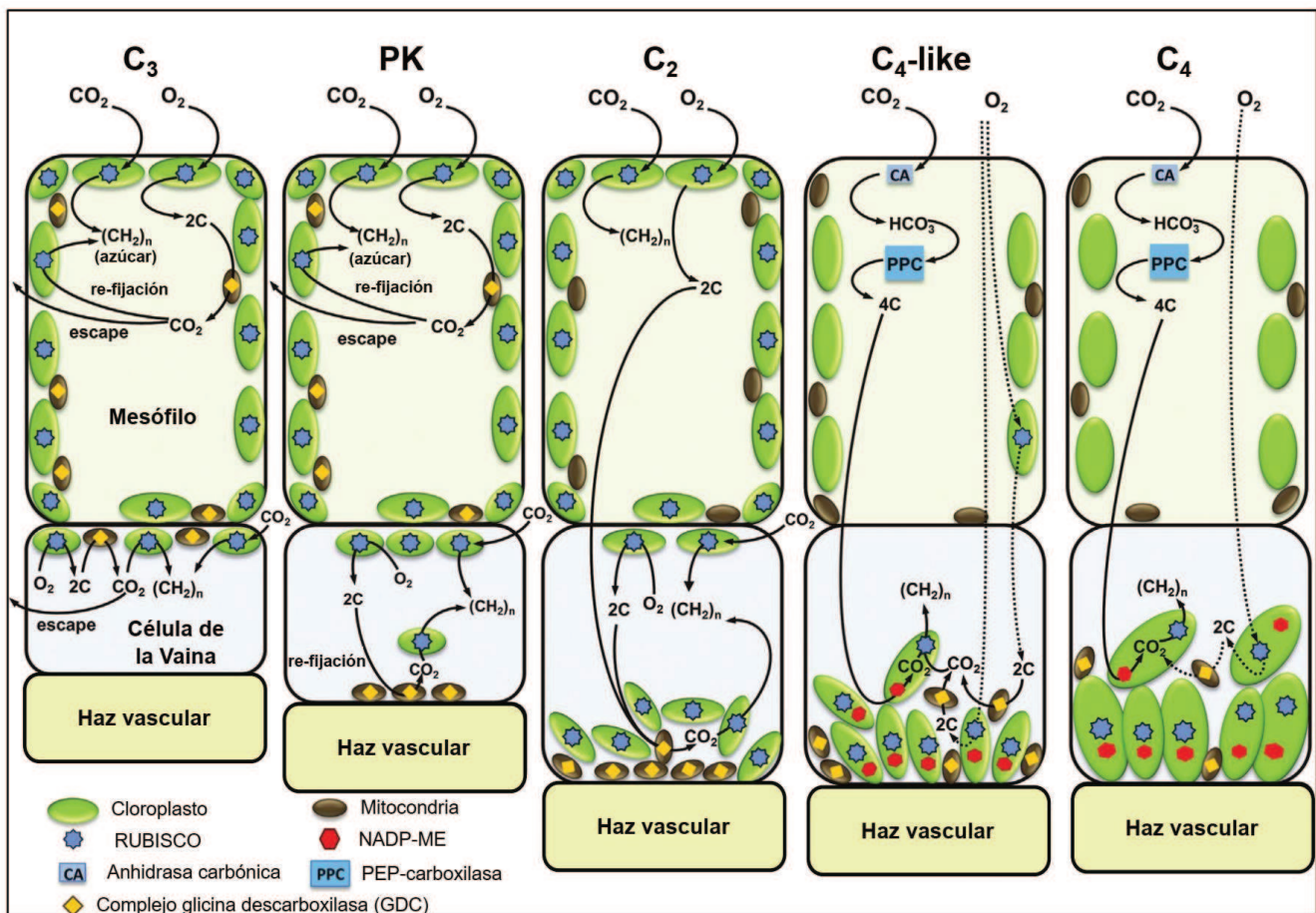


Figura 6. Modelo evolutivo para la fotosíntesis C₄ a partir de ancestros C₃ atravesando tres fases distintas llamadas Proto-Kranz (PK), fotosíntesis C₂ y fotosíntesis C₄-like (Figura modificada de Sage y col., 2014).

En dicotiledóneas, generalmente las especies C₃ cercanas a especies C₄ presentan una mayor densidad de HV que otras especies C₃. Esto llevó a proponer que el preconditionante más claro de esta fase incipiente de evolución era la densidad de HV ([32], [33]). De hecho, una alta densidad de HV probablemente resulta de una adaptación a climas áridos, ya que reduce la longitud del camino del

agua y la resistencia al flujo de agua en los puntos de evaporación, así como también una mayor respuesta a la rehidratación [34]. Sin embargo, estudios posteriores en dicotiledóneas y gramíneas sugieren que una alta densidad de HV es un rasgo que se adquirió en una etapa posterior ([30], [35]–[37]).

Las especies C_3 cercanas a especies C_4 , también suelen tener una mayor fracción del área foliar ocupada por las CV, con valores que tienden a solaparse. El tamaño celular de las CV, por otro lado, es un rasgo que habría sido adquirido en pasos posteriores ([30], [36], [37]).

En relación al tamaño del genoma, se sospecha que la presencia de genomas de gran tamaño y numerosos eventos de duplicación génica, incrementó la plasticidad genética de los ancestros C_3 . Esto promovió eventos de neofuncionalización y subfuncionalización útiles para la evolución de especies C_4 [6]. En las gramíneas, por ejemplo, un evento de duplicación del genoma hace 70 millones de años podría haberlas predispuesto a evolucionar numerosas veces hacia C_4 una vez establecida las condiciones ambientales favorables [38].

Especies C_3 Proto-Kranz (PK)

Las especies C_3 que están muy relacionadas a especies C_2 tienen un número de características que indica la presencia de algún mecanismo que reduce la fotorrespiración dentro de las CV (Figura 6). Estas características incluyen un mayor tamaño de las CV, una mayor presencia de organelas (cloroplastos, mitocondrias y una vacuola de mayor tamaño) y una organización de las mismas similar a la observada en las especies C_2 (mitocondrias posicionadas más cerca del límite con los HV y cloroplastos cercanos al límite con las células del M o envolviendo a las mitocondrias) (Figura 6; [33]). Cuando las CV cuentan con estas características se dice que están “activadas” y a la anatomía foliar se la denomina “Proto-Kranz”. Todas estas características redundan en una potencial re-captación del CO_2 fotorrespirado gracias a la relocalización de las organelas en las CV ([33], [39]).

Especies C_2

En las especies C_2 , la vía de la fotorrespiración se particiona entre M y CV, con la liberación del CO_2 ocurriendo predominantemente en las CV. Esto sucede debido a que la enzima glicina descarboxilasa (GDC), que es responsable de la liberación de CO_2 durante la fotorrespiración, está expresada exclusivamente en las mitocondrias de las CV en especies C_2 [40]. En estas especies, la

glicina formada durante la fotorrespiración en el M debe difundir o ser transportada a las CV para ser metabolizada a serina, que luego retorna al M para completar el ciclo fotorrespiratorio [41].

La optimización de la anatomía en la condición PK podría haber tenido una gran significancia para la evolución de C_2 y C_4 ya que una mayor actividad de la GDC en las CV, permitiría que una pérdida de la actividad de esta enzima en el M no sea letal [26]. Más adelante en la evolución y una vez que la GDC se localizó en la CV, la presión de selección podría haber cambiado: en lugar de tener efectos inhibitorios, la fotorrespiración se transformaría en un recurso para abastecer de CO_2 a los cloroplastos de las CV. Esto implicaría una presión de selección para generar un aumento de la superficie de contacto, un aumento de los plasmodesmos y otros mecanismos que vuelvan más eficiente el intercambio de glicina entre el M y las CV [6]. En la mayoría de las especies C_2 , la característica más distintiva es una densa agregación de cloroplastos y mitocondrias de gran tamaño a lo largo de la periferia interna de las CV (Figura 6; [26]).

Si bien las especies C_2 presentan un fenotipo menos diverso que las especies C_4 , pueden clasificarse en dos grupos con diferencias marcadas. El primero incluye a la definición clásica (C_2 tipo I) donde la reducción de la fotorrespiración ocurre solamente a través de la refijación del CO_2 en las CV, mientras que el segundo grupo además expresa débilmente a las enzimas del ciclo C_4 (como PEPC, PPK y NADP-ME). Estas especies de tipo II todavía conservan algo de expresión de RUBISCO en el M, y las enzimas C_4 presentan formas más similares a las de las especies C_3 que a las de especies C_4 [6]. La anatomía foliar suele presentar valores de rasgos fenotípicos que se solapan entre especies C_3 y C_4 , como el tamaño de las CV, la densidad de HV y el tamaño y número de organelas en las CV ([15], [33], [36], [42], [43]). Este tipo de anatomía, conservada entre los dos grupos C_2 , es denominada C_2 -Kranz [26].

Especies C_4 -like

Estudios en el género *Flaveria* indican que en este grupo existe un paso intermedio entre la evolución C_2 y C_4 denominado " C_4 -like" (Figura 6; [26]). Esta transición implica una reducción sustancial en el contenido de RUBISCO de las hojas, por la pérdida de expresión en el M, y un aumento de la eficiencia en el uso del agua debida a cambios en el control estomático desde valores C_3 a valores C_4 ([44], [45]). Sin embargo, las especies C_4 -like no son consideradas como especies C_4 desarrolladas porque la localización de RUBISCO en las CV es incompleta, la inhibición de la fotosíntesis por O_2 presenta valores intermedios entre C_3 y C_4 , y las tasas de isótopos de carbono están por debajo de los de especie C_4 [27]. Se ha propuesto que la adquisición de un ciclo C_4 completamente funcional en esta transición, es producto de cambios en los elementos promotores que incrementan la expresión y

compartimentalización de enzimas claves como RUBISCO, PEPC y CA y su posterior optimización cinética (Figura 6; [6], [46], [47]).

OBJETIVOS

Objetivo general

Estudiar los mecanismos moleculares responsables del origen y diversificación de los sistemas fotosintéticos en gramíneas empleando como grupo modelo a la subtribu Otachyriinae.

Objetivos específicos

- Reconstruir la evolución molecular de los genes SCARECROW, SHORT-ROOT, e IDD. Identificar duplicaciones y determinar las homologías entre los genes. Determinar cambios en la secuencia proteica que puedan estar relacionados con variaciones en la función.
- Caracterización de la anatomía foliar: Re-evaluar en forma comparativa, y empleando conocimientos modernos, la anatomía foliar de especies de la subtribu Otachyriinae.
- Evolución de los patrones de expresión: Estudiar el programa de desarrollo de las hojas a través de la secuenciación masiva del ARN foliar de diversas especies de Otachyriinae. Identificar cambios en los patrones de expresión que se correlacionen con cambios anatómicos y subtipos fotosintéticos considerando la evolución de las especies.

CAPÍTULO 1

Evolución molecular de las subfamilias
de reguladores transcripcionales SHR,
SCR e IDD

INTRODUCCIÓN

Familia GRAS

Las proteínas GRAS son una familia de reguladores transcripcionales específicos de plantas cuyo nombre proviene de los tres primeros miembros identificados: GIBBERELLIN-INSENSITIVE (GAI), REPRESSOR of *ga1*–3 (RGA) y SCARECROW (SCR) [48]. Estos reguladores transcripcionales se caracterizan por poseer un dominio N-terminal variable e intrínsecamente desordenado y un dominio C-terminal altamente conservado (el dominio GRAS) [49]. El dominio GRAS está conformado por 5 motivos conservados, nombrados por sus residuos aminoacídicos más característicos: VHIID, LHRI y LHRII (Región Rica en Leucina 1 y 2), PFYRE y SAW [48]. Los motivos VHIID, LHRI y LHRII están involucrados en las interacciones proteína-proteína. El motivo VHIID (junto con PFYRE y SAW) es importante en la conformación estructural del dominio GRAS y su función ([50], [51]).

Algunas proteínas GRAS poseen motivos conservados adicionales en el N-terminal, tales como los presentes en la subfamilia DELLA. Las proteínas DELLA poseen dos motivos adicionales, denominados DELLA y TVHYNP, los cuales han sido caracterizados como esenciales para su función [49]. Por otro lado, el dominio N-terminal de las proteínas GRAS, por su característica de intrínsecamente desordenado, presenta MoRFs (del inglés: *molecular recognition features*), que son pequeños segmentos aminoacídicos que representan sitios potenciales de unión proteína-proteína y pueden actuar como carnada en eventos de reconocimiento molecular durante el desarrollo de las plantas. Esta diversidad en el dominio N-terminal le otorga a las GRAS un nivel de plasticidad que puede relacionarse con su versatilidad funcional [49].

Un gran número de proteínas GRAS fueron caracterizadas funcionalmente en varias especies. En particular, se ha observado que las proteínas GRAS están involucradas en procesos de crecimiento y desarrollo tales como: transducción de señales de giberelinas y fitocromos, iniciación del meristema axilar, mantenimiento del meristema apical, organización radial de la raíz, desarrollo foliar, gametogénesis, formación de nódulos en raíces, entre otros ([52], [53], [62]–[66], [54]–[61]).

Los primeros estudios filogenéticos de la familia, realizados en *Arabidopsis thaliana* y *Oryza sativa*, clasificaron a las GRAS en ocho familias de acuerdo con sus motivos conservados y similitud de secuencias: LILIUM SCARECROW-LIKE (LISCL), SHORTROOT (SHR), PHYTOCHROME A SIGNAL TRANSDUCTION 1 (PAT1), SCARECROW-LIKE 3 (SCL3), DELLA, SCR, LATERAL SUPPRESSOR (LS) y HAIRY MERISTEM (HAM) ([48], [67]). Recientemente, se han identificado y clasificado a los miembros de esta familia en varias especies vegetales como: *Solanum lycopersicum*,

Zea mays (maíz), *Populus trichocarpa*, *Brassica rapa*, *Medicago truncatula*, *Vitis vinifera*, *Pinus radiata*, *Setaria italica* y *Capsicum annuum*, entre otros ([61], [67]–[75]). El número de subfamilias en estas especies varía entre 8 y 16, lo cual sugiere eventos de duplicación génica específicos de cada especie. Muy pocos trabajos han tratado de abordar la evolución y diversificación de las GRAS con estudios filogenéticos abarcativos en los que se incluyan más de dos especies. Tanto Engstrom [76] como Cenci y Rouard [77], muestran que la diversificación de las subfamilias GRAS precede a las angiospermas y probablemente a las embriofitas con un número de entre 12 y 16 subfamilias.

La presencia de una señal de localización nuclear y una supuesta relación entre las GRAS y las proteínas STAT (del inglés: *Signal Transducer and Activator of Transcription*), llevaron a caracterizarlas como una familia de factores de transcripción convencional ([48], [78]). Sin embargo, Zhang y colaboradores (2012) postularon, a través de análisis filogenéticos y de modelado molecular, que las GRAS se originaron en bacterias y pertenecen a la superfamilia de metiltransferasas con plegamiento de Rossmann [79]. Mientras todas las GRAS bacterianas funcionan como metilasas de moléculas pequeñas, la mayoría de las GRAS de plantas han perdido uno o más residuos de unión a S-Adenosil metionina (SAM), lo cual no permitiría su función de metilasas. Estos modelos fueron comprobados por la resolución de la estructura cristalográfica del homodímero SCL7 en *O. sativa* y el heterotrímero (SCARECROW-SHORTROOT-JACKDAW (SCR-SHR-JKD) en *A. thaliana* ([50], [51]). Mientras que SHR y SCR forman un heterodímero en una configuración cabeza a cabeza que genera una carga superficial negativa que impediría su unión al ADN, SCL7 forma un homodímero en configuración cabeza con cola que deja expuesto un posible surco de unión al ADN [51]. Solamente SCL7 de *O. sativa* y NODULATION SIGNALING PATHWAY 1 (NSP1) de *M. truncatula* han sido reportadas como de unión directa al ADN por ensayos de EMSA, mientras que no hay evidencia de unión directa para el resto de los miembros de la familia ([50], [80]). Esta información sugiere que las GRAS actúan como reguladoras transcripcionales que, salvo excepciones, necesitan de la interacción con otros factores de transcripción para realizar su función.

Subfamilias SHR y SCR

Las subfamilias SHR y SCR fueron caracterizados originalmente como participantes claves, junto a miembros de la subfamilia de *zinc finger* IDD, en el desarrollo de la endodermis en la raíz [81]. Además, se ha observado que la pérdida de función de SHR o SCR afecta la diferenciación de las células de la vaina (CV) en las hojas y de la endodermis de hipocotilos y tallos florales ([57], [82]). Esto sugiere que, a nivel general, SHR y SCR regulan el desarrollo del patrón radial del tejido basal, tanto en raíces como en tallos y hojas. De hecho, las mutantes de *shr* y *scr* en *A. thaliana*, poseen hojas más pequeñas y un retraso general en el crecimiento atribuido a la pérdida de función de SHR y SCR en el

tejido foliar [83]. Esta idea fue propuesta por primera vez en la década de 1950 cuando Katherine Esau consideró, en base a observaciones anatómicas, que las CV en las hojas son un tipo de tejido homólogo a la endodermis de raíz [66].

Estos datos, sumados a la presencia del polímero suberina, tanto en la endodermis de la raíz como en las CV en especies con fotosíntesis C_4 , y el patrón de expresión de SCR en tejidos foliares en *Z. mays*, llevaron a Slewinski y colaboradores (2012) a investigar el rol de estos genes como determinantes de la identidad de CV y la anatomía foliar Kranz ([66], [84]–[86]). Posteriormente se demostró que en *A. thaliana*, SCARECROW-LIKE 23 (SCL23, parálogo de SCR) también está involucrado en determinar la identidad y función de las CV y la endodermis en hipocotilos. En particular, se ha observado que dicho proceso de diferenciación involucra la interacción entre SCR, SCL23, SHR y factores de transcripción de la subfamilia IDD ([64], [65], [87]).

Recientemente, fue reportado que los genes SCR de *Z. mays* son requeridos para establecer y/o mantener la capacidad fotosintética en las hojas ([88]). Además de defectos en la anatomía Kranz, estas mutantes fallan en organizar la endodermis en raíz, indicando que SCR en *Z. mays* regula el patrón celular tanto en hojas como en raíces. Sin embargo, las vías SHR-SCR de las hojas y las raíces se diferencian entre ellas por la capa en la que operan: mesófilo (M) en las hojas, versus la endodermis en raíces [89]. En conjunto, estos resultados muestran que la vía SHR-SCR representa un módulo regulatorio flexible que ha sido cooptado para modelar tipos celulares en un gran rango de contextos de desarrollo, tanto en raíces como en los tallos y hojas de las plantas con flores [88].

Subfamilia IDD

Los *Zinc-Fingers* (ZF) son pequeños motivos proteicos caracterizados por la coordinación de un átomo de zinc para formar una superestructura secundaria. La primera proteína *Zinc-Finger* (ZFP) descubierta fue el factor de transcripción TFIIA de *Xenopus laevis* [90]. Desde entonces, las ZFP han sido caracterizadas en eucariotas y procariotas [91]. El origen del motivo ZF es discutido, principalmente porque el mismo varía ampliamente en estructura, con al menos ocho grupos de plegamiento ([92], [93]). La diversidad en las estructuras de los ZF implica un amplio rango de funciones desde la unión a ADN y ARN a interacciones proteína-proteína y asociación a membranas ([92], [94]).

Uno de esos grupos de plegamiento, el dominio tipo ZF Cys2-His2 (C2H2), comprende a la mayoría de las ZFP agrupándolas en una de las familias de factores de transcripción más grande en Eukarya ([95], [96]). Los miembros de esta familia tienen entre uno y cuarenta ZF C2H2, en los cuales las cisteínas y/o histidinas están coordinando a un átomo de zinc para formar una estructura peptídica

que es requerida para sus funciones específicas [96]. El número y arreglos de ZF dentro de una proteína y los niveles de conservación varían muchísimo entre los reinos eucariotas ([91], [97]).

En plantas, Englbrecht y colaboradores (2004) clasificaron a las proteínas C2H2 en cuatro grupos dependiendo del número y la distancia entre los ZF. Dos de estos subgrupos, A1 y C1, parecen ser específicos de plantas y representan alrededor del 60% de todas las C2H2 de *A. thaliana* [98]. El grupo A1 incluye proteínas con cuatro ZF organizados en tándem, mientras que los miembros del grupo C1 poseen de uno a cinco dominios ZF dispersos entre sí [98].

La subfamilia de reguladores transcripcionales INDETERMINATE DOMAIN (IDD) es la más grande del subgrupo A1. El nombre de la subfamilia proviene del gen de maíz INDETERMINATE 1 (ID1), el primer gen caracterizado de la subfamilia [99]. Los miembros de esta subfamilia comparten un dominio N-terminal (el dominio IDD propiamente dicho) que incluye dos ZF C2H2 canónicos más dos ZF C2HC (ambos motivos con un largo de 25 aminoácidos y representados por la secuencia conservada XCX [1-5]-C-X3-X5-X2-HX[3-6]-H/C) y una señal de localización nuclear (NLS), seguido de un dominio C-terminal menos conservado [98]. Las funciones de algunos de los miembros de la subfamilia han sido caracterizadas especialmente en *A. thaliana*, además de *Z. mays* y *O. sativa*. En *A. thaliana* los IDDs estudiados están implicados en diversos procesos metabólicos y de desarrollo como: desarrollo radicular ([100]–[103]), gravitropismo del tallo ([104]–[106]), desarrollo foliar ([107]–[109]), tiempo de floración [110], maduración de semillas y señalización de hormonas ([106], [111], [112]). Recientemente, Völz y colaboradores (2019) mostraron que al menos un IDD de *A. thaliana*, además de poseer un rol en el crecimiento, está implicado en la regulación de la respuesta de defensa a patógenos hemibiotrofos [113]. En maíz, sólo tres IDDs han sido estudiados: INDETERMINATE1 (ID1), el regulador maestro en floración ([99], [114], [115]) y los genes *Naked Endosperm 1* y *2* (*NKD1* y *NKD2*) involucrados en el desarrollo de semillas ([116], [117]). En arroz, varios IDDs cumplen funciones en floración, incluyendo los ortólogos de ID1 ([118]–[120]). Otros IDDs de arroz son componentes importantes en procesos de gravitropismo y arquitectura de la planta [121], metabolismo de carbohidratos [122] y respuesta a estrés abiótico como la tolerancia a nitrógeno y al frío ([123], [124]).

Existe evidencia de que la subfamilia de IDDs actúa a través de la unión de los ZFs en el dominio IDD al surco principal del ADN reconociendo la secuencia consenso “TTTGTCGTATT” en la zona promotora de los genes blanco [125]. Sin embargo, también ha sido demostrado que los IDDs pueden unir al ADN cuando dicha secuencia consenso se encuentra modificada o ausente ([87], [117], [123]). Adicionalmente las ZFPs son capaces de formar homo y heterodímeros con proteínas de diferentes familias, especialmente con miembros de la familia de reguladores transcripcionales GRAS. Tanto el dominio IDD como los motivos C-terminales son importantes para esta interacción ([51], [112], [117]).

Existe evidencia que sugiere que la interacción entre las proteínas IDD y GRAS generan centros de señalización que le confieren la identidad a importantes tejidos vegetales como la endodermis de raíz, la vaina de almidón en los tallos, y las CV en las hojas C₄ [126]. En especial, basados en datos de expresión de IDD en células del M y la vaina, Coelho y colaboradores (2018) encontraron poblaciones divergentes de IDDs expresados en estos tipos celulares morfológica y fisiológicamente diferentes. Este hecho llevó a proponer que las interacciones entre dos proteínas GRAS (SHR/SCR) y diversos IDDs en las CV y del M podrían estar disparando dos procesos distintos de diferenciación celular necesarios para la compartimentalización de la maquinaria bioquímica de las hojas C₄.

A pesar de las diversas funciones en el desarrollo vegetal, sólo un puñado de estudios se han enfocado en la filogenia y la identificación de ortólogos entre los IDDs, SHRs y SCRs. Los estudios existentes se realizaron sobre la base de secuencias de uno o unos pocos genomas de las angiospermas ([76], [77], [98], [101], [126]–[128]). Es por ello que este capítulo aborda un estudio filogenético exhaustivo sobre la evolución molecular de los genes que codifican para las proteínas IDD, SHR y SCR en las plantas verdes (Viridiplantae).

OBJETIVOS ESPECÍFICOS

Reconstruir la evolución molecular de los genes SCARECROW, SHORT-ROOT, e IDDs: Identificar duplicaciones y determinar las homologías entre los genes. Determinar cambios en la secuencia proteica que puedan estar relacionados con variaciones en la función.

MATERIALES Y MÉTODOS

Obtención y filtrado de secuencias

Las secuencias proteicas fueron obtenidas de las bases de datos Phytozome v.12.1 [129] y PlantTFDB v.4.0 [130]. Se usaron secuencias anotadas de *A. thaliana* como moldes para una búsqueda BLASTp en cada uno de los 15 genomas estudiados en la división Streptophyta: *Amborella trichopoda* (AmTr), *A. thaliana* (At), *Brachypodium distachyon* (Bd), *Coffea canephora* (Cc), *Klebsormidium nitens* (Kn), *Marchantia polymorpha* (Mp), *O. sativa* (Os), *Physcomitrella patens* (Pp), *Phoenix dactylifera* (Pd),

Picea abies (Pa), *Selaginella moellendorffii* (Sm), *Setaria italica* (Si), *Theobroma cacao* (Tc), *Vitis vinifera* (Vv) y *Z. mays* (Zm). El primer resultado de cada búsqueda por BLASTp fue usado para una nueva búsqueda BLASTp intragenómica usando una matriz BLOSUM82. Para eliminar secuencias incompletas o demasiado divergentes, las secuencias fueron escaneadas usando MEME v.4.10 (*Multiple EM for motif elicitation* [131]). La búsqueda se realizó con un máximo de 10 motivos y parámetros por defecto. Las secuencias que no contenían al menos dos motivos compartidos fueron descartadas.

Reconstrucción filogenética

Las matrices de secuencias primarias para la reconstrucción en las embriofitas de las subfamilias SHR y SCR están compuestas por 164 y 162 potenciales miembros, respectivamente.

Para los IDD, debido a la dificultad de obtener buenos alineamientos con un gran número de secuencias, se optó por seguir una metodología de análisis diferente. Para ello se decidió comenzar por reconstruir la filogenia de las angiospermas para después ir ampliando el muestreo a otros genomas hasta encontrar el origen de la subfamilia. La matriz de datos primaria de los IDD de las angiospermas está compuesta por 267 potenciales IDD más cuatro secuencias de *A. thaliana* que se usaron como *outgroups* (C2H2 no IDD: At3g48430, AtREF6; At2g23740, AtSUVR5; At4g06634, AtYY1; y At1g72050, AtTF3A).

Las secuencias proteicas se alinearon con MAFFT [132] a través de la versión online usando una estrategia E-INS-i. Los bloques conservados fueron extraídos manualmente utilizando MEGA7 [133]. Los árboles filogenéticos fueron construidos con MrBayes 3.2.6 [134] en la plataforma CIPRES Science Gateway [135] usando un modelo de sustitución fija y una corrida de 80 millones de generaciones (muestreando cada 4000 generaciones). Cuatro cadenas de Markov fueron corridas en simultáneo en dos análisis independientes empezando con un árbol aleatorio hasta que el diagnóstico de convergencia (desviación estándar de la separación de secuencias) cayó a un valor por debajo de 0,01. Las primeras 20 millones de generaciones fueron descartadas (25%) y el resto de las muestras provenientes de dos réplicas fueron combinadas. Un árbol consenso de 30.000 árboles fue construido y visualizado usando Mesquite [136].

A partir de la reconstrucción filogenética inicial, se identificó: (a) un grupo monofilético que incluye 70 secuencias proteicas de SHRs de embriofitas, (b) un grupo monofilético que incluye 38 secuencias proteicas de SCRs de embriofitas y (c) 163 secuencias proteicas de IDD en las angiospermas estudiadas. En el caso de los IDD, debido a la baja resolución interna del árbol, se decidió volver a realizar una reconstrucción filogenética incluyendo únicamente las 163 secuencias proteicas de IDD identificadas previamente, siguiendo la metodología ya descrita. Los linajes IDD se

definieron considerando las separaciones entre monocotiledóneas y dicotiledóneas como referencia para relaciones de ortología. Los linajes fueron nombrados de acuerdo al primer gen o proteína estudiada en trabajos previos [126]. En linajes donde ningún gen o proteína había sido estudiado se usó el nombre asignado al gen de *A. thaliana* para nombrar al linaje. Para los linajes que además carecían de genes de *A. thaliana*, se propusieron nuevos nombres.

Para la reconstrucción de la filogenia IDD en la división Streptophyta se empleó una matriz de datos inicial con 219 secuencias (163 secuencias IDD ya identificadas de las angiospermas, más 56 C2H2 no angiospermas) más seis secuencias C2H2 no IDD de *A. thaliana* como *outgroups* (At1g34370, AtSTOP1; At1g51220, AtWIP; At3g48430, AtREF6; At2g23740, AtSUVR5; At4g06634, AtYY1; y At1g72050, AtTF3A). La reconstrucción filogenética se realizó siguiendo la metodología descrita arriba. Como resultado de esta reconstrucción filogenética, un total de 20 secuencias proteicas de IDD fueron confirmadas en las especies no-angiospermas. Finalmente, y en base a estos resultados, un segundo análisis fue llevado a cabo utilizando las 183 secuencias IDD de la división Streptophyta más una secuencia C2H2 no IDD de *K. nitens* como *outgroup* (kfl00223_0040, KfC2H2f).

Para confirmar el origen de la subfamilia IDD al ancestro común de la división Streptophyta realizamos una búsqueda BLASTp en 23 genomas de la división Viridiplantae: *A. trichopoda* (AmTr), *Auxenochlorella protothecoides* (Ap), *Bathycoccus prasinos* (Bp), *Chara braunii* (Cb), *Chlamydomonas reinhardtii* (Cr), *Chlorella variabilis NC64A* (Cv), *Coccomyxa subellipsoidea C-169* (Cs), *Dunaliella salina* (Ds), *Gonium pectorale* (Gp), *Helicosporidium sp.* (Hsp), *K. nitens* (Kn), *M. polymorpha* (Mp), *Micromonas pusilla CCMP1545* (Mpu), *Micromonas sp. RCC299* (Msp), *Monoraphidium neglectum* (Mn), *Ostreococcus lucimarinus* (Ol), *Ostreococcus sp. RCC809* (Osp), *Ostreococcus tauri* (Ot), *P. patens* (Pp), *P. abies* (Pa), *Picochlorum sp. SENEW3* (Ps), *S. moellendorffii* (Sm) y *Volvox carteri* (Vc). Para las búsquedas BLASTp se utilizó la secuencia del gen que codifica para la proteína IDD de *K. nitens* previamente identificada y las cuatro secuencias peptídicas de IDD de *A. thaliana* usadas anteriormente (At3g50700, At5g03150, At1g68130, At1g34370). Las secuencias peptídicas se alinearon con MAFFT (Kato y col., 2017) usando la versión online y la estrategia E-INS-i. Los bloques conservados se extrajeron con MEGA7 [133]. Los árboles filogenéticos fueron reconstruidos como se describió anteriormente.

Basados en los árboles generados más arriba usando las 183 secuencias IDD de la división Streptophyta, se construyeron matrices de datos individuales para cada uno de los 10 linajes IDD identificados. Cada uno incluyó las secuencias propias de cada linaje más dos secuencias de linajes vecinos como *outgroups*. Los alineamientos y la reconstrucción filogenética se realizaron como se describió anteriormente.

Análisis de motivos conservados

Las secuencias pertenecientes a las subfamilias SHR y SCR fueron escaneadas con MEME v.4.12 [131] empleando una búsqueda de hasta 15 motivos entre 6 y 40 aa y parámetros por defecto. Dada las limitantes del programa para procesar grandes conjuntos de datos, para llevar a cabo los análisis de motivos sobre las IDD, se decidió trabajar con dos matrices de secuencias proteicas parciales identificadas en Streptophyta: (a) la matriz “IDD Streptophyta N-ter”, conteniendo todas las secuencias desde el codón de inicio hasta el final del cuarto ZF; y (b) la matriz “IDD Streptophyta C-ter”, conteniendo todas las secuencias desde el final del cuarto ZF hasta el codón de terminación. Cada matriz fue escaneada con MEME v.4.12 con una búsqueda de hasta 15 motivos entre 6 y 40 aa y parámetros por defecto. Para detectar nuevos motivos característicos de los distintos linajes de IDD, se realizó un análisis adicional empleando MEME con las secuencias completas de cada linaje IDD descrito. La búsqueda incluyó un máximo de 12 motivos de 6 a 40 aa de largo y parámetros por defecto. Sólo los motivos con un valor-E menor a $1e-50$ fueron incluidos. Para cada linaje, se construyó un esquema de cajas para mostrar la distribución y frecuencia de ocurrencia de los motivos específicos.

RESULTADOS

Evolución de la familia GRAS

Evolución de la subfamilia SHR

Para identificar los miembros de la subfamilia SHR en embriofitas, se llevó a cabo una reconstrucción filogenética utilizando 689 residuos alineados de 155 secuencias candidatas GRAS obtenidas de la búsqueda BLASTp (Tabla 1). El árbol resultante muestra que la subfamilia SHR se resuelve como un clado monofilético de 70 secuencias [Probabilidad posterior (PP)=1] dentro de la familia GRAS (Figura 7). Las secuencias más cercanas a la subfamilia corresponden a miembros de los linajes LISCL y PAT1. La subfamilia SHR está conformada por miembros de todas las especies analizadas (Tabla 1).

La reconstrucción filogenética realizada indica que la subfamilia SHR puede dividirse en tres linajes monofiléticos los que se originaron a partir de dos duplicaciones ancestrales: (a) un linaje que agrupa a todas las secuencias proteicas de SHR y (b) un linaje que incluye un grupo monofilético integrado por las secuencias de NSP1, hermano a un grupo monofilético que incluye a las secuencias proteicas de SCL32. El linaje SHR, incluye secuencias proteicas de *S. moellendorffii* y *P. patens* y de

todas las angiospermas estudiadas. A partir del análisis de la topología del árbol obtenido se observa que el linaje SHR está dividido en dos subclados monofiléticos: (a) un subclado, aquí denominado SHR1, que agrupa secuencias de angiospermas entre las que se destacan las secuencias de *A. thaliana* y *O. sativa* estudiadas funcionalmente y dos subgrupos de secuencias proteicas de gramíneas, y (b) un subclado, aquí denominado SHR2, que incluye secuencias proteicas exclusivamente de eudicotiledóneas excepto *P. dactylifera* y *A. trichopoda*.

Los linajes NSP1 y SCL32 incluyen secuencias proteicas pertenecientes a todas las especies estudiadas. La topología del árbol de consenso muestra que el linaje NSP1 es de copia única a excepción de las especies *P. patens*, y dos miembros representantes del PACMAD (*S. viridis* y *Z. mays*) las que presentan dos, tres y cuatro copias respectivamente. El linaje SCL32 presenta un evento de duplicación alrededor del origen de las espermatófitas, que generó dos linajes: SCL32-a, con miembros en todas las especies de angiospermas estudiadas; y SCL32-b el cual fue retenido en la gimnosperma *P. abies* y se encuentra presente en todas las angiospermas a excepción de *A. thaliana*.

La búsqueda de motivos conservados arrojó sólo motivos pertenecientes al dominio GRAS.

Tabla 1. Secuencias de SHR identificadas en genomas de embriofitas. Número de secuencias peptídicas GRAS de 14 genomas de embriofitas recuperadas después de una búsqueda BLASTp, un análisis MEME (Múltiple EM para elucidar motivos) y una validación final a partir de la reconstrucción filogenética.

Especies	N° Secuencias candidatas	N° Secuencias confirmadas
<i>Marchantia polymorpha (Mp)</i>	8	2
<i>Physcomitrella patens (Pp)</i>	15	9
<i>Selaginella moellendorffii (Sm)</i>	15	5
<i>Picea abies (Pa)</i>	9	2
<i>Amborella trichopoda (AmTr)</i>	9	5
<i>Arabidopsis thaliana (At)</i>	8	4
<i>Coffea canephora (Cc)</i>	10	6
<i>Theobroma cacao (Tc)</i>	10	5
<i>Vitis vinifera (Vv)</i>	10	3
<i>Phoenix dactylifera (Pd)</i>	11	5
<i>Brachypodium distachyon (Bd)</i>	12	4
<i>Oryza sativa (Os)</i>	9	5
<i>Setaria italica (Si)</i>	12	6
<i>Zea mays (Zm)</i>	17	9
Total	155	70

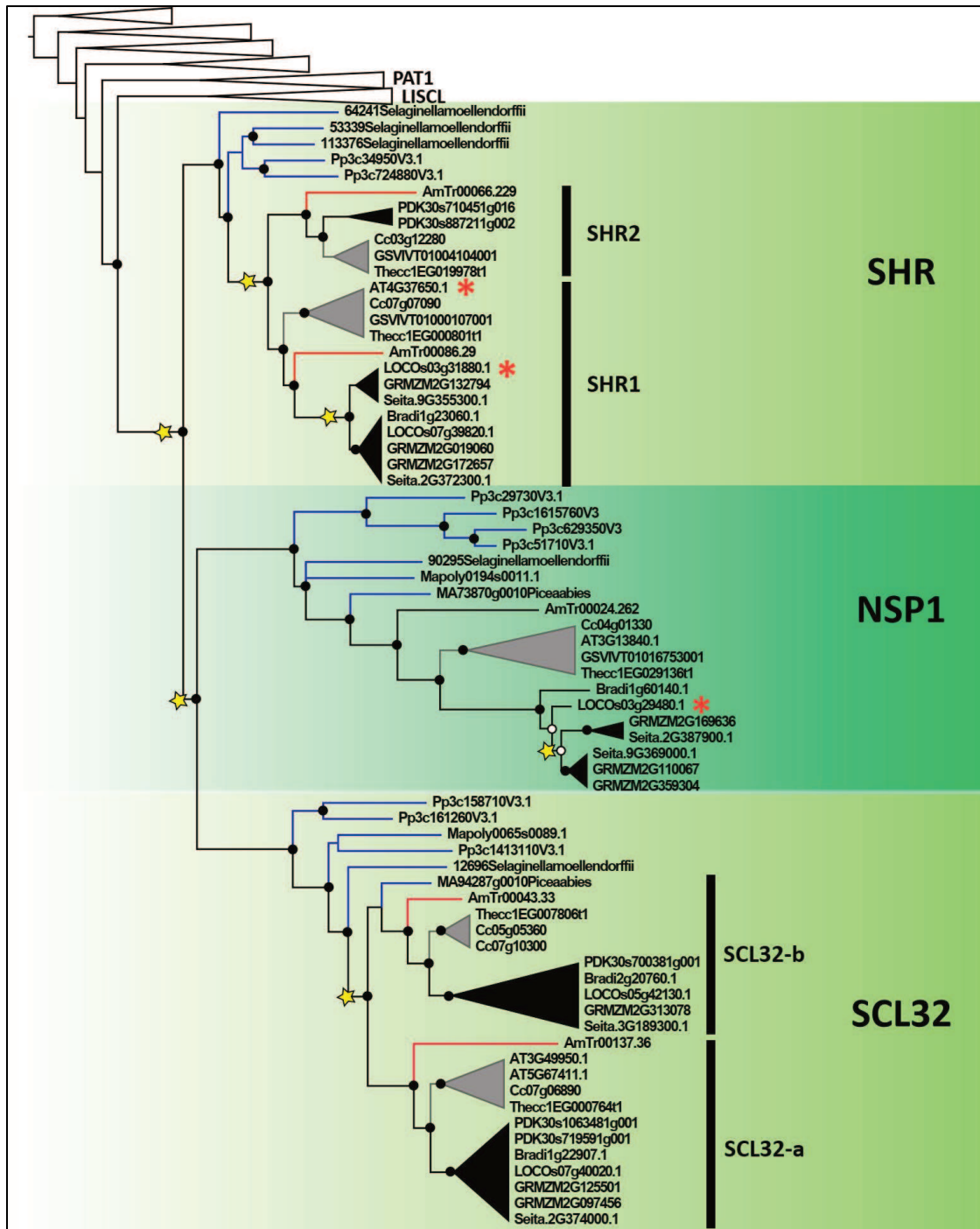


Figura 7. Evolución SHR en las embriofitas. Árbol de consenso de mayoría (N=30.000 árboles) de la subfamilia SHR usando una matriz de 155 secuencias peptídicas. Los puntos blancos indican una Probabilidad Posterior (PP) = [0,8 a 0,89], los puntos negros indican PP= [0.9 a 1]. Los asteriscos naranjas señalan a los miembros de la subfamilia que han sido estudiados funcionalmente. Las estrellas amarillas indican eventos de duplicación. Los clados de monocotiledóneas y dicotiledóneas fueron colapsados en negro y gris, respectivamente. Las secuencias pertenecientes a otras subfamilias GRAS fueron colapsadas en blanco. Las líneas rojas y azules indican secuencias de *A. trichopoda* y de especies no-angiospermas, respectivamente.

Evolución de la subfamilia SCR

Para identificar los miembros SCR de las embriofitas, realizamos una reconstrucción filogenética empleando 688 residuos alineados pertenecientes a 161 secuencias candidatas GRAS (Tabla 2). El árbol resultante muestra que 38 de estas secuencias se agrupan en un clado monofilético [Probabilidad posterior (PP)=1] dentro de la familia GRAS (Figura 8). Las secuencias más cercanas a la subfamilia SCR corresponden a miembros del linaje DLT (*Dwarf and Low Tillering*) al igual que lo reportado por Cenci y Rouard [77]. La subfamilia SCR está conformada por secuencias proteicas pertenecientes a la mayoría de las especies analizadas. No se han identificado secuencias proteicas de SCR en el genoma de *M. polymorpha* (Tabla 2).

Tabla 2. Secuencias de SCR identificadas en genomas de embriofitas. Número de secuencias peptídicas GRAS de 14 genomas de embriofitas recuperadas después de: una búsqueda BLASTp, un análisis MEME (Múltiple EM para elucidar motivos) y una validación final a partir de la reconstrucción filogenética.

Especies	N° Secuencias candidatas	N° Secuencias confirmadas
<i>Marchantia polymorpha</i> (Mp)	10	0
<i>Physcomitrella patens</i> (Pp)	12	3
<i>Selaginella moellendorffii</i> (Sm)	10	2
<i>Picea abies</i> (Pa)	10	3
<i>Amborella trichopoda</i> (AmTr)	12	4
<i>Arabidopsis thaliana</i> (At)	12	2
<i>Coffea canephora</i> (Cc)	12	3
<i>Theobroma cacao</i> (Tc)	12	3
<i>Vitis vinifera</i> (Vv)	10	1
<i>Phoenix dactylifera</i> (Pd)	13	2
<i>Brachypodium distachyon</i> (Bd)	11	3
<i>Oryza sativa</i> (Os)	12	4
<i>Setaria italica</i> (Si)	12	4
<i>Zea mays</i> (Zm)	13	4
Total	161	38

La reconstrucción filogenética realizada indica que la subfamilia SCR puede dividirse en tres linajes monofiléticos que se originaron a partir de dos duplicaciones ancestrales: por un lado, SCL23 y por otro SCR1 y 2. El linaje SCL23, está integrado por secuencias proteicas de angiospermas y es de copia única a excepción de *A. trichopoda* cuyo genoma presenta dos copias. El linaje SCR está dividido en dos sub clados monofiléticos: (a) un subclado aquí denominado SCR1, que agrupa secuencias de

todas las especies presentes en la subfamilia y presenta duplicaciones recientes en gramíneas y en *P. patens*; y (b) un subclado aquí denominado SCR2, que incluye secuencias proteicas de las angiospermas y gimnospermas estudiadas a excepción de *A. thaliana*, *V. vinifera* y *P. dactylifera*. Los SCR estudiados funcionalmente en *A. thaliana*, *O. sativa* y *Z. mays*, pertenecen al clado SCR1.

Además de los motivos conservados que constituyen el dominio GRAS en el C-terminal (rango de valores E entre $1,6e-975$ y $1,9e-90$), la búsqueda de motivos conservados arrojó la presencia de un motivo específico del linaje SCR1 (SCR1-M12), en el N-terminal que está presente en el 70% de las secuencias. Este motivo, posee un largo de 40 aa y se caracteriza por la secuencia consenso “[ID][GILM][SIM][HS][SG][GS][AP][GAN]V[SP][IV][APE]Q[LI][HQ]N[VL]REI[IL]HPCNP[GHN]LA[SA][LI][LI]E[LY]RLRSL” (valor E = $1,3e-209$).

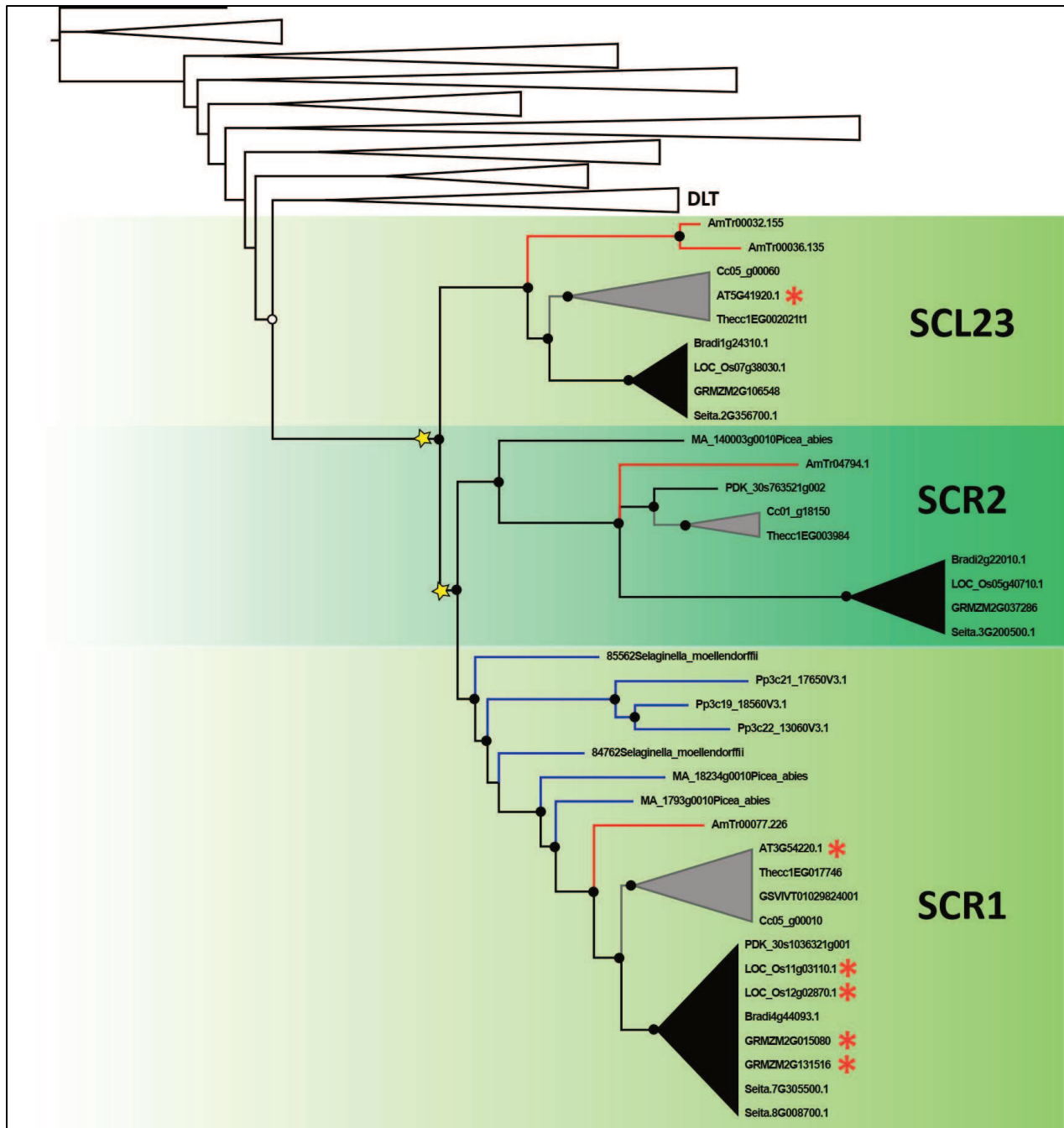


Figura 8. Evolución de SCR en las embriofitas. Árbol de consenso de mayoría ($N=30.000$ árboles) de la subfamilia SCR usando una matriz de 161 secuencias peptídicas. Los puntos blancos indican una Probabilidad Posterior (PP) = [0,8 a 0,89], los puntos negros indican PP= [0.9 a 1]. Los asteriscos naranjas señalan a los miembros de la subfamilia que han sido estudiados funcionalmente. Las estrellas amarillas indican eventos de duplicación. Los clados de monocotiledóneas y dicotiledóneas fueron colapsados en negro y gris, respectivamente. Las secuencias pertenecientes a otras subfamilias GRAS fueron colapsadas en blanco. Las líneas rojas y azules indican secuencias de *A. trichopoda* y de especies no-angiospermas, respectivamente.

Evolución de la subfamilia IDD

Evolución en las angiospermas

Para identificar los miembros IDD de las angiospermas, se realizó una reconstrucción filogenética usando 1218 residuos alineados provenientes de 267 secuencias candidatas C2H2 (Tabla 3). El árbol resultante muestra que 163 de estas secuencias se agrupan en un clado monofilético [Probabilidad posterior (PP) = 0,97] dentro de la familia C2H2 (Anexo 1, Figura S1). Se observa que, *A. trichopoda* presentó el menor número de genes (9) y las monocotiledóneas *P. dactylifera* y *Z. mays* presentaron los mayores números (24 y 22, respectivamente) (Tabla 3).

Tabla 3. Secuencias de IDD identificadas en genomas de angiospermas. Número de secuencias peptídicas C2H2 de 10 genomas de angiospermas recuperadas después de: una búsqueda BLASTp, un análisis MEME (Múltiple EM para elucidar motivos) y una validación final a partir de la reconstrucción filogenética.

Especies	Nº Secuencias candidatas	Nº Secuencias confirmadas
<i>Amborella trichopoda</i> (AmTr)	15	9
<i>Arabidopsis thaliana</i> (At)	24	16
<i>Coffea canephora</i> (Cc)	23	16
<i>Theobroma cacao</i> (Tc)	24	15
<i>Vitis vinifera</i> (Vv)	23	15
<i>Phoenix dactylifera</i> (Pd)	35	24
<i>Brachypodium distachyon</i> (Bd)	28	16
<i>Oryza sativa</i> (Os)	28	15
<i>Setaria italica</i> (Si)	27	15
<i>Zea mays</i> (Zm)	40	22
Total	267	163

Para estudiar la evolución de las angiospermas en mayor profundidad se reconstruyó una nueva filogenia con 966 residuos alineados pertenecientes a las 163 secuencias IDD validadas. El resultado de este análisis muestra que, en las angiospermas, los IDD pueden dividirse en ocho linajes monofiléticos integrados por al menos una secuencia peptídica previamente caracterizada (Figura 9): SHOOT GRAVITROPISM 5 (SG5), GAI INTERACTING FACTOR 1 (GAF1), INDETERMINATE 1 (ID1), JACKDAW (JKD), MAGPIE (MGP), NAKED ENDOSPERM (NKD), WARBLER (WRB) y WOODPECKER (WDP); y 2 linajes adicionales, que incluyen proteínas de funciones desconocidas y que están ausentes en el genoma de *A. thaliana*. Dado que estos dos últimos linajes se agrupan entre los linajes conocidos con nombres de aves (*birds*), propusimos aquí nombrarlos como CONDOR-a

(CNDa) y CONDOR-b (CNDb) (Figura 9). La mayoría de estos linajes contiene secuencias de dicotiledóneas y monocotiledóneas, excepto por los linajes ID1 y JKD que son específicos de monocotiledóneas y dicotiledóneas respectivamente. Aunque ambas, dicotiledóneas y monocotiledóneas, poseen un número similar de IDD en sus genomas, éstas últimas presentaron un mayor número de clados, sugiriendo una mayor cantidad de eventos de duplicación y retención ancestrales (17 clados de monocotiledóneas y 10 clados de dicotiledóneas) (Tabla 3; Figura 9).

Análisis en la división Streptophyta

La presencia de numerosas secuencias IDD pertenecientes a *A. trichopoda* en el árbol de las angiospermas sugiere que el origen de la subfamilia precede a las plantas con flores. Por este motivo, para investigar el origen de la subfamilia IDD, se extendió el análisis al subreino Viridiplantae. Para ello, se incluyeron secuencias adicionales de los genomas de la gimnosperma *P. abies*, la licofita *S. moellendorffii*, el musgo *P. patens* y la hepática *M. polymorpha*. Además, se incluyeron tres secuencias C2H2 identificadas en el genoma de la carofita *K. nitens*.

La reconstrucción filogenética se llevó a cabo empleando un total de 1190 residuos alineados pertenecientes a 219 secuencias peptídicas candidatas recuperadas luego de curar la matriz de secuencias con un análisis MEME. La topología del árbol filogenético muestra que 20 de las 56 secuencias candidatas C2H2 no-angiospermas pertenecen a la subfamilia IDD, incluyendo una secuencia de *K. nitens* (Anexo 1, Figura S2). En total, la subfamilia IDD consta de 183 secuencias en los genomas de las estreptofitas estudiadas (Tabla 4).

Basados en estos resultados, se construyó un nuevo árbol filogenético para representar la evolución de los IDD en la división Streptophyta empleando un total de 1258 residuos pertenecientes a las 184 secuencias (Anexo 1, Figura S3). La topología del árbol consenso mantiene la estructura en 10 linajes como se describió previamente para las angiospermas; sin embargo, las relaciones evolutivas internas están débilmente soportadas. Las nuevas secuencias de genomas no-angiospermas están distribuidas de manera asimétrica entre los linajes, excepto la única secuencia proveniente de *K. nitens* (KfIDD2L, kfl002320130) que se presenta como hermana a todas las demás IDD.

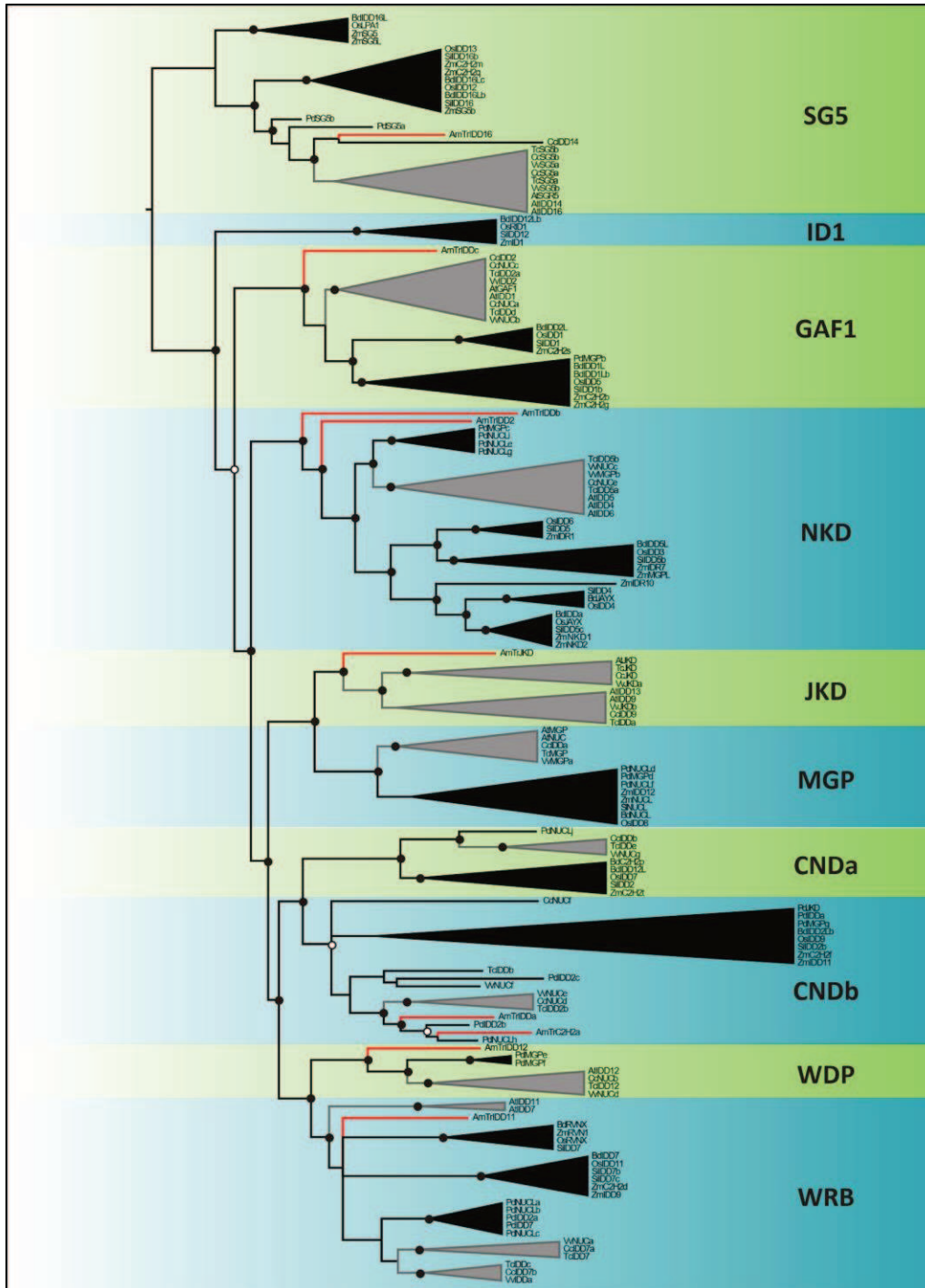


Figura 9. Evolución de los IDD en las angiospermas. Árbol de consenso de mayoría (N=30.000 árboles) de la subfamilia IDD usando una matriz de 163 secuencias peptídicas. Los puntos blancos indican una Probabilidad Posterior (PP) = [0,8 a 0,89], los puntos negros indican PP= [0.9 a 1]. Los clados de monocotiledóneas y dicotiledóneas fueron colapsados en negro y gris, respectivamente. Las secuencias de *A. trichopoda* se muestran en rojo. Abreviaturas: SHOOT GRAVITROPISM 5 (SG5), INDETERMINATE 1 (ID1), GAI INTERACTING

FACTOR 1 (GAF1), NAKED ENDOSPERM (NKD), JACKDAW (JKD), MAGPIE (MGP), CONDOR a (CNDa), CONDOR b (CNDb), WOODPECKER (WDP), WARBLER (WRB).

Tabla 4. Secuencias de IDD identificadas en las estreptofitas. Número de secuencias peptídicas C2H2 de 15 genomas de estreptofitas recuperadas luego de: una búsqueda BLASTp, un análisis MEME (Múltiple EM para elucidar motivos) y una validación final a partir de la reconstrucción filogenética.

Especies	N° Secuencias candidatas	N° Secuencias confirmadas
<i>Klebsormidium nitens (Kn)</i>	3	1
<i>Marchantia polymorpha (Mp)</i>	8	3
<i>Physcomitrella patens (Pp)</i>	19	6
<i>Selaginella moellendorffii (Sm)</i>	11	5
<i>Picea abies (Pa)</i>	16	5
<i>Amborella trichopoda (AmTr)</i>	15	9
<i>Arabidopsis thaliana (At)</i>	24	16
<i>Coffea canephora (Cc)</i>	23	16
<i>Theobroma cacao (Tc)</i>	24	15
<i>Vitis vinifera (Vv)</i>	23	15
<i>Phoenix dactylifera (Pd)</i>	35	24
<i>Brachypodium distachyon (Bd)</i>	28	16
<i>Oryza sativa (Os)</i>	28	15
<i>Setaria italica (Si)</i>	27	15
<i>Zea mays (Zm)</i>	40	22
Total	324	183

Origen de la subfamilia IDD

Para obtener una mayor evidencia sobre el origen de los IDD, realizamos una reconstrucción filogenética a partir de 769 residuos alineados de 141 secuencias candidatas pertenecientes a la familia C2H2 en genomas de las divisiones Streptophyta y Chlorophyta, obtenidas por búsquedas de BLASTp. No se han obtenido secuencias candidatas en 12 de los 16 genomas de clorófitas analizados (Figura 10).

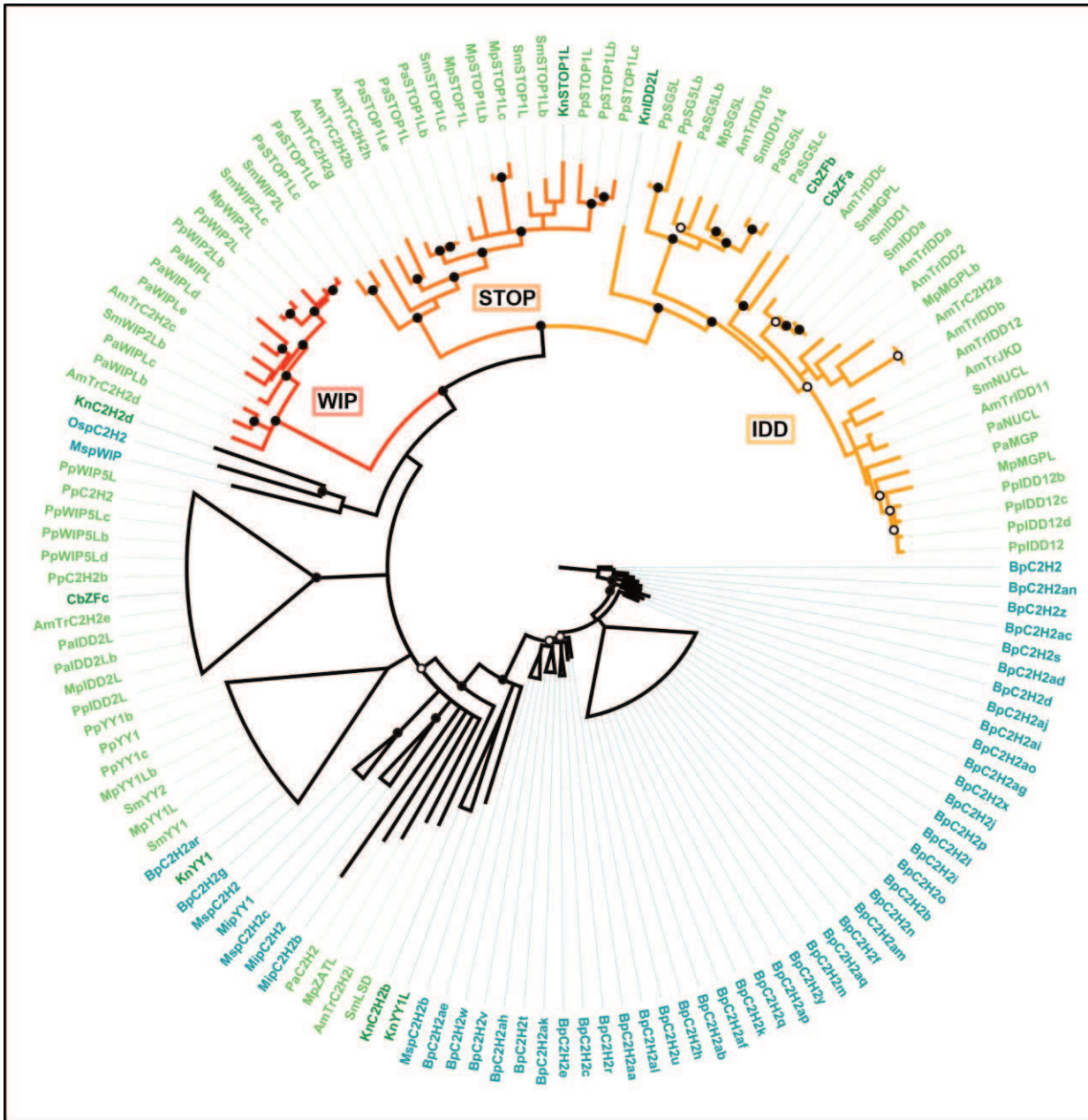


Figura 10. Origen de la subfamilia IDD. Árbol consenso (N=30.000 árboles) de 141 secuencias peptídicas C2H2 provenientes de *A. trichopoda* (AmTr), *P. abies* (Pa), *S. moellendorffii* (Sm), *P. patens* (Pp), *M. polymorpha* (Mp), *C. braunii* (Cb), *K. nitens* (Kn), *M. pusilla* (Mip), *B. prasinos* (Bp), *M. sp. RCC299* (Msp) y *O. sp. RCC809* (Osp). Los puntos blancos indican una Probabilidad Posterior (PP)= [0,8 a 0,89], los puntos negros indican PP= [0,9 a 1]. Los nombres de las secuencias pertenecientes a las divisiones Chlorophyta, Charophyta y Embryophyta están resaltados en turquesa, verde oscuro y verde claro, respectivamente.

La topología del árbol consenso muestra sólo una secuencia de *K. nitens* y dos de *C. braunii* dentro del clado perteneciente a los IDD. Todas las secuencias candidatas en la división Chlorophyta

se ubicaron por fuera del linaje de IDD. Adicionalmente el clado STOP resultó hermano a los IDD e incluye una secuencia de *K. nitens* (Kfl004330010).

Diversificación de motivos conservados

Se realizó un análisis de secuencias basado en motivos conservados para determinar las características específicas propias de la subfamilia IDD y sus linajes. Los 25 motivos encontrados están listados en la Tabla S1 (Anexo 1). Los IDD se caracterizan por poseer 6 motivos conservados HGC (del inglés: *High Global Coverage Motifs*) que corresponden a los cuatro ZF (HGC-M1 a 4), una NLS (HGC-M6) y un motivo de función desconocida (HGC-M5), que puede interpretarse como el fragmento N-terminal del primer ZF. El motivo HGC-M4, conformado por el denominado *SHR binding motif* (SHBM) [51], está presente en todas las secuencias excepto en el linaje SG5 y en la secuencia IDD de *K. nitens*. En el 23% de los IDD la señal de localización nuclear se encuentra ausente o es lo suficientemente divergente para no ser considerada como HGC-M6 en el estudio (Anexo 1, Tabla S1).

El análisis global detectó cuatro motivos MGC (del inglés: *Medium Global Coverage Motifs*) que son compartidos por la mayoría de los linajes (Anexo 1, Tabla S1). El motivo MGC-16 (conocido como "SATAL"; [112]) surgió luego de la divergencia entre los genes ID1 y fue retenido en al menos el 70% de las secuencias muestreadas. Los motivos MGC-M7 (que también puede interpretarse como el fragmento C-terminal del cuarto ZF) y MGC-M17 se originaron a partir de la división de ID1 y fue retenido en el 70% de las secuencias analizadas. Los motivos MGC-M16 y MGC-M17 están reportados como necesarios para las interacciones proteína-proteína con los reguladores transcripcionales DELLA y SCL3, ambos miembros de la familia GRAS ([111], [112], [137]). El motivo MGC-M9, el cual surgió en la base de la subfamilia, fue retenido en el 50% de las secuencias IDD. Se identificaron también 15 motivos auxiliares LGC (del inglés: *Low Global Coverage Motifs*) que son compartidos por unos pocos linajes. Si bien la mayoría de estos motivos poseen una función desconocida, la presencia de la secuencia proteica LXLXL en el LGC-M25 sugiere una función represora del tipo EAR para este segmento proteico. Este motivo está presente en el IDD de *K. nitens*, el linaje SG5 y en una secuencia miembro del linaje ID1.

Estudio de los linajes

Para obtener más información sobre la evolución de la subfamilia y poder establecer mejores relaciones de ortología entre las secuencias de las especies estudiadas, se construyeron nuevas filogenias y se llevaron a cabo búsquedas de motivos adicionales para cada uno de los 10 linajes (Figuras 11 y 12, Tabla 5).

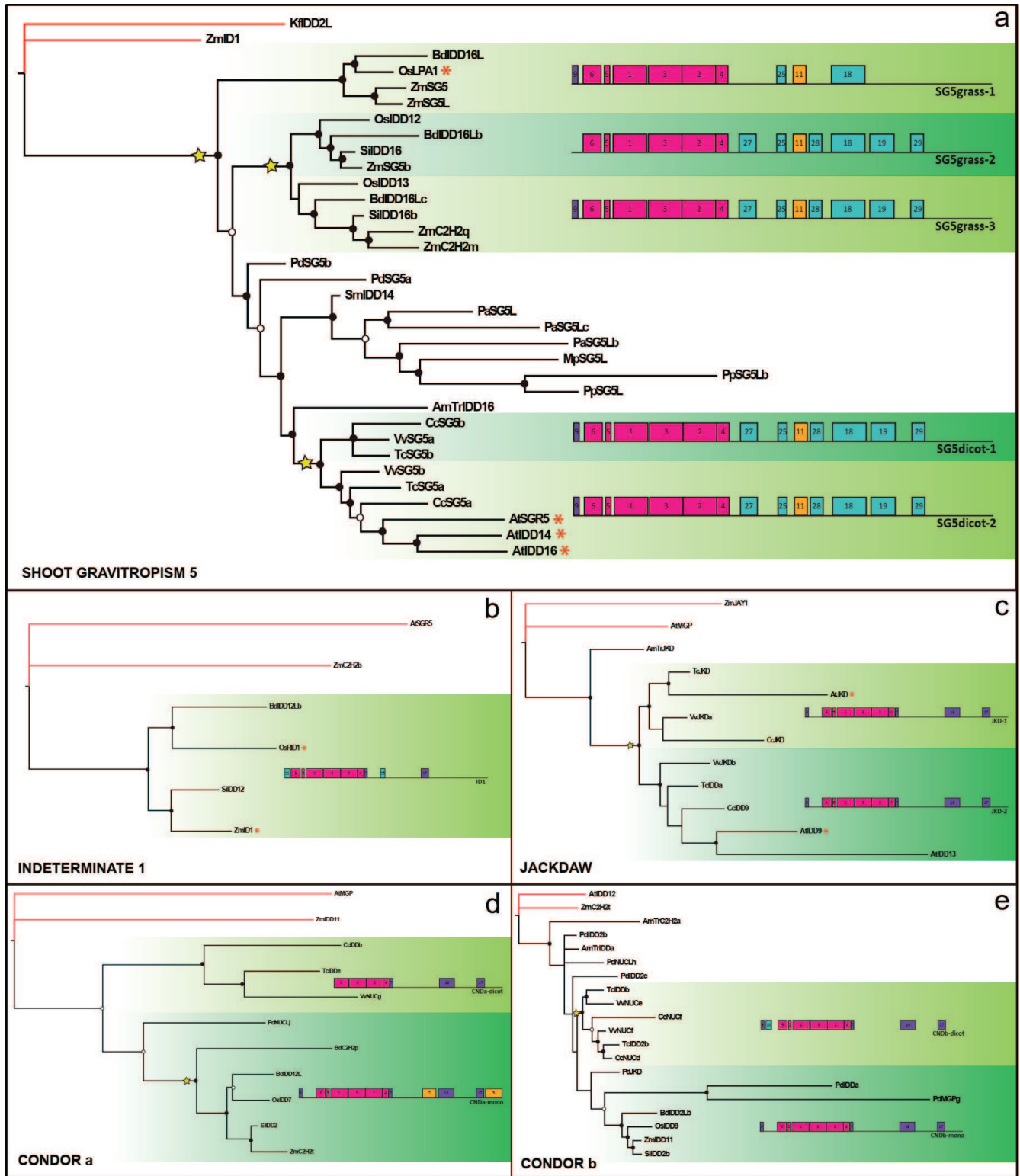


Figura 11. Diversificación de los linajes SG5, ID1, JKD y CND en streptophyta. Árboles consenso ($N=30.000$ árboles) de los linajes SG5 (a), ID1 (b), JKD (c), CNDa (d) y CNDb (e). Los puntos blancos indican una Probabilidad Posterior Bayesiana (PP) = [0,8 a 0,89], los puntos negros indican PP = [0,9 a 1]. Las secuencias

outgroups se muestran en rojo, las estrellas amarillas indican eventos de duplicación génica y los asteriscos naranjas señalan los genes con función conocida. Los motivos identificados en cada clado están representados en esquemas de cajas sobre la derecha: HGC, rosado; MGC, violeta; LGC, celeste y los específicos SG5 y CND en naranja.

El linaje SG5, reconstruido a partir de un alineamiento de 503 residuos provenientes de 34 secuencias, incluye a miembros de plantas terrestres sin semillas, gimnospermas y angiospermas. Este linaje se ha diversificado a través de varios eventos de duplicación luego de la separación entre dicotiledóneas y monocotiledóneas (Figura 11 a). Al menos dos duplicaciones ocurrieron durante la evolución de las gramíneas y una durante la evolución de las dicotiledóneas. Como resultado de estas duplicaciones se encontraron tres linajes específicos de gramíneas (SG5grass-1 a 3) y dos específicos de dicotiledóneas (SG5dicot-1 y 2). La búsqueda de motivos específicos de linaje identificó un motivo adicional (SG5-M11) conservado a lo largo del linaje (Tabla 5).

El linaje ID1 fue reconstruido a partir de un alineamiento de 525 residuos pertenecientes a 6 secuencias. El mismo está representado sólo por secuencias de gramíneas, es consistente con la evolución de las especies y contiene genes de copia única (Figura 11 b). No se detectaron motivos específicos adicionales.

El linaje JKD, reconstruido a partir de un alineamiento de 257 residuos pertenecientes a 12 secuencias, incluye unade *A. trichopoda* hermana a secuencias de dicotiledóneas, sugiriendo una pérdida de estos genes en las monocotiledóneas (Figura 11 c). La topología del árbol muestra una duplicación que pudo haber ocurrido en el origen de las dicotiledóneas. No se encontraron motivos específicos adicionales en este linaje.

El linaje CNDa, reconstruido a partir de un alineamiento de 588 residuos pertenecientes a 11 secuencias, tiene un soporte moderado (PP = 0,78) e incluye a miembros de monocotiledóneas y dicotiledóneas (Figura 11 d). El análisis detectó una duplicación en el origen de las gramíneas, la cual generó dos copias divergentes. La búsqueda de motivos detectó dos motivos característicos de las secuencias CNDa de monocotiledóneas (Tabla 5).

El linaje CNDb, reconstruido a partir de un alineamiento de 215 residuos pertenecientes a 20 secuencias, incluye miembros de monocotiledóneas, dicotiledóneas y *A. trichopoda* (Figura 11 e). La resolución interna del grupo se resolvió parcialmente. Tanto CNDa como CNDb carecen de secuencias en el genoma de *A. thaliana*. No se encontraron motivos específicos adicionales en este linaje.

En término de número de secuencias proteicas, el linaje NKD es el más numeroso luego de SG5 e incluye a dos secuencias de *A. trichopoda* en una posición basal a secuencias de monocotiledóneas y dicotiledóneas (Figura 12 a). El mismo fue reconstruido a partir de un alineamiento de 657 residuos, pertenecientes a 33 secuencias. La topología del árbol obtenido muestra una

duplicación en el origen de las angiospermas ($PP = 0.95$) que generó dos linajes de los cuales uno se conservó sólo en *A. trichopoda* y el otro fue retenido en todas las angiospermas utilizadas para este estudio. A lo largo de la evolución, NKD sufrió una duplicación en las dicotiledóneas y una en las monocotiledóneas. Una de las copias de monocotiledóneas fue retenida por *P. dactylifera*, y la segunda copia fue retenida por todas las especies de monocotiledóneas usadas para el estudio. Luego, NKD se diversificó en las gramíneas a través de dos eventos sucesivos de duplicación y retención, generando cuatro copias (NKDgrass-1 a 4). La búsqueda global MEME identificó dos motivos específicos para NKD, LGC-M22 y LGC-M30 (Anexo 1, Tabla S1). Mientras que el motivo LGC-MC22 está presente en casi todas las secuencias del linaje, LGC-M30 está ausente en los grupos NKDgrass-1 y NKDgrass-2.

El linaje WDP, reconstruido a partir de un alineamiento de 648 residuos pertenecientes a 10 secuencias, está formado por miembros de *P. abies*, *A. trichopoda* y dicotiledóneas (Figura 12 b). No se observan eventos de duplicación intra linaje y no se encontraron motivos adicionales específicos.

El linaje WRB, reconstruido a partir de un alineamiento de 782 residuos pertenecientes a 25 secuencias, incluye tanto monocotiledóneas como dicotiledóneas (Figura 12 c). La reconstrucción encontró al menos un evento de duplicación en las dicotiledóneas, cinco copias en el genoma de *P. dactylifera* agrupadas en un clado monofilético y una duplicación en el origen de las gramíneas. El análisis de motivos encontró un motivo característico (LGC-M23) que no se detectó en el clado WRBgrass-1 y en algunas secuencias de dicotiledóneas (Anexo 1, Tabla S1).

El linaje GAF1, reconstruido a partir de un alineamiento de 743 residuos pertenecientes a 24 secuencias, incluye miembros de *S. moellendorffii* y angiospermas (Figura 12 d). Un evento de duplicación dio lugar a una copia que fue retenida en las gramíneas (GAF1-a) y una segunda copia que está presente en *S. moellendorffii* y todas las angiospermas estudiadas aquí (GAF1-b). Además del motivo característico encontrado en la búsqueda global (LGC-M24), se encontraron cuatro motivos adicionales en la búsqueda específica (Tabla 5). Dos de los motivos, LGC-M24 y GAF1-M11 son específicos del clado GAF1-a. El motivo GAF1-M10, un motivo EAR no relacionado con los caracterizados para SG5, está presente en todo el linaje y fue reportado como el dominio de unión a varios miembros de la familia de co-represores TOPLESS [112].

El linaje MGP, reconstruido a partir de un alineamiento de 528 residuos pertenecientes a 27 secuencias, contiene miembros de todos los genomas estudiados aquí, excepto por *A. trichopoda* (Figura 12 e). Varias especies contienen múltiples copias de MGP en sus genomas (cuatro copias en *P. patens*, tres copias en *S. moellendorffii*, dos copias en *A. thaliana*, cinco copias en *P. dactylifera* y dos copias en *Z. mays*). Dos secuencias de *P. dactylifera* (PdMGPe y PdMGPf) presentan conflicto con la topología de árboles previos. De hecho, mientras que se ubican dentro de los WDP en la filogenia de las angiospermas, se localizan dentro de las MGP en la filogenia de las estreptofitas. Los análisis

de linaje resolvieron que estas dos secuencias pertenecen al linaje MGP. El análisis MEME detectó un motivo adicional característico de las secuencias de *P. patens* (MGP-M9, Tabla 5).

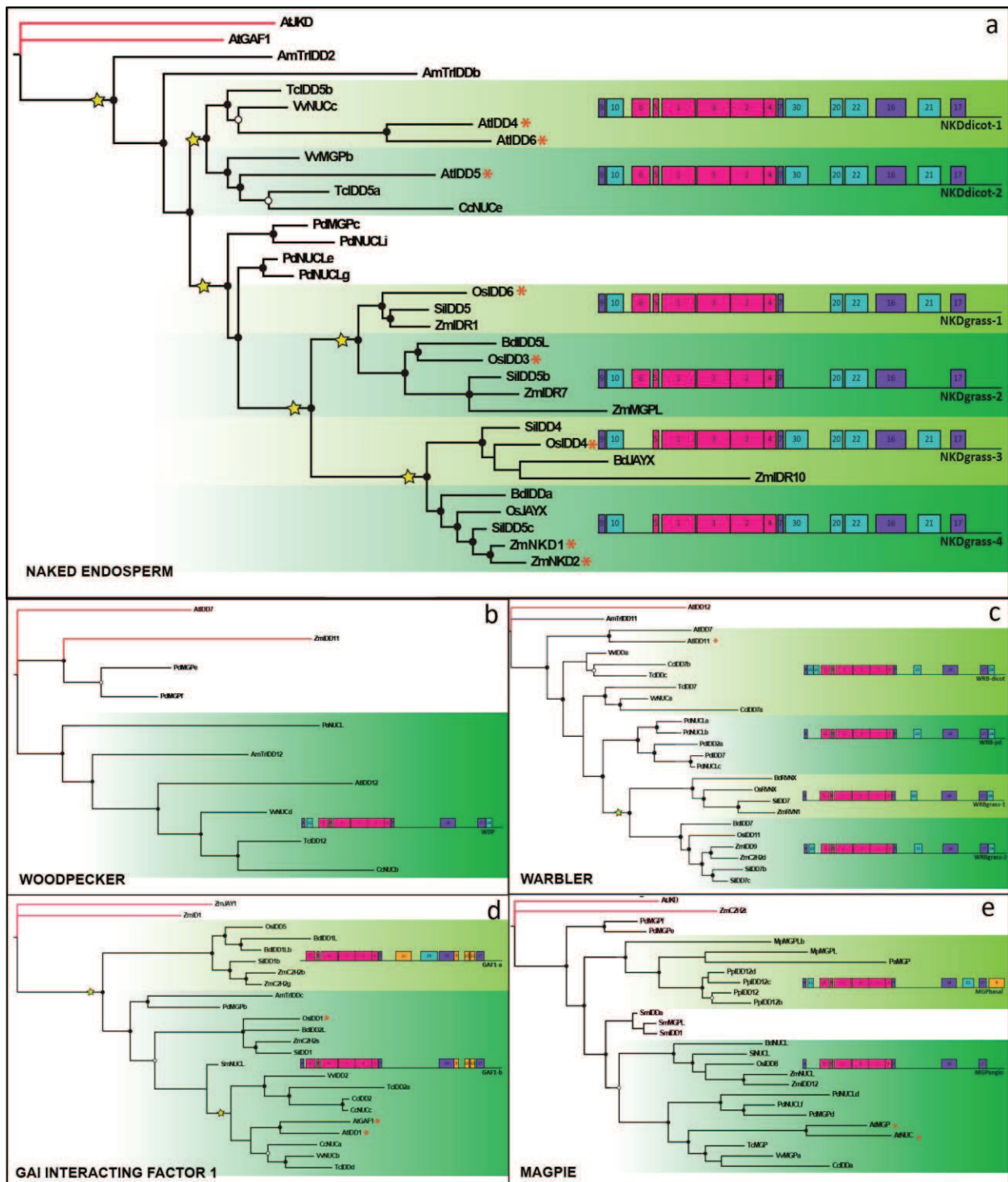


Figura 12. Diversificación de los linajes NKD, WDP, WRB, GAF1 y MGP en la división Streptophyta. Árboles consenso ($N=30.000$ árboles) de los linajes NKD (A), WDP (B), WRB (C), GAF1 (D) y MGP (E). Los puntos

blancos indican una Probabilidad Posterior Bayesiana (PP)= [0,8 a 0,89], los puntos negros indican PP= [0,9 a 1]. Las secuencias outgroups se muestran en rojo, las estrellas amarillas indican eventos de duplicación génica y los asteriscos anaranjados señalan los genes con función conocida. Los motivos identificados en cada clado están representados en esquemas de cajas sobre la derecha: HGC, rosado; MGC, violeta; LGC, celeste y los específicos GAF1 y MGP en naranja.

Tabla 5. Motivos específicos de linaje (secuencias consenso).

Motivo	Secuencia consenso	Valor E	Largo	Presencia (%)
SG5-M11	X[AT][TM][RQ]LQLSIGPX[GD]	5.6e-082	13	76.5
GAF1-M9	[SA]S[FL]L[RK][GE][LF]G[LI][AVD]SS[ST]SS	4.6e-080	15	86.4
GAF1-M10	[MS][LV][SP][AG]GLGLGL[PA][YC][DE][SG]G	1.6e-067	15	81.8
GAF1-M11	[LQ]E[DE]E[DT][DEK]QDEDED[VT]CFQEADQY[KD]DAELE[DV]SNL[LP]DKD	1.4e-065	33	27.3
GAF-M12	[SV]GL[PK][DE]LMMG[PQ][SP]	5.3e-058	11	90.9
MGP-M9	RDLA[SR]I[TA]SLG[PS]GVDLG[AV]FFNQ[RG][ED]NLR[SV][NG]ASG[HR]SPGKSWD	2.4e-052	40	17.4
CNDa-M5	[YA][ND][AETV][DS][VA][TV]AL[SN]YSD[EG]Q[QE][KP][PST][QG][EDH][HG][IM]Y[SP]YHQ[SM]S[LD][DG][SP][TAN]ALEKP[ILM]W[D	7.6e-077	40	55.6
CNDa-M7	[NH][DN][HQ][NS]SD[SMN][HC][LP]RSLSPY[AT][LF][IV][TA]R[NG]T[AST]LFSNQ[IM][SL][PS][KD][DE]	4.3e-056	33	55.6

DISCUSIÓN

Evolución de las subfamilias SHR y SCR

La literatura actual sugiere que las GRAS de plantas evolucionaron en la división Embryophyta por un evento de transferencia lateral de una metiltransferasa entre bacterias y un ancestro de las plantas terrestres. Sin embargo, hasta el día de hoy se desconoce el momento evolutivo en el que se sucedió dicho evento [79]. Nuestro análisis demuestra la ausencia de secuencias peptídicas GRAS en el alga carofita *K. nitens* (Orden *Klebsormidiales*) sin embargo, Bowman y colaboradores (2017) identificaron secuencias GRAS en el transcriptoma de otro miembro de la misma división, *Spirogyra pratensis* (Orden *Zygnematales*) [53]. Hasta el momento, con la información disponible, es posible proponer dos escenarios de evolución: (1) la familia GRAS se expandió rápidamente en el origen de las embriofitas, como la gran mayoría de las familias de reguladores transcripcionales, o (2) por el

contrario, la expansión fue gradual en la división Charophyta. A medida que se incorporen datos provenientes de la secuenciación de más especies de esta división, se podrá conocer con más precisión el origen y la diversificación de las GRAS.

A pesar de su importancia y el gran número de trabajos publicados sobre el rol de las GRAS en las plantas, el origen y la evolución de la familia ha demostrado ser difícil de abordar. En este capítulo, se utilizaron secuencias de genomas de las plantas verdes para reconstruir la evolución de dos subfamilias de las GRAS. Se clasificaron las proteínas SHR y SCR de 15 genomas en la división Embryophyta, se identificaron duplicaciones, motivos característicos y se obtuvo un marco de referencia que permitirá mejores estudios funcionales de los miembros de estas subfamilias.

La presencia de especies no angiospermas dentro de linajes internos en las subfamilias SHR y SCR sugiere que el origen de estas subfamilias se remonta, al menos, hasta el origen de la división Embryophyta. En la subfamilia SCR, luego de establecidos los linajes, no se evidencian eventos de duplicación internos, a excepción de eventos recientes específicos de especie. La mayor diversidad de la subfamilia SHR estaría dada por una serie de eventos de duplicación dentro de los tres linajes: (a) una duplicación ancestral alrededor del origen de las espermatofitas, (b) una duplicación más reciente de NSP1 durante el origen y diversificación del PACMAD de gramíneas, y (c) dos eventos consecutivos de duplicación en el origen de las angiospermas y posteriormente en las gramíneas.

Evolución de la subfamilia IDD

A pesar de poseer roles preponderantes en varios aspectos del desarrollo vegetal, el origen y la evolución de la subfamilia de reguladores transcripcionales IDD no había sido explorado. En este capítulo, se utilizaron secuencias de genomas de las plantas verdes para rastrear el origen y reconstruir la evolución de los genes IDD. Se identificaron los genes IDD de 16 genomas de las estreptofitas y se obtuvo un marco de referencia que vincula información genética y funcional de genes estudiados con especies no modelo.

El origen de los IDD se remonta al último ancestro común entre las algas carofitas y las plantas terrestres. La división Charophyta es un grupo parafilético de organismos de agua dulce que comparten un origen común con las plantas terrestres. Juntas representan al clado Phragmoplastophyta, caracterizado por la presencia de fragmoplastos durante la citocinesis ([138], [139]).

La topología de los árboles reconstruidos muestra al linaje STOP como hermano de los IDD. La identificación de una secuencia de *K. nitens* dentro de los linajes IDD y STOP y la localización de las secuencias de C2H2 de clorófitas por fuera de los clados STOP+IDD y el clado hermano WIP, sugiere que estos tres linajes provienen de una duplicación en el origen de Streptophyta. En *A. thaliana*, la

subfamilia STOP/WIP posee un menor número de miembros que la subfamilia IDD (2 STOPs y 6 WIPs, versus 16 IDDs) y sus funciones están relacionadas al desarrollo vegetal y la tolerancia a diversos estreses. Por ejemplo, AtSTOP1 y 2 son reguladores positivos de genes que responden al estrés debido a la toxicidad por aluminio y ácidos ([140], [141]). Los genes WIP de *A. thaliana* cumplen un rol en el desarrollo floral, desarrollo de HV, y acumulación de flavonoides en semillas ([142], [143]). Este solapamiento en las funciones de los linajes STOP/WIP e IDD, sumado a una historia evolutiva en común nos hace suponer que la definición de la subfamilia IDD, incluyendo al linaje SG5 y excluyendo a los linajes STOP/WIP pareciera ser una cuestión semántica debido a la forma en que estos genes se fueron estudiando en las últimas décadas.

La conquista de los hábitats terrestres por parte de las plantas requirió la evolución de casi todos los órganos vegetales y con ello, el establecimiento de nuevas redes de genes y vías de señalización. El desarrollo de nuevos tipos celulares, órganos y adaptaciones que ocurrieron durante la evolución temprana tiene que haber ido acompañado con la diversificación de genes ancestrales y el origen de nuevos [144]. De hecho, se puede observar una retención en el número de IDD durante la conquista de la tierra y una posterior expansión rápida en el número de genes y linajes durante la diversificación de las angiospermas (Tabla 3). La colonización de la tierra por parte de las plantas comenzó alrededor de 470 millones de años atrás. Por esos tiempos, cerca del origen de las embriofitas, los IDD experimentaron una duplicación que generó dos linajes divergentes: el linaje SG5 y un linaje que agrupa a los demás IDD.

El linaje SG5 fue retenido en todos los genomas aquí estudiados y no presentó duplicaciones hasta la diversificación de las monocotiledóneas y dicotiledóneas (~140–150 millones de años atrás). La presencia de múltiples copias SG5 en *P. patens* coincide con la evidencia de dos duplicaciones sucesivas de genomas completos (WGD, del inglés: *whole genome duplications*) compartidas por los musgos, pero no por hepáticas y antoceros (Lang y col., 2018). Como ninguna WGD se ha reportado para el genoma de *P. abies*, se asume que las copias SG5 pueden provenir de una duplicación génica y/o ser vestigios de un evento WGD de las gimnospermas ancestrales ([145]–[147]).

En contraste con esta aparente tasa de evolución estable de los SG5 hasta tiempos relativamente recientes, los linajes que agrupan al resto de IDD se han duplicado muchas veces en paralelo con el origen y diversificación de las plantas terrestres. Interesantemente, aunque las monocotiledóneas y dicotiledóneas presentan un número similar de genes IDD, las monocotiledóneas poseen un mayor número de clados, la mayoría de ellos específicos de gramíneas. Si bien esto puede ser un artefacto proveniente de la selección de especies para el estudio, también puede estar reflejando los eventos de WGD que dieron origen a las gramíneas ([148], [149]). El incremento en el número de genomas disponibles públicamente y la inclusión de esta información en nuevas filogenias, permitirá

en el futuro una mayor estimación de los clados IDD en monocotiledóneas y dicotiledóneas.

La expansión y diversificación de los IDD en tierra podrían estar relacionados con la gran transformación del cuerpo de las plantas durante la evolución en los ambientes terrestres. De hecho, el gravitropismo ha jugado un rol activo en el origen y la evolución de las plantas en contacto con la tierra [150]. La respuesta a la gravedad surgió junto con el origen del orden Charales y las estreptofitas y se piensa que fue determinante en establecer la forma final de las plantas. En línea con esto, las proteínas SG5 poseen un rol central en los mecanismos que controlan el gravitropismo ([104]–[106], [121]). Las proteínas SG5 fueron retenidas en los genomas de todas las embriofitas estudiadas, sugiriendo que pueden jugar un rol principal en las plantas modernas.

Otro paso importante en la terrenalización de las plantas fue el origen de órganos laterales una vez establecido el meristema apical en la separación de las briofitas y las plantas vasculares. Las raíces y sus meristemas surgieron en el origen de las plantas vasculares [151]. Varios estudios en *A. thaliana* muestran que las proteínas MGP (AtMGP y AtNUC) juegan un rol clave en la diferenciación de tejidos radiculares, actuando en conjunto con JKD (AtJKD y AtIDD9) y NKD (AtIDD4 y AtIDD6), y las proteínas SHR, SCR y SLC3 ([101], [102], [152]). Esta función se logra a través de un complejo mecanismo que involucra la activación transcripcional e interacciones proteína-proteína que actúan de forma diferencial en varios tejidos celulares. De manera similar a lo que ocurre con SG5, las proteínas MGP fueron retenidas en los genomas de todas las plantas terrestres utilizadas en este trabajo. El rol de las proteínas MGP en plantas terrestres tempranas es desconocido.

Muchos de los miembros actuales de los IDD controlan el desarrollo de las raíces, hojas, inflorescencia y semillas [126], importantes novedades morfológicas que han transformado el cuerpo de las plantas. Sin embargo, la información disponible, sugiere que las especies de plantas no vasculares estudiadas aquí carecen de genes IDD involucrados en raíces y el desarrollo de órganos laterales. El reclutamiento de los IDD por las angiospermas probablemente dependió de la participación activa en redes de regulaciones génicas (GRN, del inglés: *Gene Regulatory Networks*) con genes de otras familias. El ejemplo más documentado probablemente sean los reguladores transcripcionales de la familia GRAS. A partir de los resultados obtenidos en este trabajo, se identificó que los GRAS se han diversificado activamente durante la especiación de mono y dicotiledóneas, con un patrón similar al observado en la evolución de los IDD [77]. Además, la interacción entre proteínas IDD y GRAS se reportó numerosas veces ([51], [87], [100], [101], [112], [137], [152]). De hecho, los IDD y GRAS han cobrado relevancia en los últimos años debido a sus roles en el control de los límites entre tejidos ([103], [108], [117], [153]). En este escenario, la co-evolución de los IDD con otra familia de reguladores transcripcionales podrían ser parte de un paso relevante en la conquista de la tierra y la diversificación de las angiospermas.

Hasta la fecha, todos los reportes indican que los IDD están presentes en el núcleo y funcionan como factores de transcripción. Sin embargo, el motivo de la señal de localización nuclear está ausente en algunas de las secuencias IDD validadas en este trabajo. Esto implica que algunos IDD pueden contener otro tipo de NLS, no detectada por el análisis de motivos, y/o algunos IDD podrían haber desarrollado nuevas funciones fuera del núcleo a lo largo de la evolución. La expansión y diversificación de la subfamilia durante el curso de la conquista de la tierra fue acompañada por la evolución de nuevos motivos conservados a lo largo del C-terminal. La presencia y conservación de tales motivos en los linajes IDD sugiere un rol en las funciones de las proteínas.

El linaje SG5 presenta el dominio C-terminal más divergente entre los IDD. Entre los motivos conservados del linaje, se han encontrado dos motivos represores EAR, los cuales fueron caracterizados para OsLPA1 [121]. En línea con esto, Qi y col. (2019) mostraron que AtIDD16, un miembro de SG5dicot-2, regula negativamente la iniciación estomática a través de la represión en *trans* de SPEECHLESS, sugiriendo una actividad general represora para el linaje SG5. Uno de estos motivos EAR (LGC-M25) está presente en la secuencia de *K. nitens*, lo cual podría sugerir un rol ancestral represor para los IDD. Sin embargo, está documentado que algunos SG5 de *A. thaliana* pueden cumplir ambas funciones, como represores y activadores transcripcionales directos ([109], [154]).

Ha sido documentado que la actividad de los represores EAR depende del número de residuos L en el motivo y del número de motivos EAR consecutivos, y pueden ser atenuados por motivos activadores cercanos ([155][156]). En línea con esto, OsLPA1 y sus ortólogos en el subclado SG5grass-1, carecen de cuatro de los motivos que flanquean a los EAR del resto de las SG5 (LGC-M19, LGC-M27, LGC-M28 y LGC-M29). Esto podría sugerir un rol divergente entre las proteínas SG5.

Adicionalmente, las GAF1 presentan un motivo EAR no relacionado con aquellos en SG5. De hecho, GAF1 ha sido reportado como un represor al unirse a varios miembros de la familia de co-represores TOPLESS cuando las proteínas DELLA se encuentran ausentes. En presencia de DELLA, GAF1 se une a ellas por medio del motivo SATAL (MGC-M126) y actúa como activador [112].

En cuanto a la unión de los IDD con los miembros de la familia GRAS, Hirano y colaboradores (2017) mostraron que la interacción ocurre a través del motivo de unión a SHR al final del cuarto ZF [51]. Este motivo aparece conservado en todos los IDD, excepto en la secuencia de *K. nitens* y en el linaje SG5. Además, las interacciones físicas entre IDD y las DELLA se ven comprometidas cuando los motivos MGC-M16 (SATAL) y MGM-M17 están ausentes. Esto estaría mostrando que los IDD presentan cierta flexibilidad a la hora de interactuar con otras proteínas, lo cual podría ser un promotor de la diversificación y neo-funcionalización de los linajes de la subfamilia. Esta evolución de los motivos auxiliares podría tener un efecto corriente abajo de las proteínas GRAS en diferentes tejidos, al cambiar la población de IDD involucrados en la interacción.

CAPÍTULO 2

Estudio de la anatomía foliar y la
fisiología en Otachyriinae

INTRODUCCIÓN

La modificación de la anatomía foliar como uno de los primeros pasos en la evolución de la fotosíntesis C₄

A pesar de su complejidad, la fotosíntesis C₄ evolucionó alrededor de 70 veces de forma independiente en las angiospermas. Sin embargo, hasta el día de hoy, se conoce que la mayoría de los orígenes C₄ se agrupan en unos pocos linajes de plantas. Esto hace suponer que ciertos linajes tienen características que aumentan la capacidad de desarrollar una fotosíntesis C₄ ([11], [31]).

En un estudio realizado con 157 especies de gramíneas, donde se analizaron los rasgos fenotípicos de las especies teniendo en cuenta su historia evolutiva, Christin y colaboradores (2013) demostraron que sólo algunos rasgos fenotípicos individuales presentan diferencias significativas entre especies C₃ y C₄ (fracción ocupada por la CVE y CVI, número de células del M entre HV, tamaño celular de CVI). En términos generales, se encontraron diferencias más significativas entre especies C₃ del clado BEP y especies C₃ del clado PACMAD que entre especies C₃ y C₄ del clado PACMAD, donde los valores mostraron una tendencia a solaparse [30]. En dicotiledóneas, Lauterbach y colaboradores (2019) encontraron una tendencia similar en la tribu Tribuloideae (Familia Zygophyllaceae) [37]. En este trabajo se observó que las especies C₄ de la tribu presentaron una mayor fracción del área ocupada por las CV. Notoriamente, también observaron que las especies C₃ próximas evolutivamente a las C₄, presentaban un área ocupada por CV con dimensiones similares a las especies C₄. Ambos trabajos sugieren que muchos de los rasgos anatómicos, como la distancia entre HV, los tamaños celulares y el área ocupada por las CV, son rasgos ancestrales en el camino evolutivo hacia el establecimiento de una anatomía Kranz y una fotosíntesis C₄.

Además, se encontró que, la densidad de HV fue sólo significativamente mayor para las especies C₄ en relación a las especies C₃ [37]. En el mismo sentido Lundgren y colaboradores (2019) establecieron que, en *Alloteropsis semialata* (Familia Poaceae), el único rasgo compartido por todas las poblaciones C₄ fue una mayor densidad de HV [36]. Estos resultados demuestran que el incremento en la densidad de HV no sería un pre-condicionante anatómico necesario para impulsar el cambio hacia una anatomía C₄, como se postuló anteriormente [26].

Estos trabajos remarcan la importancia de la selección de especies a estudiar y su contexto evolutivo a la hora de tratar de establecer marcadores fenotípicos importantes para la evolución de la fotosíntesis C₄.

La subtribu Otachyriinae como modelo de estudio de la evolución de la fotosíntesis C₄

En las gramíneas, todas las especies que realizan fotosíntesis C₄ o C₂, se encuentran en el clado PACMAD. Dentro del PACMAD, la subfamilia Panicoideae posee alrededor de 3.300 especies, lo que la hace la segunda subfamilia más grande de las gramíneas. Dentro de esta subfamilia se pueden distinguir tres tribus: Andropogoneae, Paniceae y Paspalae. Así mismo, la tribu Paspalae se puede dividir en tres subtribus: Otachyriinae, Arthropogoninae, y Paspalinae [157].

Los estudios sistemáticos más recientes, basados en marcadores nucleares y cloroplastídicos, muestran que la subtribu Otachyriinae es un grupo monofilético integrado por 35 especies distribuidas en siete géneros ([158], [159]). Las especies de la subtribu están ampliamente distribuidas en regiones húmedas de los trópicos de América, Asia y Australia. Morfológicamente, las especies de la subtribu se caracterizan por presentar inflorescencias cuyas espiguillas se organizan unilateralmente sobre las ramas. Las espiguillas presentan la gluma inferior más corta que la gluma superior y lemma inferior, y un antecio superior membranoso a endurecido ([158], [159]).

En relación al tipo de fotosíntesis, la subtribu incluye un género C₄, algunas especies C₂, PK y especies C₃ (Figura 13). La reconstrucción de caracteres ancestrales sugiere que el ancestro común de la subtribu utilizaba la vía fotosintética C₃. La condición C₄ habría evolucionado tempranamente en el clado *Anthaenantia*, mientras que la condición intermedia C₂ habría evolucionado más tarde, previo a la diversificación del género *Steinchisma* [159]. Recientemente se han documentado eventos de alopoliploidía y de transferencia lateral de genes (TLG) que dificultan la clasificación interna de la subtribu ([158], Acosta y Reinheimer, com. pers.). La diversidad de formas fotosintéticas que presentan las especies de la subtribu Otachyriinae convierten a este linaje en un verdadero laboratorio para responder preguntas múltiples sobre la evolución y la optimización de la fotosíntesis en gramíneas.

El género monofilético *Anthaenantia*, integrado por 4 especies, es hermano del resto de la subtribu. Estudios filogenéticos indican que *A. lanata*, la única especie sudamericana de *Anthaenantia*, es hermana del resto de las especies norteamericanas y centroamericanas *A. rufa*, *A. texana* y *A. villosa* [158]. Morfológicamente el género se reconoce por su inflorescencia contraída con espiguillas densamente pilosas las que carecen de gluma inferior. El antecio superior es cartilaginoso cuya lemma no alcanza la longitud de la pálea superior. Se estima que este linaje surgió hace unos 7,2 millones de años, lo que en términos de evolución C₄ podría considerarse relativamente reciente [24]. La tetraploidía de este género parece haberse originado por autoduplicación ([158], [159]).

denominado *Hymenachne New World* integrado por especies Sudamericanas. Este género es reconocido por su inflorescencia abierta a contraída, espiguillas glabras, dispuestas unilateralmente en las ramas, con el antecio superior membranoso y su lema cubriendo parcialmente la pálea en el ápice. Varias especies de este género fueron reportadas como tetraploides ([158], [159]).

El género monofilético *Rugoloa* está integrado por 3 especies distribuidas en América. Es reconocido por presentar una inflorescencia laxa a contraída, con las espiguillas dispuestas unilateralmente en las ramas, espiguillas pilosas o glabras, y el antecio superior membranoso endurecido. Si bien *R. pilosa* es diploide, hay evidencias de alopoliploidía dentro de este género ([158], [159]).

Los géneros *Otachyrium*, *Plagiantha* y *Steinchisma* conforman un clado monofilético con la mayoría de las especies diploides (incluida *S. hians*). Característicamente las especies de *Otachyrium* tienen inflorescencias laxas con espiguillas dispersas y solitarias con ambas glumas más cortas que la espiguilla, dejando así expuesto el antecio superior, el que es endurecido, liso y brillante; la lemma superior es oscura a la madurez [160]. El género *Plagiantha*, integrado por una única especie, y *Steinchisma* integrado por 8 especies, difieren de *Otachyrium* al tener las espiguillas adpresas en las ramas de la inflorescencia, con la gluma superior de longitud igual o mayor que el antecio superior cartilaginoso y papiloso. El género *Steinchisma* se caracteriza por presentar una inflorescencia laxa a contraída, con espiguillas glabras no dispuestas unilateralmente, pálea inferior expandida y endurecida en la madurez, lema inferior nervada y el antecio superior cartilaginoso con papilas compuestas en toda su superficie ([158], [159]).

Varios trabajos se han enfocado en definir los miembros de la subtribu Otachyriinae, así como sus relaciones taxonómicas y su historia evolutiva ([158], [159]). Sin embargo, ningún estudio se centró en el desarrollo foliar en esta subtribu y muy pocos reportan la caracterización de la anatomía foliar ([15], [30], [161]–[166]).

OBJETIVOS ESPECÍFICOS

Caracterización de la anatomía foliar: Re-evaluar en forma comparativa, y empleando conocimientos modernos, la anatomía foliar de especies de la subtribu Otachyriinae.

MATERIALES Y MÉTODOS

Selección de especies

Para llevar a cabo los estudios fisiológicos y anatómicos se seleccionaron 4 especies de la subtribu Otachyriinae: *Hymenachne amplexicaulis* (Rudge) Nees (C₃), *Rugoloa pilosa* (Sw.) Zuloaga (PK), *Steinchisma hians* (Elliott) Nash (C₃-C₄ intermedia) y *Anthraenantia lanata* (Kunt) Benth (C₄). La selección de especies se basó principalmente en la proximidad filogenética, vía fotosintética y disponibilidad de material. Semillas o rizomas de las especies estudiadas fueron colectados a campo. Los ejemplares fueron depositados en los herbarios Instituto Botánica Darwinion (SI) y Arturo Ragonese (SF). A continuación, se listan los *vouchers* de colección: *Hymenachne amplexicaulis* (Rudge) Nees Prochetto y Reinheimer 1 (SF), *Rugoloa pilosa* (Sw.) Zuloaga s/n (SI), *Steinchisma hians* (Elliott) Nash Marino s/n (SF), *Anthraenantia lanata* (Kunt) Benth Acosta (SI).

Estudio fisiológico

Las tasas netas de asimilación de CO₂ fueron medidas en hojas jóvenes, completamente expandidas usando el equipo de intercambio de gases LI-COR 6400XT (LI-COR Biosciences). Las mediciones fueron realizadas a 27°C y con una irradiación de 1500 μmol fotones m⁻² s⁻¹. Las curvas de respuesta al CO₂ fueron realizadas por triplicado con los siguientes niveles de CO₂ en la cámara: 400, 200, 100, 50, 40, 500, 1000, 1500, 2000 y 2200 ppm.

Condiciones de cultivo y toma de muestra

Individuos de cuatro especies de Otachyriinae con diferentes subtipos fotosintéticos (*H. amplexicaulis*, *R. pilosa*, *S. hians* y *A. lanata*) fueron crecidas, a partir de semilla o rizoma, en cámaras de cultivos a 27°C bajo condiciones de día largo (16 horas de luz y 8 horas de oscuridad). Para cada especie, se colectó la 5ta hoja joven de 10 individuos. Para estudiar el gradiente de desarrollo foliar, cada hoja fue dividida en dos secciones, tomando como referencias la zona de transición fuente-sumidero. Las secciones fueron luego divididas en cuatro fragmentos de igual longitud y rotulados F1 a F8 desde la base a la punta de la hoja (Anexo 1, Figura S4). Las réplicas se mantuvieron pareadas a lo largo del análisis.

Preparación y análisis de las muestras para microscopía óptica

Los fragmentos frescos fueron colocados en tacos con agarosa de bajo punto de fusión al 5% en PBS 0,05 M a 50°C y dejados hasta solidificación a 4°C por al menos 1 hora. Se obtuvieron cortes transversales de 100 µm de espesor con un micrótopo de cuchilla vibrante (Leica VT1000 S) y se montaron en portaobjetos con PBS 0,05 M. Las secciones fueron visualizadas y fotografiadas en el momento bajo microscopía fluorescencia y luz blanca (Nikon Eclipse E200). El procesamiento de las imágenes y las mediciones anatómicas fueron realizadas a través del software FIJI [167].

Análisis estadístico

Para determinar la significancia entre las diferencias en los caracteres fenotípicos se utilizaron métodos no paramétricos contemplando muestras pareadas (para comparaciones intra especies) y no pareadas (para comparaciones entre especies) con el programa Graphpad Prism (versión 6.0.1 para Windows, GraphPad Software, San Diego, California USA, www.graphpad.com). Para las muestras pareadas se utilizó el test de Friedman y el test de comparaciones múltiples de Dunn. Para las muestras no pareadas se utilizó el test de Kruskal-Wallis y el test de comparaciones múltiples de Dunn. Se consideraron significativos los valores p menores a 0,05.

RESULTADOS

Aspectos cualitativos

Las especies *H. amplexicaulis*, *R. pilosa* y *S. hians* presentan una anatomía foliar típica de gramíneas C₃, con haces vasculares (HV) organizados linealmente y rodeados por dos capas de células de la vaina: la más pequeña formada por células de la vaina interna (CVI) y la más grande formada por células de la vaina externa (CVE) (Figura 14). En estas especies, si bien los cloroplastos se observan mayormente en las células del mesófilo, algunos pueden observarse en las CVE en las especies *H. amplexicaulis* y *R. pilosa*. En *S. hians* se observan cloroplastos con una distribución centripeta en las CVE, algo característico de las especies C₂. Por el contrario, en la especie C₄ (*A. lanata*) se observa una distribución no lineal de los haces vasculares (HV) que están rodeados por una única capa de células de la vaina (CVI). Se observa un gradiente en la cantidad de cloroplastos desde los haces vasculares, con las CVI con la mayor densidad de cloroplastos, las células del mesófilo en contacto con las CVI con una menor densidad y las células del mesófilo más alejadas, con la menor densidad (Figura 15). Además, en las especies C₃ se observaron grandes espacios intercelulares: (1) en *H. amplexicaulis*

se observa la presencia de aerénquima; (2) en *R. pilosa* los espacios intercelulares son interpretadas como el colapso de células fusoides (Figura 14).

En términos generales se observó una maduración de las hojas desde la base a la punta, evidenciado por una mayor lignificación de los haces vasculares, un aumento en la cantidad de haces vasculares secundarios, una mayor diferenciación entre células de la vaina y mesófilo con respecto al número de cloroplastos, y una disminución en la cantidad de tejido de sostén.

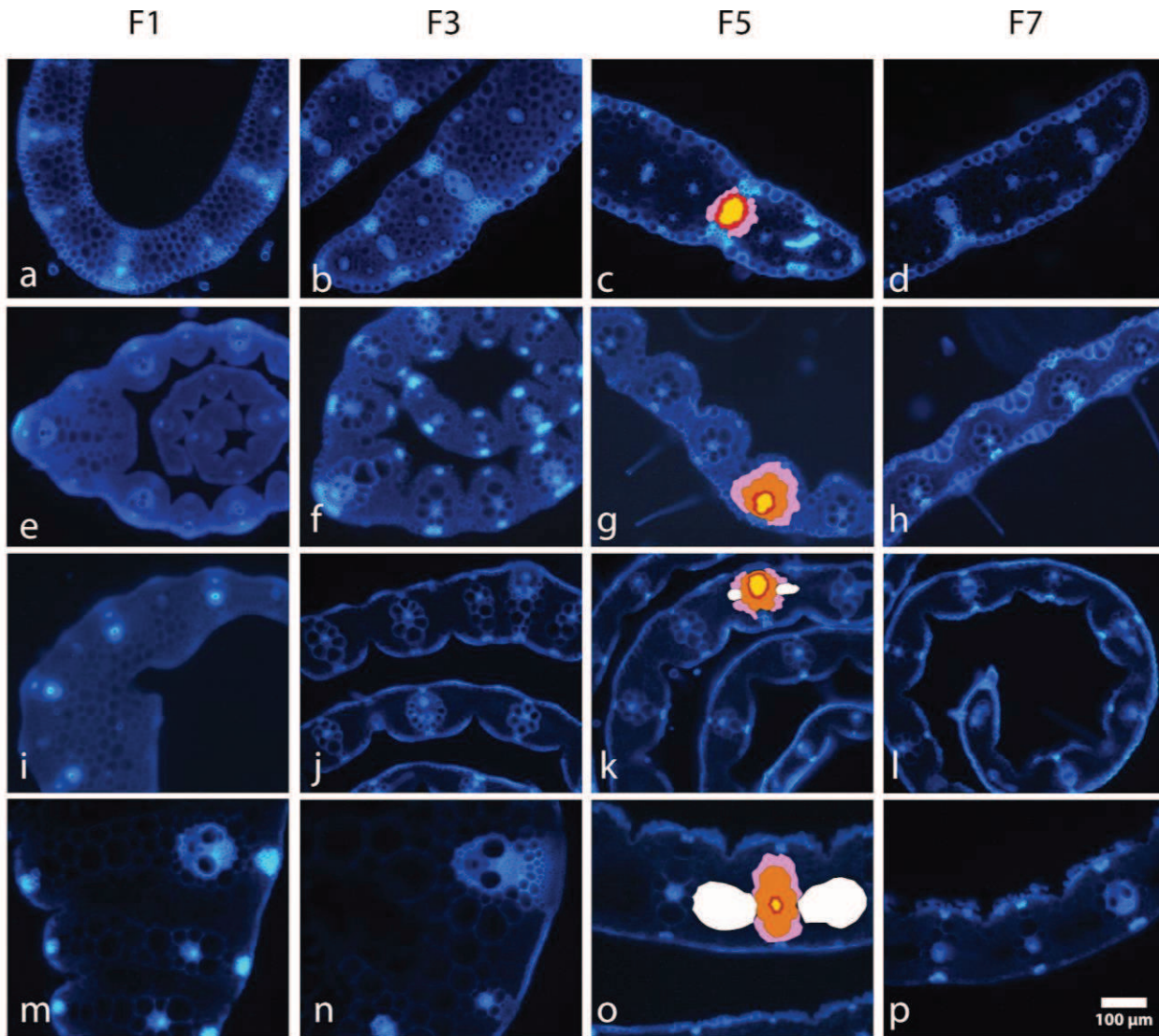


Figura 14. Gradiente de maduración en hojas de *Otacyriinae*. Fragmentos F1, F3, F5 y F7 de la 5ta hoja joven. *A. lanata* (a-d), *S. hians* (e-h), *R. pilosa* (i-l) y *H. amplexicaulis* (m-p). En las imágenes del F5 se señalan diferentes tipos celulares: HV, amarillo; CVI, rojo; CVE, naranja; M en contacto con las CV, rosado; cavidades intercelulares, blanco.

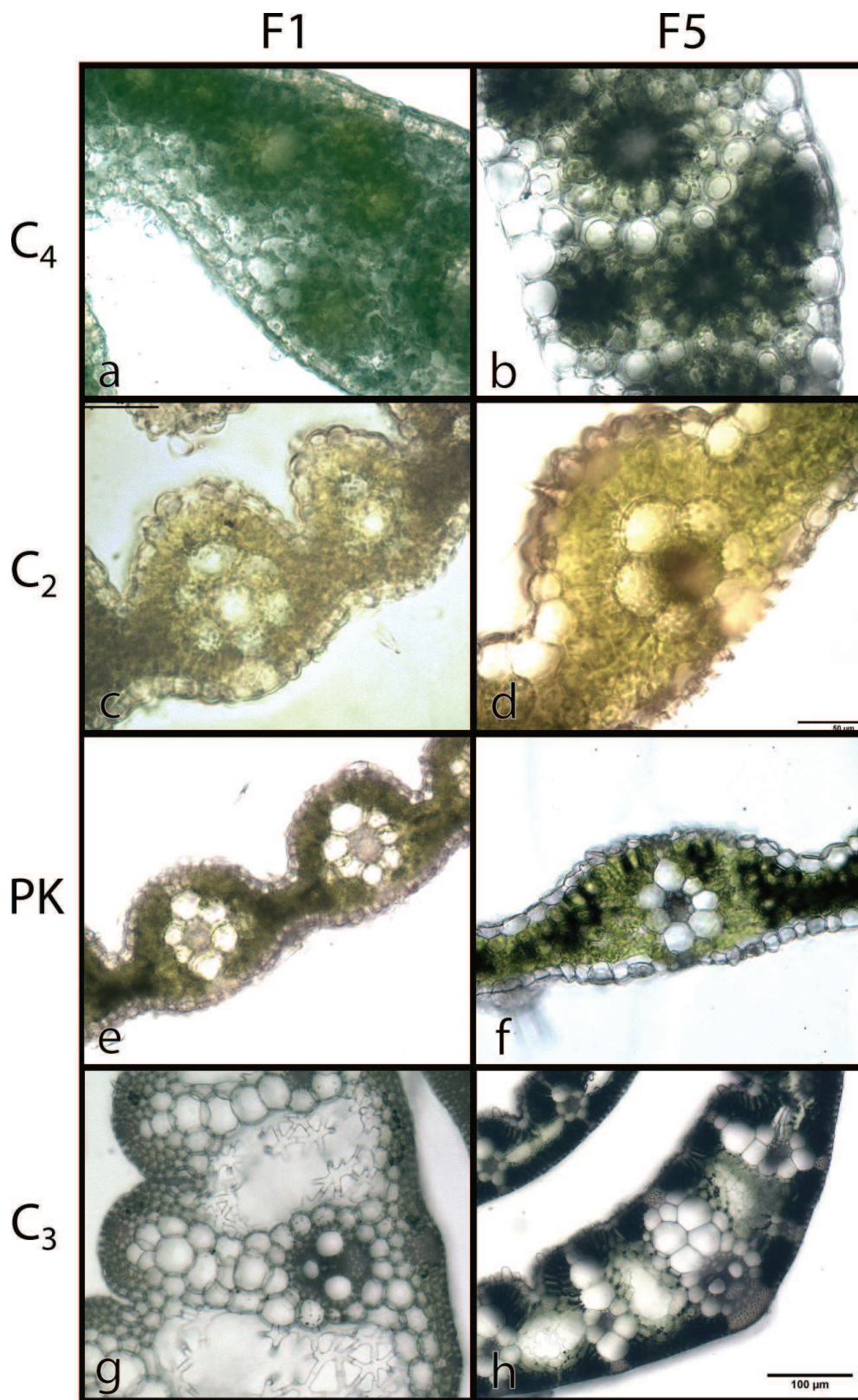


Figura 15. Localización de cloroplastos en cortes transversales de hojas de Otachyriinae. Fragmentos F1 y F5 de la 5ta hoja joven de *A. lanata* (a-b), *S. hians* (c-d), *R. pilosa* (e-f) y *H. amplexicaulis* (g-h).

Estudio de rasgos fenotípicos asociados al desarrollo

Parámetros del desarrollo foliar asociados a los haces vasculares

Los haces vasculares, al estar íntimamente relacionados a las CVI y CVE, juegan un papel fundamental en la estructura de la anatomía foliar. En esta sección estudiamos cómo varían los rasgos fenotípicos a lo largo del desarrollo y cómo se diferencian estos patrones en especies con diferentes subtipos fotosintéticos.

La distancia entre los HV tiende a aumentar a lo largo del desarrollo para las especies C₃, PK y C₂, mientras que se reduce en la especie C₄ (Figura 16 a-d). Este aumento en la distancia genera una disminución en la densidad de haces a medida que las hojas maduran en *R. pilosa* y *S. hians* (Figura 16 f y g). Sin embargo, en *H. amplexicaulis*, la reducción del tejido aerenquimático (que presenta su mayor extensión en los fragmentos más inmaduros) compensa este efecto hasta revertir esta tendencia en el último fragmento estudiado (Figura 16 e). En este sentido, en todas las especies los cambios más marcados se evidencian en la transición F1-F3.

El área transversal ocupada por los haces vasculares (Área HV %; Figura 17 p-s) muestra un patrón descendente a lo largo del gradiente de maduración para todas las especies, pero con una mayor intensidad en las especies C₂ y C₄. Este fenómeno está explicado tanto por la reducción en el tamaño de los HV primarios como a la aparición de nuevos HV secundarios de menor tamaño (Figura 17 i-o). La reducción en el porcentaje de HV primarios es muy sutil y gradual, aunque estadísticamente significativa, para *H. amplexicaulis*. Para las demás especies esta transición es menos gradual, pero con un salto en el porcentaje de HV primarios en la transición F1-F3, especialmente en *R. pilosa* y *A. lanata*.

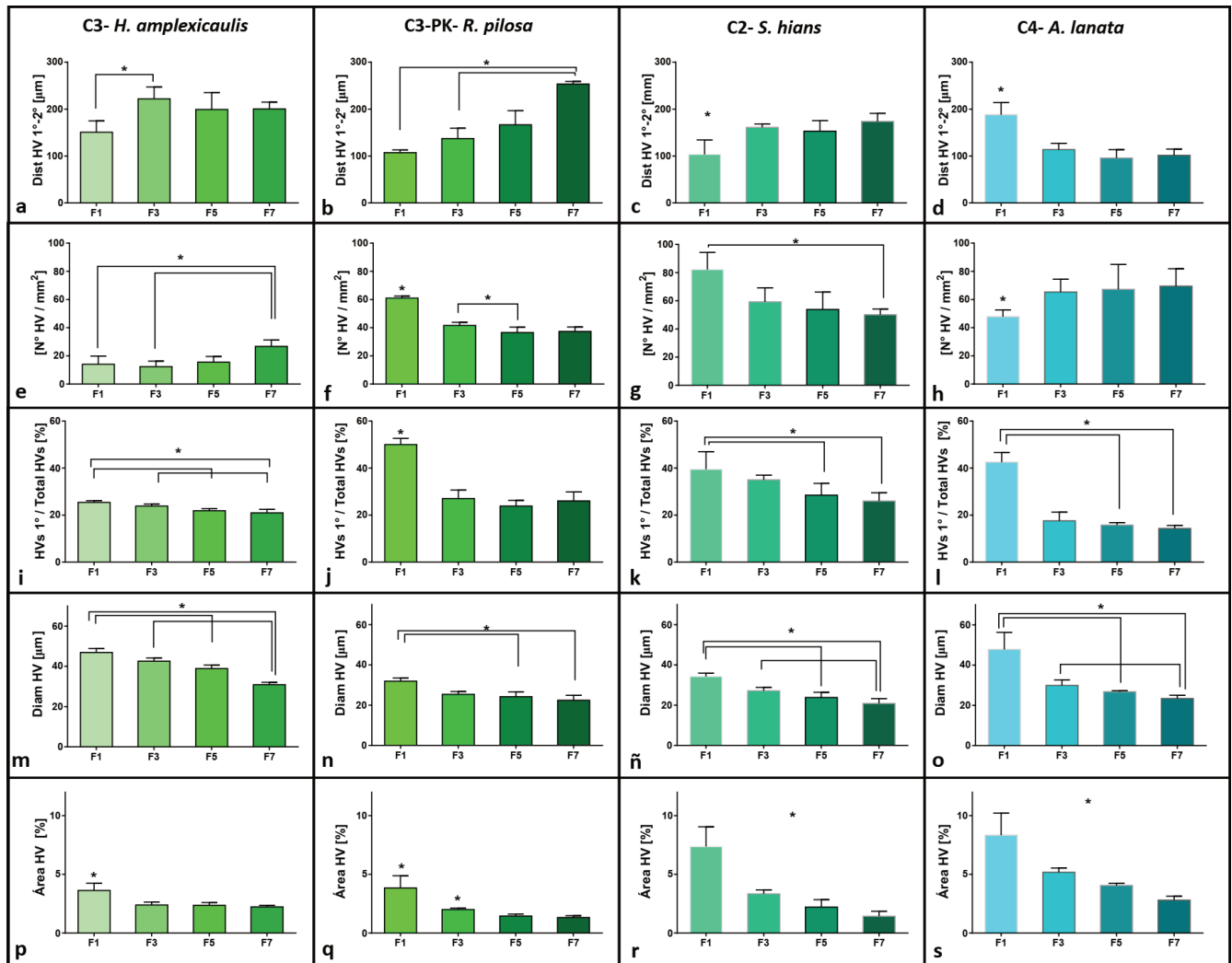


Figura 16. Parámetros del desarrollo foliar relacionados con los haces vasculares (HV). Distancia entre el haz vascular central y sus haces secundarios contiguos (a-d). Densidad de haces vasculares (e-h). Porcentaje de haces vasculares primarios sobre el total de haces (i-j). Diámetro promedio de los haces vasculares (m-o). Fracción del tejido foliar ocupada por haces vasculares (p-s).

Parámetros del desarrollo foliar asociados a las células de la vaina y del mesófilo

Una de las características de la anatomía Kranz es la mayor relevancia que juegan las células de la vaina con respecto a las especies C₃ en términos de tamaño y distancia entre sí [6]. Para verificar este fenómeno en estas especies, se midió tanto el tamaño de las CVI y CVE como la fracción de tejido que ocupan dentro de la hoja.

En *H. amplexicaulis* ambos, el tamaño de la CVI como la fracción que ocupa en el tejido (Figura 17 a y e), son constantes a lo largo del gradiente estudiado. Estos resultados sugieren que la

diferenciación de las CVI forma parte de una etapa previa al muestreo realizado en este trabajo. En *R. pilosa* tampoco se observan diferencias en el tamaño celular de las CVI, pero sí en la fracción del área que ocupan, la cual tiende a disminuir a medida que la hoja madura (Figura 17 b y f). En *S. hians* se da un proceso similar al de *R. pilosa*; sin embargo, en esta especie los datos obtenidos sugieren un aumento del tamaño celular de las CVI en los fragmentos maduros F3 y F5 (Figura 17 c y g). En la especie *C4* se observa un incremento significativo inicial en el tamaño de las CVI; sin embargo, este incremento no se traduce en cambios significativos en la fracción ocupada por las CVI a lo largo del desarrollo (Figura 17 d y h).

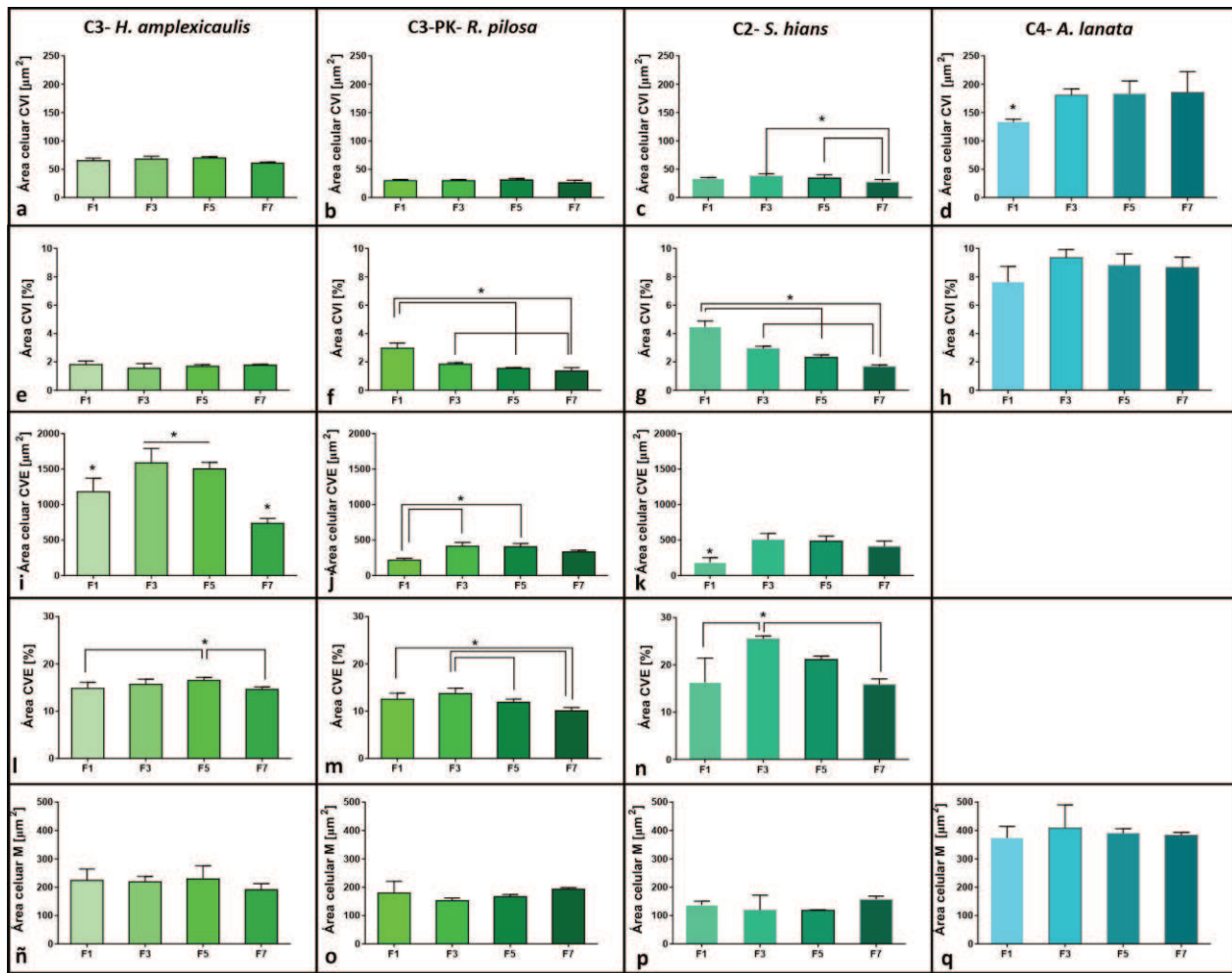


Figura 17. Parámetros del desarrollo foliar relacionados con las células de la vaina y del mesófilo. Tamaño celular promedio de las CVI (a-d). Fracción del tejido foliar ocupada por las CVI (e-h). Tamaño celular promedio de las CVE (i-k). Fracción del tejido foliar ocupada por las CVE (l-n). Tamaño celular promedio de las células del mesófilo (ñ-q).

En todas las especies estudiadas, las CVE presentan un patrón de desarrollo característico que implica tanto un aumento del tamaño celular como de la fracción que ocupan en los fragmentos 3 y 5 (Figura 17 i-n). En *S. hians* y *A. lanata*, el tamaño celular de las CVI se encuentra ligado al desarrollo foliar en el gradiente estudiado, mientras que el tamaño promedio de las células del mesófilo no se ve significativamente alterado en ninguna de las especies (Figura 17 ñ-q).

Correlación entre parámetros del desarrollo foliar

Para comprender la relación entre los cambios en los parámetros fenotípicos y el momento del desarrollo en el que se suceden dichos cambios se realizó un análisis de correlación entre las 40 muestras de cada especie (4 fragmentos y 10 réplicas por fragmento). Las correlaciones más significativas (valor $p < 0,001$) se muestran en la figura 18.

En *H. amplexicaulis* se evidencia una relación inversa entre la densidad de HV y los tamaños de las células (CVI, CVE y M), a su vez correlacionados con el tamaño total de la hoja (Figura 18 a). Mientras que el tamaño de las células del M correlaciona con el tamaño de las CVE en *H. amplexicaulis*, en el resto de las especies este rasgo fenotípico no correlaciona con ninguno de los caracteres de tamaños celulares. Esto podría indicar un desacople en el desarrollo del mesófilo con respecto de los otros tipos celulares en las especies PK, C₂ y C₄ vs la especie C₃.

Al desglosar el parámetro de densidad de HV, en densidad de HV primarios y secundarios, se observa que en *A. lanata* éste depende exclusivamente de la densidad de HV secundarios (Figura 18 d), mientras que en las demás especies la densidad de HV correlaciona con ambos tipos de HV. Algo similar se observa con respecto al tamaño de la hoja que correlaciona negativamente con la densidad de HV primarios y secundarios en todas las especies, excepto en *A. lanata* donde sólo correlaciona con la densidad de HV primarios. Sumado a esto, se observa que mientras en las especies no C₄ las densidades primarias y secundarias de HV tienden a correlacionar con otros parámetros en el mismo sentido, en la especie C₄ las densidades primarias y secundarias correlacionan de manera opuesta en varios de ellos (Figura 18 d). Todo esto podría estar indicando que el desarrollo de los HV secundarios es independiente al de los HV primarios en la especie C₄.

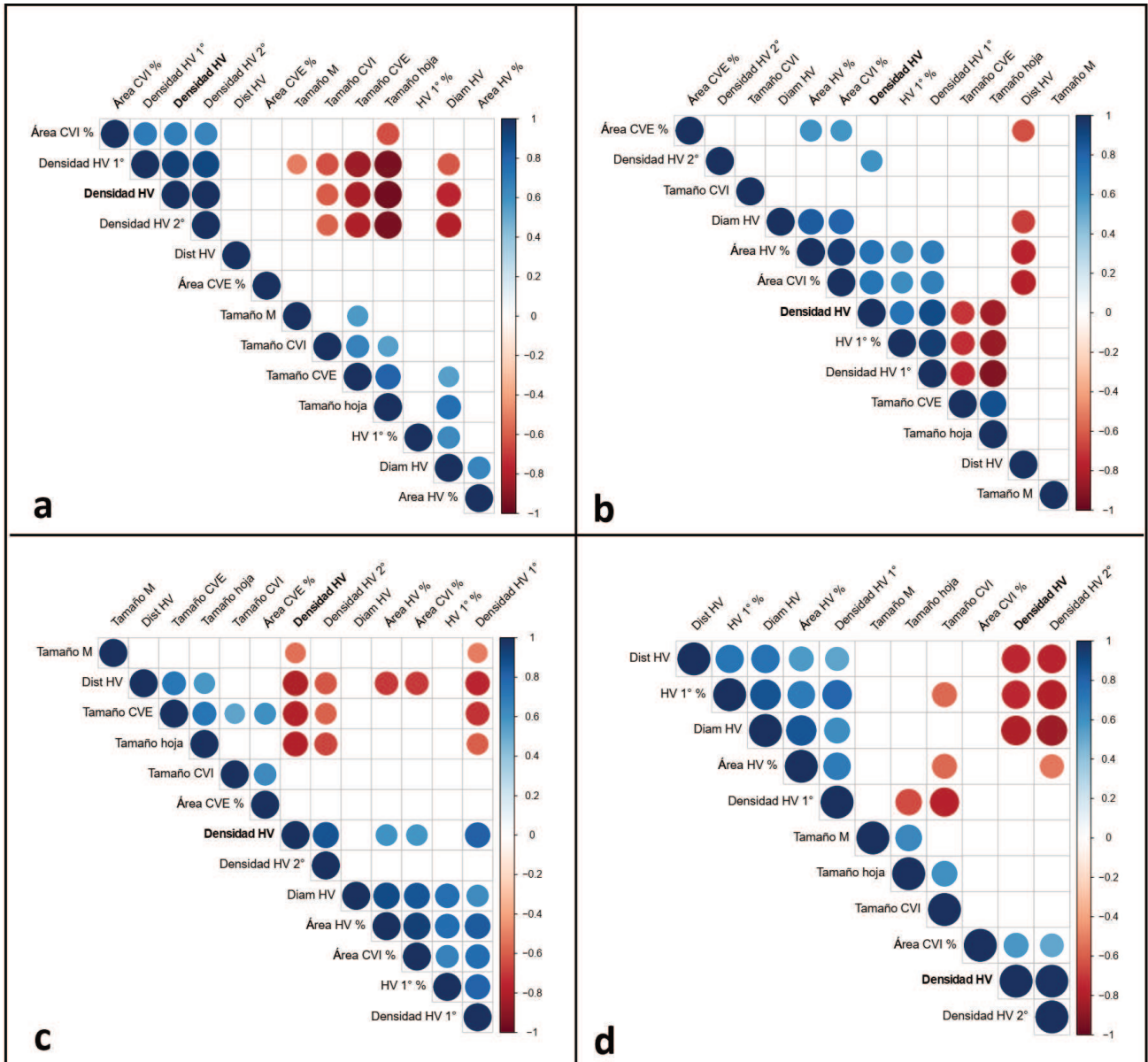


Figura 18. Correlación entre parámetros del desarrollo foliar. Correlación de spearman entre rasgos fenotípicos para 40 muestras. *H. amplexicaulis* (a), *R. pilosa* (b), *S. hians* (c) y *A. lanata* (d). Se muestran sólo las correlaciones con un valor p menor a 0,001. El tamaño de los círculos es proporcional a la significancia y la intensidad del color al valor de la correlación.

Comparación de rasgos fenotípicos entre especies

En la figura 19 se muestran los resultados de comparar los parámetros asociados a los HV en los F5 entre las cuatro especies. Como se esperaba, *A. lanata* presentó una menor distancia

entre HV y por lo tanto una mayor densidad de HV. A su vez, se observa un gradiente ascendente en la densidad de HV desde la especie C₃, pasando por las especies PK y C₂ (Figura 19 a-b).

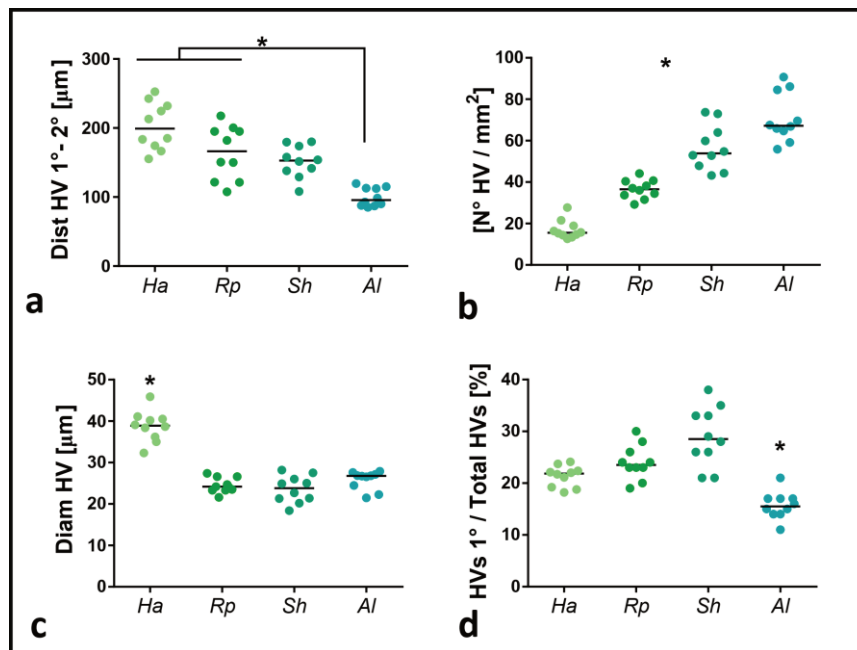


Figura 19. Comparación de rasgos fenotípicos asociados a los haces vasculares (HV) en F5. Distancia entre el HV central y sus HV contiguos (a), densidad de HV (b), Diámetro promedio de HV (c) y Fracción de HV primarios sobre el total de HV (d). Abreviaturas: *H. amplexicaulis* (Ha), *R. pilosa* (Rp), *S. hians* (Sh), *A. lanata* (Al).

El tamaño promedio de los HV resultó ser significativamente mayor en *H. amplexicaulis* que en el resto de las especies (Figura 19 c), mientras que la proporción de HV primarios sobre el total de HV resultó ser significativamente menor en *A. lanata* (Figura 19 d).

Con respecto a los tamaños celulares, la especie C₄ presentó los mayores tamaños promedio de CVI y de M (Figura 20 a y c). En cuanto al tamaño de la CVE, *H. amplexicaulis* posee un tamaño promedio superior a las otras dos especies no C₄ (Figura 20 b). La fracción foliar ocupada por las CVI resultó ser significativamente mayor en *A. lanata* (Figura 20 d) y la ocupada por la CVE presentó el mayor porcentaje en *S. hians* y el menor en *R. pilosa* (Figura 20 e).

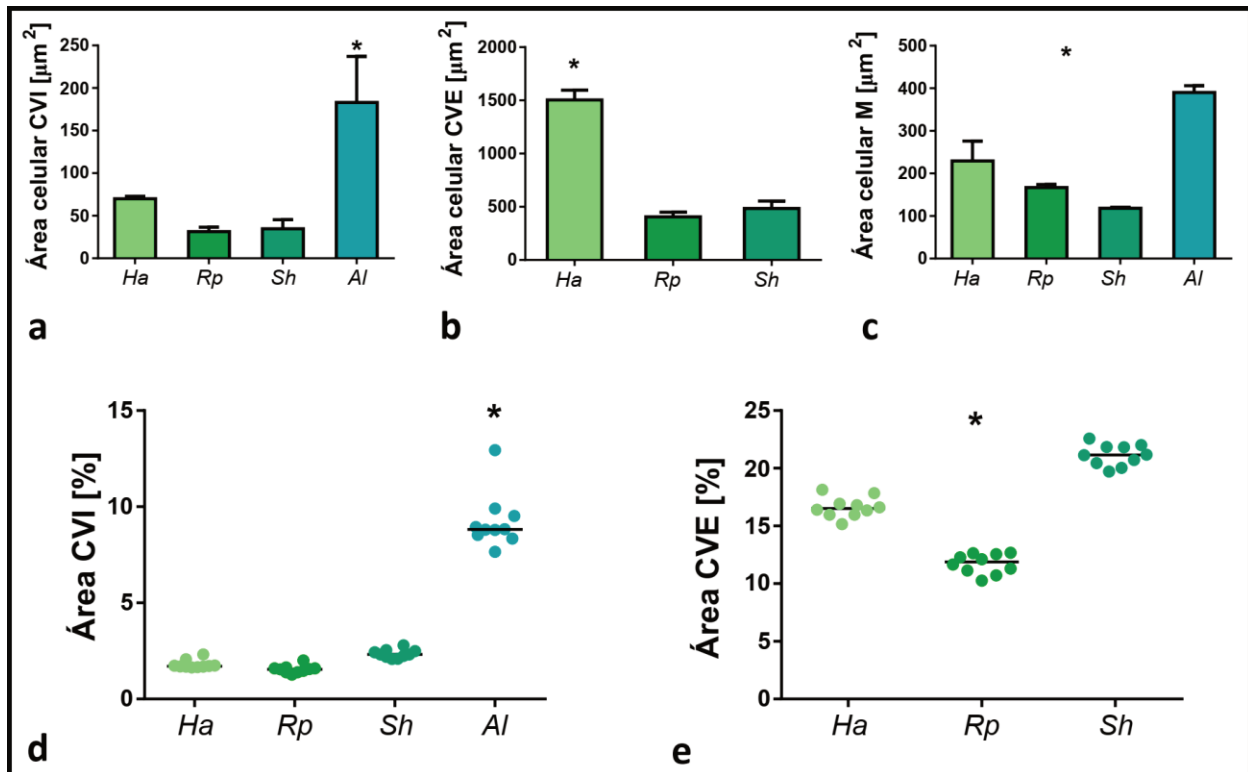


Figura 20. Comparación de rasgos fenotípicos asociados a las CV y el M. Tamaño celular promedio de las CVI (a), CVE (b) y M (c). Fracción ocupada por las CVI (d) y CVE (e). Abreviaturas: *H. amplexicaulis* (Ha), *R. pilosa* (Rp), *S. hians* (Sh), *A. lanata* (Al).

La necesidad de mantener un transporte significativo de metabolitos entre CV y M en las especies C_4 y C_2 , hace suponer la necesidad de una adaptación anatómica que facilite el transporte. Además de la densidad y tamaño de los plasmodesmos, esto se puede conseguir aumentando la superficie de contacto entre ambos tipos celulares. El parámetro de superficie S_b , definido por Pengelly y colaboradores (2010) se calcula a partir de un corte transversal midiendo el perímetro de las CV dentro del espacio comprendido entre dos haces y dividiéndolo por la distancia entre haces [168]. Sin embargo, este parámetro no tiene en cuenta la organización no lineal de los haces en *A. lanata*, por lo que se decidió sumar una forma alternativa de ese parámetro, que llamamos S_b^* . Este nuevo parámetro es calculado como el perímetro de todas las células de la vaina en contacto con el mesófilo dividido el área transversal de la hoja. Los resultados muestran que *S. hians* y *A. lanata* poseen superficies de contacto, estimadas como S_b y S_b^* , significativamente mayores a las especies C_3 . Así mismo, este patrón se observa en el área de contacto entre las CV y los HV, aquí denominadas S_v y S_v^* , y entre las CVI y CVE, aquí denominadas S_{ib} y S_{ib}^* (Figura 21).

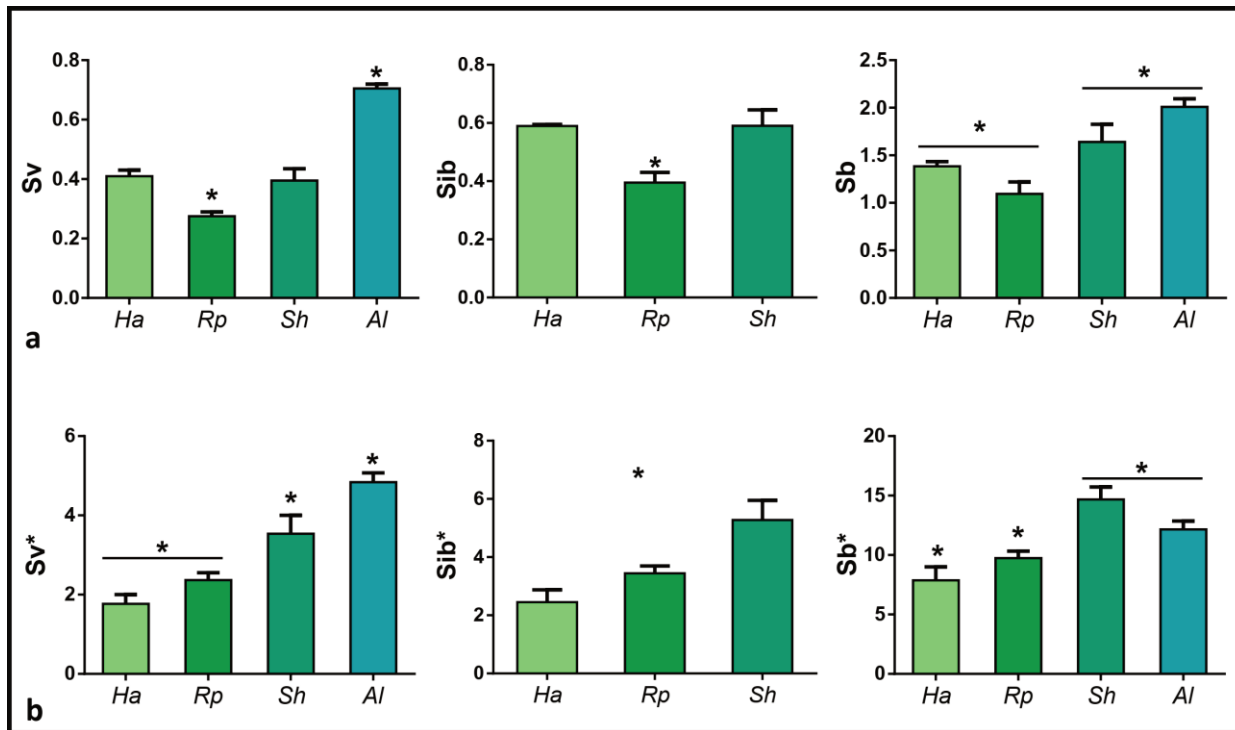


Figura 21. Comparación de superficies de contacto entre HV y CV (S_v y S_v^*), entre CVI y CVE (S_{ib} y S_{i^*}) y entre CV y M (S_b y S_b^*). Superficies de contacto normalizadas a la distancia entre HV (a) y al área foliar transversal (b). Abreviaturas: *H. amplexicaulis* (Ha), *R. pilosa* (Rp), *S. hians* (Sh), *A. lanata* (Al).

Fisiología de las especies estudiadas

Las curvas de respuesta de fotosíntesis versus CO_2 (Asimilación vs Concentración de CO_2 en el espacio intercelular de la hoja o A vs C_i) muestran el carácter C_4 de *A. lanata* evidenciado por una pendiente inicial de la curva mayor que la de las especies C_3 (Figura 22). Los puntos de compensación de CO_2 aparentes (PC de CO_2), es decir la concentración de CO_2 donde la tasa de fotosíntesis iguala a la tasa respiración en términos de asimilación de carbono, muestran una diferencia significativa entre *A. lanata* y *S. hians* con respecto a las especies C_3 (Tabla 6). Si bien se podría esperar un menor PC de CO_2 para *R. pilosa* con respecto a *H. amplexicaulis*, debido a su carácter Proto-Kranz [39], el mayor valor puede deberse a dos razones: por un lado una razón intrínseca relacionada con el solapamiento de los rasgos fenotípicos de las especies perteneciente o cercanas a un linaje C_4 ; y por otro lado a una dificultad en poder obtener mediciones precisas en condiciones de muy bajo CO_2 con el equipamiento disponible.

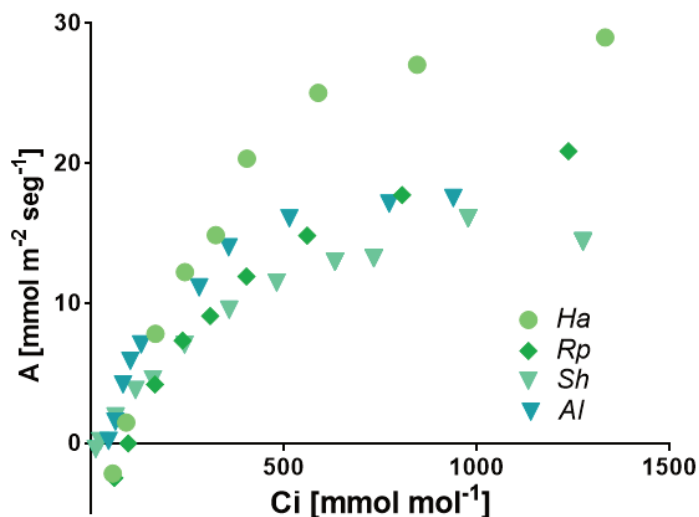


Figura 22. Curvas A vs Ci. Curvas de respuesta al CO₂ para las especies estudiadas. El ensayo está realizado a 27°C y 1500 μmol fotones m⁻²s⁻¹ con tres individuos por especie. Abreviaturas: *H. amplexicaulis* (Ha), *R. pilosa* (Rp), *S. hians* (Sh), *A. lanata* (Al).

En las especies C₄ el flujo de los ácidos C₄ desde las células del M hacia las CV tiene que ser igual o superior a la tasa de asimilación de CO₂. Esto puede ser calculado como la asimilación de CO₂ (A) dividido el área de contacto entre células de la vaina y el mesófilo (Sb). En las especies no C₄, los parámetros A/Sb y A/Sb* pueden interpretarse como una estimación del flujo necesario si la vía C₄ fuera instalada sin modificaciones anatómicas extras, con las CVE como compartimento Kranz, mientras que los parámetros A/Sib y A/Sib* representan la estimación si el compartimento Kranz fuera las CVI.

Tabla 6. Parámetros fisiológicos relacionados a la fotosíntesis en hojas maduras de *A. lanata*, *S. hians*, *R. pilosa* y *H. amplexicaulis*. Para los parámetros de asimilación por superficie BS, se asume que en la especie C₄ el mínimo flujo de los ácidos C₄ a través de los PD debe ser igual o mayor a la tasa de asimilación de CO₂. Para el resto de las especies, los valores indican los flujos requeridos si se instalara una vía C₄ sin cambios anatómicos. Ambos valores están normalizados a la especie C₄.

Especie	A [mmol m ⁻² seg ⁻¹]	PC de CO ₂ [ppm]	Asimilación por superficie de CVE		Asimilación por superficie de CVI	
			A/Sb	A/Sb*	A/Sib	A/Sib*
<i>H. amplexicaulis</i>	31,36	72,80	2,59	2,69	6,16	9,01
<i>R. pilosa</i>	23,21	92,40	2,40	1,69	6,45	4,74
<i>S. hians</i>	19,10	16,40	1,29	0,93	3,69	2,51
<i>A. lanata</i>	17,82	26,75	1,00	1,00	1,00	1,00

En la tabla 6, estos parámetros se muestran relativizados a los valores de *A. lanata*. Los resultados muestran que de instalarse una vía C_4 en *H. amplexicaulis*, el flujo entre M y CVE debería ser alrededor de 2,59 ó 2,69 veces el flujo que posee *A. lanata*. Mientras que, si la vía se instalara en la CVI, los flujos deberían ser más intensos todavía. La especie Proto-Kranz y especialmente la especie C_2 poseen valores más bajos de estos parámetros, evidenciando el impacto que tienen las modificaciones anatómicas en la posibilidad de establecer una vía C_4 en sus hojas.

DISCUSIÓN

Existen diferencias estructurales en la anatomía foliar desde el inicio del gradiente de desarrollo

El estudio de la anatomía foliar a través del análisis de secciones de hojas jóvenes, demostró la existencia de un gradiente de desarrollo foliar en aspectos cualitativos y cuantitativos. Las especies estudiadas presentan similitudes en diversos parámetros anatómicos foliares analizados. Desde la base al ápice de la hoja se destacan: (a) un incremento gradual en la lignificación de HV primarios, (b) un aumento gradual en la cantidad de HV secundarios, (c) una disminución de los tejidos de sostén, (d) un aumento de la presencia de cloroplastos y (e) una mayor diferenciación entre CV y M. La formación de los HV primarios y sus CV asociadas parece formar parte de una etapa anterior en la ontogenia foliar.

Resulta interesante remarcar que, a partir de este trabajo, también se destacan diferencias sustanciales de rasgos anatómicos entre las especies estudiadas. Estas diferencias se observan desde el F1. La especie C_4 presentó una mayor variación de los aspectos cualitativos de diferenciación desde el F1 al F7. En particular, la mayoría de los rasgos fenotípicos presentaron un cambio importante entre el F1 y F3, con excepción de la fracción de tejido ocupada por CVI y el tamaño de las células del mesófilo. Sin embargo, los datos obtenidos sugieren que la anatomía foliar de esta especie varió levemente en términos estructurales a lo largo del gradiente de desarrollo en comparación con las otras especies.

Las especies C_3 , PK y C_2 presentan una gran proporción de tejido de sostén en la zona del HV central, que se va reduciendo notablemente a medida que la hoja madura. Esto produce una distancia significativamente mayor entre el lado adaxial y abaxial en esa zona de la hoja. En cambio, en la especie C_4 esta distancia entre los lados abaxial y adaxial es más constante tanto entre la zona del haz principal y los extremos, como entre los diferentes fragmentos. Así mismo, las células del M en estas especies, a pesar de no presentar tamaños significativamente diferentes entre estados del desarrollo, presentan un cambio en la morfología de las células desde una forma cilíndrica en el F1 a una forma más rectangular o hexagonal alargada en los F5 y F7. El tamaño de las células del M también parece ser una característica asociada a cada especie, siendo las células del M de la especie C_4 las más grandes de todas.

Las hojas de *H. amplexicaulis* a pesar de presentar cambios muy graduales en los caracteres fenotípicos cuantitativos, muestran un cambio estructural notorio durante el gradiente de desarrollo. Este cambio estructural se centra en tres características principales tales como forma de la hoja, disminución de los tejidos de sostén y modificación de la forma de las células del M. A su vez, esta especie es la única donde se observa una correlación entre el tamaño de las células del mesófilo con el tamaño de las CV.

Las especies *R. pilosa* y *S. hians*, son las que presentan la mayor similitud en términos de estructura foliar (forma de la hoja, distribución de los HV, proporción de tejido de sostén, etc.) pero se evidencia un mayor grado de “carácter Kranz” en la especie C_2 dado por la localización de cloroplastos dentro de la CVE, la densidad de los HV, la fracción ocupada por las CVE y las superficies de contacto entre los diferentes tipos celulares.

Estudios previos demostraron que no hay una transición fisiológica C_3 - C_4 durante el desarrollo foliar, sino que el ciclo C_4 está activo desde el principio ([169], [170]). En línea con estos trabajos, los datos aquí generados muestran que la existencia de este gradiente de desarrollo no implica una localización tisular diferencial de los cloroplastos a lo largo del mismo, en especial en la especie C_4 . Sin embargo, para confirmar la disposición de los cloroplastos y su estructura con mayor precisión se debería recurrir a un análisis con microscopía de transmisión.

La comunicación entre CV y M como indicador del grado “Kranz” de una especie

Se ha sugerido que una mayor fracción ocupada por las CV en especies C_4 en comparación con especies C_3 , es uno de los rasgos que distinguen a la anatomía C_4 ([24], [28], [30], [37]). En este

trabajo, las comparaciones anatómicas se realizaron sobre una especie C_4 de tipo NADP-ME que no posee CVE. Hasta el día de hoy no existe evidencia suficiente que permita identificar el momento en el que la selección natural optó por la CVI o CVE como compartimento Kranz y el momento y las razones por las cuales las especies C_4 NADP-ME perdieron las CVE. En relación a esto, se piensa que las especies C_4 que utilizan a la CVI como compartimento Kranz provienen de un ancestro C_4 con CVE [30]. Incluso, el rol de la vía C_2 como un estado de transición o una vía paralela a la vía C_4 , también es materia de debate ([41], [165], [171]). Por ejemplo, el tamaño de las CVI en la especie C_4 es hasta 5,8 veces más grande que en el resto de las especies, pero las CVE en las especies no C_4 son entre 2 y 8 veces más grandes que las CVI de la C_4 . Del mismo modo, la fracción transversal ocupada por las CVI de la especie C_4 es menor a cualquiera de las fracciones CVE. Estos resultados coinciden con lo reportado por Christin y colaboradores y por Lauterbach y colaboradores ([30], [37]).

Si bien los rasgos fenotípicos cuantitativos asociados a los HV también están condicionados por las diferencias estructurales entre las especies, en la especie C_4 son notablemente diferentes (configuración espacial, proporción $1^\circ/2^\circ$, tamaños). Esto sugiere, en línea con lo obtenido por otros grupos de trabajo, que la densidad de HV probablemente no es una precondition en un ancestro C_3 que promueve el origen de la fotosíntesis C_4 , sino que es un rasgo que evolucionó en una fase posterior ([36], [37]).

La importancia del tamaño de la CV como rasgo Kranz viene dada por el supuesto de que la CV debe ser lo suficientemente grande como para albergar una gran cantidad de cloroplastos que realicen el ciclo de CBB y así mantener una buena eficiencia fotosintética [28]. Sin embargo, recientemente se ha destacado la importancia del transporte de ácidos C_4 por medio de plasmodesmos entre mesófilo y CV como un limitante de este proceso ([12], [13]). En este sentido, se ha demostrado que las especies C_4 poseen una mayor densidad de plasmodesmos en las interfases entre CV y M, como así también una mayor superficie de contacto entre ambos tipos celulares, mientras que los volúmenes de las CV no son significativamente diferentes entre especies C_3 y C_4 [12]. Dada la falta de información de densidad de plasmodesmos en especies intermedias, el lugar de este evento en la evolución de la fotosíntesis C_4 es desconocido.

En las especies estudiadas aquí, se identificó un gradiente ascendente de los parámetros de superficie de contacto desde la especie C_3 hasta la especie C_4 . Al cruzar estos datos con información fisiológica de tasas de asimilación de CO_2 , se puede observar que la instalación de una vía C_4 en las especies C_3 requeriría de un aumento previo de la superficie de contacto y/o un aumento en la densidad de plasmodesmos para garantizar que este rasgo no condicione la capacidad fotosintética. Sin embargo, la especie C_2 pareciera tener una superficie de contacto similar a la especie C_4 . Esto no sería

una sorpresa si se tiene en cuenta que el ciclo C_2 implica también un mecanismo de concentración de CO_2 en el sitio de acción de RUBISCO.

Resulta interesante destacar que *S. hians* y *A. lanata*, logran incrementar la superficie de contacto en parte a través de estrategias diferentes. Si bien *S. hians* presenta CVE con un tamaño similar a las de *R. pilosa* y mucho menor que la de *H. amplexicaulis*, su alta superficie de contacto se explica por un aumento de la densidad de HV y una disminución en la cantidad de células del mesófilo entre los HV. Comparando a *S. hians* con *A. lanata* vemos que esta última posee CVI que son menos de la mitad del tamaño que las CVE de la primera. Sin embargo, *A. lanata* posee una mayor densidad de HV, en especial de HV secundarios, logrando una superficie de contacto similar con CV más pequeñas. Una tendencia similar se observó en el género *Flaveria* y otras especies de gramíneas que usan a la CVI como compartimento Kranz ([30], [172], [173]). La determinación de las densidades y tamaños de los plasmodesmos en estas especies ayudarían a corroborar estas suposiciones.

Finalmente, es importante remarcar la gran variabilidad natural de los rasgos fenotípicos cuantificados entre individuos de una misma especie. Esto sugiere la necesidad de trabajar con un gran número de réplicas para obtener resultados que representen más fielmente a la población estudiada a la hora de establecer marcadores fenotípicos de importancia en el establecimiento de la anatomía Kranz.

CAPÍTULO 3

Análisis de la transcriptómica del gradiente foliar en Otachyriinae

INTRODUCCIÓN

Ensamblado de transcriptomas de novo y el gradiente de maduración

En la última década, la secuenciación de transcriptomas (RNAseq) ha demostrado ser una técnica poderosa para entender la versatilidad de los mecanismos moleculares y abordar diversas preguntas biológicas ([86], [174], [175],[176]–[179]). En particular, ante la ausencia de genomas de referencia para organismos no modelos, la secuenciación de ARN es usada para reconstruir y cuantificar transcriptomas completos en un proceso llamado ensamblaje de novo [180]. Para esto, existe una gran diversidad de métodos y programas que resultan más o menos eficientes en función del tipo de muestra, la especie estudiada y el análisis posterior requerido [181].

Además del tradicional análisis de expresión diferencial, los ensayos de transcriptómica pueden usarse para reconstruir las redes de expresión de los genes y obtener información adicional sobre las interacciones entre los mismos. El análisis de redes de co-expresión de genes ponderadas (WGCNA, del inglés: *Weighted Gene Co-expression Network Analysis*) es un tipo de análisis de la biología de sistemas que describe la correlación entre genes en un gran número de muestras de un microarreglo o un RNAseq [182]. Puede ser usado para encontrar grupos de genes altamente relacionados (módulos) y resumir sus comportamientos usando parámetros específicos que explican sus diferencias de expresión y conectividad en la red. Este enfoque permite identificar módulos de genes asociados con determinados tipos celulares, especies y rasgos fenotípicos, así como también identificar miembros que pueden cumplir un rol regulatorio en procesos biológicos ([182], [183]).

Las hojas en desarrollo de las gramíneas proveen un gradiente de sucesivos estados de desarrollo, de la base a la punta, que puede servir como fuente de datos comparables entre sistemas biológicos. A lo largo de este gradiente, se puede observar la aparición de eventos celulares, fisiológicos y moleculares que se corresponden con estados de desarrollo distintivos [86], [178]. Desde la base hasta la punta de la hoja se organizan zonas celulares de división, expansión/elongación y diferenciación. Estas zonas señalan la transición desde un tejido respiratorio (o tejido sumidero) a uno fotosintético (o tejido fuente) y la aparición o desaparición de otras características anatómicas y actividades bioquímicas que corresponden a los estados de desarrollo de la hoja madura ([86], [178], [184]). Hasta el momento, estudios que aprovechen este gradiente foliar de las gramíneas fueron realizados solamente en *Z. mays* ([86], [178]).

Para que los enfoques sistémicos sean exitosos en entender la regulación génica de la fotosíntesis C₄ deberían cumplir determinados criterios: (a) el muestreo debe incluir estados de desarrollo durante los cuales queden determinadas las características anatómicas, (b) las secuenciaciones deben poseer una buena profundidad y cobertura para poder capturar la mayor cantidad de los componentes interactuantes posibles, (c) la toma de muestras y las condiciones ambientales donde se lleva a cabo deben ser rigurosas y reproducibles y (d) la información proveniente de diferentes sistemas debe venir de materiales biológicos que presenten el mismo estado de desarrollo [184].

Los rasgos foliares que caracterizan a las especies C₄ parecen ser resultados de ajustes espaciales y cuantitativos de rasgos C₃ sucedidos a lo largo del desarrollo foliar [184]. Esto significa que es más probable que la vía C₄ haya evolucionado por medio de la regulación de genes existentes, y no a través de la evolución de nuevos genes [175]. Esto es consistente con el gran número de linajes independientes en los cuales las condiciones para la fisiología C₄ fueron alcanzadas durante la evolución.

Hasta el momento no existen estudios que aborden la transcriptómica foliar de especies de Otachyriinae. Recientemente, Carvalho y colaboradores (2019) ensamblaron los transcriptomas de tres especies de gramíneas incluyendo el de *H. amplexicaulis* [185]. Sin embargo, la reconstrucción se realizó con fines de establecer la presencia/ausencia de genes de la vía C₄ y las relaciones filogenéticas entre las especies utilizadas. Además, al ser reconstruido por medio de la secuenciación de una sola biblioteca, el transcriptoma no ofrece datos sobre el desarrollo foliar [185].

Por todos estos motivos, se propuso entender los mecanismos moleculares responsables de la diversificación de los subtipos fotosintéticos en gramíneas partiendo del estudio de datos provenientes de RNAseq en especies no modelo. Para ello se construyeron los transcriptomas de novo para tres especies de la subtribu Otachyriinae considerando el gradiente foliar. Luego, se investigó la dinámica transcripcional, la expresión de los genes a lo largo del gradiente foliar y se identificaron diferencias y similitudes entre el desarrollo foliar de las especies.

OBJETIVOS ESPECÍFICOS

Evolución de los patrones de expresión: Estudiar el programa de desarrollo de las hojas a través de la secuenciación masiva del ARN foliar de diversas especies de Otachyriinae. Identificar cambios

en los patrones de expresión que se correlacionen con cambios anatómicos y subtipos fotosintéticos considerando la evolución de las especies

MATERIALES Y MÉTODOS

Condiciones de cultivo y toma de muestras

Individuos de 3 especies de Otachyriinae con diferentes subtipos fotosintéticos (*H. amplexicaulis*, *R. pilosa* y *A. lanata*) fueron crecidos en cámaras de cultivo a 27°C bajo condiciones de día largo (16 horas de luz y 8 horas de oscuridad). Para cada especie, se colectó la 5° hoja joven de 90 individuos, en el momento que la misma alcanzó en longitud el tamaño de la 4° hoja (Anexo 1, Figura S4). Cada hoja fue dividida en dos secciones, tomando como referencia la zona de transición fuente-sumidero. Las secciones fueron luego divididas en cuatro fragmentos de igual longitud y rotulados F1 a F8 desde la base a la punta de la hoja. Las muestras fueron organizadas en 9 grupos de 10 individuos. Los grupos fueron aleatorizados para formar 3 réplicas de 30 individuos cada una totalizando cuatro fragmentos por especie con tres réplicas cada uno (F1, F3, F5 y F7). Las réplicas se mantuvieron pareadas a lo largo del análisis.

Extracción de ARN y secuenciación

La extracción de RNA total se realizó con el reactivo Tripure (Sigma), siguiendo el protocolo del fabricante. Para purificar el RNA y remover la contaminación de ADN en las muestras se llevó a cabo una precipitación con LiCl [186].

Las bibliotecas de ARNm y la secuenciación masiva se realizaron empleando el servicio *The High-Throughput Sequencing and Genotyping Unit* de la Universidad de Illinois en Urbana-Champaign (IL, USA). Las bibliotecas fueron construidas con el kit “*TruSeq Stranded mRNA*” (Illumina) siguiendo las instrucciones del fabricante y posteriormente secuenciadas con la plataforma Illumina con lectura “pair-ends” (PE) y fragmentos de 250 pb. Luego de la secuenciación se obtuvieron 36 bibliotecas con 25 a 40 millones de lecturas en cada una.

Ensamblaje de novo de transcriptomas

La calidad de las lecturas PE crudas de cada biblioteca fue examinada usando la herramienta FastQC v. 0.11.5 [187]. Las lecturas de baja calidad fueron recortadas y filtradas con el programa

Trimmomatic v. 0.38 [188]. Luego de este proceso, entre el 3% y el 7% de las lecturas fueron descartadas.

El ensamblaje de novo de los transcriptomas se realizó con el paquete Trinity v.2.4.0 [189]. La proporción de lecturas mapeadas al ensamblado fue evaluada con Bowtie2 v. 2.3.2 [190]. Para reducir la probabilidad de obtener transcriptos espurios y atenuar la redundancia de transcriptos, los *contigs* fueron filtrados utilizando tres métodos. Primero, las isoformas expresadas débilmente fueron removidas de acuerdo a sus valores de expresión. Para ello, se obtuvieron valores de lecturas normalizadas TPM (transcriptos por millón de kilo bases. Del inglés: *Transcripts Per Kilobase Million*) con SALMON v. 0.14.1 [191], y las isoformas débilmente expresadas se removieron con el script de Trinity, *filter_low_expr_transcripts.pl* con el parámetro “*-highest_iso_only*”. Segundo, un conjunto de transcriptos representativos no-redundantes fue generado con el paquete CD-Hit v. 4.6.6, con un nivel de identidad de 95% [192]. Finalmente, se usó Transdecoder v. 5.1.0 para identificar todas las posibles regiones codificantes en los transcriptos ensamblados, y luego se filtró seleccionando el mejor marco abierto de lectura (ORF) por transcripto [193]. Los transcriptos con ORFs menores a 100 pb de longitud fueron removidos de posteriores análisis.

La evaluación del grado de completitud del transcriptoma se realizó con la herramienta BUSCO v.3 para obtener el porcentaje de ortólogos de copia única representados en el conjunto de datos de monocotiledóneas (3278 genes) [194].

Anotación de genes

Para la correcta identificación de secuencias ortólogas se utilizó Orthofinder v. 1.0.6, con seis proteomas de referencia [195]. Dos proteomas ensamblados de novo de la misma subtribu, *S. hians* y *Steinchisma laxa*, sin publicar aún, fueron provistos por el laboratorio del Dr. Studer. Cuatro proteomas de gramíneas fueron obtenidos de la base de datos de Phytozome v. 12.1.6 [129]: *Zea mays*, *Sorghum bicolor*, *Setaria viridis* y *Oryza sativa*. Para cada transcripto, se extrajo la información funcional de su ortólogo en *S. bicolor*, *S. viridis* o *Z. mays*.

Análisis de expresión diferencial

La cuantificación de los transcriptos fue realizada con la herramienta de Trinity “*align and estimate abundance*”, por medio del mapeo de las lecturas de cada réplica biológica contra su transcriptoma ensamblado respectivo. Para ello, se utilizó SALMON como método para estimar la abundancia y la herramienta “Ptr” del paquete Trinity para el control de calidad de las réplicas. Las matrices de expresión génicas fueron construidas con la herramienta “*abundance estimates to matrix*”

del paquete Trinity y fueron importadas a R [196] donde se utilizaron para el análisis de expresión diferencial con el paquete *EdgeR* [197]. Para filtrar los Transcritos Poco Expresados (TPE) en cada especie, se mantuvieron aquellos genes que estaban expresados al menos 1 cpm (lecturas por millón, del inglés: *counts per million*) en la mitad de las bibliotecas. Para estimar los genes expresados diferencialmente (GED) se calculó su tasa de cambio logarítmica (logFC, del inglés: *logarithmic fold change*) entre pares y la significancia de cambios de expresión se determinó usando test-t de Student. Los valores p se ajustaron con el método Benjamini-Hochberg para corregir la tasa de descubrimiento falso (FDR, del inglés: *False discovery rate*). Un gen fue considerado expresado diferencialmente si presentó un valor p ajustado $< 0,001$ y $\logFC > 2$.

Para estudiar las diferencias de expresión entre las especies, se estimó el número de GED entre fragmentos foliares homólogos utilizando los ortólogos inferidos por Orthofinder. Dado que en algunos casos los ortogrupos (OG) presentaron diferente número de copias entre especies, se optó por sumar los valores de expresión de los transcritos y combinar la información de la anotación. Para su tratamiento posterior (expresión diferencial, enriquecimiento de Gene Ontology y redes de co-expresión) los OG fueron considerados análogos de genes. La obtención de los ortogrupos expresados diferencialmente (OED) entre especies se realizó con edgeR de manera similar a la descrita en el párrafo anterior.

Análisis de enriquecimiento de términos de Gene Ontology (GO)

El análisis de enriquecimiento de términos de ontología génica (GO) se realizó con el paquete de R TopGO [198]. Los términos GO para cada transcritos fueron obtenidos de la base de datos de Phytozome. Para buscar términos enriquecidos que sean estadísticamente significativos se calcularon valores p utilizando el algoritmo “elim”.

Análisis de redes de co-expresión de genes ponderado (WGCNA)

El paquete de R WGCNA (*Weighted Gene Co-expression Network Analysis*) fue usado para la construcción de las redes [182]. El análisis construye redes usando el coeficiente de correlación de Pearson como una medida de la co-expresión de genes, el cual es elevado a una potencia para crear

una matriz de adyacencias. Para la construcción de la red se utilizó la matriz de expresión de OG. Para obtener una red libre de escala se usó una potencia de 8 al construir la matriz de adyacencias. La distancia de solapamiento topológico calculada a partir de la matriz de adyacencias es luego clasificada con el método de agrupamiento jerárquico de vinculación promedio. Los módulos fueron definidos usando la función “*cutTreeStatic*” con un tamaño mínimo de módulo de 30 miembros y una altura de corte dinámica. Un umbral de distancia de 0,3 en el valor eigengen del módulo (EM) fue usada para unir módulos muy similares. La calidad de los módulos se evaluó mediante tres factores: (1), observando la similitud en los patrones de expresión de los OG por medio de *heatmaps*; (2), evaluando los promedios de conectividad intramodular (k_{im}) de cada módulo y (3), evaluando la correlación entre el parámetro de membresía al módulo (MM) y la k_{im} para cada OG. Los OG con valores de MM en el primer decil del módulo fueron considerados OG centrales o *hubs*. Para determinar su importancia los *hubs* se ordenaron por valores decrecientes de k_{im} .

RESULTADOS

Obtención de los transcriptomas

Filtrado y control de calidad

El ensamblaje de novo de los transcriptomas fue llevado a cabo utilizando el paquete Trinity. Luego del primer ensamblado se detectó un gran número de transcritos por gen (Tabla 7). Dados los objetivos del trabajo y la imposibilidad de determinar si las isoformas tienen sentido biológico o son artefactos del ensamblador, se decidió continuar conservando sólo las isoformas más expresadas. La tasa de mapeo luego de la eliminación de las isoformas disminuyó a 90,6%, 90,7% y 87,0% para *H. amplexicaulis*, *R. pilosa* y *A. lanata* respectivamente.

Tabla 7. Estadísticas de los primeros ensamblajes de novo utilizando Trinity. El ensamblado se realizó combinando las 12 bibliotecas para cada especie.

	<i>H. amplexicaulis</i>	<i>R. pilosa</i>	<i>A. lanata</i>
N° de transcritos	245892	164363	355460
N° de genes	157690	69387	167531
Largo promedio	1378	1783	1018
Contig N50	2623	2766	1653

Tasa de mapeo SALMON [%]	97,5	96,9	95,2
Transcriptos > 1 TPM	34121	32165	32126

En segundo lugar, se usó CD-HIT para eliminar secuencias redundantes y transdecoder para identificar a todas las potenciales regiones codificantes en los transcriptos ensamblados. Las características de las versiones finales de los transcriptomas ensamblados se muestran en la Tabla 8. Al filtrar las secuencias por medio de la selección del mejor ORF, el número de secuencias en los transcriptomas se redujo aproximadamente a un tercio, y la tasa de mapeo de lecturas cayó desde alrededor del 90% al 60% (Tablas 7 y 8).

Para evaluar el contenido de secuencias duplicadas y el nivel de completitud de los transcriptomas se realizó una búsqueda con BUSCO (Evaluación comparativa de ortólogos universales de copia única. Del inglés: *Benchmarking Universal Single-Copy Orthologs*) con el conjunto de datos de genes de copia única de las monocotiledóneas. Los resultados muestran bajos niveles de secuencias ausentes (11,0 a 12,7%, Tabla 8) y muy bajos de secuencias duplicadas (1,7% a 1,8%, Tabla 8) para los tres transcriptomas.

El número de genes codificantes de proteínas en las versiones finales de los transcriptomas varía entre las especies: 29370 para *R. pilosa*, 50890 para *A. lanata*, y 56064 genes para *H. amplexicaulis*. Si no se tienen en cuenta los transcriptos poco expresados (menor a 1 TPM), el número de genes y las diferencias entre especies se reducen notablemente: 17509 para *R. pilosa*, 18318 para *H. amplexicaulis* y 21798 para *A. lanata* (Tabla 8)

Tabla 8. Estadísticas de los transcriptomas ensamblados de novo utilizando Trinity, luego del filtrado de isoformas, secuencias redundantes y secuencias sin marcos abiertos de lectura. El ensamblado se realizó combinando las 12 bibliotecas para cada especie.

	<i>H. amplexicaulis</i>	<i>R. pilosa</i>	<i>A. lanata</i>
N° de transcriptos codificantes para proteínas	56064	29370	50890
Largo promedio	911.9	1029.5	787.7
Contig N50	1233	1377	1059
Tasa de mapeo SALMON [%]	62.7	61.6	55.9
Transcriptos > 1 TPM	18318	17509	21798
BUSCO	C:85.7%[S:83.9%, D:1.8%],F:3.3%,M:11.0%	C:83.4%[S:81.7%, D:1.7%],F:4.1%,M:12.5%	C:82.8%[S:81.1%, D:1.7%],F:4.5%,M:12.7%

Anotación funcional

Para generar la anotación funcional de los transcriptos, primero se realizó la identificación de ortólogos con el programa Orthofinder, utilizando los tres proteomas ensamblados de novo y seis proteomas de referencia: *S. hians*, *S. laxa*, *Z. mays*, *S. bicolor*, *S. viridis* y *O. sativa*. El árbol de especies reconstruido por Orthofinder a partir de las secuencias peptídicas correlaciona con lo reportado en la literatura (Figura 23, Figura 13; Acosta y col., 2019). Alrededor del 81% de los genes fueron localizados en ortogrupos (OG) y un total de 2246 OG de copia única fueron detectados entre los nueve proteomas.

Para importar información funcional a las secuencias de las especies analizadas se utilizó la información correspondiente al ortólogo de *S. bicolor*, cuando éste estaba disponible. En su ausencia, se usó la anotación del ortólogo en *Z. mays* o *S. viridis*. Un total de 25314 (44,8%), 21710 (73,9%) y 27147 (53.3%) de genes fueron anotados para *H. amplexicaulis*, *R. pilosa* y *A. lanata* respectivamente (Tabla 9).

A pesar de las diferencias en los tamaños de los transcriptomas, el número de transcriptos anotados fue similar (Figura 23). Para explorar esto, se contó el número de secuencias luego de filtrar los Transcriptos Poco Expresados (TPE), y la proporción de ellos anotados o pertenecientes a ortogrupos. La Tabla 9 muestra que la gran mayoría de las secuencias que fueron retenidas luego de filtrar los TPE son transcriptos anotados que pertenecen a ortogrupos. Además, los tres transcriptomas presentan un número similar de transcriptos anotados y totales.

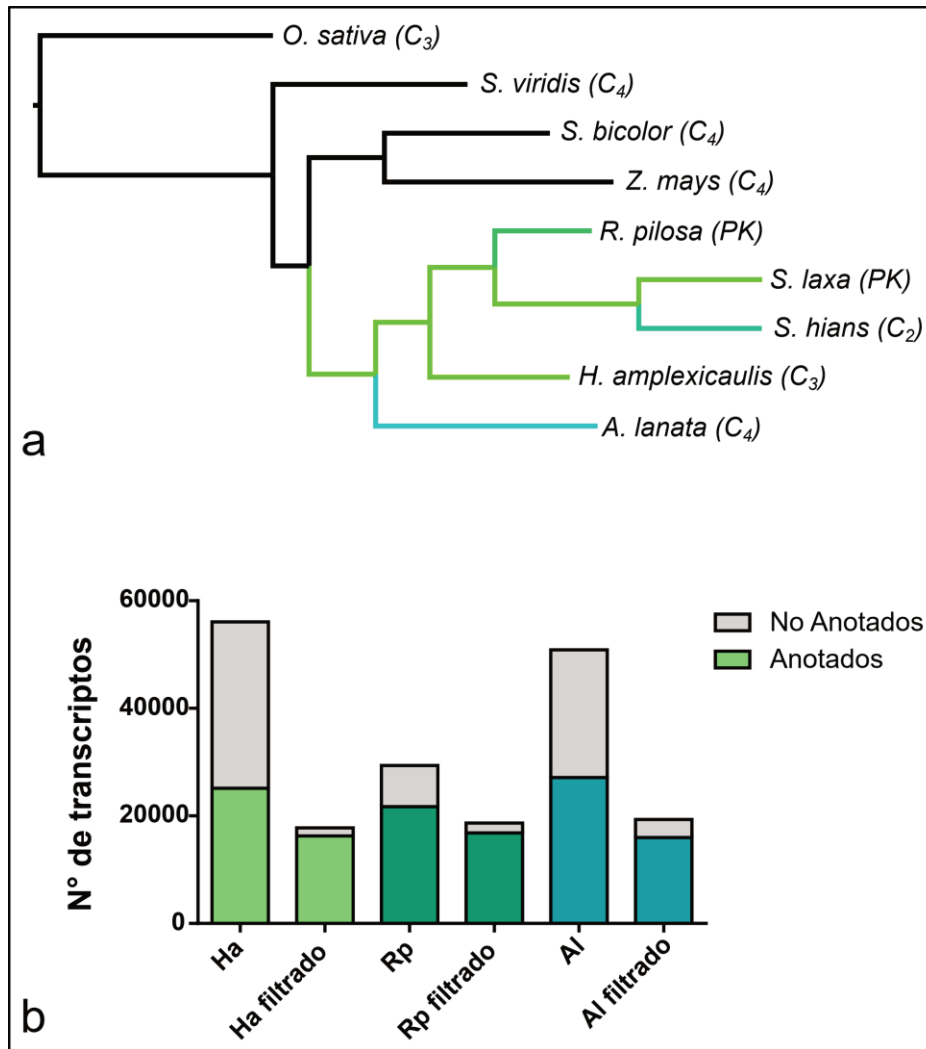


Figura 23. Características de los transcriptomas ensamblados de novo. Árbol filogenético de especies generado por Orthofinder (a) Tamaño de los transcriptomas antes y después de filtrar a los TPE (b). Abreviaturas: *H. amplexicaulis* (Ha), *R. pilosa* (Rp), *A. lanata* (Al).

Tabla 9: Estadísticas de la anotación para los ensamblajes de novo de los transcriptomas de *H. amplexicaulis*, *R. pilosa* y *A. lanata*, antes y después de filtrar los TPE (los transcritos con al menos 1 cpm en 6 bibliotecas fueron conservados).

	<i>H. amplexicaulis</i>	<i>R. pilosa</i>	<i>A. lanata</i>
N° de transcritos de ARNm	56064	29370	50890
N° de transcritos en ortogrupos	43331 [77,3%]	25727 [87,6%]	38863 [76,4%]
N° de transcritos anotados	25134 [44,8%]	21710 [73,9%]	27147 [53,3%]

Después de filtrar TPE			
N° de transcritos de ARNm	17752	18656	19300
N° de transcritos en ortogrupos	17314 [97,5%]	18146 [97,3%]	18115 [93,9%]
N° de transcritos anotados	16269 [91,6%]	16849 [90,3%]	15999 [82,9%]

Expresión diferencial de genes

La dinámica transcripcional durante el desarrollo

Para estudiar la dinámica de los transcriptomas durante el desarrollo foliar, se realizaron comparaciones de a pares entre muestras de diferentes estados de desarrollo, dentro de una especie, estimando el número de genes expresados diferencialmente (GED). Luego de filtrados los TPE, se estimó el logFC para cada transcripto entre las transiciones del desarrollo (Figura 24). La transición entre el fragmento 1 y el fragmento 3 presentó la mayor cantidad de GED en todas las especies, mientras que en las demás transiciones los GED tienden a decrecer. En general, cuando se comparan fragmentos individuales, se observó que la especie C₃ presentó siempre un menor número de GED que la especie C₄. La especie PK presentó valores intermedios, similar al máximo C₄ en la transición F1-F3 y al mínimo C₃ en las transiciones F3-F5 y F5-F7. Estos resultados sugieren que los cambios en el programa de desarrollo son más dinámicos en el transcriptoma de C₄ que en el C₃. A su vez, el patrón de expresión de la especie PK, es similar al de la especie C₄ durante la primera transición y similar al de la especie C₃ en las transiciones F3-F5 y F5-F7 (Figura 24 a-c).

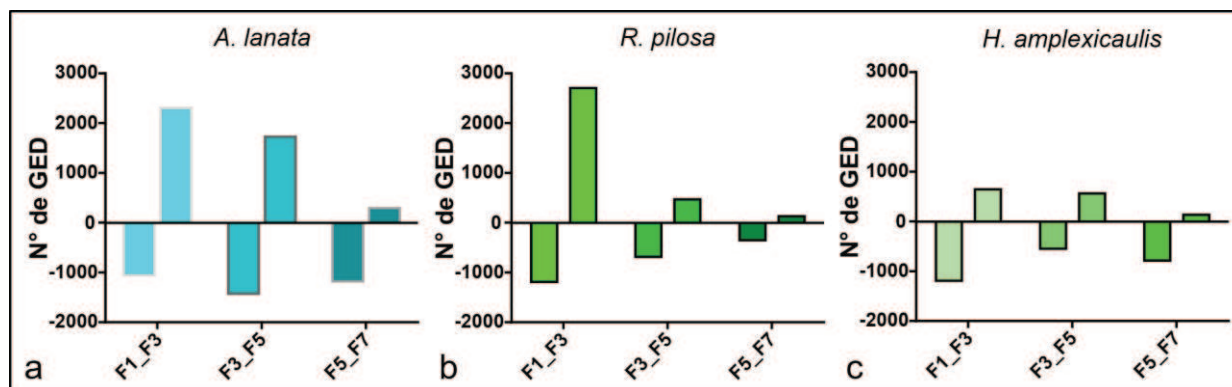


Figura 24 Análisis de expresión diferencial a lo largo del gradiente foliar. Número de genes expresados diferencialmente (GED) entre diferentes estados de desarrollo en *A. lanata* (a), *R. pilosa* (b) y *H. amplexicaulis* (c).

Con el objetivo de obtener información sobre los procesos biológicos relacionados a los GED en cada transición se realizó un análisis de ontología de genes (GO) con TopGO. En general, las tres especies presentaron vías de desarrollo conservadas, con diferencias en el comienzo y la duración de algunos de esos procesos. Los términos relacionados a la fotosíntesis, como por ejemplo la respuesta a la luz, fotosistemas, membrana tilacoide y otros, están enriquecidos en genes sobreexpresados en los fragmentos 3 vs 1 para todas las especies. En *A. lanata*, además, aparecen en genes sobreexpresados en el fragmento 5 vs el 3. Los términos relacionados al crecimiento y el desarrollo celular (términos GO asociados a replicación del ADN, modificación de la pared celular, procesos de desarrollo y división celular) están presentes en GED hacia abajo entre los fragmentos 1 y 3 para las tres especies. Además, la especie *C₄* es la única que presenta términos relacionados a traducción y biogénesis de ribosomas (Anexo 1, Hoja de datos suplementaria 1)

En la transición del F3-F5 muchos de los procesos asociados a GED caen en las categorías de óxido-reducción, metabolismo de carbohidratos y respuesta al estrés oxidativo. En particular, *H. amplexicaulis* posee transcritos sobreexpresados en el F5 en procesos de fosforilación de proteínas y respuesta de defensa.

Otros procesos que aparecen notablemente en las transiciones F3-F5 y F5-F7 en las 3 especies incluyen la respuesta a auxinas, la modificación de la pared celular y el transporte transmembrana (Anexo 1, Hoja de datos suplementaria 1).

Las diferencias entre subtipos fotosintéticos

Para estudiar las diferencias de expresión entre las especies, se estimó el número de OED entre fragmentos foliares homólogos utilizando los ortólogos inferidos por Orthofinder. En la mayoría de los casos (70,3% en *A. lanata*, 78,8% en *R. pilosa* y 62,8% en *H. amplexicaulis*), los ortogrupos construidos son de copia única dentro de una especie. Para el resto de los ortogrupos, se sumaron los valores de expresión de los transcritos y se combinó la información de la anotación. Como resultado, un total de 12.723 ortogrupos fueron analizados luego de filtrar por TPE.

El análisis de componentes principales (PCA, del inglés: *Principal Component Analysis*) muestra que las réplicas de cada estadio se agrupan juntas (Figura 25). En él, la Componente Principal 1 separa a las muestras por estado de desarrollo y explica un 35,6 % de la variabilidad, mientras que la Componente Principal 2 las separa por especie y explica un 17,1% de la variabilidad, con las réplicas de un mismo fragmento y especie agrupadas.

El número de OED entre las especies C_3 y la especie C_4 tiende a incrementarse a medida que las hojas maduran, indicando un patrón de desarrollo divergente, mientras que cuando se comparan las dos especies C_3 , el número de OED se mantiene relativamente constante a lo largo del desarrollo foliar (Figura 26). Al analizar el número de OED entre la especie C_4 y la especie PK se observa que la diferencia en el número de OED es baja en el F1 (1937 OED) y se incrementa sustancialmente en fragmentos más maduros (79,4%). Por otro lado, al comparar el número de OED entre la especie C_4 y la especie C_3 se observa una diferencia sustancial en genes expresados en el F1 (3022 OED) y se incrementa moderadamente en fragmentos más maduros (24,2%).

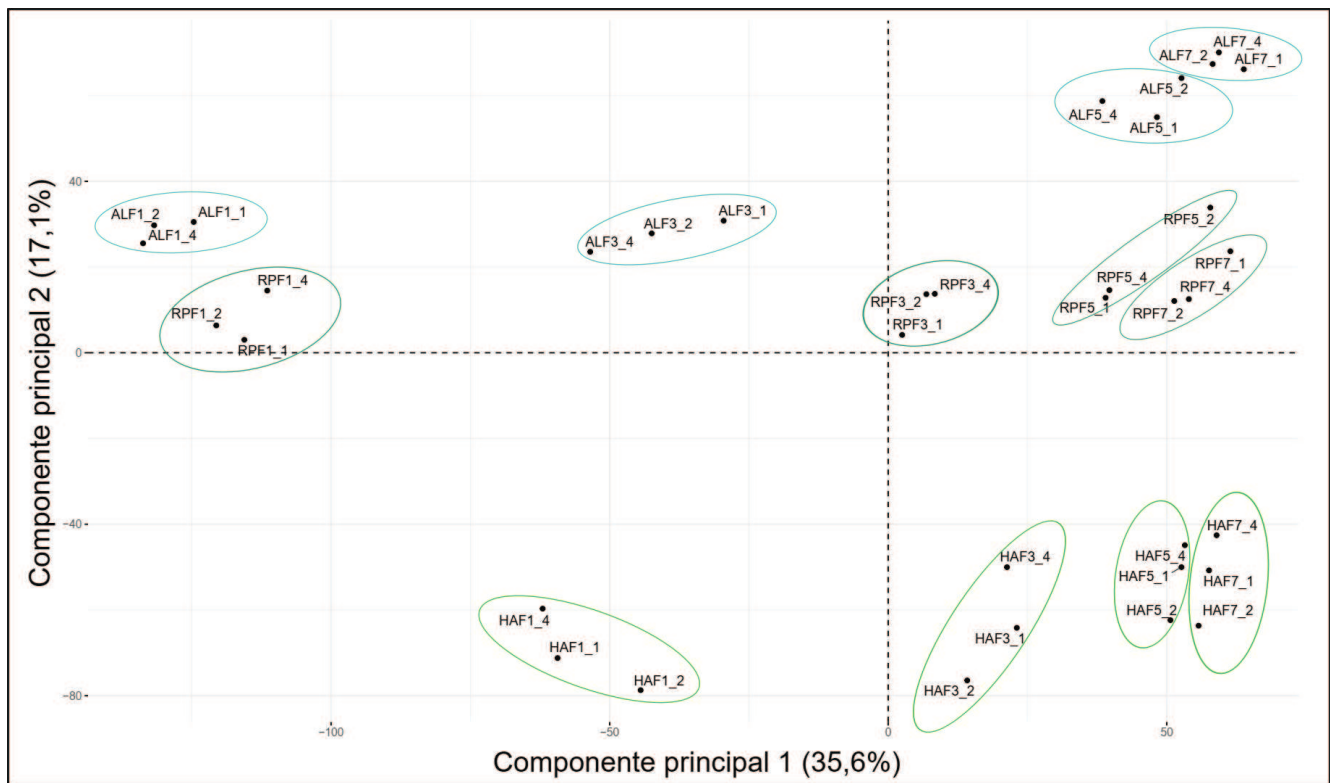


Figura 25. Análisis de componentes principales para los valores de expresión de los transcriptomas en *A. lanata*, *R. pilosa* y *H. amplexicaulis*. Las elipses encierran a las réplicas de un mismo fragmento y especie. Abreviaturas: *H. amplexicaulis* (Ha), *R. pilosa* (Rp), *A. lanata* (Al).

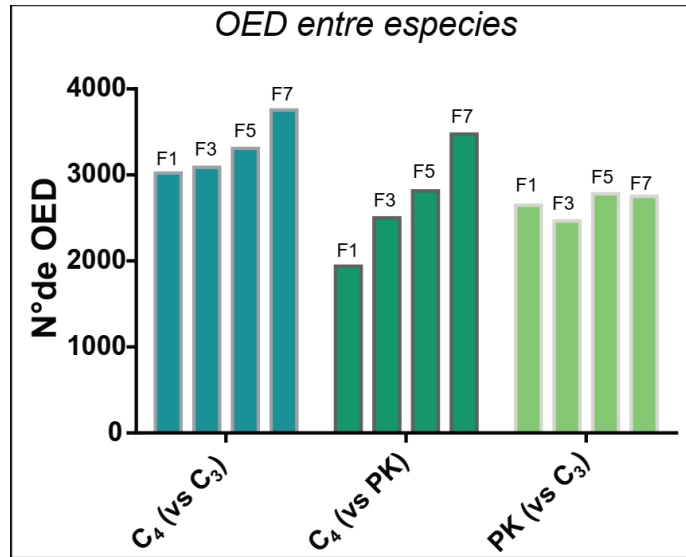


Figura 26 Análisis de expresión diferencial entre especies. Número de OED entre el mismo estado de desarrollo en *A. lanata* vs *H. amplexicaulis* (C₄ vs C₄), *A. lanata* vs *R. pilosa* (C₄ vs PK) y *R. pilosa* vs *H. amplexicaulis* (PK vs C₃).

En general, el número de OED entre especies fueron 6077, 5245 y 4902 para los pares *A. lanata*-*H. amplexicaulis*, *A. lanata*-*R. pilosa*, y *R. pilosa*-*H. amplexicaulis*, respectivamente. Estas diferencias coinciden con las relaciones filogenéticas entre las especies y los subtipos fotosintéticos.

Análisis de GO para los OED entre las especies

Para entender qué genes y procesos se encuentran expresados diferencialmente entre las especies, se realizó un análisis de GO. Los términos con los valores *p* más bajos se encuentran listados en la hoja de datos suplementaria 2 (Anexo 1). Ciertos términos relacionados a la fotosíntesis (GO:0009765, GO:0009521 y GO:0042651) se encuentran enriquecidos en OED en *A. lanata* vs las especies C₃, particularmente subexpresados en el F1 y sobreexpresados en el F7. En línea con esto, los mismos términos aparecen enriquecidos en transcritos sobreexpresados en *R. pilosa* vs *H. amplexicaulis* en los F5 y F7. Esto podría sugerir que la maquinaria fotosintética tiene un desarrollo más tardío en las hojas de la especie C₄ y se mantiene activa durante más tiempo, en comparación con las especies C₃. Al mismo tiempo, en la especie PK, los términos asociados a la fotosíntesis están sobreexpresados en los fragmentos maduros en comparación con la especie C₃.

Una tendencia similar se observa al analizar términos asociados a procesos metabólicos del ADN como replicación, reparación del ADN, ensamblado de la cromatina, entre otros (GO:0006260, GO:0006281, GO:0071103, GO:0006323, GO:0006333, GO:0032200, GO:0000793). Estos procesos

aparecen enriquecidos en los OG sobreexpresados en los F1 y F3 de las hojas de la especie C₄ y PK en comparación a fragmentos equivalentes en la especie C₃. Posteriormente, en el F3, estos términos aparecen enriquecidos en los OG sobreexpresados en la especie C₄ en comparación con las especies PK y C₃.

Otros términos que aparecen consistentemente, en especial en los OED de *H. amplexicaulis*, son los relacionados a la fosforilación de proteínas (GO:0006468) y la respuesta a estrés oxidativo (GO:000697) sobreexpresados y subexpresados, respectivamente.

Ortogrupos ortogrupos involucrados en procesos fotosintéticos

Para estudiar los procesos relacionados a la fotosíntesis, se construyeron listas de ortogrupos de acuerdo a la literatura y a términos de GO previamente identificados (Anexo 1, Hoja de datos suplementaria 3). A partir de esta lista se estimó, para cada estado de desarrollo, el porcentaje de OED entre especies (Figura 27). Como se esperaba, las diferencias entre la especie C₄ y las especies C₃ en los ortogrupos de fotorrespiración y el CCBB se observan desde el comienzo del gradiente de desarrollo (Figura 27 a y b). Con respecto a *A. lanata*, el 21% y 29% de los ortogrupos relacionados a la fotorrespiración están sobreexpresados en los F1 de *R. pilosa* y *H. amplexicaulis* respectivamente. Estos valores son mayores para el CCBB, con un 33% y 48% de los ortogrupos sobreexpresados en las especies C₃.

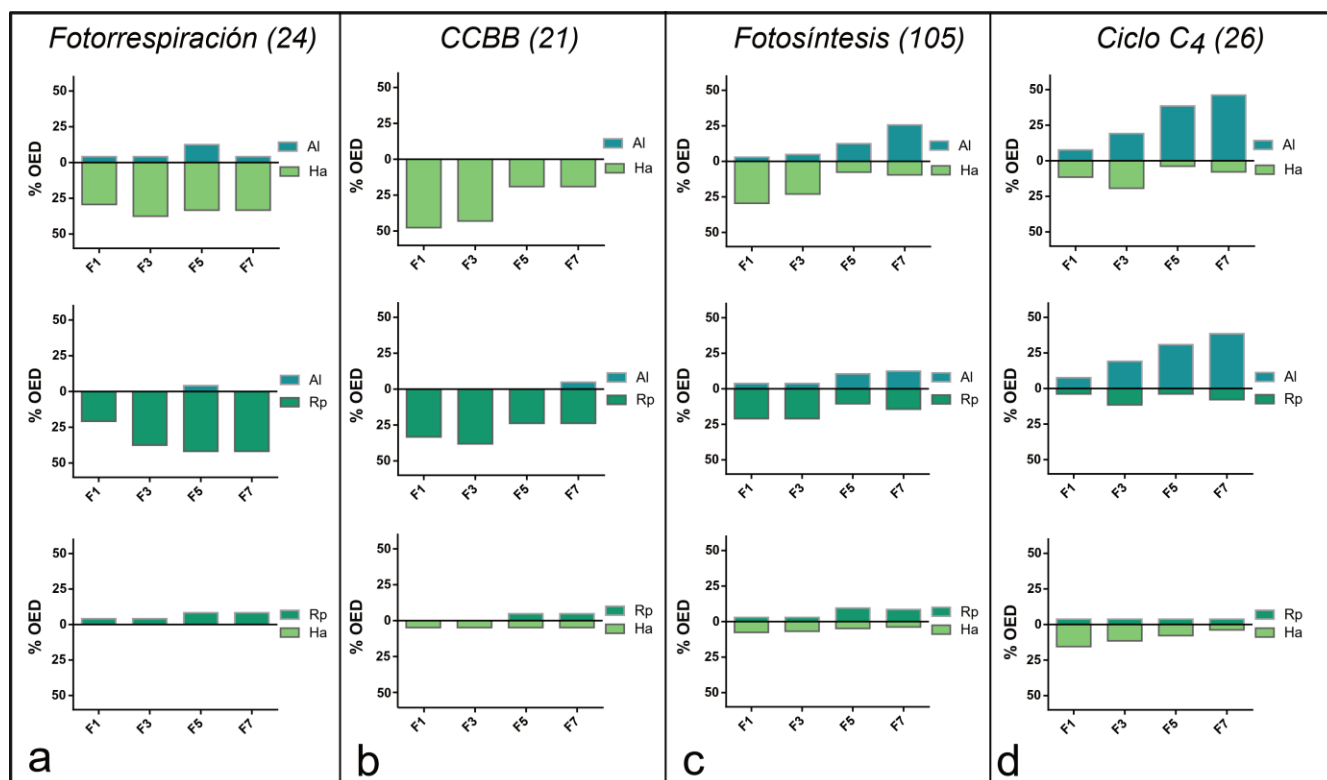


Figura 27. Porcentaje de ortogrupos expresados diferencialmente (% OED) en procesos relacionados con la fotosíntesis entre cada par de especies y en diferentes estados de desarrollo. Ortogrupos involucrados en los procesos de fotorrespiración (a), ciclo de CBB (b), fotosíntesis (c) y ciclo C₄ (d). Los números entre paréntesis indican la cantidad de ortogrupos presentes en cada categoría. Abreviaturas: *H. amplexicaulis* (Ha), *R. pilosa* (Rp), *A. lanata* (Al).

Los miembros del grupo de fotosíntesis muestran un patrón similar en el F1, con más del 25% de los ortogrupos sobreexpresados en las especies C₃. Sin embargo, esta tendencia se reduce o incluso se invierte a medida que la hoja madura (Figura 27 c). En los ortogrupos pertenecientes al ciclo C₄, sorprendentemente existen OG sobreexpresados en las especies C₃ y PK en comparación con la especie C₄. Este patrón es más notorio en los F1 y F3. En el F1, un 11,5% de los OG están sobreexpresados en la especie C₃, y en el F3 un 19,2%. Al mismo tiempo, existe una mayor cantidad de OG sobreexpresados en *H. amplexicaulis* en comparación con *R. pilosa*.

Ortogramos involucrados en la formación de la vasculatura y expansión de la pared celular

Teniendo en cuenta que las diferencias en el desarrollo entre las especies podrían ser consecuencia de diferencias en los tiempos de los procesos de desarrollo, se buscaron otros marcadores de maduración. Es por ello que se decidió analizar el comportamiento de la expresión de OG relacionados con dos procesos involucrados en la determinación de la estructura de la anatomía foliar: la formación de HV y la expansión de la pared celular (Figura 28).

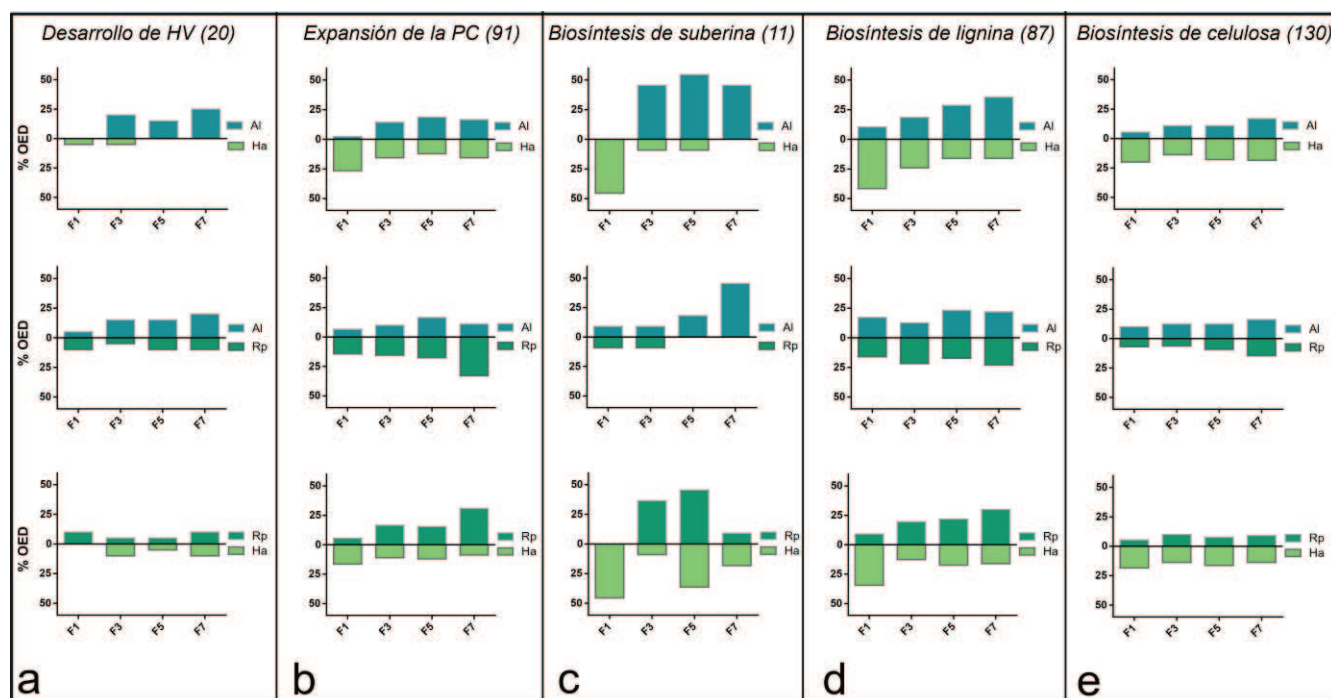


Figura 28. Porcentaje de ortogramos expresados diferencialmente (% OED) en procesos relacionados al desarrollo foliar entre cada par de especies y en diferentes estados de desarrollo. Ortogramos involucrados en los procesos de: desarrollo de HV (a), expansión de la pared celular (b), biosíntesis de suberina (c), biosíntesis de lignina (d), y biosíntesis de celulosa (e). Los números entre paréntesis indican la cantidad de ortogramos presentes en cada categoría. Abreviaturas: *H. amplexicaulis* (Ha), *R. pilosa* (Rp), *A. lanata* (Al).

Al estudiar los ortogramos involucrados en la regulación de la formación de HV, las diferencias en el número de OG sobreexpresados en *A. lanata* presentan una tendencia ascendente hacia fragmentos más maduros (Figura 28 a). Esto sucede particularmente desde el F3, en correlación con el aumento en la densidad de HV en la especie C₄. Un patrón similar se observa con los ortogramos

involucrados en la expansión de la pared celular, donde la mayoría de los ortogrupos sobreexpresados pertenecen a *H. amplexicaulis* en el F1, y a *A. lanata* y *R. pilosa* en el F7 (Figura 28 b).

Al estudiar más detalladamente los procesos involucrados en el desarrollo de la pared celular, se observa que algunos de los ortogrupos involucrados en la biosíntesis de celulosa se encuentran sobreexpresados, pero sin una tendencia clara (Figura 28 e). Por el contrario, los OG relacionados con dos componentes de las paredes celulares de los HV y de las CVI, como la suberina y la lignina, muestran una tendencia a estar sobreexpresados en las especies C₄ y PK vs la especie C₃ (Figura 28 c y d).

Patrones de expresión de genes

Marcadores de zonas fuente/sumidero

Para confirmar el gradiente de maduración a lo largo de las hojas muestreadas, se exploraron los patrones de expresión de genes previamente descritos como marcadores de zonas fuente y sumidero ([86], [178], [199], [200]). Los niveles de expresión normalizados de ortogrupos marcadores de zonas sumidero, como las moduladoras del ciclo celular ciclinas D4 y D6 (CycD4 y CycD6) y la enzima sacarosa sintasa (SUS) son máximos en el F1, con un patrón descendente hacia el F7 (Figura 29). Por el contrario, los niveles de expresión de los marcadores de zonas fuente, como los transportadores de sacarosa SWEET y STP1 y la enzima nitrato reductasa (NR1) presentan valores mínimos en el F1, con un patrón ascendente hacia el F7.

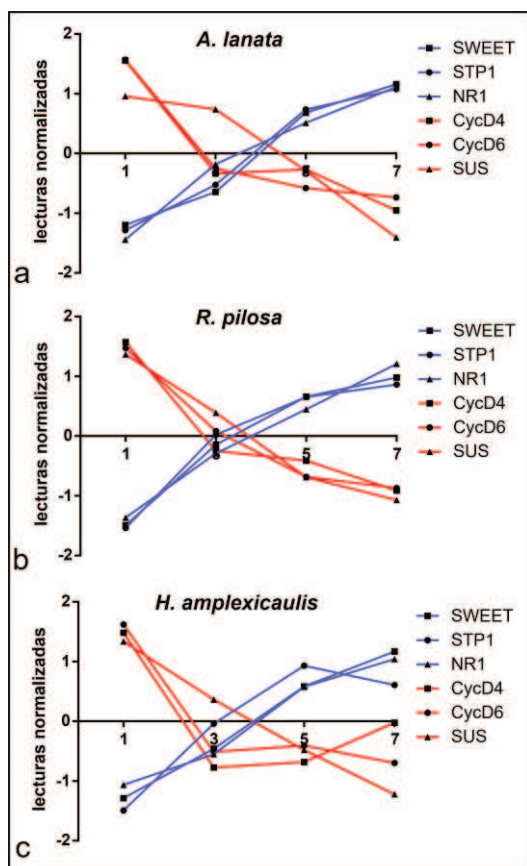


Figura 29. Nivel de expresión normalizados de marcadores foliares de zonas fuente/sumidero para *A. lanata* (a), *R. pilosa* (b) y *H. amplexicaulis* (c). Abreviaturas: transportador de sacarosa SWEET (SWEET), transportador de sacarosa STP1 (STP1), nitrato reductasa (NR1), ciclina D4 (CycD4), ciclina D6 (CycD6), sacarosa sintasa (SUS).

Expresión de enzimas involucradas en el ciclo C₄

Al comparar los patrones de expresión de las enzimas del ciclo C₄ durante el desarrollo foliar, se observaron similitudes entre las tres especies (Figura 30, Anexo 1 Figura S5). En *A. lanata*, la mayoría de las enzimas claves del ciclo C₄ presenta un patrón ascendente a lo largo del desarrollo, en sintonía con el desarrollo de la fotosíntesis. Interesantemente, 5 de las enzimas (PEPC2-3, PPDK, CA, PEP-CK y ALA-AT) poseen un patrón de expresión conservado entre las tres especies (cuatro con un patrón ascendente y una con un mínimo en el F3). Además, dos de ellas, NADP-ME y aspartato aminotransferasa (ASP-AT5), presentan un patrón conservado entre las especies C₄ y PK (Figura 30 a).

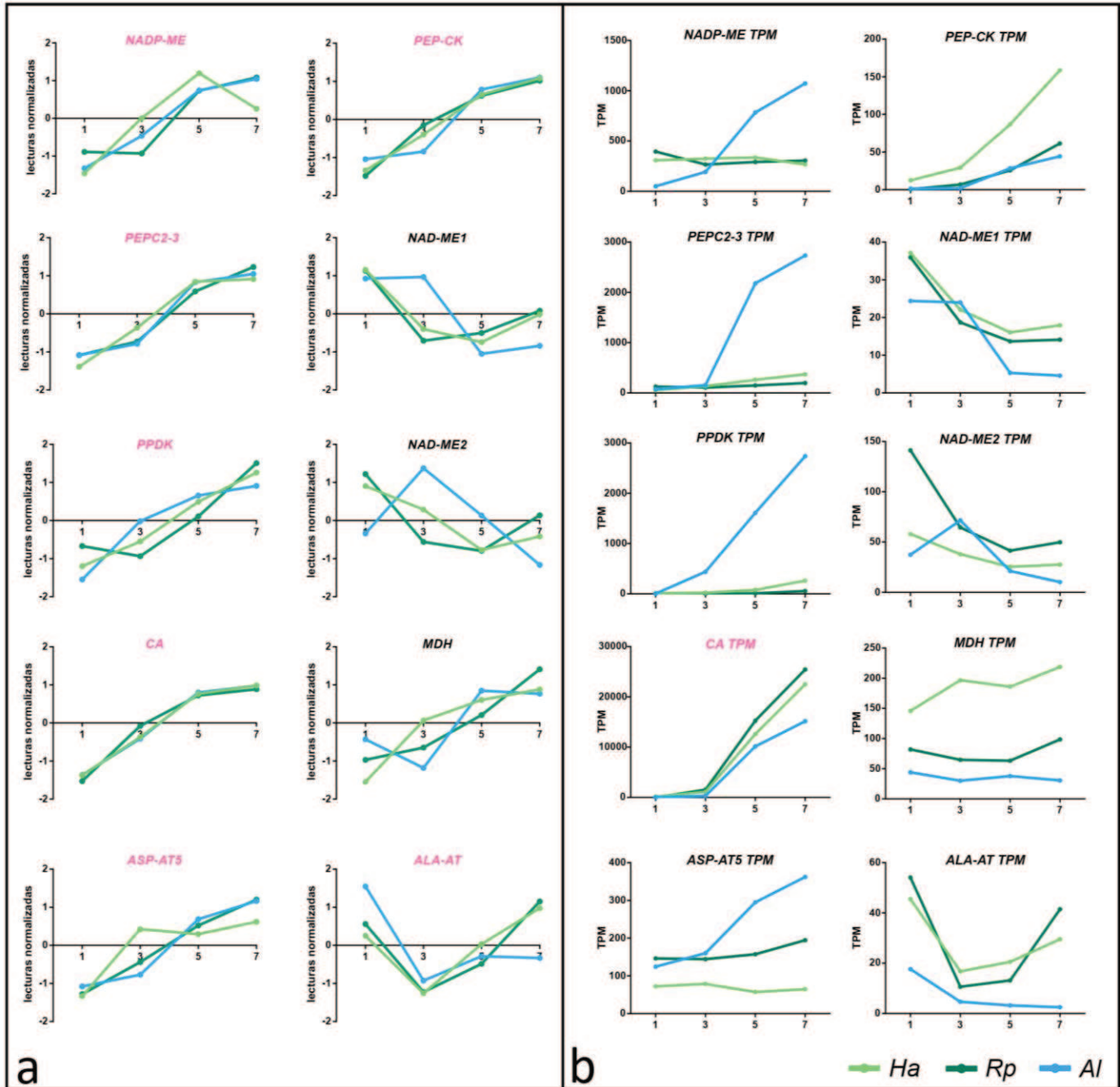


Figura 30. Patrones de expresión de ortogrupos que codifican para enzimas del ciclo C_4 durante el desarrollo. Promedios de valores de expresión de los ortogrupos en los fragmentos 1, 3, 5 y 7 para *H. amplexicaulis* (verde claro), *R. pilosa* (verde oscuro) y *A. lanata* (celestes). Lecturas normalizadas dentro de cada especie (a). Valores de expresión expresados en TPM (b). Los nombres en rosado indican los genes con patrones de expresión conservados entre la especie C_4 y al menos una de las especies C_3 . Abreviaturas: NADP enzima málica (NADP-ME), PEP carboxilasa (PEPC2-3), piruvato fosfato dikinasa (PPDK), anhídrida carbónica (CA), aspartato aminotransferasa (ASP-AT5), fosfoenolpiruvato carboxiquinasa (PEP-CK), NAD enzima málica (NAD-ME1 y

NAD-ME 2), malato deshidrogenasa (*MDH*), alanina aminotransferasa (*ALA-AT*), *H. amplexicaulis* (*Ha*), *R. pilosa* (*Rp*), *A. lanata* (*Al*).

Al tener en cuenta los niveles de expresión, se destaca que el ortogrupo de la anhidrasa carbónica (*CA*) presenta valores de TPM similares en las tres especies (Figura 30 b). Así mismo, el nivel de expresión del ortogrupo asociado a la PEP-CK es similar entre las especies *C₄* y *PK* (Figura 30 b). En los fragmentos maduros, las enzimas descarboxilantes asociadas a los subtipos *NAD-ME* y *PEP-CK* están menos expresadas en *A. lanata* que en las especies *C₃*. La especie *C₄* también posee niveles más bajos de expresión en las enzimas malato deshidrogenasa (*MDH*), alanina aminotransferasa (*ALA-AT*) y *CA*; mientras que los valores de TPM de las enzimas *PEC2-3*, *NADP-ME*, *PPDK* Y *ASP-AT5* son sustancialmente más altos en fragmentos maduros.

Expresión de genes involucrados en la fotorrespiración

Para estudiar los patrones de expresión de los ortogrupos involucrados en proceso de fotorrespiración, se analizaron los ortogrupos que codifican para las subunidades del complejo glicina descarboxilasa (*GDC*) (Figura 31). En total se detectaron seis ortogrupos: uno para la subunidad *L* (*GDC L*), uno para la subunidad *P* (*GDC P*), dos para las subunidades *H* (*GDC H* y *Hb*) y *T* (*GDC T* y *Tb*).

Todos los ortogrupos presentaron un patrón de expresión muy conservado entre las 3 especies (Figura 31 a). Por el contrario, los niveles de expresión se encuentran moderadamente conservados en un solo ortogrupo (*GDC Tb*). En el resto de los ortogrupos, se observa una tendencia de la especie *C₄* en tener niveles más bajos de expresión (Figura 31 b).

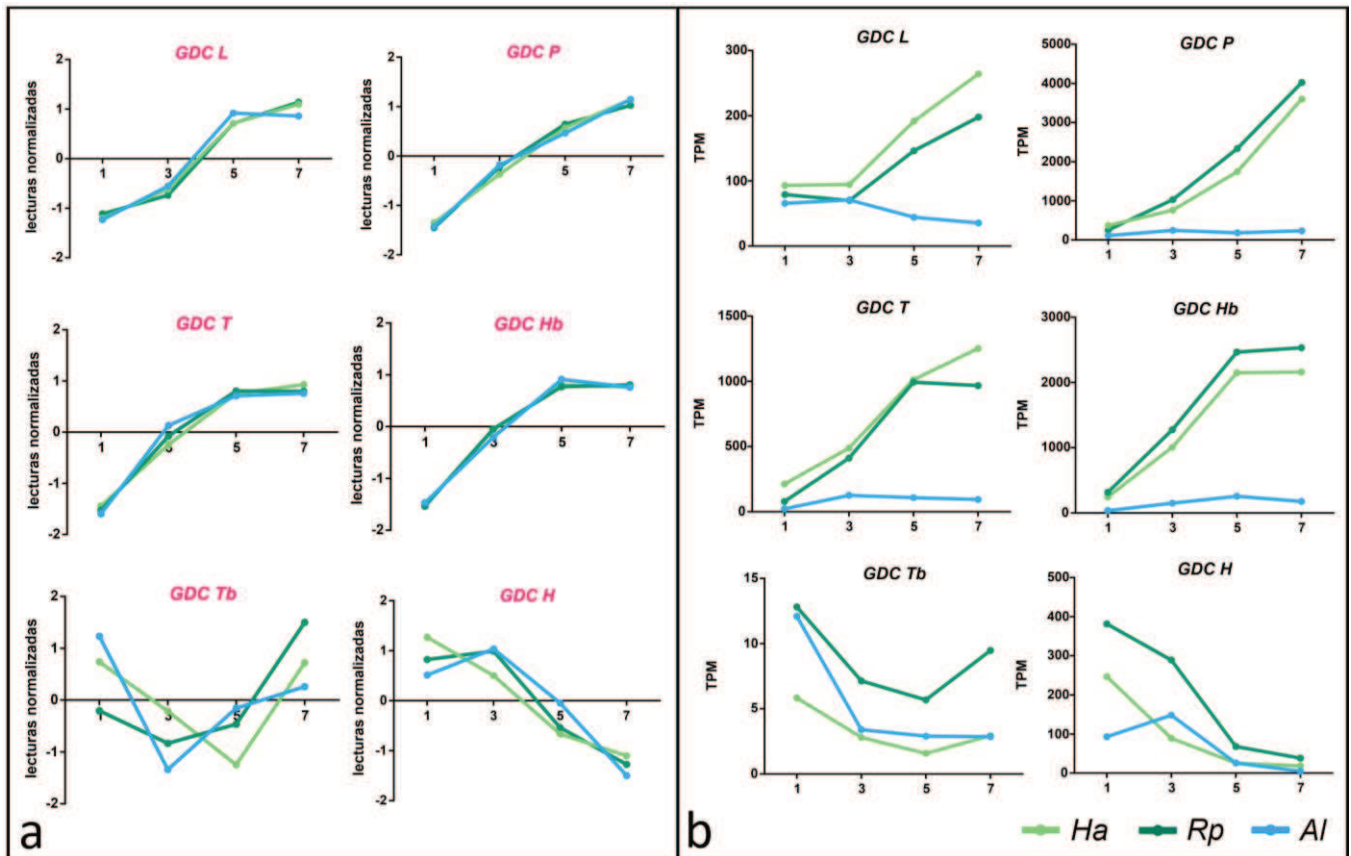


Figura 31. Patrones de expresión durante el desarrollo de los ortogrupos que codifican para las subunidades L, P, H, y T de la GDC. Promedios de valores de expresión de los ortogrupos en los fragmentos 1, 3, 5 y 7 para *H. amplexicaulis* (verde claro), *R. pilosa* (verde oscuro) y *A. lanata* (celeste). Lecturas normalizadas dentro de cada especie (a). Valores de expresión expresados en TPM (b). Los nombres en rosado indican los genes con patrones de expresión conservados entre la especie C_4 y al menos una de las especies C_3 . Abreviaturas: *H. amplexicaulis* (Ha), *R. pilosa* (Rp), *A. lanata* (Al).

Expresión de reguladores transcripcionales SHR y SCR

El ensamblado de los transcriptomas identificó un solo transcripto de SCR en las especies C_3 , y cuatro copias en la especie C_4 . Sin embargo, las cuatro copias resultaron pertenecer al mismo ortogrupo al cual pertenecen los SCR1 de *O. sativa* y *Z. mays*. El parólogo SCR2 fue sólo detectado en *H. amplexicaulis* en niveles muy bajos en dos muestras (< 1 TPM). El tercer miembro del linaje, SCL23, fue encontrado como copia única en las tres especies.

El ortogrupo SHR pertenece al linaje SHR1 e incluye una copia en la especie C_4 y dos en las especies C_3 . El ortogrupo SCL32 es de copia única en las tres especies y pertenece al linaje SCL32b.

El linaje NSP1 presentó cuatro copias en *A. lanata*, tres en *H. amplexicaulis* y dos en *R. pilosa*, todas pertenecientes a un mismo ortogrupo. Mientras que el linaje SCL32a resultó de copia única. Ambos, NSP1 y SCL32a no pasaron el filtro de expresión.

Los ortogrupos SCR y SHR presentan patrones de expresión conservados a lo largo del gradiente foliar para las tres especies. Mientras que SCL32 se muestra conservado principalmente entre *H. amplexicaulis* y *A. lanata* (Figura 32 a). Además, solo *R. pilosa* muestra un patrón de expresión conservado entre SCR y su parálogo SCL23 (Figura 32 a).

Siendo SHR un regulador positivo de la expresión de SCR, es llamativo que la especie *C₄* presente un nivel de expresión de SHR igual o menor al de las otras especies, y mayor para SCR, especialmente al principio del desarrollo (Figura 32 b).

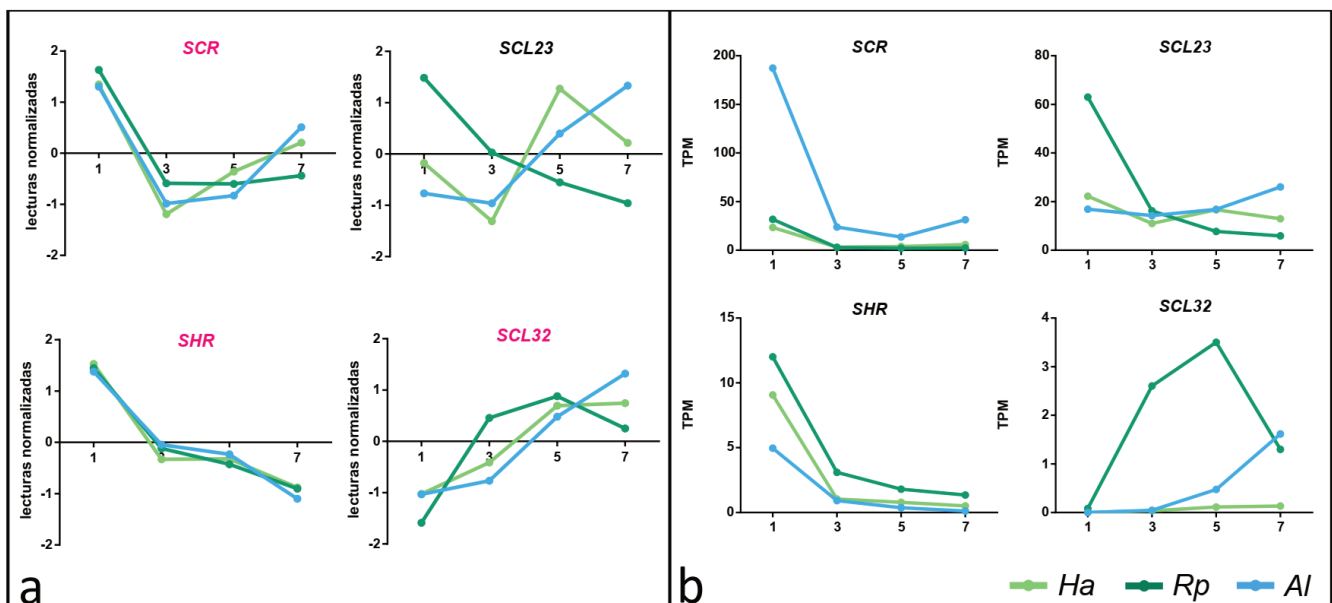


Figura 32. Patrones de expresión durante el desarrollo de ortogrupos que codifican para los reguladores transcripcionales de las subfamilias SCR y SHR. Promedios de valores de expresión de los ortogrupos en los fragmentos 1, 3, 5 y 7 para *H. amplexicaulis* (verde claro), *R. pilosa* (verde oscuro) y *A. lanata* (celeste). Lecturas normalizadas dentro de cada especie (a). Valores de expresión expresados en TPM (b). Los nombres en rosado indican los genes con patrones de expresión conservados entre la especie *C₄* y al menos una de las especies *C₃*. Abreviaturas: *H. amplexicaulis* (Ha), *R. pilosa* (Rp), *A. lanata* (Al).

Expresión de reguladores transcripcionales IDD

Se detectaron miembros de los IDD en nueve ortogrupos pertenecientes a cuatro linajes (2 ortogrupos de SG5, 1 de GAF1, 4 de NKD y 2 de WRB) y un ortogrupo de linaje indeterminado (IDD-i) al estar formado por 3 secuencias de *Z. mays* pertenecientes a 3 linajes diferentes (NKD, ID1 y MGP) (Figura 33).

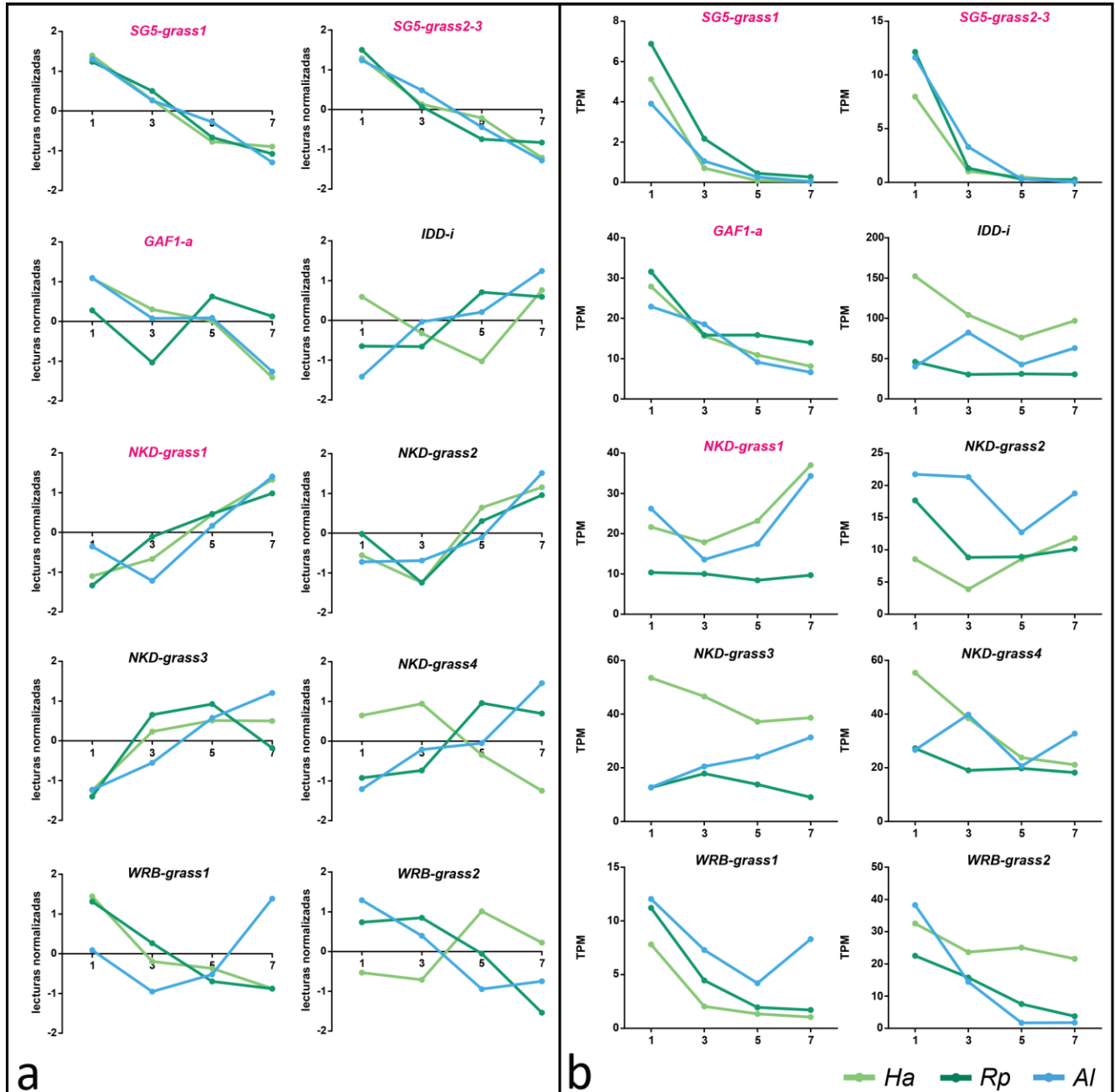


Figura 33. Patrones de expresión durante el desarrollo de ortogrupos que codifican para los reguladores transcripcionales de la subfamilia IDD. Promedios de valores de expresión de los ortogrupos en los fragmentos 1, 3, 5 y 7 para *H. amplexicaulis* (verde claro), *R. pilosa* (verde oscuro) y *A. lanata* (celeste). Lecturas normalizadas dentro de cada especie (a). Valores de expresión expresados en TPM (b). Los nombres en rosado indican los genes con patrones de expresión conservados entre la especie C_4 y al menos una de las especies C_3 . Abreviaturas: *H. amplexicaulis* (Ha), *R. pilosa* (Rp), *A. lanata* (Al).

Los ortogrupos pertenecientes al linaje SG5 pertenecen a los clados SG5-grass1 y SG5-grass2-3. El SG5-grass1 es de copia única en las 3 especies, mientras que el SG5-grass2-3 presenta dos copias en *R. pilosa*. Ambos ortogrupos, junto con GAF1-a, presentan conservados tanto los patrones como los niveles de expresión, en un gradiente descendente hacia la punta de la hoja. El ortogrupo GAF1-a también es de copia simple en las tres especies.

Los ortogrupos correspondiente al linaje NKD son todos de copia única en las tres especies, aunque sólo uno de ellos (NKD-grass1) presenta un patrón y niveles de expresión conservados (entre *A. lanata* y *H. amplexicaulis*).

En el linaje WRB, el ortogrupo WRB-grass1 es de copia única entre las tres especies, mientras que WRB-grass2 contiene dos copias en *A. lanata* y *H. amplexicaulis*.

Finalmente, el ortogrupo denominado IDD-i (IDD indeterminado) está constituido por múltiples copias (2 en *A. lanata*, 3 en *R. pilosa* y 6 en *H. amplexicaulis*) y no se presenta conservado entre las especies.

Redes de co-expresión (WGCNA)

Identificación de módulos de ortogrupos co-expresantes

El conjunto de ortogrupos utilizado en el análisis de expresión diferencial entre especies, se usó para la construcción de una red de co-expresión ponderada. Un total de 12701 OG fueron distribuidos en 17 módulos (M1 a M17) mientras que 22 OG fueron agrupados en un Módulo 0 al no pertenecer a ninguno de los 17 (Figura 34). Si bien todos los módulos presentaron buenos parámetros de correlación entre la membresía al módulo (MM) y la conectividad intramodular (k_{im}), los módulos M1, M2, M3, M5 y M6 presentaron valores promedios de K_{im}/K_{total} mayores a 0,5. Esto significa que estos cinco módulos poseen una mayor calidad o una mayor probabilidad de ser reales que el resto de los módulos.

Para cada módulo se calculó el valor eigengen del módulo (EM), que puede entenderse como la componente principal que explica la variación a lo largo de diferentes muestras [182]. Es decir, un

ortogrupo artificial que simboliza el comportamiento del módulo. Considerando que la detección de los módulos implicó una red que no tiene en cuenta las interacciones negativas entre ortogrupos, se realizó un análisis de correlación de los EM para identificar módulos con comportamientos opuestos (Anexo 1, Figura S6). El análisis arrojó seis pares de módulos opuestos con correlaciones mayores a 0,8 (valor $p < 0,01$): M1-M6, M2-M5, M3-13, M4-M10, M7-M11 y M8-M15. Estas correlaciones negativas indican que miembros de un módulo podrían ser responsables de regular negativamente a miembros del módulo opuestos. Es decir, cada uno de estos pares de módulos podrían considerarse supra módulos que intervienen en los mismos procesos celulares.

La mayoría de los módulos presentaron patrones de expresión asociados al desarrollo en las tres especies (ascendentes, descendentes o con un máximo en un fragmento particular). Sin embargo, los pares de módulos M3-M13 y M7-M11 agruparon a OG expresados diferencialmente en *H. amplexicaulis* y *R. pilosa*, respectivamente (Figura 34).

Para profundizar en las relaciones entre los procesos involucrados en la vía C_4 , la fotorrespiración y los reguladores transcripcionales IDD, SHR y SCR, se identificó la pertenencia de estos OG a cada módulo (Figura 34). Notablemente, la mayoría de ellos formaron parte de los pares M1-M6 y M2-M5. El resto pertenece a los pares M3-M13, M4-M10 y a los módulos M7 y M15. El único OG que no pudo ser clasificado en ningún módulo fue el correspondiente a la enzima descarboxilante NADP-ME, indicando que el patrón de expresión del mismo es demasiado único como para formar parte de un módulo co-expresante entre las tres especies.

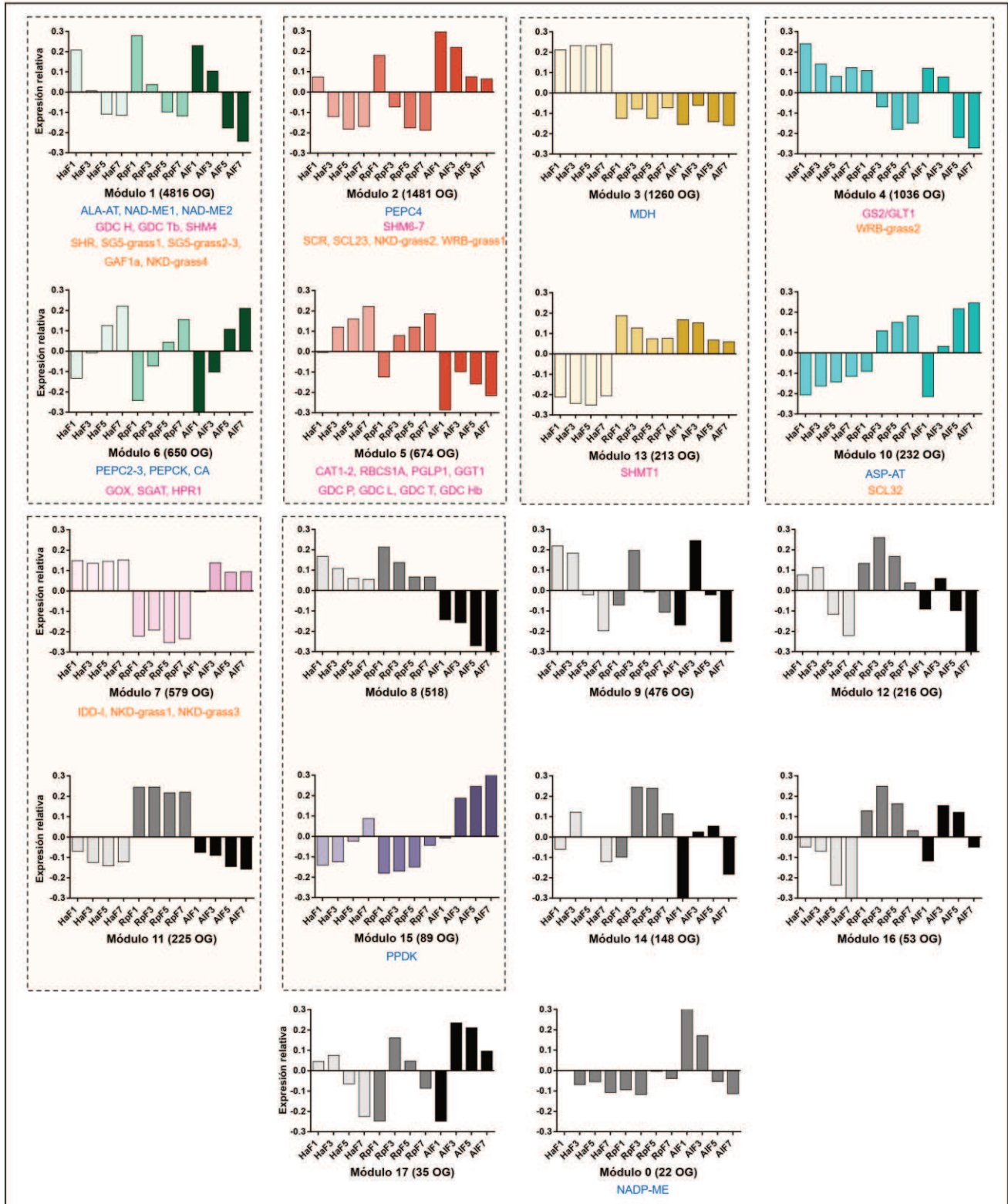


Figura 34. Módulos de co-expresión comunes a las tres especies encontrados mediante WGCNA. Para cada módulo se presentan los promedios de los valores de expresión del módulo (EM) en cada fragmento y especie. Entre paréntesis se indica el número de ortogrupos en cada módulo. El módulo 0, es en realidad un módulo

artificial donde se agrupan los ortogrupos que no pertenecen a ningún módulo. Los módulos opuestos, determinados por valores altos de correlación negativos ($cor > 0,8$; $valor\ p < 0,01$) se muestran juntos en rectángulos con línea punteada. Los ortogrupos correspondientes a genes involucrados en el ciclo C_4 , la fotorrespiración y los reguladores transcripcionales *IDD*, *SHR* y *SCR*, se listan en su correspondiente módulo en azul, rosado y naranja, respectivamente. Abreviaturas: NADP enzima málica (NADP-ME), PEP carboxilasa (PEPC 2-3, 4), Piruvato fosfato dikinasa (PPDK), anhidrasa carbónica (CA), aspartato aminotransferasa (ASP-AT), fosfoenolpiruvato carboxiquinasa (PEP-CK), NAD enzima málica (NAD-ME1 y 2), malato deshidrogenasa (MDH), alanina aminotransferasa (ALA-AT), subunidad pequeña de RUBISCO (RBCS1A), fosfoglicolato fosfatasa (PGLP1), glicolato oxidasa (GOX), catalasa (CAT1-2), glutamato-glioxilato amino transferasa (GGT1), serina-glioxilato aminotransferasa (SGAT), subunidades del complejo glicina descarboxilasa (GDC H, Hb, L, P, T, Tb), serina hidroximetiltransferasa (SHMT1, SHM4, SHM6-7), hidroxipiruvato reductasa (HPR1), glutamato sintetasa (GS2/GLT1), *H. amplexicaulis* (Ha), *R. pilosa* (Rp), *A. lanata* (Al).

Co-expresión de los reguladores transcripcionales *IDD*, *SHR* y *SCR*

Los pares de módulos M1-M6 y M2-M5 fueron estudiados en mayor profundidad por ser los que reúnen a la mayor cantidad de genes involucrados en los procesos del ciclo C_4 , la fotorrespiración y los reguladores transcripcionales *IDD*, *SHR*, y *SCR*. Ambos pares de módulos presentan patrones de expresión similares, siendo el par M1-M6 el que demuestra estar más conservado entre las especies.

La caracterización de los módulos incluyó un análisis de GO para estimar las posibles funciones biológicas de los mismos, la identificación de los factores de transcripción de cada uno y el lugar que ocupan los ortogrupos de interés previamente mencionados en la red (centro o periferia). En el M1 se encontraron 247 OG anotados como miembros de familias de factores de transcripción. Las familias MYB, bHLH y C2H2 fueron las que presentaron el mayor número de FT con 31, 23 y 22, respectivamente (Figura 35 a). En el M6 se encontraron 47 FT y las familias con mayor número fueron bHLH, CO-like y HD-ZIP con 6, 4 y 4 miembros, respectivamente (Figura 35 b). Entre ellos, 15 y 3 FT se encontraron dentro de los OG *hubs* del M1 y del M6, respectivamente. Uno de los FT *hubs* resultó ser un OG miembro de los *IDD*, SG5-grass2-3.

El análisis de GO reveló que el M1 está fuertemente asociado a procesos activos durante el desarrollo como el transporte intracelular de proteínas, reparación del ADN y procesamiento del ARNm (Figura 35 c). Por otro lado, el M6 está asociado a la fotosíntesis y el metabolismo de carbohidratos (Figura 35 d). Los *hubs* en M1 y M6 son los OG que presentaron un valor de MM superior a 0,93 y 0,92 respectivamente (Figura 35 e y f).

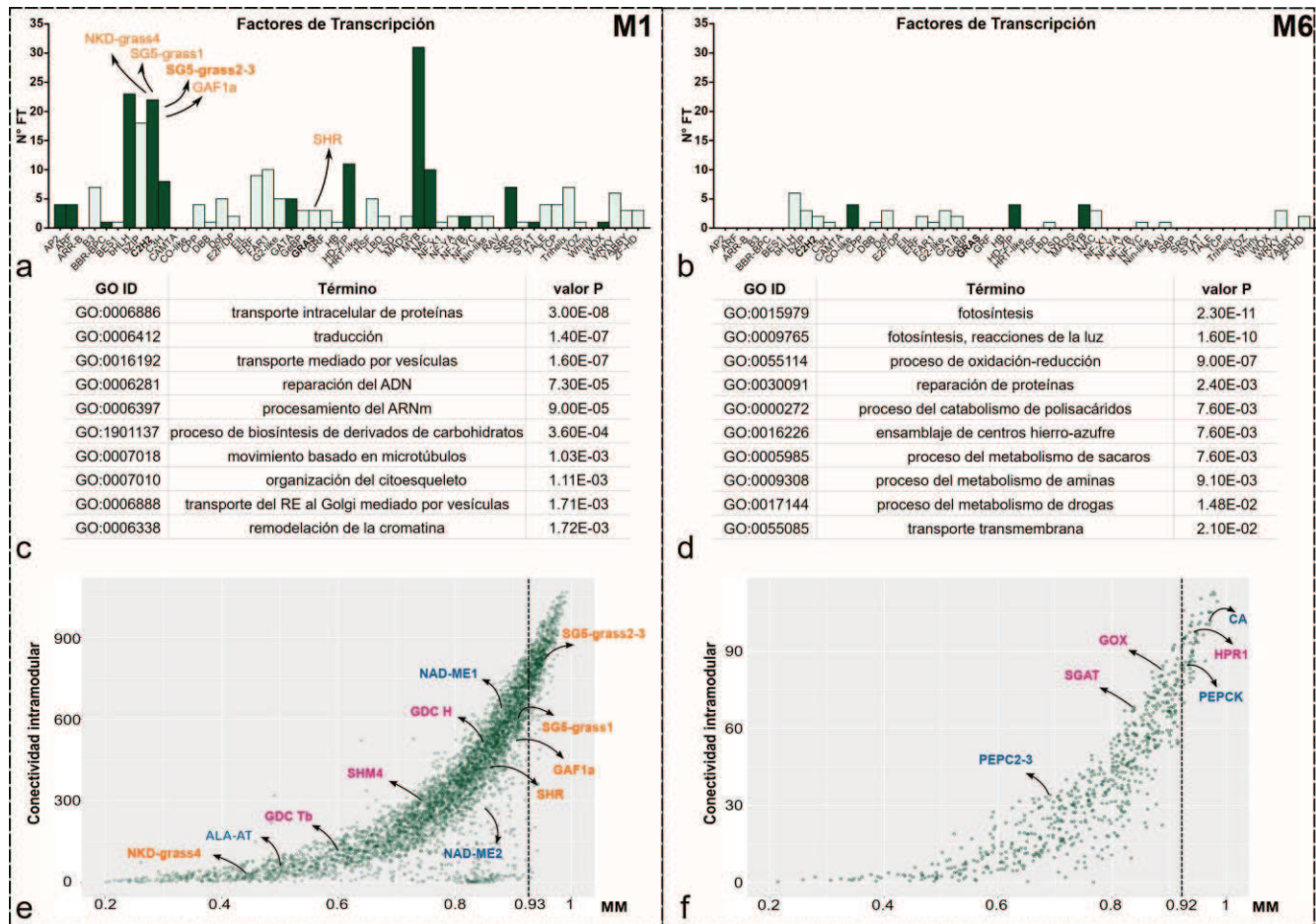


Figura 35. Características distintivas de los módulos opuestos M1 (a, c y e) y M6 (b, d y f). Número de factores de transcripción pertenecientes al módulo distribuidos por familias (a-b). Las barras más oscuras destacan a las familias que contienen al menos un ortogrupo hub. Análisis de enriquecimiento de términos de GO en procesos biológicos: primeros diez términos de GO enriquecidos en el módulo (c-d). Gráficos de dispersión de los parámetros de Membresía de Módulo (MM) y Conectividad intramodular para cada ortogrupo (e-f). La barra vertical punteada indica el valor de MM a partir del cual un ortogrupo es considerado hub. Los ortogrupos correspondientes al ciclo C_4 , la fotorrespiración y los reguladores transcripcionales IDD y SHR se señalan en azul, rosado y naranja, respectivamente. Abreviaturas: PEP carboxilasa (PEPC 2-3), anhidrasa carbónica (CA), fosfoenolpiruvato carboxiquinasa (PEP-CK), NAD enzima málica (NAD-ME1 y 2), alanina aminotransferasa (ALA-AT), glicolato oxidasa (GOX), serina-glioxilato aminotransferasa (SGAT), subunidades del complejo glicina descarboxilasa (GDC H, Tb), serina hidroximetiltransferasa (SHM4), hidroxipiruvato reductasa (HPR1).

En el M2 se encontraron 92 OG anotados como miembros de familias de factores de transcripción. Las familias ERF, bZIP y GRAS fueron las que presentaron el mayor número de FT con 9, 6 y 6 miembros, respectivamente (Figura 36 a). En el M5 se encontraron 32 FT y las familias con mayor número fueron bHLH y WRKY con 6 y 4 miembros, respectivamente (Figura 36 b). Entre ellos,

seis FT se encontraron dentro de los OG *hubs* del M2, mientras que ninguno de los perteneciente al M5, resultó ser *hub*.

El análisis de GO reveló que el M2 está asociado a procesos relacionados con la proteólisis, el transporte de derivados de carbohidratos, el empaquetamiento del ADN y la regulación del ciclo celular (Figura 36 c). Por otro lado, el M5 aparece asociado al metabolismo de la molécula de señalización guanosina tetrafosfato, traducción, procesos de óxido reducción y metabolismo de carbohidratos (Figura 36 d).

Los *hubs* en M2 y M5 son los OG que presentaron un valor de MM superior a 0,91 (Figura 35 e y f). Entre ellos se encuentran el regulador transcripcional SCR y tres subunidades del complejo GDC.

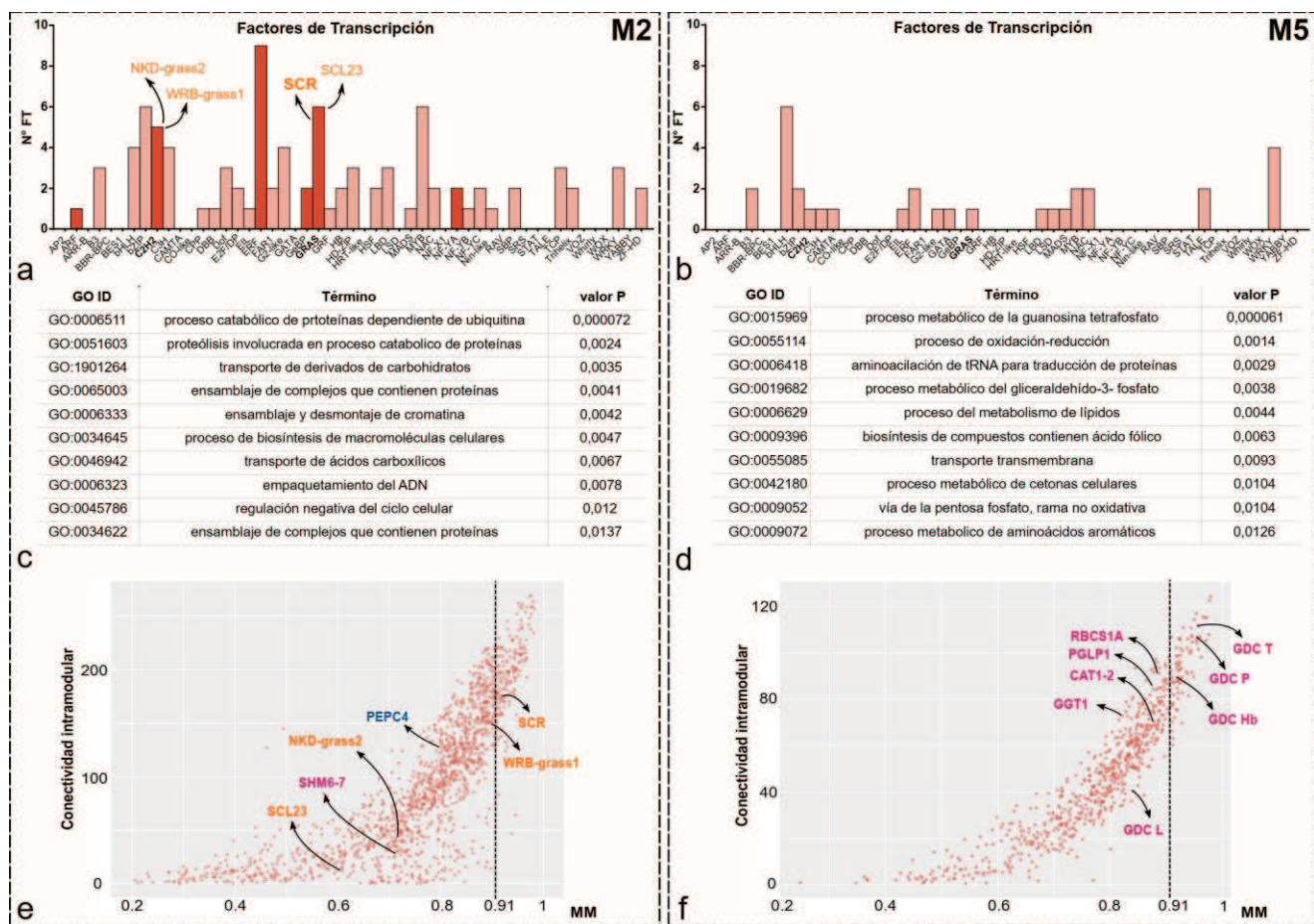


Figura 36. Características distintivas de los módulos opuestos M2 (a, c y e) y M5 (b, d y f). Número de factores de transcripción pertenecientes al módulo distribuidos por familias (a-b). Las barras más oscuras destacan a las familias que contienen al menos un ortogrupo *hub*. Análisis de enriquecimiento de términos de GO en procesos biológicos: primeros diez términos de GO enriquecidos en el módulo (c-d). Gráficos de dispersión de los parámetros de Membresía de Módulo (MM) y Conectividad intramodular para cada ortogrupo (e-f). La barra

vertical punteada indica el valor de MM a partir del cual un ortogrupo es considerado hub. Los ortogrupos correspondientes al ciclo C_4 , la fotorrespiración y los reguladores transcripcionales IDD y SHR se señalan en azul, rosado y naranja, respectivamente. Abreviaturas: PEP carboxilasa (PEPC 4), subunidad pequeña de RUBISCO (RBCS1A), fosfoglicolato fosfatasa (PGLP1), catalasa (CAT1-2), glutamato-glioxilato amino transferasa (GGT1), subunidades del complejo glicina descarboxilasa (GDC Hb, L, P, T), serina hidroximetiltransferasa (SHM6-7).

DISCUSIÓN

La calidad de los transcriptomas

El ensamblado de novo de los transcriptomas con Trinity generó transcriptomas iniciales con un gran número de genes, y alrededor del doble de transcriptos. Sin embargo, cuando se compara con los transcriptomas de otras especies de gramíneas con genomas secuenciados el número de genes es sustancialmente menor. El transcriptoma de *Z. mays* contiene 63480 genes y 88760 transcriptos (cuando se tienen en cuenta a las isoformas); el de *S. bicolor* contiene 34129 genes y 47121 transcriptos; y el de *S. viridis* contiene 38334 genes y 52459 transcriptos [129]. Estos números implican un aumento del número de transcriptos con respecto al de genes del 39,8%, 38,1% y 36,8% para *Z. mays*, *S. bicolor* y *S. viridis*, respectivamente. Por el contrario, el aumento del número de transcriptos con respecto al número de genes es de 55,9 %, 136,8% y 112,1% para *H. amplexicaulis*, *R. pilosa* y *A. lanata*, respectivamente. Estos resultados sugieren que una gran cantidad de las isoformas pueden ser artefactos del ensamblador, ya sea producto de errores de la secuenciación o transcriptos incompletos en genes poco expresados. Así mismo, también podrían indicar la presencia de pre mRNA en transcriptos altamente expresados. El filtrado de isoformas generó una reducción en el número de secuencias del orden del 50%, pero la tasa de mapeo apenas cayó del 95-97% al 87-91%. Esta hipótesis, como muchas otras relacionadas a la calidad de los transcriptomas, podrá verificarse cuando los genomas de estas especies se secuencien.

Al analizar los conjuntos de datos filtrando todas aquellas secuencias que no contuvieran al menos un ORF completo o parcial se observó una reducción sustancial en el número de genes. En este paso se eliminaron $\frac{2}{3}$ de las secuencias de *H. amplexicaulis* y *A. lanata* y la mitad de las secuencias de *R. pilosa*. Sin embargo, la tasa de mapeo tuvo una reducción del 87-91% al 56-63%. Estas secuencias eliminadas podrían ser ARN no codificantes, y secuencias de elementos transponibles remanentes de la fase de construcción de las bibliotecas, ya que la misma implica un enriquecimiento de ARNm en una muestra de ARN total. En un ensayo similar en maíz, Li y colaboradores (2010) determinaron que el 84% de las lecturas mapearon con genes codificantes de proteínas, mientras que

el resto lo hizo con intrones, regiones intergénicas, elementos transponibles y uniones empalme entre exones [86].

El análisis de calidad de los transcriptomas realizado con BUSCO sugiere dos cosas interesantes. Por un lado, el número de secuencias duplicadas es muy bajo (1,7 a 1,8%), lo que implica transcriptomas poco redundantes. Por otro lado, al utilizar solamente muestras de tejidos foliares, se esperaría encontrar una pequeña cantidad de transcriptos involucrados en los procesos de germinación, el desarrollo radicular o el desarrollo floral, redundando en niveles de completitud moderados. Sin embargo, a pesar de estos motivos, el bajo número de secuencias no encontradas (11,0 a 12,7%) sugiere transcriptomas con un buen nivel de completitud. Esto podría deberse a dos motivos: (1) a una estrategia de muestreo que logró seleccionar una variedad de estados de desarrollo y (2) a la cantidad total de lecturas utilizadas en el ensamblado (alrededor de 420 millones de lecturas por especie), incrementando el nivel de detección de transcriptos muy poco expresados. Finalmente, al analizar los patrones de expresión diferencial, se vio que la mayoría de las secuencias que fueron eliminadas por encontrarse muy poco expresadas, no pertenecían a ningún ortogrupo. Además, luego del filtrado por valores de expresión, tanto los tamaños de los transcriptomas como el porcentaje de las secuencias anotadas se homogeneizó notablemente. Una evaluación más profunda de las secuencias descartadas permitiría entender mejor las razones de estos cambios.

Diferentes dinámicas entre las especies

Como era esperable, los fragmentos inmaduros presentaron una mayor dinámica transcripcional, medida como el número de GED en una determinada transición del desarrollo. Así mismo, en *A. lanata* el número de GED es elevado (con una tendencia decreciente) a lo largo del gradiente foliar, lo que sugiere un proceso de desarrollo más lento o que se prolonga más en el tiempo. En *H. amplexicaulis* también se observa una tendencia decreciente en el número de GED, pero con valores iniciales mucho más bajos que en la especie *C₄*. Notablemente, *R. pilosa* presenta un patrón de desarrollo intermedio entre las dos especies con un incremento en la dinámica de la expresión en la primera transición del desarrollo y un marcado descenso en las siguientes transiciones. Mientras que los estudios sistemáticos en Otachyriinae resuelven a las especies *H. amplexicaulis* y *R. pilosa* como muy próximas evolutivamente, en términos de GED el transcriptoma de *R. pilosa* es más similar al de *A. lanata* en los fragmentos inmaduros de la hoja.

Al analizar los GED se observa que la mayoría de ellos se encuentran enriquecidos en procesos de desarrollo como el establecimiento de la maquinaria fotosintética, la formación de HV, la remodelación de la pared celular y hormonas del desarrollo como las auxinas. Sin embargo, en casi todas las transiciones aparecen GED involucrados en la respuesta a estrés oxidativo.

Cuando se analizan los OED entre especies, se puede notar que las dos especies C_3 presentan un menor número de OED entre ellas, que con respecto a la especie C_4 . Estos resultados se correlacionan con las relaciones filogenéticas entre las 3 especies y los subtipos fotosintéticos que poseen. Sin embargo, esta diferencia no es constante a lo largo del desarrollo. En el F1, el menor número de OED se observa entre el par *A. lanata-R. pilosa*. Esta información, junto con los resultados que muestran una mayor dinámica transcripcional para los estados menos desarrollados, sugiere que ambas especies comparten el camino del desarrollo en estados tempranos, mientras que las diferencias se acrecientan acompañando el desarrollo madurativo de las hojas. Esto se sustenta además al observar que entre los 10 procesos más significativos en los GED entre los F1 y F3, un 50% son compartidos por ambas especies, mientras que los pares *H. amplexicaulis-A. lanata* y *H. amplexicaulis-R. pilosa* comparten el 30% y el 35%, respectivamente. En la transición siguiente, F3-F5, el par *H. amplexicaulis-R. pilosa* comparte el 60% de los términos de GO en sus GED, mientras que los pares *R. pilosa-A. lanata* y *H. amplexicaulis-A. lanata* comparten sólo el 10% de los términos cada uno. Estos resultados podrían estar indicando la predominancia de procesos divergentes entre las especies hacia el final del desarrollo.

Al analizar la expresión de genes biológicamente relacionados, como los involucrados en la fotosíntesis o el CCBB, se han podido identificar tendencias generalizadas entre las tres especies a lo largo del desarrollo foliar. Estos ortogrupos pasan de estar sobreexpresados en las especies C_3 y PK en fragmentos inmaduros a estar sobreexpresados en la especie C_4 en fragmentos maduros. Esto es notable para el ciclo C_4 , donde no solo hay muy pocos ortogrupos sobreexpresados en el fragmento 1 de la especie C_4 , sino que incluso hay un número considerable de ortogrupos sobreexpresados en las especies C_3 y PK. Los ortogrupos relacionados a procesos de formación de HV y pared celular, también muestran un mayor retraso en la expresión en el sentido C_4 .

En conjunto, los resultados sugieren que hay una diferencia en los momentos en que los procesos de desarrollo se llevan a cabo. En términos de fotosíntesis, luego de la primera transición, el aparato fotosintético en *H. amplexicaulis* parece estar completamente desarrollado. En cambio, en *A. lanata* esto sucede en el fragmento 5 y *R. pilosa* pareciera presentar valores intermedios. Esto podría explicar las similitudes observadas entre los transcriptomas entre estas dos especies.

Conservación de patrones de expresión

Para analizar el grado de conservación de los procesos de la fotosíntesis C_4 , se estudiaron puntualmente los patrones de expresión de ortogrupos claves. La mayoría de los patrones de expresión de las enzimas asociadas al subtipo NADP-ME de la vía C_4 se encuentran conservados entre la especie C_4 y las especies C_3 . Esta conservación en término de los patrones de expresión entre especies C_3 y

C₄ también se observa en otros casos. El grado de conservación de los patrones de expresión entre especies C₃ y C₄ también ha sido caracterizado en otros grupos taxonómicos. En particular, al comparar la variación en la expresión de genes asociados al ciclo C₄ en el proceso de de-etiolación de hojas, Xu y colaboradores (2016) observaron que los patrones de expresión de 7 de un total de 15 genes presentaban patrones de expresión conservados entre arroz y maíz [201]. A partir del análisis realizado en este trabajo, se identificó que 14 de los 15 genes estudiados previamente [201] muestran un patrón de expresión conservado a lo largo del gradiente foliar entre las especies estudiadas. Estos resultados sugieren un mayor grado de conservación de la vía C₄ en Otachyriinae. Sin embargo, a partir de los datos obtenidos en este trabajo, se ha identificado que existen diferencias en los niveles de expresión de la especie C₄ con respecto al resto de las especies analizadas. En *A. lanata* se han determinado niveles de expresión de los genes que codifican para enzimas de la vía C₄ varias veces superiores en los fragmentos más maduros de las hojas en comparación con las especies C₃ estudiadas. Las excepciones a esta tendencia son los genes que codifican para las enzimas CA, MDH y ALA-AT que presentan valores similares o más bajos en la especie C₄. Este fenómeno podría darse por diferencias en los niveles de compartimentalización entre M y CV entre las especies.

Al estudiar los patrones de expresión de los ortogrupos que codifican para las subunidades de la GDC, se nota una marcada conservación de los mismos. Los tres ortogrupos más expresados en las especies C₃ (P, Hb y T) presentan un patrón ascendente correlacionado con el desarrollo fotosintético. Dichos ortogrupos se encuentran subexpresados en la especie C₄ vs las otras. En *Flaveria*, se ha postulado que una duplicación ancestral en el gen GDC P, permitió la restricción de la expresión de una de las copias en las CV mientras que la otra fue reduciendo su expresión hasta perderse [202]. La presencia de una sola copia, podría indicar que el proceso evolutivo fue diferente en Otachyriinae. Una mejor identificación de las isoformas y de su expresión a nivel de tipo celular, podrían ayudar a discernir esta trayectoria evolutiva.

Los resultados de la expresión de genes del ciclo C₄ y del ciclo fotorrespiratorio sugieren que la conservación de los patrones de expresión podría ser un prerrequisito para la evolución de la fotosíntesis C₄, al menos en la subtribu Otachyriinae. En este escenario, la instalación de estas nuevas vías en las especies C₃ requeriría de un menor número cambios en los patrones de expresión durante el desarrollo.

Los patrones de expresión descendientes durante el desarrollo foliar identificados por SHR, SCR y la mitad de los IDD detectados coinciden con su función como reguladores transcripcionales en procesos de desarrollo. En hojas de *A. thaliana*, SHR se expresa principalmente en la vasculatura y es un activador de la transcripción tanto de SCR como SCL23 ([64], [83]). A pesar de que *A. lanata* presenta alrededor del doble de la fracción del área foliar ocupada por las HV, los niveles de expresión

de SHR en la misma son similares o inferiores a las demás especies. Al mismo tiempo no parece haber una relación directa entre los niveles de expresión de SHR y de sus posibles genes blanco SCR y SCL23.

En hojas de *Z. mays*, SCR se expresa más intensamente en las células precursoras del M (células del meristema de los primordios foliares) y, si bien no parece ser necesario para otorgarle la identidad a las CV, posee un efecto en la acumulación de NADP-ME en esas células ([88], [89]). Consecuentemente con esto, SCR está sobreexpresado en *A. lanata* vs las especies C₃. La determinación de la cantidad real del número de copias en *A. lanata* y su patrón de expresión tisular podrían ayudar a entender mejor su función y su relación con SHR y SCL23.

Un gran número de IDD está expresado en las hojas de las tres especies, cuatro de ellos con patrones de expresión muy conservados y en varios casos con niveles de transcritos similares. En *A. thaliana*, miembros de los linajes SG5, NKD y WRB están relacionados al desarrollo foliar [126]. Aquí, los SG5 y WRB poseen un patrón de expresión descendiente que los hace compatibles con una función de ese tipo. Los ortogrupos que no están conservados, no presentan una tendencia clara que permita realizar inferencias sobre sus funciones en el desarrollo foliar ni relación con otros genes como SHR y SCR.

El análisis de co-expresión detectó un gran número de módulos de ortogrupos conservados entre las tres especies. Si bien la mayoría de los módulos identificados están asociados al desarrollo, el análisis también detectó algunos relacionados a cada una de las especies, como los pares M3-M13, M7-M11 y M8-M15. Al mismo tiempo, ninguno de los módulos detectados está asociado a una réplica en particular, lo cual coincide con el análisis de componentes principales y significa que las diferencias observadas no están relacionadas a las réplicas biológicas.

La presencia de una gran cantidad de genes asociados a la fotorrespiración y al ciclo C₄ en dos "supramódulos" (M1-M6 y M2-M5) con patrones de expresión similares implican cambios sutiles de regulación entre las especies C₃, PK y C₄. Al mismo tiempo, estos resultados muestran que los reguladores transcripcionales SHR, SCR y SCL23 se co-expresan con varios miembros de la subfamilia IDD en hojas de gramíneas. La existencia de ambos pares de módulos donde los reguladores transcripcionales IDD, SHR y SCR, coexisten junto con genes involucrados en la fotorrespiración y la vía C₄, sugiere una regulación en común entre ellos. Mientras que el par M1-M6 presenta un patrón de expresión muy conservado entre las especies, el par M2-M5 muestra una divergencia entre las especies C₃ y la especie C₄. El *hub* SCR, junto con SCL23, NKD-grass2 y WRB-grass1, y el *hub* SG5-grass2-3 junto con SG5-grass1 y SHR podrían estar involucrados en la regulación de los genes asociados a la fotosíntesis C₄.

Notablemente, la principal enzima descarboxilante de la vía C_4 en *A. lanata*, NADP-ME, no resultó co-expresarse con ninguno de los módulos encontrados. La ausencia de un módulo que contenga a esta enzima en las tres especies sugiere la existencia de un evento que podría haber alterado la conectividad y/o expresión del gen que codifica para la NADP-ME de manera radical, en algún momento de la evolución de la fotosíntesis C_4 .

CONCLUSIONES GENERALES

Los estudios filogenéticos en las subfamilias SHR y SCR permitieron identificar eventos de duplicación específicos en ambas familias y determinar de una manera más correcta las relaciones de ortología entre las especies. El origen de ambas subfamilias precede al origen de las embriofitas y parece remontarse al origen mismo de los GRAS como reguladores transcripcionales.

En cuanto a los IDD, los resultados indican que surgieron de un ancestro común de las estreptofitas. Una vez en tierra, experimentaron una rápida radiación que acompañó las modificaciones clave a nivel morfológico y fisiológico que requirieron las plantas para el proceso de terrestrialización. Posteriormente, fueron reclutados por las angiospermas, donde se diversificaron en número de copias mientras adquirieron una gran heterogeneidad en su dominio C-terminal y patrones de expresión. Su evolución molecular sugiere un patrón dinámico de sub- y neo-funcionalización, que puede ser interpretado como un paso relevante en la conquista de la tierra por parte de las plantas y su posterior especiación.

El estudio de la anatomía foliar a través del análisis de secciones de hojas jóvenes, demostró la existencia de un gradiente de desarrollo foliar con diferencias marcadas entre las especies con diferentes subtipos fotosintéticos. Sin embargo, desde el F1, se observaron diferencias estructurales entre las especies que podrían condicionar su desarrollo posterior. La especie C₃ presentó una mayor disparidad estructural entre fragmentos, con un desarrollo de rasgos fenotípicos cuantitativos más sutil a lo largo del gradiente. La especie C₄ presentó una menor disparidad estructural y una mayor diversidad en cuanto a rasgos cuantitativos en el gradiente de desarrollo.

Los resultados mostrados en Capítulo 2 en conjunto con análisis reportados en la bibliografía sugieren que la superficie de contacto entre CV y M, es un rasgo importante a la hora de permitir la posibilidad de iniciar y sostener una vía C₄. Los parámetros de superficie de contacto entre CV y M propuestos en este trabajo (*Sb*, *Sb**, *Sib* y *Sib**) resultaron ser de utilidad dado que contemplan el tamaño de las CV, y su superficie activa para el transporte de metabolitos en la vía C₄.

Por medio del ensamblado de novo se lograron transcriptomas con un gran nivel de completitud, un bajo nivel de secuencias redundantes y una gran cantidad de transcriptos anotados. Una estimación más correcta del número de isoformas permitirá, en un futuro, mejorar sustancialmente la calidad de los mismos.

A través de un análisis de expresión diferencial se encontraron diferencias significativas entre las especies, con respecto a las dinámicas transcripcionales. Una mayor dinámica transcripcional

durante el desarrollo, se asocia con diferencias en la aparición y duración de los procesos de maduración de las hojas. La especie C₄ estudiada se asoció a un mayor número de transcritos expresados diferencialmente a lo largo del desarrollo, mientras que las especies C₃ se asociaron con un menor número. A pesar de esto, la especie PK presenta transcritos expresados diferencialmente que son similares (en cuanto a calidad y cantidad) a los de la especie C₄, en los fragmentos más inmaduros de las hojas.

Al analizar la expresión de grupos de genes biológicamente relacionados, como los involucrados en la fotosíntesis o el CCBB, se observan tendencias claras a lo largo del desarrollo. Estos ortogrupos pasan de estar más expresados en las especies C₃ y PK en fragmentos más inmaduros a estar más expresados en la especie C₄ en fragmentos más maduros. Esto es notable para el ciclo C₄, donde no solo hay muy pocos ortogrupos sobreexpresados en el fragmento 1 de la especie C₄, sino que incluso hay un número considerable de ortogrupos sobreexpresados en las especies C₃ y PK. Los ortogrupos relacionados a procesos de formación de HV y pared celular, también muestran un mayor retraso en la expresión en el sentido C₄.

Con respecto al ciclo C₄, la mayoría de los patrones de expresión de las enzimas asociadas al subtipo NADP-ME se encuentran conservados a lo largo del desarrollo entre la especie C₄ y las especies C₃. La diferencia entre las especies C₃ y la especie C₄, radica en que los niveles de expresión en esta última son varias veces superiores en los fragmentos de hoja maduros. Estos resultados sugieren que la conservación de los patrones de expresión de las enzimas C₄ podría ser un prerrequisito, en la subtribu Otachyriinae, para la evolución de la fotosíntesis C₄.

BIBLIOGRAFÍA

- [1] MP Johnson (2016) Photosynthesis. *Essays Biochem.* 60 (3):255–273.
- [2] T Shikanai, H Yamamoto (2017) Contribution of Cyclic and Pseudo-cyclic Electron Transport to the Formation of Proton Motive Force in Chloroplasts. *Mol. Plant* 10 (1):20–29.
- [3] TB Schreier, JM Hibberd (2019) Variations in the Calvin-Benson cycle: Selection pressures and optimization? *J. Exp. Bot.* 70 (6):1697–1701.
- [4] RF Sage, DA Way, DS Kubien (2008) Rubisco, Rubisco activase, and global climate change, in *Journal of Experimental Botany*, 2008.
- [5] E Carmo-Silva, JC Scales, PJ Madgwick, MAJ Parry (2015) Optimizing Rubisco and its regulation for greater resource use efficiency, *Plant, Cell and Environment*. 2015.
- [6] RF Sage, TL Sage, F Kocacinar (2012) Photorespiration and the evolution of C₄ photosynthesis. *Annu. Rev. Plant Biol.* 63:19–47.
- [7] MD Hatch, CR Slack (1966) Photosynthesis by sugar-cane leaves. A new carboxylation reaction and the pathway of sugar formation. *Biochem. J.* 101 (1):103–111.
- [8] D Kumar, EA Kellogg (2018) Getting closer: vein density in C₄ leaves, *New Phytologist*, 221 (3). :1260–1267, 2018.
- [9] RF Sage, RK Monson (1999) *C₄ Plant biology*, 105 (2). .
- [10] GE Edwards, E V. Voznesenskaya (2011) Chapter 4 C₄ Photosynthesis: Kranz Forms and Single-Cell C₄ in Terrestrial Plants, En: *C₄ Photosynthesis and Related CO₂ Concentrating Mechanisms*, (Chapter 4), p. :29–61.
- [11] RF Sage, PA Christin, EJ Edwards (2011) The C₄ plant lineages of planet Earth. *J. Exp. Bot.* 62 (9):3155–3169.
- [12] FR Danila, WP Quick, RG White, RT Furbank, S von Caemmerer (2016) The Metabolite Pathway between Bundle Sheath and Mesophyll: Quantification of Plasmodesmata in Leaves of C₃ and C₄ Monocots. *Plant Cell* .
- [13] FR Danila, WP Quick, RG White, S Kelly, S Von Caemmerer, RT Furbank (2018) Multiple mechanisms for enhanced plasmodesmata density in disparate subtypes of C₄ grasses. *J. Exp. Bot.* 69 (5):1135–1145.
- [14] M Stata, TL Sage, TD Rennie, R Khoshravesh, S Sultmanis, Y Khaikin, M Ludwig, RF Sage (2014) Mesophyll cells of C₄ plants have fewer chloroplasts than those of closely related C₃ plants. *Plant Cell Environ.* 37 (11):2587–2600.

- [15] Y Hatakeyama, O Ueno (2017) Intracellular position of mitochondria in mesophyll cells differs between C3 and C4 grasses. *J. Plant Res.* 130 (5):885–892.
- [16] Y Wang, A Bräutigam, APM Weber, XG Zhu (2014) Three distinct biochemical subtypes of C4 photosynthesis? A modelling analysis. *J. Exp. Bot.* 65 (13):3567–3578.
- [17] O Ghannoum, J Evans, S von Caemmerer (2011) Chapter 8: Nitrogen and Water Use Efficiency of C 4 Plants, En: *C4 photosynthesis and related CO2 concentrating mechanisms*, p. :129–146.
- [18] MA Hernández-Prieto, C Foster, A Watson-Lazowski, O Ghannoum, M Chen (2019) Comparative analysis of thylakoid protein complexes in the mesophyll and bundle sheath cells from C 3 , C 4 and C 3 –C 4 Paniceae grasses. *Physiol. Plant.* 166 (1):134–147.
- [19] G Schuster, I Ohad, B Martineau, WC Taylor (1985) Differentiation and development of bundle sheath and mesophyll thylakoids in maize. Thylakoid polypeptide composition, phosphorylation, and organization of photosystem II. *J. Biol. Chem.* 260 (21):11866–11873.
- [20] O Ghannoum, JR Evans, SC Wah, TJ Andrews, JP Conroy, S Von Caemmerer (2005) Faster Rubisco is the key to superior nitrogen-use efficiency in NADP-malic enzyme relative to NAD-malic enzyme C4 grasses. *Plant Physiol.* 137 (2):638–650.
- [21] M Huber, R Caballero (2011) The early Eocene equable climate problem revisited. *Clim. Past* 7 (2):603–633.
- [22] JC McElwain, KJ Willis, R Lupia (2005) Cretaceous CO2 Decline and the Radiation and Diversification of Angiosperms. *A Hist. Atmos. CO2 Its Eff. Plants, Anim. Ecosyst.* (1967):133–165.
- [23] LM Gerhart, JK Ward (2010) Plant responses to low [CO2] of the past. *New Phytol.* 188 (3):674–695.
- [24] RF Sage (2017) A portrait of the C4 photosynthetic family on the 50th anniversary of its discovery: Species number, evolutionary lineages, and Hall of Fame, *Journal of Experimental Botany*, 68 (2). :e11–e28, 2017.
- [25] G Besnard, AM Muasya, F Russier, EH Roalson, N Salamin, PA Christin (2009) Phylogenomics of C4 photosynthesis in sedges (Cyperaceae): Multiple appearances and genetic convergence. *Mol. Biol. Evol.* 26 (8):1909–1919.
- [26] RF Sage, R Khoshravesh, TL Sage (2014) From proto-Kranz to C4 Kranz: Building the bridge to C 4 photosynthesis, *Journal of Experimental Botany*, 65 (13). Oxford University Press, :3341–3356, 2014.
- [27] RK Monson, S Rawsthorne, JI Centre (2000) Chapter 22 Intermediate Plants. *Physiology* :533–550.
- [28] RF Sage, RF Sage (2004) The evolution of C 4 photosynthesis. 161:341–370.
- [29] BP Williams, IG Johnston, S Covshoff, JM Hibberd (2013) Phenotypic landscape inference reveals multiple evolutionary paths to C4 photosynthesis. *Elife* 2:1–19.

- [30] PA Christin, CP Osborne, DS Chatelet, JT Columbus, G Besnard, TR Hodkinson, LM Garrison, MS Vorontsova, EJ Edwards (2013) Anatomical enablers and the evolution of C4 photosynthesis in grasses. *Proc. Natl. Acad. Sci. U. S. A.* 110 (4):1381–1386.
- [31] PA Christin, CP Osborne (2013) The recurrent assembly of C4 photosynthesis, an evolutionary tale, *Photosynthesis Research*, 117 (1–3). :163–175, 2013.
- [32] R Muhaidat, RF Sage, NG Dengler (2007) Diversity of Kranz anatomy and biochemistry in C4 eudicots. *Am. J. Bot.* 94 (3):362–381.
- [33] R Muhaidat, TL Sage, MW Frohlich, NG Dengler, RF Sage (2011) Characterization of C3-C4 intermediate species in the genus *Heliotropium* L. (Boraginaceae): Anatomy, ultrastructure and enzyme activity. *Plant, Cell Environ.* 34 (10):1723–1736.
- [34] C Scoffoni, M Rawls, A Mckown, H Cochard, L Sack (2011) Decline of leaf hydraulic conductance with dehydration: Relationship to leaf size and venation architecture. *Plant Physiol.* 156 (2):832–843.
- [35] H Griffiths, G Weller, LFM Toy, RJ Dennis (2013) You're so vein: Bundle sheath physiology, phylogeny and evolution in C3 and C4 plants. *Plant, Cell Environ.* 36 (2):249–261.
- [36] MR Lundgren, LT Dunning, JK Olofsson, JJ Moreno-Villena, JW Bouvier, TL Sage, R Khoshravesh, S Sultmanis, M Stata, BS Ripley, MS Vorontsova, G Besnard, C Adams, N Cuff, A Mapaura, ME Bianconi, CM Long, PA Christin, CP Osborne (2019) C4 anatomy can evolve via a single developmental change. *Ecol. Lett.* 22 (2):302–312.
- [37] M Lauterbach, R Zimmer, AC Alexa, S Adachi, R Sage, T Sage, T MacFarlane, M Ludwig, G Kadereit (2019) Variation in leaf anatomical traits relates to the evolution of C4 photosynthesis in Tribuloideae (Zygophyllaceae). *Perspect. Plant Ecol. Evol. Syst.* 39.
- [38] X Wang, U Gowik, H Tang, JE Bowers, P Westhoff, AH Paterson (2009) Comparative genomic analysis of C4 photosynthetic pathway evolution in grasses. *Genome Biol.* 10 (6).
- [39] TL Sage, FA Busch, DC Johnson, PC Friesen, CR Stinson, M Stata, S Sultmanis, BA Rahman, S Rawsthorne, RF Sage (2013) Initial events during the evolution of C4 photosynthesis in C3 species of Flaveria. *Plant Physiol.* 163 (3):1266–1276.
- [40] S Rawsthorne, CM Hylton, AM Smith, HW Woolhouse (1988) Photorespiratory metabolism and immunogold localization of photorespiratory enzymes in leaves of C3 and C3-C4 intermediate species of *Moricandia*. *Planta* 173 (3):298–308.
- [41] RF Sage, RK Monson, JR Ehleringer, S Adachi, RW Pearcy (2018) Some like it hot: the physiological ecology of C4 plant evolution. *Oecologia* 187 (4):941–966.
- [42] MR Lundgren, PA Christin, EG Escobar, BS Ripley, G Besnard, CM Long, PW Hattersley, RP Ellis, RC Leegood, CP Osborne (2016) Evolutionary implications of C3–C4 intermediates in the grass *Alloteropsis*

semialata. *Plant Cell Environ.* 39 (9):1874–1885.

- [43] Y Yorimitsu, A Kadosono, Y Hatakeyama, T Yabiku, O Ueno (2019) Transition from C3 to proto-Kranz to C3–C4 intermediate type in the genus *Chenopodium* (Chenopodiaceae). *J. Plant Res.* 132 (6):839–855.
- [44] S-H Cheng, BD Moore, GE Edwards, MSB Ku (1988) Photosynthesis in *Flaveria brownii*, a C4-Like Species. *Plant Physiol.* 87 (4):867–873.
- [45] TE Huxman, RK Monson (2003) Stomatal responses of C3, C3-C4 and C4 *Flaveria* species to light and intercellular CO2 concentration: Implications for the evolution of stomatal behaviour. *Plant, Cell Environ.* 26 (2):313–322.
- [46] U Gowik, S Schulze, M Saladié, V Rolland, SK Tanz, P Westhoff, M Ludwig (2017) A MEM1-like motif directs mesophyll cell-specific expression of the gene encoding the C4 carbonic anhydrase in *Flaveria*. *J. Exp. Bot.* 68 (2):311–320.
- [47] M Ludwig (2013) Evolution of the C4 photosynthetic pathway: Events at the cellular and molecular levels. *Photosynth. Res.* 117 (1–3):147–161.
- [48] LD Pysh, JW Wysocka-Diller, C Camilleri, D Bouchez, PN Benfey (1999) The GRAS gene family in *Arabidopsis*: Sequence characterization and basic expression analysis of the SCARECROW-LIKE genes. *Plant J.* 18 (1):111–119.
- [49] X Sun, WT Jones, EHA Rikkerink (2012) GRAS proteins: The versatile roles of intrinsically disordered proteins in plant signalling. *Biochem. J.* 442 (1):1–12.
- [50] S Li, Y Zhao, Z Zhao, X Wu, L Sun, Q Liu, Y Wu (2016) Crystal Structure of the GRAS Domain of SCARECROW-LIKE7 in *Oryza sativa*. *Plant Cell*.
- [51] Y Hirano, M Nakagawa, T Suyama, K Murase, M Shirakawa, S Takayama, TP Sun, T Hakoshima (2017) Structure of the SHR-SCR heterodimer bound to the BIRD/IDD transcriptional factor JKD. *Nat. Plants* 3.
- [52] J Peng, P Carol, DE Richards, KE King, RJ Cowling, GP Murphy, NP Harberd (1997) The *Arabidopsis* GAI gene defines a signaling pathway that negatively regulates gibberellin responses. *Genes Dev.* 11 (23):3194–3205.
- [53] JL Bowman, T Kohchi, KT Yamato, J Jenkins, S Shu, K Ishizaki, S Yamaoka, R Nishihama, Y Nakamura, F Berger, C Adam, SS Aki, F Althoff, T Araki, MA Arteaga-Vazquez, S Balasubramanian, K Barry, D Bauer, CR Boehm, L Briginshaw, J Caballero-Perez, B Catarino, F Chen, S Chiyoda, M Chovatia, KM Davies, M Delmans, T Demura, T Dierschke, L Dolan, AE Dorantes-Acosta, DM Eklund, SN Florent, E Flores-Sandoval, A Fujiyama, H Fukuzawa, B Galik, D Grimanelli, J Grimwood, U Grossniklaus, T Hamada, J Haseloff, AJ Hetherington, A Higo, Y Hirakawa, HN Hundley, Y Ikeda, K Inoue, S ichiro Inoue, S Ishida, Q Jia, M Kakita, T Kanazawa, Y Kawai, T Kawashima, M Kennedy, K Kinose, T Kinoshita, Y Kohara, E Koide, K Komatsu, S Kopschke, M Kubo, J Kyojuka, U Lagercrantz, SS Lin, E Lindquist, AM Lipzen, CW Lu, E De Luna, RA Martienssen, N Minamino, M Mizutani, M Mizutani, N Mochizuki, I Monte,

R Mosher, H Nagasaki, H Nakagami, S Naramoto, K Nishitani, M Ohtani, T Okamoto, M Okumura, J Phillips, B Pollak, A Reinders, M Rövekamp, R Sano, S Sawa, MW Schmid, M Shirakawa, R Solano, A Spunde, N Suetsugu, S Sugano, A Sugiyama, R Sun, Y Suzuki, M Takenaka, D Takezawa, H Tomogane, M Tsuzuki, T Ueda, M Umeda, JM Ward, Y Watanabe, K Yazaki, R Yokoyama, Y Yoshitake, I Yotsui, S Zachgo, J Schmutz (2017) Insights into Land Plant Evolution Garnered from the *Marchantia polymorpha* Genome. *Cell* 171 (2):287-304.e15.

- [54] T Greb, O Clarenz, E Schafer, R Herrero, G Schmitz, K Theres (2003) Molecular analysis of the LATERAL SUPPRESSOR gene in *Arabidopsis*. *Genes Dev.* 17:1175–1187.
- [55] X Li, Q Qian, Z Fu, Y Wang, G Xiong, D Zeng, X Wang, X Liu, S Teng, F Hiroshi, M Yuan, D Luo, B Han, J Li (2003) Control of tillering in rice. *Nature* 422 (6932):618–621.
- [56] J Stuurman, F Jäggi, C Kuhlemeier (2002) Shoot meristem maintenance is controlled by a GRAS-gene mediated signal from differentiating cells. *Genes Dev.* 16 (17):2213–2218.
- [57] Y Helariutta, H Fukaki, J Wysocka-Diller, K Nakajima, J Jung, G Sena, MT Hauser, PN Benfey (2000) The SHORT-ROOT gene controls radial patterning of the *Arabidopsis* root through radial signaling. *Cell* 101 (5):555–567.
- [58] C Bolle, C Koncz, NH Chua (2000) PAT1, a new member of the GRAS family, is involved in phytochrome A signal transduction. *Genes Dev.* 14 (10):1269–1278.
- [59] K Morohashi, M Minami, H Takase, Y Hotta, K Hiratsuka (2003) Isolation and characterization of a novel GRAS gene that regulates meiosis-associated gene expression. *J. Biol. Chem.* 278 (23):20865–20873.
- [60] M Mayrose, SK Ekengren, S Melech-Bonfil, GB Martin, G Sessa (2006) A novel link between tomato GRAS genes, plant disease resistance and mechanical stress response. *Mol. Plant Pathol.* 7 (6):593–604.
- [61] HS Ma, D Liang, P Shuai, XL Xia, WL Yin (2010) The salt-and drought-inducible poplar GRAS protein SCL7 confers salt and drought tolerance in *Arabidopsis thaliana*. *J. Exp. Bot.* 61 (14):4011–4019.
- [62] H Cui, Y Hao, D Kong (2012) SCARECROW has a SHORT-ROOT-independent role in modulating the sugar response. *Plant Physiol.* 158 (4):1769–1778.
- [63] P Kaló, C Gleason, A Edwards, J Marsh, RM Mitra, S Hirsch, J Jakab, S Sims, SR Long, J Rogers, GB Kiss, JA Downie, GED Oldroyd (2005) Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. *Science* (80-.). 308 (5729):1786–1789.
- [64] EK Yoon, S Dhar, MH Lee, JH Song, SA Lee, G Kim, S Jang, JW Choi, JE Choe, JH Kim, MM Lee, J Lim (2016) Conservation and Diversification of the SHR-SCR-SCL23 Regulatory Network in the Development of the Functional Endodermis in *Arabidopsis* Shoots. *Mol. Plant* .
- [65] H Cui, D Kong, X Liu, Y Hao (2014) SCARECROW, SCR-LIKE 23 and SHORT-ROOT control bundle

sheath cell fate and function in *Arabidopsis thaliana*. *Plant J.* .

- [66] TL Slewinski, AA Anderson, C Zhang, R Turgeon (2012) Scarecrow plays a role in establishing Kranz anatomy in maize leaves. *Plant Cell Physiol.* .
- [67] C Tian, P Wan, S Sun, J Li, M Chen (2004) Genome-wide analysis of the GRAS gene family in rice and *Arabidopsis*. *Plant Mol. Biol.* 54:519–532.
- [68] D Abarca, A Pizarro, I Hernández, C Sánchez, SP Solana, A del Amo, E Carneros, C Díaz-Sala (2014) The GRAS gene family in pine: Transcript expression patterns associated with the maturation-related decline of competence to form adventitious roots. *BMC Plant Biol.* 14 (1):1–19.
- [69] W Huang, Z Xian, X Kang, N Tang, Z Li (2015) Genome-wide identification, phylogeny and expression analysis of GRAS gene family in tomato. *BMC Plant Biol.* 15 (1):1–18.
- [70] J Lu, T Wang, Z Xu, L Sun, Q Zhang (2014) Genome-wide analysis of the GRAS gene family in *Prunus mume*. *Mol. Genet. Genomics* 290 (1):303–317.
- [71] XM Song, TK Liu, WK Duan, QH Ma, J Ren, Z Wang, Y Li, XL Hou (2014) Genome-wide analysis of the GRAS gene family in chinese cabbage (*brassica rapa ssp. pekinensis*). *Genomics* .
- [72] H Liu, J Qin, H Fan, J Cheng, L Li, Z Liu (2017) Genome-wide identification, phylogeny and expression analyses of SCARECROW-LIKE(SCL) genes in millet (*Setaria italica*). *Physiol. Mol. Biol. Plants* .
- [73] J Grimplet, P Agudelo-Romero, RT Teixeira, JM Martinez-Zapater, AM Fortes (2016) Structural and Functional Analysis of the GRAS Gene Family in Grapevine Indicates a Role of GRAS Proteins in the Control of Development and Stress Responses. *Front. Plant Sci.* 7.
- [74] H Zhang, Y Cao, C Shang, J Li, J Wang, Z Wu, L Ma, T Qi, C Fu, Z Bai, B Hu (2017) Genome-wide characterization of GRAS family genes in *Medicago truncatula* reveals their evolutionary dynamics and functional diversification. *PLoS One* 12 (9):1–22.
- [75] B Liu, Y Sun, J Xue, X Jia, R Li (2018) Genome-wide characterization and expression analysis of GRAS gene family in pepper (*Capsicum annuum L.*). *PeerJ* 2018 (5).
- [76] EM Engstrom (2011) Phylogenetic analysis of GRAS proteins from moss, lycophyte and vascular plant lineages reveals that GRAS genes arose and underwent substantial diversification in the ancestral lineage common to bryophytes and vascular plants. *Plant Signal. Behav.* 6 (6):850–854.
- [77] A Cenci, M Rouard (2017) Evolutionary Analyses of GRAS Transcription Factors in Angiosperms. *Front. Plant Sci.* 8 (March):1–15.
- [78] DE Richards, J Peng, NP Harberd (2000) Plant GRAS and metazoan STATs: One family? *BioEssays* 22 (6):573–577.
- [79] D Zhang, LM Iyer, L Aravind (2012) Bacterial GRAS domain proteins throw new light on gibberellic acid response mechanisms. *Bioinformatics* .

- [80] S Hirsch, J Kim, A Muñoz, AB Heckmann, JA Downie, GED Oldroyd (2009) GRAS Proteins Form a DNA Binding Complex to Induce Gene Expression during Nodulation Signaling in *Medicago truncatula*. *Plant Cell* 21 (2):545–557.
- [81] H Cui, MP Levesque, T Vernoux, JW Jung, AJ Paquette, KL Gallagher, JY Wang, I Blilou, B Scheres, PN Benfey (2007) An evolutionarily conserved mechanism delimiting SHR movement defines a single layer of endodermis in plants. *Science* (80-.). 316 (5823):421–425.
- [82] H Fukaki, J Wysocka-Diller, T Kato, H Fujisawa, PN Benfey, M Tasaka (1998) Genetic evidence that the endodermis is essential for shoot gravitropism in *Arabidopsis thaliana*. *Plant J.* 14 (4):425–430.
- [83] S Dhondt, F Coppens, F De Winter, K Swarup, RMH Merks, D Inze, MJ Bennett, GTS Beemster (2010) SHORT-ROOT and SCARECROW Regulate Leaf Growth in *Arabidopsis* by Stimulating S-Phase Progression of the Cell Cycle. *PLANT Physiol.* .
- [84] RF Evert, W Eschrich, W Heyser (1977) Distribution and structure of the plasmodesmata in mesophyll and bundle-sheath cells of *Zea mays* L. *Planta* 136 (1):77–89.
- [85] J Lim, JW Jung, CE Lim, M Lee, BJ Kim, WB Bruce, PN Benfey (2006) NIH Public Access. 59 (4):619–630.
- [86] P Li, L Ponnala, N Gandotra, L Wang, Y Si, SL Tausta, TH Kebrom, N Provar, R Patel, CR Myers, EJ Reidel, R Turgeon, P Liu, Q Sun, T Nelson, TP Brutnell (2010) The developmental dynamics of the maize leaf transcriptome. *Nat. Genet.* 42 (12):1060–1067.
- [87] A Kobayashi, S Miura, A Kozaki (2017) INDETERMINATE DOMAIN PROTEIN binding sequences in the 5'-untranslated region and promoter of the SCARECROW gene play crucial and distinct roles in regulating SCARECROW expression in roots and leaves. *Plant Mol. Biol.* 94 (1–2).
- [88] TE Hughes, JA Langdale (2020) SCARECROW gene function is required for photosynthetic development in maize. *Plant Direct* 4 (9):1–10.
- [89] TE Hughes, O V. Sedelnikova, H Wu, PW Becraft, JA Langdale (2019) Redundant SCARECROW genes pattern distinct cell layers in roots and leaves of maize. *Development* 146 (14):dev177543.
- [90] JS Hanas, DJ Hazuda, DF Bogenhagen, FY Wu, CW Wu (1983) *Xenopus* transcription factor A requires zinc for binding to the 5 S RNA gene. *J. Biol. Chem.* 258 (23):14120–14125.
- [91] H Takatsuji (1998) Zinc-finger transcription factors in plants, *Cellular and Molecular Life Sciences*, 54 (6). :582–596, 1998.
- [92] SS Krishna, I Majumdar, N V. Grishin (2003) Structural classification of zinc fingers, *Nucleic Acids Research*, 31 (2). :532–550, 2003.
- [93] G Malgieri, M Palmieri, L Russo, R Fattorusso, P V. Pedone, C Isernia (2015) The prokaryotic zinc-finger: Structure, function and comparison with the eukaryotic counterpart, *FEBS Journal*, 282 (23). :4480–4496,

2015.

- [94] JH Laity, BM Lee, PE Wright (2001) Zinc finger proteins: New insights into structural and functional diversity, *Current Opinion in Structural Biology*, 11 (1). :39–46, 2001.
- [95] R Tupler, G Perini, MR Green (2001) Expressing the human genome, *Nature*, 409 (6822). :832–833, 2001.
- [96] KJ Brayer, DJ Segal (2008) Keep your fingers off my DNA: Protein-protein interactions mediated by C2H2 zinc finger domains, *Cell Biochemistry and Biophysics*. 2008.
- [97] SA Wolfe, L Nekludova, CO Pabo (2002) DNA Recognition by Cys 2 His 2 Zinc Finger Proteins. *Annu. Rev. Biophys. Biomol. Struct.* 29 (1):183–212.
- [98] CC Englbrecht, H Schoof, S Böhm (2004) BMC Genomics. *BMC Genomics* 17:1–17.
- [99] J Colasanti, Z Yuan, V Sundaresan (1998) The indeterminate gene encodes a zinc finger protein and regulates a leaf-generated signal required for the transition to flowering in maize. *Cell* 93 (4):593–603.
- [100] D Welch, H Hassan, I Blilou, R Immink, R Heidstra, B Scheres (2007) Arabidopsis JACKDAW and MAGPIE zinc finger proteins delimit asymmetric cell division and stabilize tissue boundaries by restricting SHORT-ROOT action. *Genes Dev.* 21 (17):2196–2204.
- [101] H Ogasawara, R Kaimi, J Colasanti, A Kozaki (2011) Activity of transcription factor JACKDAW is essential for SHR/SCR-dependent activation of SCARECROW and MAGPIE and is modulated by reciprocal interactions with MAGPIE, SCARECROW and SHORT ROOT. *Plant Mol. Biol.* 77 (4–5):489–499.
- [102] MA Moreno-Risueno, R Sozzani, GG Yardimci, JJ Petricka, T Vernoux, I Blilou, J Alonso, CM Winter, U Ohler, B Scheres, PN Benfey (2015) Transcriptional control of tissue formation throughout root development. *Science* (80-.). 350 (6259):426–430.
- [103] Y Long, W Smet, A Cruz-Ramírez, B Castelijn, W de Jonge, AP Mähönen, BP Bouchet, GS Perez, A Akhmanova, B Scheres, I Blilou (2015) Arabidopsis BIRD Zinc Finger Proteins Jointly Stabilize Tissue Boundaries by Confining the Cell Fate Regulator SHORT-ROOT and Contributing to Fate Specification. *Plant Cell* 27 (4):1185–1199.
- [104] MT Morita, K Sakaguchi, SI Kiyose, K Taira, T Kato, M Nakamura, M Tasaka (2006) A C2H2-type zinc finger protein, SGR5, is involved in early events of gravitropism in Arabidopsis inflorescence stems. *Plant J.* 47 (4):619–628.
- [105] M Tanimoto, R Tremblay, J Colasanti (2008) Altered gravitropic response, amyloplast sedimentation and circumnutation in the Arabidopsis shoot gravitropism 5 mutant are associated with reduced starch levels. *Plant Mol. Biol.* 67 (1–2):57–69.
- [106] D Cui, J Zhao, Y Jing, M Fan, J Liu, Z Wang, W Xin, Y Hu (2013) The Arabidopsis IDD14, IDD15, and

IDD16 Cooperatively Regulate Lateral Organ Morphogenesis and Gravitropism by Promoting Auxin Biosynthesis and Transport. *PLoS Genet.* 9 (9).

- [107] T Liu, BJ Reinhart, E Magnani, T Huang, R Kerstetter, MK Barton (2012) Of blades and branches: Understanding and expanding the arabidopsis ad/abaxial regulatory network through target gene identification. *Cold Spring Harb. Symp. Quant. Biol.* 77:31–45.
- [108] BJ Reinhart, T Liu, NR Newell, E Magnani, T Huang, R Kerstetter, S Michaels, MK Barton (2013) Establishing a Framework for the Ad/Abaxial Regulatory Network of Arabidopsis: Ascertaining Targets of Class III HOMEODOMAIN LEUCINE ZIPPER and KANADI Regulation. *Plant Cell* 25 (9):3228–3249.
- [109] SL Qi, QF Lin, XJ Feng, HL Han, J Liu, L Zhang, S Wu, J Le, E Blumwald, XJ Hua (2019) IDD16 negatively regulates stomatal initiation via trans-repression of SPCH in Arabidopsis. *Plant Biotechnol. J.* 17 (7):1446–1457.
- [110] PJ Seo, J Ryu, SK Kang, CM Park (2011) Modulation of sugar metabolism by an INDETERMINATE DOMAIN transcription factor contributes to photoperiodic flowering in Arabidopsis. *Plant J.* 65 (3):418–429.
- [111] JA Feurtado, D Huang, L Wicki-Stordeur, LE Hemstock, MS Potentier, EWT Tsang, AJ Cutler (2011) The Arabidopsis C2H2 Zinc Finger INDETERMINATE DOMAIN1/ENHYDROUS Promotes the Transition to Germination by Regulating Light and Hormonal Signaling during Seed Maturation. *Plant Cell* 23 (5):1772–1794.
- [112] J Fukazawa, H Teramura, S Murakoshi, K Nasuno, N Nishida, T Ito, M Yoshida, Y Kamiya, S Yamaguchi, Y Takahashi (2014) DELLAs Function as Coactivators of GAI-ASSOCIATED FACTOR1 in Regulation of Gibberellin Homeostasis and Signaling in Arabidopsis. *Plant Cell* 26 (7):2920–2938.
- [113] R Völz, SK Kim, J Mi, AA Rawat, A Veluchamy, KG Mariappan, N Rayapuram, JM Daviere, P Achard, I Blilou, S Al-Babili, M Benhamed, H Hirt (2019) INDETERMINATE-DOMAIN 4 (IDD4) coordinates immune responses with plant-growth in Arabidopsis thaliana. *PLoS Pathog.* 15 (1).
- [114] V Coneva, T Zhu, J Colasanti (2007) Expression differences between normal and indeterminate1 maize suggest downstream targets of ID1, a floral transition regulator in maize. *J. Exp. Bot.* 58 (13):3679–3693.
- [115] AYM Wong, J Colasanti (2007) Maize floral regulator protein INDETERMINATE1 is localized to developing leaves and is not altered by light or the sink/source transition. *J. Exp. Bot.* 58 (3):403–414.
- [116] G Yi, AK Neelakandan, BC Gontarek, E Vollbrecht, PW Becraft (2015) The naked endosperm Genes Encode Duplicate INDETERMINATE Domain Transcription Factors Required for Maize Endosperm Cell Patterning and Differentiation. *Plant Physiol.* 167 (2):443–456.
- [117] BC Gontarek, AK Neelakandan, H Wu, PW Becraft (2016) NKD Transcription Factors Are Central Regulators of Maize Endosperm Development. *Plant Cell* 28 (12):2916–2936.

- [118] K Matsubara, U Yamanouchi, Z-X Wang, Y Minobe, T Izawa, M Yano (2008) Ehd2, a rice ortholog of the maize INDETERMINATE1 gene, promotes flowering by up-regulating Ehd1. *Plant Physiol.* 148 (3):1425–35.
- [119] C Wu, C You, C Li, T Long, G Chen, ME Byrne, Q Zhang (2008) RID1, encoding a Cys2/His2-type zinc finger transcription factor, acts as a master switch from vegetative to floral development in rice. *Proc. Natl. Acad. Sci.* 105 (35):12915–12920.
- [120] L Deng, L Li, S Zhang, J Shen, S Li, S Hu, Q Peng, J Xiao, C Wu (2017) Suppressor of rid1 (SID1) shares common targets with RID1 on florigen genes to initiate floral transition in rice. *PLoS Genet.* 13 (2).
- [121] S Wu, KL Gallagher (2013) Intact microtubules are required for the intercellular movement of the SHORT-ROOT transcription factor. *Plant J.* .
- [122] P Huang, H Yoshida, K Yano, S Kinoshita, K Kawai, E Koketsu, M Hattori, S Takehara, J Huang, K Hirano, RL Ordonio, M Matsuoka, M Ueguchi-Tanaka (2018) OsIDD2, a zinc finger and INDETERMINATE DOMAIN protein, regulates secondary cell wall formation. *J. Integr. Plant Biol.* 60 (2):130–143.
- [123] YH Xuan, RA Priatama, J Huang, B Il Je, JM Liu, SJ Park, HL Piao, DY Son, JJ Lee, SH Park, KH Jung, TH Kim, C deok Han (2013) Indeterminate domain 10 regulates ammonium-mediated gene expression in rice roots. *New Phytol.* 197 (3):791–804.
- [124] M Dou, S Cheng, B Zhao, Y Xuan, M Shao (2016) The indeterminate domain protein ROC1 regulates chilling tolerance via activation of DREB1B/CBF1 in rice. *Int. J. Mol. Sci.* 17 (3).
- [125] A Kozaki, S Hake, J Colasanti (2004) The maize ID1 flowering time regulator is a zinc finger protein with novel DNA binding properties. *Nucleic Acids Res.* 32 (5):1710–1720.
- [126] CP Coelho, P Huang, DY Lee, TP Brutnell (2018) Making Roots, Shoots, and Seeds: IDD Gene Family Diversification in Plants, *Trends in Plant Science*, 23 (1). :66–78, 2018.
- [127] S Fan, D Zhang, L Xing, S Qi, L Du, H Wu, H Shao, Y Li, J Ma, M Han (2017) Phylogenetic analysis of IDD gene family and characterization of its expression in response to flower induction in *Malus*. *Mol. Genet. Genomics* 292 (4):755–771.
- [128] M Kumar, DT Le, S Hwang, PJ Seo, HU Kim (2019) Role of the INDETERMINATE DOMAIN genes in plants. *Int. J. Mol. Sci.* 20 (9).
- [129] DM Goodstein, S Shu, R Howson, R Neupane, RD Hayes, J Fazo, T Mitros, W Dirks, U Hellsten, N Putnam, DS Rokhsar (2012) Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Res.* 40 (D1):1178–1186.
- [130] J Jin, F Tian, DC Yang, YQ Meng, L Kong, J Luo, G Gao (2017) PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45 (D1):D1040–D1045.

- [131] TL Bailey, M Boden, FA Buske, M Frith, CE Grant, L Clementi, J Ren, WW Li, WS Noble (2009) MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Res.* 37 (SUPPL. 2).
- [132] K Katoh, J Rozewicki, KD Yamada (2018) MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* 20 (4):1160–1166.
- [133] S Kumar, G Stecher, K Tamura (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* 33 (7):1870–4.
- [134] F Ronquist, M Teslenko, P Van Der Mark, DL Ayres, A Darling, S Höhna, B Larget, L Liu, MA Suchard, JP Huelsenbeck (2012) Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61 (3):539–542.
- [135] MA Miller, W Pfeiffer, T Schwartz (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *2010 Gatew. Comput. Environ. Work. GCE 2010* :1–8.
- [136] W Maddison, D Maddison (2018) Mesquite: a modular system for evolutionary analysis. Version 3.51. 2018.
- [137] H Yoshida, K Hirano, T Sato, N Mitsuda, M Nomoto, K Maeo, E Koketsu, R Mitani, M Kawamura, S Ishiguro, Y Tada, M Ohme-Takagi, M Matsuoka, M Ueguchi-Tanaka (2014) DELLA protein functions as a transcriptional activator through the DNA binding of the INDETERMINATE DOMAIN family proteins. *Proc. Natl. Acad. Sci.* 111 (21):7861–7866.
- [138] RE Timme, TR Bachvaroff, CF Delwiche (2012) Broad phylogenomic sampling and the sister lineage of land plants. *PLoS One* 7 (1).
- [139] K Hori, F Maruyama, T Fujisawa, T Togashi, N Yamamoto, M Seo, S Sato, T Yamada, H Mori, N Tajima, T Moriyama, M Ikeuchi, M Watanabe, H Wada, K Kobayashi, M Saito, T Masuda, Y Sasaki-Sekimoto, K Mashiguchi, K Awai, M Shimojima, S Masuda, M Iwai, T Nobusawa, T Narise, S Kondo, H Saito, R Sato, M Murakawa, Y Ihara, Y Oshima-Yamada, K Ohtaka, M Satoh, K Sonobe, M Ishii, R Ohtani, M Kanamori-Sato, R Honoki, D Miyazaki, H Mochizuki, J Umetsu, K Higashi, D Shibata, Y Kamiya, N Sato, Y Nakamura, S Tabata, S Ida, K Kurokawa, H Ohta (2014) Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. *Nat. Commun.* 5 (May):3978.
- [140] M Sagasser, GH Lu, K Hahlbrock, B Weisshaar (2002) A. thaliana Transparent testa 1 is involved in seed coat development and defines the WIP subfamily of plant zinc finger proteins. *Genes Dev.* 16 (1):138–149.
- [141] I Appelhagen, G Huet, GH Lu, G Strompen, B Weisshaar, M Sagasser (2010) Weird fingers: Functional analysis of WIP domain proteins. *FEBS Lett.* 584 (14):3116–3122.
- [142] Y Sawaki, S Iuchi, Y Kobayashi, Y Kobayashi, T Ikka, N Sakurai, M Fujita, K Shinozaki, D Shibata, M Kobayashi, H Koyama (2009) STOP1 Regulates Multiple Genes That Protect Arabidopsis from Proton and Aluminum Toxicities. *PLANT Physiol.* 150 (1):281–294.

- [143] Y Kobayashi, Y Ohyama, Y Kobayashi, H Ito, S Iuchi, M Fujita, CR Zhao, T Tanveer, M Ganesan, M Kobayashi, H Koyama (2014) STOP2 activates transcription of several genes for Al- and low pH-tolerance that are regulated by STOP1 in arabidopsis. *Mol. Plant* 7 (2):311–322.
- [144] ND Pires, L Dolan (2012) Morphological evolution in land plants: New designs with old genes. *Philos. Trans. R. Soc. B Biol. Sci.* 367 (1588):508–518.
- [145] Y Jiao, NJ Wickett, S Ayyampalayam, AS Chanderbali, L Landherr, PE Ralph, LP Tomsho, Y Hu, H Liang, PS Soltis, DE Soltis, SW Clifton, SE Schlarbaum, SC Schuster, H Ma, J Leebens-Mack, CW dePamphilis (2011) Ancestral polyploidy in seed plants and angiosperms. *Nature* 473:97.
- [146] B Nystedt, NR Street, A Wetterbom, A Zuccolo, Y-C Lin, DG Scofield, F Vezzi, N Delhomme, S Giacomello, A Alexeyenko, R Vicedomini, K Sahlin, E Sherwood, M Elfstrand, L Gramzow, K Holmberg, J Hällman, O Keech, L Klasson, M Koriabine, M Kucukoglu, M Källér, J Luthman, F Lysholm, T Niittylä, Å Olson, N Rilakovic, C Ritland, JA Rosselló, J Sena, T Svensson, C Talavera-López, G Theißen, H Tuominen, K Vanneste, Z-Q Wu, B Zhang, P Zerbe, L Arvestad, R Bhalerao, J Bohlmann, J Bousquet, R Garcia Gil, TR Hvidsten, P de Jong, J MacKay, M Morgante, K Ritland, B Sundberg, S Lee Thompson, Y Van de Peer, B Andersson, O Nilsson, PK Ingvarsson, J Lundeberg, S Jansson (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature* 497:579.
- [147] C Ruprecht, R Lohaus, K Vanneste, M Mutwil, Z Nikoloski, Y Van De Peer, S Persson (2017) Revisiting ancestral polyploidy in plants. *Sci. Adv.* 3 (7):1–7.
- [148] EA Kellogg, JL Bennetzen (2004) The evolution of nuclear genome structure in seed plants. *Am. J. Bot.* 91 (10):1709–1725.
- [149] EA Kellogg (2006) Progress and challenges in studies of the evolution of development. *J. Exp. Bot.* 57 (13):3505–3516.
- [150] D Hodick (1994) Negative gravitropism in *Chara protonemata*: A model integrating the opposite gravitropic responses of protonemata and rhizoids. *Planta* 195 (1):43–49.
- [151] CJ Harrison (2014) Development and genetics in the evolution of land plant body plans. *Philos. Trans. R. Soc. B-Biological Sci.* 372 (1713):20150490.
- [152] Y Long, J Goedhart, M Schneijderberg, I Terpstra, A Shimotohno, BP Bouchet, A Akhmanova, TWJ Gadella, R Heidstra, B Scheres, I Blilou (2015) SCARECROW-LIKE23 and SCARECROW jointly specify endodermal cell fate but distinctly control SHORT-ROOT movement. *Plant J.* .
- [153] C Drapek, EE Sparks, PN Benfey (2017) Uncovering Gene Regulatory Networks Controlling Plant Cell Differentiation, *Trends in Genetics*, 33 (8). :529–539, 2017.
- [154] PJ Seo, MJ Kim, JY Ryu, EY Jeong, CM Park (2011) Two splice variants of the IDD14 transcription factor competitively form nonfunctional heterodimers which may regulate starch metabolism. *Nat. Commun.* 2 (1).

- [155] X Wu, D Tang, M Li, K Wang, Z Cheng (2012) Loose Plant Architecture1, an INDETERMINATE DOMAIN Protein Involved in Shoot Gravitropism, Regulates Plant Architecture in Rice. *PLANT Physiol.* 161 (1):317–329.
- [156] J Ke, H Ma, X Gu, A Thelen, JS Brunzelle, J Li, HE Xu, K Melcher (2015) Structural basis for recognition of diverse transcriptional repressors by the TOPLESS family of corepressors. *Sci. Adv.* 1 (6):e1500107.
- [157] O Morrone, L Aagesen, MA Scataglini, DL Salariao, SS Denham, MA Chemisquy, SM Sede, LM Giussani, EA Kellogg, FO Zuloaga (2012) Phylogeny of the Paniceae (Poaceae: Panicoideae): Integrating plastid DNA sequences and morphology into a new classification. *Cladistics* 28 (4):333–356.
- [158] JM Acosta, FO Zuloaga, R Reinheimer (2019) Nuclear phylogeny and hypothesized allopolyploidization events in the Subtribe Otachyriinae (Paspaleae, Poaceae). *Syst. Biodivers.* 17 (3):277–294.
- [159] JM Acosta, MA Scataglini, R Reinheimer, FO Zuloaga (2014) A phylogenetic study of subtribe Otachyriinae (Poaceae, Panicoideae, Paspaleae). *Plant Syst. Evol.* 300 (10):2155–2166.
- [160] T Sendulsky, TR Soderstrom (1984) Revision of the South American genus *Otachyrium* (Poaceae, Panicoideae) /. *Revis. South Am. genus Otachyrium (Poaceae, Panicoideae)* / (57):1–24.
- [161] C. Holaday, A.S., Black (1981) Comparative Characterization of Phosphoenolpyruvate. *Plant Physiol.* 67 (1981):330–334.
- [162] RH Brown, JH Bouton, L Rigsby, M Rigler (1983) Photosynthesis of Grass Species Differing in Carbon Dioxide Fixation Pathways. *Plant Physiol.* 71 (2):425–431.
- [163] GE EDWARDS, MSB KU (1987) *Biochemistry of C3–C4 Intermediates*. ACADEMIC PRESS, INC.,
- [164] MR Lundgren, CP Osborne, PA Christin (2014) Econstructing Kranz anatomy to understand C₄ evolution, *Journal of Experimental Botany*. 2014.
- [165] R Khoshravesh, CR Stinson, M Stata, FA Busch, RF Sage, M Ludwig, TL Sage (2016) C₃-C₄ intermediacy in grasses: Organelle enrichment and distribution, glycine decarboxylase expression, and the rise of C₂ photosynthesis. *J. Exp. Bot.* 67 (10):3065–3078.
- [166] TD Leandro, TM Rodrigues, LG Clark, VL Scatena (2018) Fusoid cells in the grass family Poaceae (Poales): A developmental study reveals homologies and suggests new insights into their functional role in young leaves. *Ann. Bot.* 122 (5):833–848.
- [167] J Schindelin, I Arganda-Carreras, E Frise, V Kaynig, M Longair, T Pietzsch, S Preibisch, C Rueden, S Saalfeld, B Schmid, JY Tinevez, DJ White, V Hartenstein, K Eliceiri, P Tomancak, A Cardona (2012) Fiji: An open-source platform for biological-image analysis. *Nat. Methods* 9 (7):676–682.
- [168] JLL Pengelly, XRR Sirault, Y Tazoe, JR Evans, RT Furbank, S Von Caemmerer (2010) Growth of the C₄ dicot *Flaveria bidentis*: Photosynthetic acclimation to low light through shifts in leaf anatomy and biochemistry. *J. Exp. Bot.* 61 (14):4109–4122.

- [169] O Ghannoum, K Siebke, S von Caemmerer, JP Conroy (1998) The photosynthesis of young *Panicum* C4 leaves is not C3-like. *Plant. Cell Environ.* 21:1123–1131.
- [170] NK Koteyeva, E V. Voznesenskaya, AB Cousins, GE Edwards (2014) Differentiation of C4 photosynthesis along a leaf developmental gradient in two *Cleome* species having different forms of Kranz anatomy. *J. Exp. Bot.* 65 (13):3525–3541.
- [171] M Stata, TL Sage, RF Sage (2019) Mind the gap: the evolutionary engagement of the C4 metabolic cycle in support of net carbon assimilation, *Current Opinion in Plant Biology*, 49. Elsevier Ltd, :27–34, 01-Jun-2019.
- [172] AD McKown, NG Dengler (2010) Vein patterning and evolution in C₄ plants. *Botany* .
- [173] AD McKown, NG Dengler (2009) Shifts in leaf vein density through accelerated vein formation in C₄ *Flaveria* (Asteraceae). *Ann. Bot.* 104 (6):1085–1098.
- [174] M Lauterbach, H Schmidt, K Billakurthi, T Hankeln, P Westhoff, U Gowik, G Kadereit (2017) De novo Transcriptome Assembly and Comparison of C3, C3-C4, and C4 Species of Tribe Salsoleae (Chenopodiaceae). *Front. Plant Sci.* .
- [175] JJ Moreno-Villena, LT Dunning, CP Osborne, PA Christin (2018) Highly Expressed Genes Are Preferentially Co-Opted for C₄ Photosynthesis. *Mol. Biol. Evol.* .
- [176] A Bräutigam, S Schliesky, C K ulahoglu, CP Osborne, APM Weber (2014) Towards an integrative model of C4 photosynthetic subtypes: Insights from comparative transcriptome analysis of NAD-ME, NADP-ME, and PEP-CK C4 species. *J. Exp. Bot.* 65 (13):3579–3593.
- [177] C K ulahoglu, AK Denton, M Sommer, J Ma , S Schliesky, TJ Wrobel, B Berckmans, E Gongora-Castillo, C Robin Buell, R Simon, L De Veylder, A Bräutigam, APM Weber (2014) Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. *Plant Cell* 26 (8):3243–3260.
- [178] TR Pick, A Bräutigam, U Schl uter, AK Denton, C Colmsee, U Scholz, H Fahnenstich, R Pieruschka, U Rascher, U Sonnewald, APM Weber (2011) Systems analysis of a maize leaf developmental gradient redefines the current C₄ model and provides candidates for regulation. *Plant Cell* 23 (12):4208–4220.
- [179] Y Zhu, L Chen, C Zhang, P Hao, X Jing, X Li (2017) Global transcriptome analysis reveals extensive gene remodeling, alternative splicing and differential transcription profiles in non-seed vascular plant *Selaginella moellendorffii*. *BMC Genomics* 18 (Suppl 1):1–15.
- [180] SME Sahraeian, M Mohiyuddin, R Sebra, H Tilgner, PT Afshar, KF Au, N Bani Asadi, MB Gerstein, WH Wong, MP Snyder, E Schadt, HYK Lam (2017) Gaining comprehensive biological insight into the transcriptome by performing a broad-spectrum RNA-seq analysis. *Nat. Commun.* 8 (1):1–14.
- [181] M H lzer, M Marz (2019) De novo transcriptome assembly: A comprehensive cross-species comparison

of short-read RNA-Seq assemblers. *Gigascience* 8 (5):1–16.

- [182] P Langfelder, S Horvath (2008) WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* 9.
- [183] P Langfelder, R Luo, MC Oldham, S Horvath (2011) Is my network module preserved and reproducible? *PLoS Comput. Biol.* 7 (1).
- [184] T Nelson (2011) The grass leaf developmental gradient as a platform for a systems understanding of the anatomical specialization of C 4 leaves, *Journal of Experimental Botany*, 62 (9). :3039–3048, May-2011.
- [185] DS Carvalho, A V. Nishimwe, JC Schnable (2020) IsoSeq transcriptome assembly of C3 panicoid grasses provides tools to study evolutionary change in the Panicoideae. *Plant Direct* 4 (2).
- [186] SE Walker, J Lorsch (2013) *RNA purification - Precipitation methods*, 1st ed.530. Elsevier Inc.,
- [187] S Andrews (2010) FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010.
- [188] AM Bolger, M Lohse, B Usadel (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30 (15):2114–2120.
- [189] MG. Grabherr, N Brian J. Haas, Moran Yassour Joshua Z. Levin, Dawn A. Thompson, Ido Amit, Xian Adiconis, Lin Fan, Raktima Raychowdhury, Qiandong Zeng, Zehua Chen, Evan Mauceli, Nir Hacohen, Andreas Gnirke, Nicholas Rhind, Federica di Palma, Bruce W., and AR Friedman (2013) Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* 29 (7):644–652.
- [190] B Langmead, SL Salzberg (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9 (4):357–359.
- [191] R Patro, G Duggai, M Love, R Irizarry, C Kingsford (2017) Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference. *Nat. Methods* 14 (4):417–419.
- [192] L Fu, B Niu, Z Zhu, S Wu, W Li (2012) CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28 (23):3150–3152.
- [193] BJ Haas, A Papanicolaou, M Yassour, M Grabherr, D Philip, J Bowden, MB Couger, D Eccles, B Li, MD Macmanes, M Ott, J Orvis, N Pochet, F Strozzi, N Weeks, R Westerman, T William, CN Dewey, R Henschel, RD Leduc, N Friedman, A Regev (2013) *De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity*, 8 (8) . .
- [194] RM Waterhouse, M Seppey, FA Simao, M Manni, P Ioannidis, G Klioutchnikov, E V. Kriventseva, EM Zdobnov (2018) BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evol.* 35 (3):543–548.
- [195] DM Emms, S Kelly (2019) OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20:238.

- [196] R core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria., 2013.
- [197] MD Robinson, DJ McCarthy, GK Smyth (2009) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26 (1):139–140.
- [198] A Alexa, J Rahnenfuhrer (2020) topGO: Enrichment Analysis for Gene Ontology. 2020.
- [199] C Masclaux, MH Valadier, N Brugière, JF Morot-Gaudry, B Hirel (2000) Characterization of the sink/source transition in tobacco (*Nicotiana tabacum* L.) shoots in relation to nitrogen management and leaf senescence. *Planta* 211 (4):510–518.
- [200] J Lee, X Dong, K Choi, H Song, H Yi, Y Hur (2019) Identification of source-sink tissues in the leaf of Chinese cabbage (*Brassica rapa* ssp. *pekinensis*) by carbohydrate content and transcriptomic analysis. *Genes and Genomics* 42 (1):13–24.
- [201] J Xu, A Bräutigam, APM Weber, XG Zhu (2016) Systems analysis of cis-regulatory motifs in C4 photosynthesis genes using maize and rice leaf transcriptomic data during a process of de-etiolation. *J. Exp. Bot.* 67 (17):5105–5117.
- [202] S Schulze, P Westhoff, U Gowik (2016) Glycine decarboxylase in C3, C4 and C3-C4 intermediate species, *Current Opinion in Plant Biology*. 2016.

Anexo 1: Tablas, figuras e información suplementaria

Tabla S1: 25 motivos encontrados en los IDD

Motivos presentes en todos los linajes (HGC)					
Nº	Motivo	Valor E	Largo	Presencia (%)	
1	[IV]ALSP[KR]TL[ML]ATNR[FY][VI]CE[IV]C[NG]KGFQR[DE]QNLQLHRRGHN LPW	1.1e-7008	40	94.5 (99.5)	
2	GEKKWKC[DE][KR]CSK[RK]YAVQSDWKAHSA[TI]CGTREY[RK]CDCGT[LV] F	9.2e-6810	40	97.3	
3	[TS]SKEV[RK][KR][RK]VYVCPEP[TS]CVHHDP[SA]RALGDLTGIKKH[FY][CS] RKH	5.2e-6253	40	97.8 (99.5)	
4	SRRDSFITHRAFCDA	3.4e-2086	15	92.3	
5	[TN]PDPDAEV	2.6e-876	8	89.6	
6	SXSXXXXPXXXX[VA]K[KR]KRNLPG	2.3e-839	21	77.0	
Motivos ausentes en al menos un linaje (MGC)					
Nº	Motivo	Valor E	Largo	Presencia (%)	Ausente en
16	XXP[AS][AP]S[PA][HQ]MSATALLQKAAQMG[AS]T[ST]S[NS]SS	7.8e-1681	29	69.4	SG5, ID1
17	GGG[DK]GLT[RL]DFLG[VL]GG	7.2e-761	15	70.5	SG5
7	LA[EQ]ESA[RK][ALV]	4.1e-473	8	73.8	SG5
9	M[SA]NL[TS]S[AS][SA]	5.6e-154	8	51.9	ID1
Motivos presentes solo en algunos pocos linajes (LGC)					
Nº	Motivo	Valor E	Largo	Presencia (%)	Ausente en
18	EKA[AY]A[DE]EARA[QE]A[KR]RQ[A]JELAE[QR]E[LF]A[NS]A[KR]R[IM]R[QE] QA[QR][AV]EL[ES][KR]A	1.0e-603	40	15.3	SG5
19	[HQ][AV]L[RK][DE][HQ]A[VTI][RK]Q[IV][NSD][AS]T[LI]L[QE]ITC[HO][AS]C[RK]][HQ][KQ]F[RQ]A	1.7e-321	29	12.0	SG5
21	XQD[EQ][NS]EA[NH]L[QN]DLMNSL[AS][GAN]GG	4.0e-115	20	14.2	NKD, JKD, MGP

20	FHGLMQLP[DE][LQ]Q	4.9e-097	11	13.7	NKD, WRB
10	FFG[IL]R[DE][EG][DQ]Q[QH][QD]Q[QI][QK][QP]	3.6e-086	15	15.3	NKD, UIDDb, WRB
11	[GE]EASVSSG[NT]R[AT][ED]G[AFT][NS]	4.0e-082	15	12.6	WDP, ID1, WRB, MGP
27	[PQA][PA]A[SC][LS]SRTASS[TAP]SPSSDX[NV][FI]S	5.4e-080	21	9.3	SG5
22	[SA][AS][AN][AP][NG]L[FL]NL[GS]FF[SA][NG][NS][SG][NGS][SGT]S	4.9e-077	19	14.8	NKD
28	RLKEEAREQLR[LQ]AMA	6.4e-070	15	9.3	SG5
23	[PQ][QK][QE][QH][CQ]N[YF][AS]M[KR][PT]E[IM]PPW[PL]A[CM]Q[GYP]	1.9e-066	21	6.6	WRB
24	T[PS]QSRSLRDLIGVDPTFLCLAIGAPS[SL][LF][FP]	1.3e-063	29	3.3	GAF1
29	[AV]MSY[MV]SS[VA][VT]TEG[EG][DG][EA]	1.3e-063	15	9.3	SG5
25	HNLELQL[LM]P[PS]S	9.5e-061	11	14.2	SG5, ID1
30	[HG][LA][YT][GN][AS][SG][NS]M[AS]L[GS]LSQV[GN]S[QH][LI][TS]S	1.2e-059	21	8.7	NKD
26	FSH[RG][DE]IL[GNS][LM]AG[FL]D[SP]	1.6e-058	14	9.8	WRB, WDP

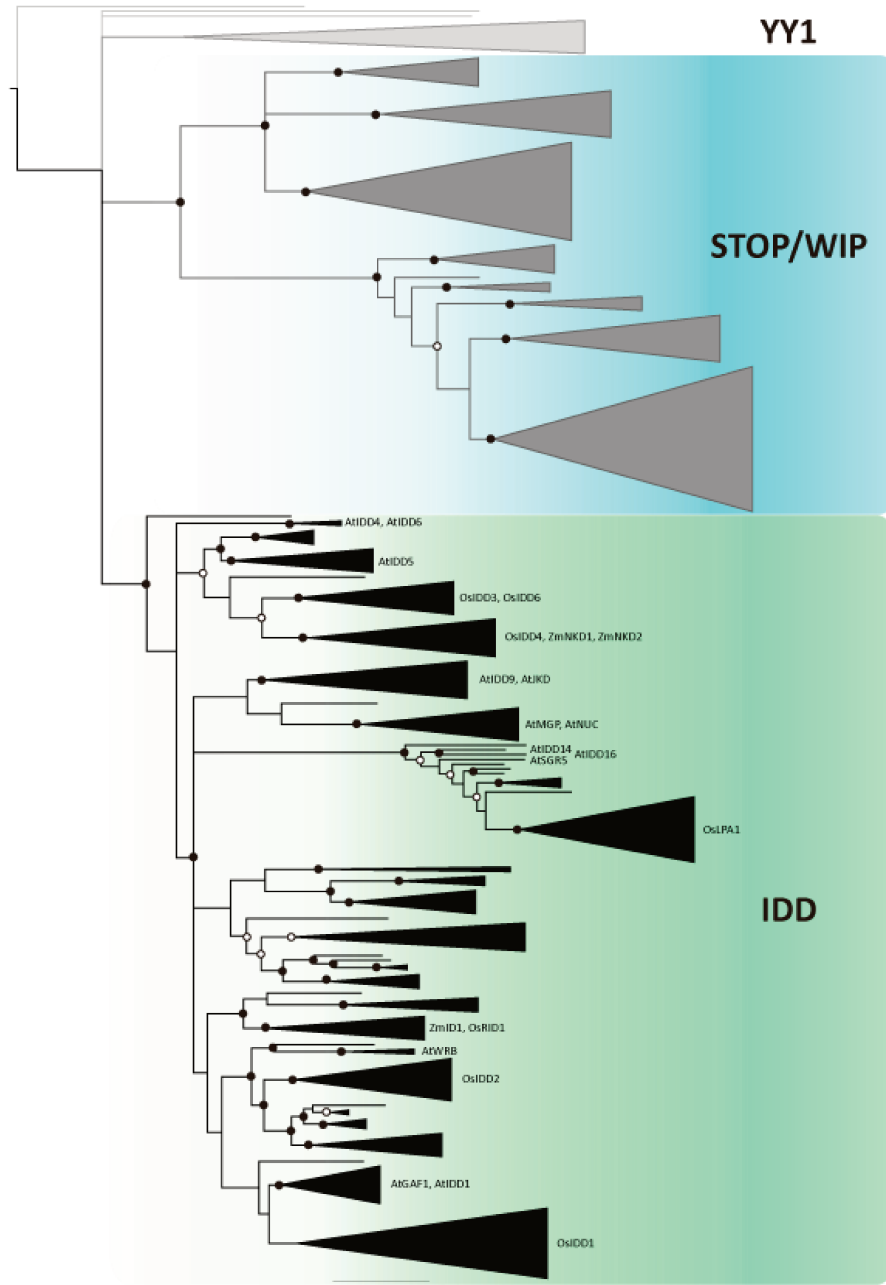


Figura S1. Reconstrucción filogenética e identificación de los IDD en angiospermas. Árbol de consenso de mayoría (N=30.000 árboles) de las potenciales IDD usando una matriz de 267 secuencias peptídicas. Los puntos blancos indican una Probabilidad Posterior (PP) = [0,8 a 0,89], los puntos negros indican PP= [0.9 a1]. Las IDD previamente caracterizadas se nombran en su clado correspondiente.

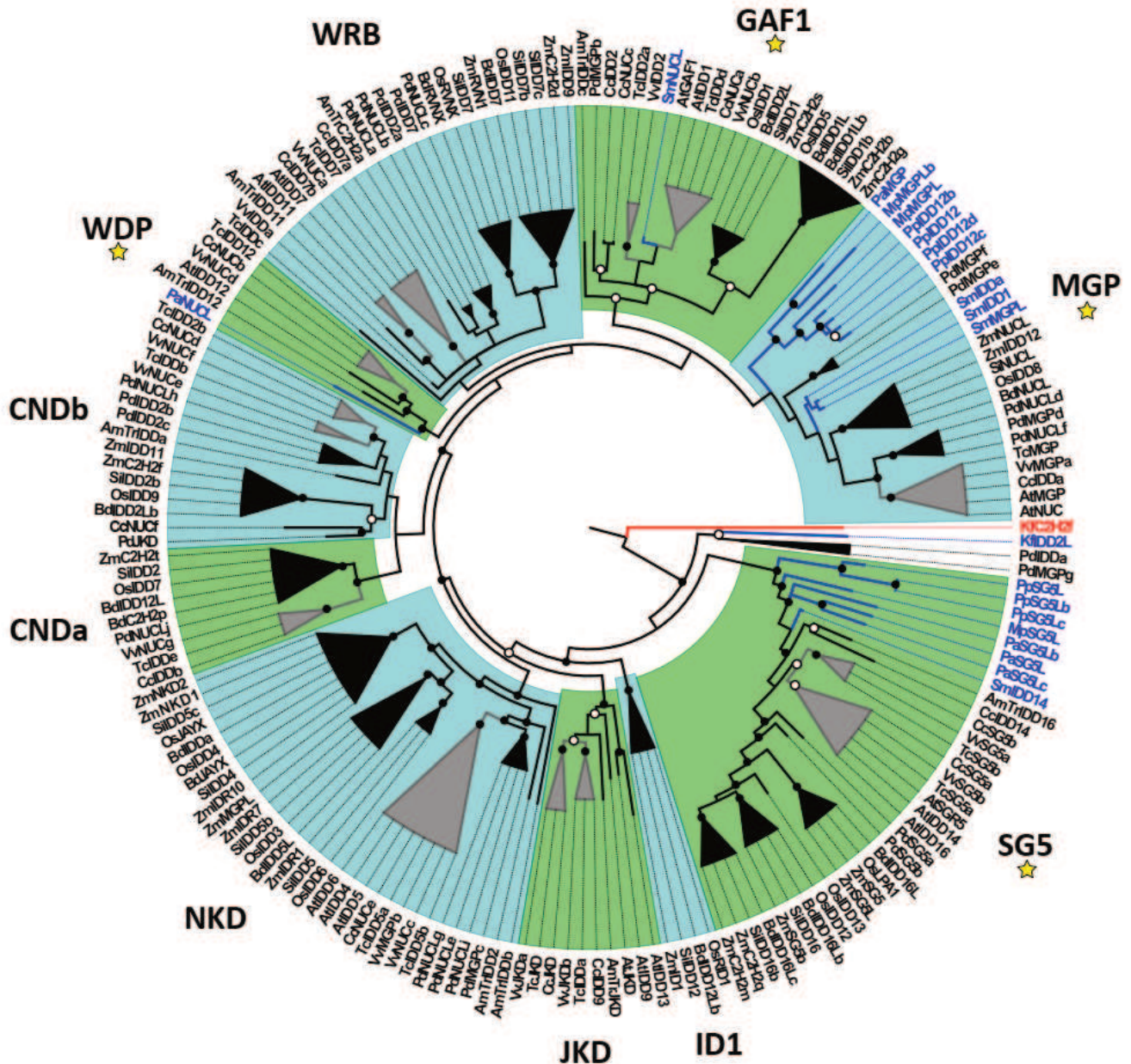


Figura S3. Reconstrucción filogenética de IDD en embriofitas. Árbol de consenso de mayoría (N=30.000 árboles) de la subfamilia IDD usando una matriz de 184 secuencias peptídicas. Los puntos blancos indican una Probabilidad Posterior (PP) = [0,8 a 0,89], los puntos negros indican PP= [0.9 a 1]. Los clados de monocotiledóneas y dicotiledóneas fueron colapsados en negro y gris, respectivamente. Las secuencias de especies no-angiospermas se muestran en azul y la secuencia outgroup C2H2 de *K. nitens*, en rojo. Los linajes que contienen secuencias de no-angiospermas están señalados con una estrella amarilla. Abreviaturas: SHOOT GRAVITROPISM 5 (SG5), INDETERMINATE 1 (ID1), GAI INTERACTING FACTOR 1 (GAF1), NAKED ENDOSPERM (NKD), JACKDAW (JKD), MAGPIE (MGP), CONDOR a (CNDa), CONDOR b (CNDb), WOODPECKER (WDP), WARBLER (WRB).

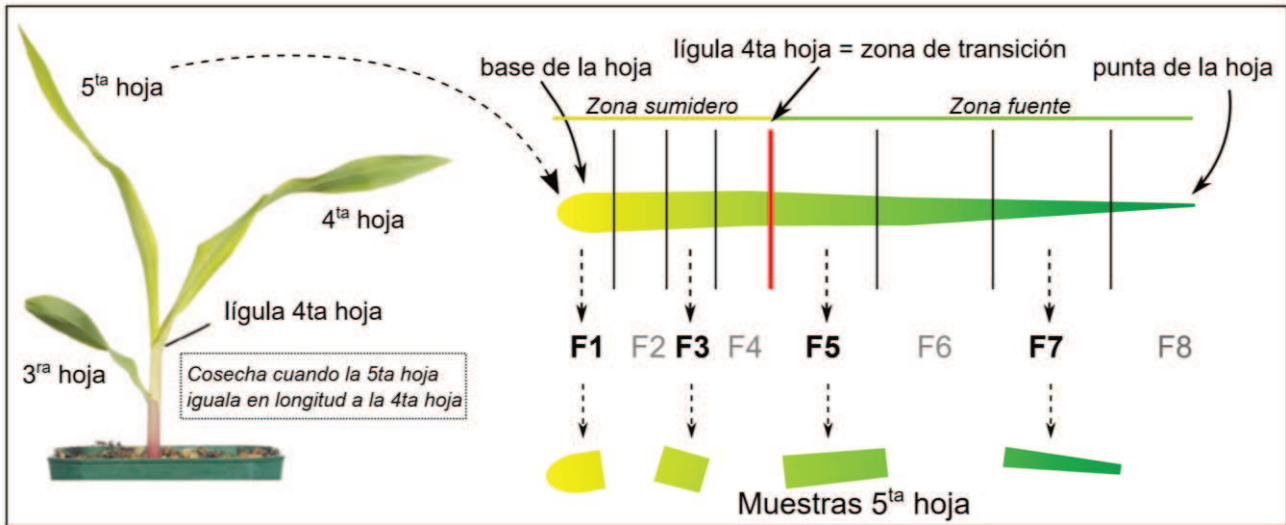


Figura S4. Esquema de la toma de muestras de la 5^{ta} hoja joven para los estudios de anatomía foliar y los ensayos de transcriptómica. La toma de las muestras se realizó cuando la 5^{ta} hoja igualó en longitud a la 4^{ta} hoja. La zona de la 5^{ta} hoja en contacto con la lígula de la hoja 4^{ta} sirvió de referencia para la zona de transición fuente/sumidero. Cada zona fue dividida en 4 fragmentos de igual longitud y rotulados 1 a 8 desde la base a la punta de la hoja. Para los ensayos se utilizaron los fragmentos 1, 3, 5 y 7.

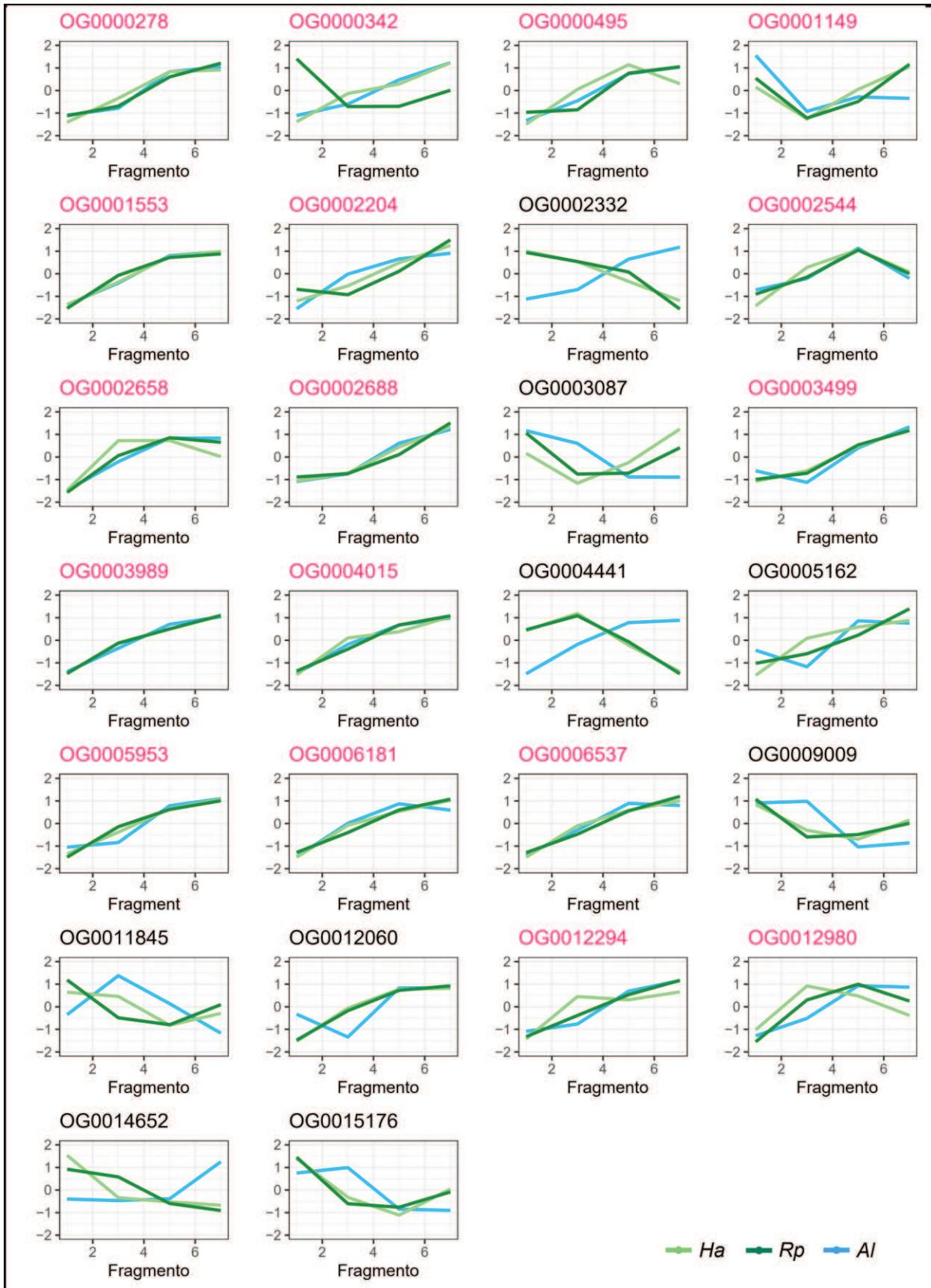


Figura S5. Patrones de expresión de ortogrupos que codifican para enzimas del ciclo C4 durante el desarrollo. Promedios de valores de expresión de los ortogrupos en los fragmentos 1, 3, 5 y 7 para *H. amplexicaulis* (verde claro), *R. pilosa* (verde oscuro) y *A. lanata* (celeste). Lecturas normalizadas dentro de cada especie. Los nombres en rosado indican los OG con patrones de expresión conservados entre la especie C4 y al menos una de las especies C3.

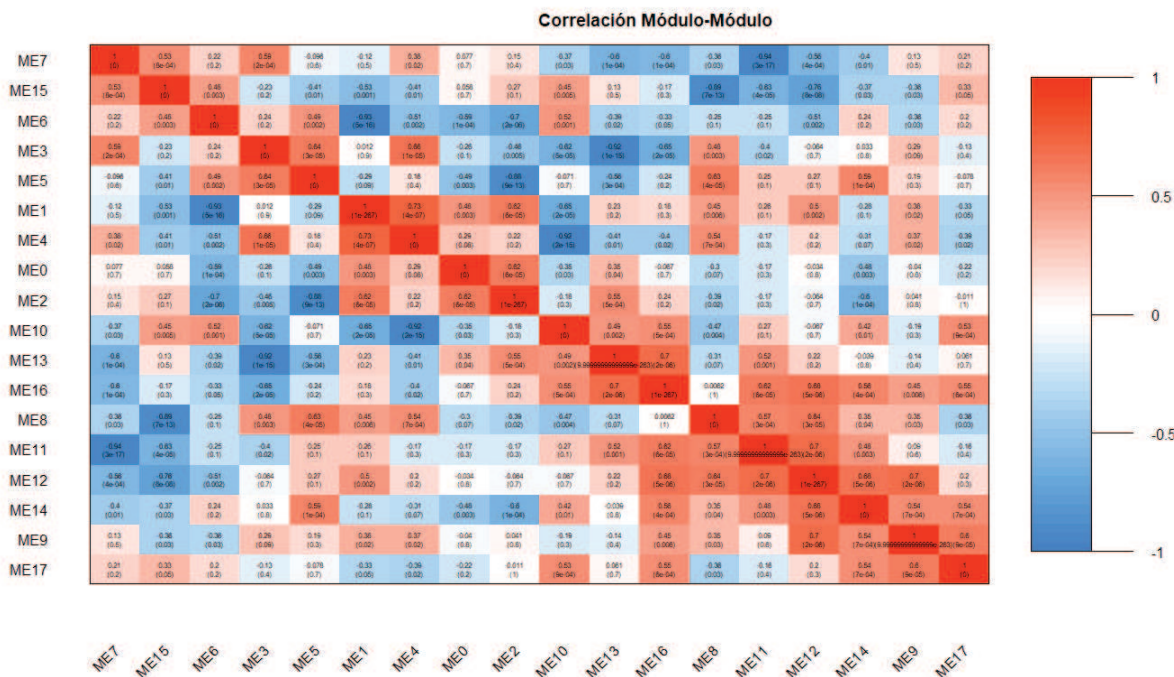


Figura S6. Correlación entre módulos de co-expresión para identificar módulos opuestos. Correlación de Pearson entre los valores EM para cada módulo. En cada rectángulo se muestra el valor de la correlación y su valor p.

Hoja de datos suplementaria 1. Términos de GO con los valores p más bajos para diferencias en el gradiente foliar. Se obtuvieron los 15 términos con valores más bajos para cada categoría: procesos biológicos (BP), función molecular (MF) y componente celular (CC). Sólo se muestran los términos que superaron el valor de corte (valor p < 0,05). Para cada término se indica: Identificador, nombre, N° de términos significativos, N° de términos esperados, valor p y categoría del término, respectivamente.

A. lanata F1-F3 Abajo

- 1,"GO:0006412", "translation", 252,81, 16.57, <1e-30, "BP"
- 2,"GO:0042254", "ribosome biogenesis", 39, 13, 2.56, 1.9e-05, "BP"
- 3,"GO:0032502", "developmental process", 36, 9, 2.37, 4e-04, "BP"
- 4,"GO:0006334", "nucleosome assembly", 17, 6, 1.12, 0.00052, "BP"
- 5,"GO:0007018", "microtubule-based movement", 32, 8, 2.1, 0.00084, "BP"
- 6,"GO:0006261", "DNA-dependent DNA replication", 5, 3, 0.33, 0.00255, "BP"

7,"GO:0006260","DNA replication",35,10,2.3,0.00262,"BP"
8,"GO:0051301","cell division",10,4,0.66,0.00281,"BP"
9,"GO:0009451","RNA modification",31,7,2.04,0.00333,"BP"
10,"GO:0006364","rRNA processing",24,6,1.58,0.00379,"BP"
11,"GO:0006979","response to oxidative stress",77,12,5.06,0.00409,"BP"
12,"GO:0009306","protein secretion",2,2,0.13,0.00431,"BP"
13,"GO:0000070","mitotic sister chromatid segregation",6,3,0.39,0.00486,"BP"
14,"GO:0007049","cell cycle",61,12,4.01,0.00853,"BP"
15,"GO:0006323","DNA packaging",20,8,1.32,0.01206,"BP"
16,"GO:0003735","structural constituent of ribosome",162,74,10.83,NA,"MF"
17,"GO:0008017","microtubule binding",45,12,3.01,2.6e-05,"MF"
18,"GO:0004003","ATP-dependent DNA helicase activity",11,5,0.74,0.00043,"MF"
19,"GO:0003777","microtubule motor activity",32,8,2.14,0.00095,"MF"
20,"GO:0003677","DNA binding",623,61,41.64,0.00133,"MF"
21,"GO:0016884","carbon-nitrogen ligase activity, with gl...",15,5,1,0.00224,"MF"
22,"GO:0003676","nucleic acid binding",1127,117,75.33,0.00323,"MF"
23,"GO:0003723","RNA binding",221,26,14.77,0.00329,"MF"
24,"GO:0004601","peroxidase activity",76,12,5.08,0.00425,"MF"
25,"GO:0005524","ATP binding",927,81,61.96,0.00581,"MF"
26,"GO:0046983","protein dimerization activity",100,14,6.68,0.00639,"MF"
27,"GO:0016762","xyloglucan:xyloglucosyl transferase acti...",19,5,1.27,0.00694,"MF"
28,"GO:0003968","RNA-directed 5'-3' RNA polymerase activi...",4,2,0.27,0.02444,"MF"
29,"GO:0030983","mismatched DNA binding",10,3,0.67,0.02501,"MF"
30,"GO:0019843","rRNA binding",12,3,0.8,0.04151,"MF"
31,"GO:0005840","ribosome",160,75,14.83,NA,"CC"
32,"GO:0000786","nucleosome",23,10,2.13,1.5e-05,"CC"
33,"GO:0032040","small-subunit processome",7,4,0.65,0.002,"CC"
34,"GO:0043229","intracellular organelle",890,146,82.51,0.007,"CC"
35,"GO:0005871","kinesin complex",32,8,2.97,0.0071,"CC"
36,"GO:0000796","condensin complex",2,2,0.19,0.0086,"CC"
37,"GO:0034708","methyltransferase complex",2,2,0.19,0.0086,"CC"
38,"GO:0005788","endoplasmic reticulum lumen",2,2,0.19,0.0086,"CC"
39,"GO:0015935","small ribosomal subunit",2,2,0.19,0.0086,"CC"
40,"GO:0048046","apoplast",19,5,1.76,0.0258,"CC"
41,"GO:0005618","cell wall",34,7,3.15,0.0327,"CC"
42,"GO:0030312","external encapsulating structure",34,7,3.15,0.0327,"CC"

A. lanata F1-F3 Arriba

1,"GO:0015979","photosynthesis",53,39,7.63,1.7e-14,"BP"
2,"GO:0055114","oxidation-reduction process",661,154,95.2,4.3e-11,"BP"
3,"GO:0009765","photosynthesis, light harvesting",13,12,1.87,8.4e-10,"BP"
4,"GO:0005975","carbohydrate metabolic process",339,82,48.83,1e-05,"BP"
5,"GO:0030001","metal ion transport",89,26,12.82,0.00022,"BP"
6,"GO:0006979","response to oxidative stress",77,23,11.09,0.00034,"BP"

7,"GO:0030244","cellulose biosynthetic process",18,8,2.59,0.00205,"BP"
8,"GO:0006032","chitin catabolic process",8,5,1.15,0.00235,"BP"
9,"GO:0016998","cell wall macromolecule catabolic proces...",9,5,1.3,0.00466,"BP"
10,"GO:0051704","multi-organism process",30,10,4.32,0.00711,"BP"
11,"GO:0048544","recognition of pollen",27,9,3.89,0.0105,"BP"
12,"GO:0044706","multi-multicellular organism process",27,9,3.89,0.0105,"BP"
13,"GO:0009856","pollination",27,9,3.89,0.0105,"BP"
14,"GO:0009875","pollen-pistil interaction",27,9,3.89,0.0105,"BP"
15,"GO:0008037","cell recognition",27,9,3.89,0.0105,"BP"
16,"GO:0020037","heme binding",220,76,29.38,3.3e-16,"MF"
17,"GO:0005506","iron ion binding",171,60,22.84,2.2e-13,"MF"
18,"GO:0016705","oxidoreductase activity, acting on paire...",163,55,21.77,1.4e-11,"MF"
19,"GO:0004553","hydrolase activity, hydrolyzing O-glycos...",203,56,27.11,8.9e-07,"MF"
20,"GO:0010333","terpene synthase activity",18,10,2.4,2.7e-05,"MF"
21,"GO:0005507","copper ion binding",25,12,3.34,2.9e-05,"MF"
22,"GO:0004601","peroxidase activity",76,23,10.15,9.1e-05,"MF"
23,"GO:0005509","calcium ion binding",65,20,8.68,2e-04,"MF"
24,"GO:0042132","fructose 1,6-bisphosphate 1-phosphatase ...",4,4,0.53,0.00032,"MF"
25,"GO:0016747","transferase activity, transferring acyl ...",132,32,17.63,0.00046,"MF"
26,"GO:0016760","cellulose synthase (UDP-forming) activit...",16,8,2.14,0.00047,"MF"
27,"GO:0004190","aspartic-type endopeptidase activity",51,16,6.81,0.00067,"MF"
28,"GO:0016491","oxidoreductase activity",783,177,104.57,0.00087,"MF"
29,"GO:0003824","catalytic activity",4132,678,551.83,0.00142,"MF"
30,"GO:0016758","transferase activity, transferring hexos...",284,61,37.93,0.00158,"MF"
31,"GO:0016020","membrane",1129,191,125.5,1.3e-15,"CC"
32,"GO:0009654","photosystem II oxygen evolving complex",20,14,2.22,6.6e-10,"CC"
33,"GO:0019898","extrinsic component of membrane",21,13,2.33,2.6e-08,"CC"
34,"GO:0009538","photosystem I reaction center",5,5,0.56,1.6e-05,"CC"
35,"GO:0009522","photosystem I",10,9,1.11,0.00063,"CC"
36,"GO:0042651","thylakoid membrane",23,17,2.56,0.00117,"CC"
37,"GO:0044436","thylakoid part",35,27,3.89,0.00385,"CC"
38,"GO:0034357","photosynthetic membrane",34,26,3.78,0.02887,"CC"
39,"GO:0009521","photosystem",33,25,3.67,0.02909,"CC"
40,"GO:0009523","photosystem II",23,16,2.56,0.03112,"CC"

A. *Ianata* F3-F5 Abajo

1,"GO:0007018","microtubule-based movement",32,18,2.32,3.9e-13,"BP"
2,"GO:0006979","response to oxidative stress",77,16,5.59,9.8e-05,"BP"
3,"GO:0000226","microtubule cytoskeleton organization",23,10,1.67,9.9e-05,"BP"
4,"GO:0007010","cytoskeleton organization",57,21,4.14,0.00017,"BP"
5,"GO:0042545","cell wall modification",14,6,1.02,0.00026,"BP"
6,"GO:0006891","intra-Golgi vesicle-mediated transport",3,3,0.22,0.00038,"BP"
7,"GO:0016192","vesicle-mediated transport",112,21,8.13,0.00069,"BP"
8,"GO:0006468","protein phosphorylation",711,73,51.58,0.00095,"BP"

9,"GO:0032012","regulation of ARF protein signal transdu...",5,3,0.36,0.00339,"BP"
10,"GO:0006694","steroid biosynthetic process",23,6,1.67,0.00491,"BP"
11,"GO:0051225","spindle assembly",6,3,0.44,0.00643,"BP"
12,"GO:0030036","actin cytoskeleton organization",25,6,1.81,0.00761,"BP"
13,"GO:0000910","cytokinesis",9,3,0.65,0.0229,"BP"
14,"GO:0051301","cell division",10,3,0.73,0.03099,"BP"
15,"GO:0033036","macromolecule localization",156,18,11.32,0.03255,"BP"
16,"GO:0008017","microtubule binding",45,24,3.28,2.8e-16,"MF"
17,"GO:0003777","microtubule motor activity",32,18,2.33,4.6e-13,"MF"
18,"GO:0005507","copper ion binding",25,11,1.82,4.8e-07,"MF"
19,"GO:0005524","ATP binding",927,98,67.54,5.8e-05,"MF"
20,"GO:0030599","pectinesterase activity",14,6,1.02,0.00026,"MF"
21,"GO:0004601","peroxidase activity",76,15,5.54,0.00032,"MF"
22,"GO:0004672","protein kinase activity",706,73,51.44,0.00109,"MF"
23,"GO:0003854","3-beta-hydroxy-delta5-steroid dehydrogen...",20,6,1.46,0.00233,"MF"
24,"GO:0003774","motor activity",41,22,2.99,0.00235,"MF"
25,"GO:0004857","enzyme inhibitor activity",34,8,2.48,0.00252,"MF"
26,"GO:0005086","ARF guanyl-nucleotide exchange factor ac...",5,3,0.36,0.00344,"MF"
27,"GO:0022804","active transmembrane transporter activit...",99,15,7.21,0.00507,"MF"
28,"GO:0005096","GTPase activator activity",18,5,1.31,0.00779,"MF"
29,"GO:0008194","UDP-glycosyltransferase activity",96,14,6.99,0.00931,"MF"
30,"GO:0016887","ATPase activity",146,19,10.64,0.00939,"MF"
31,"GO:0005871","kinesin complex",32,18,2.3,2.1e-13,"CC"
32,"GO:0030126","COPI vesicle coat",7,4,0.5,0.00076,"CC"
33,"GO:0017119","Golgi transport complex",4,3,0.29,0.00139,"CC"
34,"GO:0016459","myosin complex",9,4,0.65,0.00245,"CC"
35,"GO:0005856","cytoskeleton",90,33,6.48,0.00591,"CC"
36,"GO:0005618","cell wall",34,7,2.45,0.00901,"CC"
37,"GO:0015630","microtubule cytoskeleton",61,24,4.39,0.00916,"CC"
38,"GO:0030118","clathrin coat",7,3,0.5,0.01034,"CC"
39,"GO:0005815","microtubule organizing center",7,3,0.5,0.01034,"CC"
40,"GO:0005819","spindle",7,3,0.5,0.01034,"CC"
41,"GO:0044430","cytoskeletal part",81,29,5.83,0.01115,"CC"
42,"GO:0048475","coated membrane",30,9,2.16,0.01937,"CC"
43,"GO:0030117","membrane coat",30,9,2.16,0.01937,"CC"
44,"GO:0005875","microtubule associated complex",42,21,3.02,0.02286,"CC"
45,"GO:0070652","HAUS complex",4,2,0.29,0.02807,"CC"

A. lanata F3-F5 Arriba

1,"GO:0015979","photosynthesis",53,37,6,7.3e-15,"BP"
2,"GO:0009765","photosynthesis, light harvesting",13,12,1.47,4.7e-11,"BP"
3,"GO:0055114","oxidation-reduction process",661,128,74.77,6.1e-11,"BP"
4,"GO:0000272","polysaccharide catabolic process",8,6,0.9,4.7e-05,"BP"
5,"GO:0006855","drug transmembrane transport",27,11,3.05,8.4e-05,"BP"

6,"GO:0006873","cellular ion homeostasis",5,4,0.57,0.00074,"BP"
7,"GO:0003333","amino acid transmembrane transport",25,8,2.83,0.0048,"BP"
8,"GO:0006535","cysteine biosynthetic process from serin...",4,3,0.45,0.00528,"BP"
9,"GO:0009733","response to auxin",16,6,1.81,0.00604,"BP"
10,"GO:0055085","transmembrane transport",363,68,41.06,0.00605,"BP"
11,"GO:0009605","response to external stimulus",13,5,1.47,0.0108,"BP"
12,"GO:0051704","multi-organism process",30,8,3.39,0.01561,"BP"
13,"GO:0030091","protein repair",6,3,0.68,0.02216,"BP"
14,"GO:0015985","energy coupled proton transport, down el...",11,4,1.24,0.02793,"BP"
15,"GO:0015986","ATP synthesis coupled proton transport",11,4,1.24,0.02793,"BP"
16,"GO:0016705","oxidoreductase activity, acting on paire...",163,57,15.59,2.6e-19,"MF"
17,"GO:0005506","iron ion binding",171,58,16.35,6.7e-19,"MF"
18,"GO:0020037","heme binding",220,65,21.04,1.6e-17,"MF"
19,"GO:0016758","transferase activity, transferring hexos...",284,57,27.16,3.3e-08,"MF"
20,"GO:0003700","DNA-binding transcription factor activit...",288,53,27.54,1.8e-06,"MF"
21,"GO:0016161","beta-amylase activity",8,6,0.76,1.8e-05,"MF"
22,"GO:0015238","drug transmembrane transporter activity",30,11,2.87,5.6e-05,"MF"
23,"GO:0042132","fructose 1,6-bisphosphate 1-phosphatase ...",4,4,0.38,8.3e-05,"MF"
24,"GO:0016491","oxidoreductase activity",783,144,74.87,0.00014,"MF"
25,"GO:0015297","antiporter activity",40,12,3.82,0.00024,"MF"
26,"GO:0005215","transporter activity",411,79,39.3,0.00024,"MF"
27,"GO:0008171","O-methyltransferase activity",27,9,2.58,0.00061,"MF"
28,"GO:0005216","ion channel activity",33,10,3.16,0.00072,"MF"
29,"GO:0016655","oxidoreductase activity, acting on NAD(P...)",7,4,0.67,0.00229,"MF"
30,"GO:0022839","ion gated channel activity",21,7,2.01,0.00249,"MF"
31,"GO:0016020","membrane",1129,157,105.68,7.7e-11,"CC"
32,"GO:0009654","photosystem II oxygen evolving complex",20,10,1.87,3.4e-06,"CC"
33,"GO:0009538","photosystem I reaction center",5,5,0.47,6.9e-06,"CC"
34,"GO:0019898","extrinsic component of membrane",21,9,1.97,5e-05,"CC"
35,"GO:0009522","photosystem I",10,9,0.94,0.00032,"CC"
36,"GO:0009523","photosystem II",23,13,2.15,0.00072,"CC"
37,"GO:0042651","thylakoid membrane",23,13,2.15,0.00072,"CC"
38,"GO:0031012","extracellular matrix",2,2,0.19,0.00872,"CC"
39,"GO:0005576","extracellular region",30,8,2.81,0.03939,"CC"

A. lanata F5-F7 Abajo

1,"GO:0055114","oxidation-reduction process",661,91,45.59,1.3e-11,"BP"
2,"GO:0006979","response to oxidative stress",77,21,5.31,2.6e-08,"BP"
3,"GO:0030244","cellulose biosynthetic process",18,10,1.24,5.8e-08,"BP"
4,"GO:0006073","cellular glucan metabolic process",48,21,3.31,1.9e-06,"BP"
5,"GO:0007018","microtubule-based movement",32,11,2.21,5e-06,"BP"
6,"GO:0005975","carbohydrate metabolic process",339,59,23.38,2.8e-05,"BP"
7,"GO:0042545","cell wall modification",14,5,0.97,0.0018,"BP"
8,"GO:0000910","cytokinesis",9,4,0.62,0.0021,"BP"

9,"GO:0000226","microtubule cytoskeleton organization",23,6,1.59,0.0038,"BP"
10,"GO:0010338","leaf formation",2,2,0.14,0.0047,"BP"
11,"GO:0050826","response to freezing",2,2,0.14,0.0047,"BP"
12,"GO:0009832","plant-type cell wall biogenesis",6,3,0.41,0.0056,"BP"
13,"GO:0006284","base-excision repair",14,4,0.97,0.0128,"BP"
14,"GO:0006555","methionine metabolic process",3,2,0.21,0.0136,"BP"
15,"GO:0044036","cell wall macromolecule metabolic proces...",10,3,0.69,0.0271,"BP"
16,"GO:0020037","heme binding",220,48,15.45,6.2e-13,"MF"
17,"GO:0004553","hydrolase activity, hydrolyzing O-glycos...",203,43,14.26,3e-11,"MF"
18,"GO:0008017","microtubule binding",45,17,3.16,3.3e-09,"MF"
19,"GO:0004601","peroxidase activity",76,22,5.34,5.6e-09,"MF"
20,"GO:0016762","xyloglucan:xyloglucosyl transferase acti...",19,11,1.33,8.4e-09,"MF"
21,"GO:0016760","cellulose synthase (UDP-forming) activit...",16,10,1.12,1.5e-08,"MF"
22,"GO:0005506","iron ion binding",171,30,12.01,2.3e-06,"MF"
23,"GO:0003777","microtubule motor activity",32,11,2.25,6.2e-06,"MF"
24,"GO:0016705","oxidoreductase activity, acting on paire...",163,27,11.45,2.2e-05,"MF"
25,"GO:0016747","transferase activity, transferring acyl ...",132,23,9.27,4e-05,"MF"
26,"GO:0008725","DNA-3-methyladenine glycosylase activity",5,4,0.35,0.00011,"MF"
27,"GO:0005507","copper ion binding",25,8,1.76,0.00021,"MF"
28,"GO:0008171","O-methyltransferase activity",27,8,1.9,0.00038,"MF"
29,"GO:0010333","terpene synthase activity",18,6,1.26,0.00104,"MF"
30,"GO:0003824","catalytic activity",4132,395,290.24,0.00126,"MF"
31,"GO:0005618","cell wall",34,16,1.71,6.2e-13,"CC"
32,"GO:0048046","apoplast",19,11,0.96,1.8e-10,"CC"
33,"GO:0005871","kinesin complex",32,11,1.61,1.7e-07,"CC"
34,"GO:0005874","microtubule",14,4,0.71,0.0041,"CC"
35,"GO:0031225","anchored component of membrane",7,2,0.35,0.0448,"CC"
36,"GO:0008352","katanin complex",1,1,0.05,0.0504,"CC"
37,"GO:0005885","Arp2/3 protein complex",8,2,0.4,0.0578,"CC"
38,"GO:0044450","microtubule organizing center part",2,1,0.1,0.0983,"CC"
39,"GO:0009331","glycerol-3-phosphate dehydrogenase compl...",4,1,0.2,0.187,"CC"
40,"GO:0044430","cytoskeletal part",81,18,4.08,0.2027,"CC"
41,"GO:0005815","microtubule organizing center",7,1,0.35,0.3041,"CC"
42,"GO:0005819","spindle",7,1,0.35,0.3041,"CC"
43,"GO:0005856","cytoskeleton",90,18,4.54,0.3123,"CC"
44,"GO:0015629","actin cytoskeleton",25,2,1.26,0.3618,"CC"
45,"GO:0005875","microtubule associated complex",42,12,2.12,0.377,"CC"

A. lanata F5-F7 Arriba

1,"GO:0030418","nicotianamine biosynthetic process",2,2,0.03,0.00029,"BP"
2,"GO:0035434","copper ion transmembrane transport",3,2,0.05,0.00086,"BP"
3,"GO:0055114","oxidation-reduction process",661,19,11.34,0.01551,"BP"
4,"GO:0062012","regulation of small molecule metabolic p...",1,1,0.02,0.01716,"BP"
5,"GO:0062014","negative regulation of small molecule me...",1,1,0.02,0.01716,"BP"

6,"GO:0045980","negative regulation of nucleotide metabo...",1,1,0.02,0.01716,"BP"
7,"GO:0046493","lipid A metabolic process",1,1,0.02,0.01716,"BP"
8,"GO:0006140","regulation of nucleotide metabolic proce...",1,1,0.02,0.01716,"BP"
9,"GO:0009245","lipid A biosynthetic process",1,1,0.02,0.01716,"BP"
10,"GO:1901269","lipooligosaccharide metabolic process",1,1,0.02,0.01716,"BP"
11,"GO:0045936","negative regulation of phosphate metabol...",1,1,0.02,0.01716,"BP"
12,"GO:1901271","lipooligosaccharide biosynthetic process",1,1,0.02,0.01716,"BP"
13,"GO:0010563","negative regulation of phosphorus metabo...",1,1,0.02,0.01716,"BP"
14,"GO:0009733","response to auxin",16,2,0.27,0.02989,"BP"
15,"GO:0031497","chromatin assembly",17,2,0.29,0.03351,"BP"
16,"GO:0005506","iron ion binding",171,13,2.55,1.4e-06,"MF"
17,"GO:0016705","oxidoreductase activity, acting on paire...",163,12,2.43,5e-06,"MF"
18,"GO:0020037","heme binding",220,13,3.28,2.2e-05,"MF"
19,"GO:0030410","nicotianamine synthase activity",2,2,0.03,0.00022,"MF"
20,"GO:0005375","copper ion transmembrane transporter act...",3,2,0.04,0.00065,"MF"
21,"GO:0016758","transferase activity, transferring hexos...",284,11,4.23,0.00329,"MF"
22,"GO:0004857","enzyme inhibitor activity",34,3,0.51,0.01381,"MF"
23,"GO:0008759","UDP-3-O-[3-hydroxymyristoyl] N-acetylglu...",1,1,0.01,0.01491,"MF"
24,"GO:0019213","deacetylase activity",1,1,0.01,0.01491,"MF"
25,"GO:0015035","protein disulfide oxidoreductase activit...",44,3,0.66,0.02747,"MF"
26,"GO:0015036","disulfide oxidoreductase activity",44,3,0.66,0.02747,"MF"
27,"GO:0005345","purine nucleobase transmembrane transpor...",2,1,0.03,0.0296,"MF"
28,"GO:0008519","ammonium transmembrane transporter activ...",2,1,0.03,0.0296,"MF"
29,"GO:0015205","nucleobase transmembrane transporter act...",2,1,0.03,0.0296,"MF"
30,"GO:0030246","carbohydrate binding",80,4,1.19,0.03119,"MF"
31,"GO:0016020","membrane",1129,19,13.21,0.017,"CC"
32,"GO:0048046","apoplast",19,2,0.22,0.02,"CC"
33,"GO:0000786","nucleosome",23,2,0.27,0.029,"CC"
34,"GO:0032993","protein-DNA complex",23,2,0.27,0.029,"CC"
35,"GO:0044815","DNA packaging complex",25,2,0.29,0.033,"CC"
36,"GO:0005576","extracellular region",30,2,0.35,0.047,"CC"
37,"GO:0016021","integral component of membrane",501,10,5.86,0.049,"CC"

R. pilosa F1-F3 Abajo

1,"GO:0006260","DNA replication",45,28,2.96,2.7e-15,"BP"
2,"GO:0007018","microtubule-based movement",45,16,2.96,9.9e-09,"BP"
3,"GO:0032502","developmental process",40,12,2.63,5.7e-06,"BP"
4,"GO:0007093","mitotic cell cycle checkpoint",6,4,0.39,0.00025,"BP"
5,"GO:0030071","regulation of mitotic metaphase/anaphase...",6,4,0.39,0.00025,"BP"
6,"GO:0000077","DNA damage checkpoint",3,3,0.2,0.00028,"BP"
7,"GO:0006270","DNA replication initiation",4,3,0.26,0.00107,"BP"
8,"GO:0007076","mitotic chromosome condensation",4,3,0.26,0.00107,"BP"
9,"GO:1901991","negative regulation of mitotic cell cycl...",5,3,0.33,0.00255,"BP"

10, "GO:0006334", "nucleosome assembly", 16,5,1.05,0.00285, "BP"
 11, "GO:0051301", "cell division", 11,4,0.72,0.00418, "BP"
 12, "GO:0006261", "DNA-dependent DNA replication", 8,7,0.53,0.00421, "BP"
 13, "GO:0022402", "cell cycle process", 37,15,2.43,0.00424, "BP"
 14, "GO:0006269", "DNA replication, synthesis of RNA primer", 2,2,0.13,0.00431, "BP"
 15, "GO:0006275", "regulation of DNA replication", 2,2,0.13,0.00431, "BP"
 16, "GO:0008017", "microtubule binding", 55,18,3.56,4.5e-09, "MF"
 17, "GO:0003777", "microtubule motor activity", 45,16,2.91,8.4e-09, "MF"
 18, "GO:0003677", "DNA binding", 673,71,43.54,1.9e-05, "MF"
 19, "GO:0003682", "chromatin binding", 14,6,0.91,0.00014, "MF"
 20, "GO:0005524", "ATP binding", 962,90,62.23,0.00014, "MF"
 21, "GO:0004553", "hydrolase activity, hydrolyzing O-glycos...", 224,28,14.49,0.00056, "MF"
 22, "GO:0008408", "3'-5' exonuclease activity", 15,5,0.97,0.00194, "MF"
 23, "GO:0005515", "protein binding", 1657,145,107.19,0.00199, "MF"
 24, "GO:0045300", "acyl-[acyl-carrier-protein] desaturase a...", 5,3,0.32,0.00244, "MF"
 25, "GO:0003887", "DNA-directed DNA polymerase activity", 11,4,0.71,0.00396, "MF"
 26, "GO:0004748", "ribonucleoside-diphosphate reductase act...", 2,2,0.13,0.00418, "MF"
 27, "GO:0017056", "structural constituent of nuclear pore", 2,2,0.13,0.00418, "MF"
 28, "GO:0098772", "molecular function regulator", 89,13,5.76,0.00451, "MF"
 29, "GO:0140097", "catalytic activity, acting on DNA", 72,14,4.66,0.00511, "MF"
 30, "GO:0016762", "xyloglucan:xyloglucosyl transferase acti...", 19,5,1.23,0.00605, "MF"
 31, "GO:0005871", "kinesin complex", 45,16,3.23,2.6e-08, "CC"
 32, "GO:0005634", "nucleus", 446,63,32.05,8e-07, "CC"
 33, "GO:0000786", "nucleosome", 18,9,1.29,1.1e-06, "CC"
 34, "GO:0030896", "checkpoint clamp complex", 3,3,0.22,0.00036, "CC"
 35, "GO:0005618", "cell wall", 35,8,2.51,0.00264, "CC"
 36, "GO:0044427", "chromosomal part", 46,20,3.31,0.00308, "CC"
 37, "GO:0005680", "anaphase-promoting complex", 5,3,0.36,0.00327, "CC"
 38, "GO:0000796", "condensin complex", 2,2,0.14,0.00513, "CC"
 39, "GO:0042555", "MCM complex", 2,2,0.14,0.00513, "CC"
 40, "GO:0044452", "nucleolar part", 2,2,0.14,0.00513, "CC"
 41, "GO:0048046", "apoplast", 19,5,1.37,0.00917, "CC"
 42, "GO:0031234", "extrinsic component of cytoplasmic side ...", 3,2,0.22,0.01467, "CC"
 43, "GO:0005834", "heterotrimeric G-protein complex", 3,2,0.22,0.01467, "CC"
 44, "GO:0098562", "cytoplasmic side of membrane", 3,2,0.22,0.01467, "CC"
 45, "GO:1905360", "GTPase complex", 3,2,0.22,0.01467, "CC"

R. pilosa F1-F3 Arriba

1, "GO:0055114", "oxidation-reduction process", 754,218,131.43,3.7e-17, "BP"
 2, "GO:0015979", "photosynthesis", 48,37,8.37,8.7e-12, "BP"
 3, "GO:0009765", "photosynthesis, light harvesting", 14,13,2.44,1.5e-09, "BP"
 4, "GO:0006979", "response to oxidative stress", 93,33,16.21,2e-05, "BP"
 5, "GO:0009832", "plant-type cell wall biogenesis", 7,5,1.22,0.0025, "BP"

6,"GO:0042737","drug catabolic process",13,7,2.27,0.0031,"BP"
7,"GO:0048544","recognition of pollen",27,11,4.71,0.0037,"BP"
8,"GO:0006633","fatty acid biosynthetic process",45,15,7.84,0.0072,"BP"
9,"GO:0044264","cellular polysaccharide metabolic proces...",46,15,8.02,0.009,"BP"
10,"GO:0044036","cell wall macromolecule metabolic proces...",12,6,2.09,0.0098,"BP"
11,"GO:0010215","cellulose microfibril organization",6,4,1.05,0.0102,"BP"
12,"GO:0009664","plant-type cell wall organization",6,4,1.05,0.0102,"BP"
13,"GO:0071668","plant-type cell wall assembly",6,4,1.05,0.0102,"BP"
14,"GO:0070726","cell wall assembly",6,4,1.05,0.0102,"BP"
15,"GO:0030198","extracellular matrix organization",6,4,1.05,0.0102,"BP"
16,"GO:0020037","heme binding",258,115,40.78,7.3e-29,"MF"
17,"GO:0005506","iron ion binding",196,88,30.98,1.6e-22,"MF"
18,"GO:0016705","oxidoreductase activity, acting on paire...",193,87,30.51,2e-22,"MF"
19,"GO:0016758","transferase activity, transferring hexos...",312,83,49.32,4.6e-07,"MF"
20,"GO:0004601","peroxidase activity",92,34,14.54,5.8e-07,"MF"
21,"GO:0016491","oxidoreductase activity",898,252,141.95,4e-06,"MF"
22,"GO:0016747","transferase activity, transferring acyl ...",153,48,24.19,1.4e-05,"MF"
23,"GO:0004553","hydrolase activity, hydrolyzing O-glycos...",224,58,35.41,5.8e-05,"MF"
24,"GO:0003824","catalytic activity",4498,881,711.02,0.00033,"MF"
25,"GO:0005507","copper ion binding",34,14,5.37,0.00035,"MF"
26,"GO:0004144","diacylglycerol O-acyltransferase activit...",4,4,0.63,0.00062,"MF"
27,"GO:0004185","serine-type carboxypeptidase activity",37,14,5.85,0.00095,"MF"
28,"GO:0010333","terpene synthase activity",22,10,3.48,0.00098,"MF"
29,"GO:0030247","polysaccharide binding",26,11,4.11,0.00114,"MF"
30,"GO:0008171","O-methyltransferase activity",23,10,3.64,0.00149,"MF"
31,"GO:0016020","membrane",1185,208,137.04,2.7e-17,"CC"
32,"GO:0009654","photosystem II oxygen evolving complex",12,8,1.39,9.5e-06,"CC"
33,"GO:0009538","photosystem I reaction center",5,5,0.58,2e-05,"CC"
34,"GO:0019898","extrinsic component of membrane",14,7,1.62,0.00043,"CC"
35,"GO:0009522","photosystem I",10,9,1.16,0.00075,"CC"
36,"GO:0042651","thylakoid membrane",15,11,1.73,0.00142,"CC"
37,"GO:0048046","apoplast",19,7,2.2,0.00377,"CC"
38,"GO:0031225","anchored component of membrane",8,4,0.93,0.00839,"CC"
39,"GO:0009579","thylakoid",29,21,3.35,0.01996,"CC"
40,"GO:0044436","thylakoid part",29,21,3.35,0.01996,"CC"
41,"GO:0009523","photosystem II",15,10,1.73,0.03521,"CC"
42,"GO:0031012","extracellular matrix",3,2,0.35,0.03693,"CC"
43,"GO:0005618","cell wall",35,8,4.05,0.04182,"CC"
44,"GO:0030312","external encapsulating structure",35,8,4.05,0.04182,"CC"

R. pilosa F3-F5 Abajo

1,"GO:0006979","response to oxidative stress",93,17,3.37,2.6e-08,"BP"
2,"GO:0055114","oxidation-reduction process",754,51,27.33,4.7e-06,"BP"

3,"GO:0009733","response to auxin",24,5,0.87,0.0014,"BP"
4,"GO:0030244","cellulose biosynthetic process",16,4,0.58,0.0022,"BP"
5,"GO:0006633","fatty acid biosynthetic process",45,6,1.63,0.0053,"BP"
6,"GO:0071555","cell wall organization",21,4,0.76,0.0062,"BP"
7,"GO:0005975","carbohydrate metabolic process",360,25,13.05,0.0107,"BP"
8,"GO:0042546","cell wall biogenesis",15,3,0.54,0.0155,"BP"
9,"GO:0032501","multicellular organismal process",57,6,2.07,0.0165,"BP"
10,"GO:0010215","cellulose microfibril organization",6,2,0.22,0.0178,"BP"
11,"GO:0071668","plant-type cell wall assembly",6,2,0.22,0.0178,"BP"
12,"GO:0070726","cell wall assembly",6,2,0.22,0.0178,"BP"
13,"GO:0043062","extracellular structure organization",6,2,0.22,0.0178,"BP"
14,"GO:0009664","plant-type cell wall organization",6,2,0.22,0.0178,"BP"
15,"GO:0030198","extracellular matrix organization",6,2,0.22,0.0178,"BP"
16,"GO:0020037","heme binding",258,30,9.51,1.8e-08,"MF"
17,"GO:0004601","peroxidase activity",92,17,3.39,3.1e-08,"MF"
18,"GO:0080019","fatty-acyl-CoA reductase (alcohol-formin...",4,3,0.15,0.00019,"MF"
19,"GO:0004553","hydrolase activity, hydrolyzing O-glycos...",224,19,8.25,0.00059,"MF"
20,"GO:0016747","transferase activity, transferring acyl ...",153,16,5.64,0.00109,"MF"
21,"GO:0016788","hydrolase activity, acting on ester bond...",311,22,11.46,0.00248,"MF"
22,"GO:0005506","iron ion binding",196,15,7.22,0.00578,"MF"
23,"GO:0005507","copper ion binding",34,5,1.25,0.0076,"MF"
24,"GO:0003885","D-arabinono-1,4-lactone oxidase activity",4,2,0.15,0.00773,"MF"
25,"GO:0004144","diacylglycerol O-acyltransferase activit...",4,2,0.15,0.00773,"MF"
26,"GO:0003824","catalytic activity",4498,209,165.72,0.00795,"MF"
27,"GO:0050660","flavin adenine dinucleotide binding",50,6,1.84,0.00967,"MF"
28,"GO:0016491","oxidoreductase activity",898,62,33.09,0.01007,"MF"
29,"GO:0016705","oxidoreductase activity, acting on paire...",193,14,7.11,0.01181,"MF"
30,"GO:0016760","cellulose synthase (UDP-forming) activit...",14,3,0.52,0.01333,"MF"
31,"GO:0016020","membrane",1185,46,31.24,9.2e-05,"CC"
32,"GO:0030312","external encapsulating structure",35,4,0.92,0.012,"CC"
33,"GO:0005618","cell wall",35,4,0.92,0.012,"CC"
34,"GO:0031225","anchored component of membrane",8,2,0.21,0.017,"CC"

R. pilosa F3-F5 Arriba

1,"GO:0055114","oxidation-reduction process",754,45,24.21,1.9e-05,"BP"
2,"GO:0005975","carbohydrate metabolic process",360,22,11.56,0.0025,"BP"
3,"GO:0051704","multi-organism process",31,5,1,0.0028,"BP"
4,"GO:0042737","drug catabolic process",13,3,0.42,0.0074,"BP"
5,"GO:0009875","pollen-pistil interaction",27,4,0.87,0.0101,"BP"
6,"GO:0008037","cell recognition",27,4,0.87,0.0101,"BP"
7,"GO:0048544","recognition of pollen",27,4,0.87,0.0101,"BP"
8,"GO:0044706","multi-multicellular organism process",27,4,0.87,0.0101,"BP"
9,"GO:0009856","pollination",27,4,0.87,0.0101,"BP"
10,"GO:0000003","reproduction",31,4,1,0.0165,"BP"

11, "GO:0022414", "reproductive process", 31, 4, 1, 0.0165, "BP"
 12, "GO:0019318", "hexose metabolic process", 18, 3, 0.58, 0.0186, "BP"
 13, "GO:0009069", "serine family amino acid metabolic proce...", 7, 2, 0.22, 0.0194, "BP"
 14, "GO:0016052", "carbohydrate catabolic process", 49, 5, 1.57, 0.0197, "BP"
 15, "GO:0044282", "small molecule catabolic process", 20, 3, 0.64, 0.0249, "BP"
 16, "GO:0020037", "heme binding", 258, 27, 7.4, 4.5e-09, "MF"
 17, "GO:0016705", "oxidoreductase activity, acting on paire...", 193, 22, 5.54, 2.8e-08, "MF"
 18, "GO:0005506", "iron ion binding", 196, 22, 5.62, 3.7e-08, "MF"
 19, "GO:0016758", "transferase activity, transferring hexos...", 312, 21, 8.95, 0.00024, "MF"
 20, "GO:0008171", "O-methyltransferase activity", 23, 5, 0.66, 0.00041, "MF"
 21, "GO:0004553", "hydrolase activity, hydrolyzing O-glycos...", 224, 16, 6.43, 7e-04, "MF"
 22, "GO:0010333", "terpene synthase activity", 22, 4, 0.63, 0.00322, "MF"
 23, "GO:0003824", "catalytic activity", 4498, 170, 129.08, 0.00328, "MF"
 24, "GO:0022857", "transmembrane transporter activity", 323, 18, 9.27, 0.0053, "MF"
 25, "GO:0008061", "chitin binding", 5, 2, 0.14, 0.00774, "MF"
 26, "GO:0033897", "ribonuclease T2 activity", 5, 2, 0.14, 0.00774, "MF"
 27, "GO:0030246", "carbohydrate binding", 74, 6, 2.12, 0.01914, "MF"
 28, "GO:0004611", "phosphoenolpyruvate carboxykinase activi...", 8, 2, 0.23, 0.02048, "MF"
 29, "GO:0016161", "beta-amylase activity", 8, 2, 0.23, 0.02048, "MF"
 30, "GO:0004185", "serine-type carboxypeptidase activity", 37, 4, 1.06, 0.02079, "MF"
 31, "GO:0016020", "membrane", 1185, 31, 21.16, 0.0016, "CC"
 32, "GO:0016021", "integral component of membrane", 525, 16, 9.38, 0.0144, "CC"
 33, "GO:0031224", "intrinsic component of membrane", 533, 16, 9.52, 0.0166, "CC"
 34, "GO:0005576", "extracellular region", 32, 3, 0.57, 0.0183, "CC"
 35, "GO:0005618", "cell wall", 35, 3, 0.62, 0.0233, "CC"
 36, "GO:0030312", "external encapsulating structure", 35, 3, 0.62, 0.0233, "CC"
 37, "GO:0042555", "MCM complex", 2, 1, 0.04, 0.0354, "CC"
 38, "GO:0048046", "apoplast", 19, 2, 0.34, 0.0439, "CC"

R. pilosa F5-F7 Abajo

1, "GO:0055114", "oxidation-reduction process", 754, 43, 15.73, 1.7e-10, "BP"
 2, "GO:0006979", "response to oxidative stress", 93, 12, 1.94, 4e-07, "BP"
 3, "GO:0005975", "carbohydrate metabolic process", 360, 21, 7.51, 0.00019, "BP"
 4, "GO:0030244", "cellulose biosynthetic process", 16, 3, 0.33, 0.00407, "BP"
 5, "GO:0008152", "metabolic process", 4678, 117, 97.57, 0.01127, "BP"
 6, "GO:1905392", "plant organ morphogenesis", 1, 1, 0.02, 0.02086, "BP"
 7, "GO:1905393", "plant organ formation", 1, 1, 0.02, 0.02086, "BP"
 8, "GO:0070589", "cellular component macromolecule biosynt...", 1, 1, 0.02, 0.02086, "BP"
 9, "GO:0009834", "plant-type secondary cell wall biogenesi...", 1, 1, 0.02, 0.02086, "BP"
 10, "GO:0009965", "leaf morphogenesis", 1, 1, 0.02, 0.02086, "BP"
 11, "GO:0044038", "cell wall macromolecule biosynthetic pro...", 1, 1, 0.02, 0.02086, "BP"
 12, "GO:0010383", "cell wall polysaccharide metabolic proce...", 1, 1, 0.02, 0.02086, "BP"
 13, "GO:0010410", "hemicellulose metabolic process", 1, 1, 0.02, 0.02086, "BP"
 14, "GO:0010338", "leaf formation", 1, 1, 0.02, 0.02086, "BP"

15, "GO:0009653", "anatomical structure morphogenesis", 1, 1, 0.02, 0.02086, "BP"
 16, "GO:0020037", "heme binding", 258, 29, 5.27, 3.3e-14, "MF"
 17, "GO:0005506", "iron ion binding", 196, 18, 4, 8.5e-08, "MF"
 18, "GO:0004601", "peroxidase activity", 92, 12, 1.88, 3.1e-07, "MF"
 19, "GO:0016705", "oxidoreductase activity, acting on paire...", 193, 17, 3.94, 3.6e-07, "MF"
 20, "GO:0003824", "catalytic activity", 4498, 135, 91.89, 0.00021, "MF"
 21, "GO:0005507", "copper ion binding", 34, 5, 0.69, 0.00058, "MF"
 22, "GO:0004553", "hydrolase activity, hydrolyzing O-glycos...", 224, 13, 4.58, 0.00065, "MF"
 23, "GO:0008146", "sulfotransferase activity", 14, 3, 0.29, 0.00258, "MF"
 24, "GO:0016760", "cellulose synthase (UDP-forming) activit...", 14, 3, 0.29, 0.00258, "MF"
 25, "GO:0016491", "oxidoreductase activity", 898, 49, 18.34, 0.00481, "MF"
 26, "GO:0004190", "aspartic-type endopeptidase activity", 59, 5, 1.21, 0.0069, "MF"
 27, "GO:0016157", "sucrose synthase activity", 9, 2, 0.18, 0.01359, "MF"
 28, "GO:0047134", "protein-disulfide reductase activity", 1, 1, 0.02, 0.02043, "MF"
 29, "GO:0016668", "oxidoreductase activity, acting on a sul...", 1, 1, 0.02, 0.02043, "MF"
 30, "GO:0005092", "GDP-dissociation inhibitor activity", 1, 1, 0.02, 0.02043, "MF"
 31, "GO:0005618", "cell wall", 35, 3, 0.37, 0.0055, "CC"
 32, "GO:0016020", "membrane", 1185, 18, 12.6, 0.0231, "CC"

R. pilosa F5-F7 Arriba

1, "GO:0006811", "ion transport", 239, 9, 2.89, 0.0022, "BP"
 2, "GO:0055085", "transmembrane transport", 416, 12, 5.03, 0.0038, "BP"
 3, "GO:0034220", "ion transmembrane transport", 89, 4, 1.08, 0.0222, "BP"
 4, "GO:0072351", "tricarboxylic acid biosynthetic process", 2, 1, 0.02, 0.024, "BP"
 5, "GO:0030417", "nicotianamine metabolic process", 2, 1, 0.02, 0.024, "BP"
 6, "GO:0030418", "nicotianamine biosynthetic process", 2, 1, 0.02, 0.024, "BP"
 7, "GO:0030001", "metal ion transport", 100, 4, 1.21, 0.0323, "BP"
 8, "GO:0003333", "amino acid transmembrane transport", 25, 2, 0.3, 0.0361, "BP"
 9, "GO:0006865", "amino acid transport", 25, 2, 0.3, 0.0361, "BP"
 10, "GO:0006812", "cation transport", 154, 5, 1.86, 0.0379, "BP"
 11, "GO:0072488", "ammonium transmembrane transport", 4, 1, 0.05, 0.0475, "BP"
 12, "GO:0015696", "ammonium transport", 4, 1, 0.05, 0.0475, "BP"
 16, "GO:0015075", "ion transmembrane transporter activity", 185, 7, 1.84, 0.0024, "MF"
 17, "GO:0016740", "transferase activity", 1843, 29, 18.38, 0.005, "MF"
 18, "GO:0015318", "inorganic molecular entity transmembrane...", 168, 6, 1.68, 0.0065, "MF"
 19, "GO:0016758", "transferase activity, transferring hexos...", 312, 8, 3.11, 0.0122, "MF"
 20, "GO:0046873", "metal ion transmembrane transporter acti...", 51, 3, 0.51, 0.0141, "MF"
 21, "GO:0020037", "heme binding", 258, 7, 2.57, 0.0142, "MF"
 22, "GO:0046906", "tetrapyrrole binding", 260, 7, 2.59, 0.0147, "MF"
 23, "GO:0030410", "nicotianamine synthase activity", 2, 1, 0.02, 0.0198, "MF"
 24, "GO:0016781", "phosphotransferase activity, paired acce...", 2, 1, 0.02, 0.0198, "MF"
 25, "GO:0050242", "pyruvate, phosphate dikinase activity", 2, 1, 0.02, 0.0198, "MF"
 26, "GO:0008168", "methyltransferase activity", 162, 5, 1.62, 0.0226, "MF"

27,"GO:0016741","transferase activity, transferring one-c...",170,5,1.7,0.0271,"MF"
 28,"GO:0008324","cation transmembrane transporter activit...",115,4,1.15,0.0276,"MF"
 29,"GO:0016757","transferase activity, transferring glyco...",377,8,3.76,0.0336,"MF"
 30,"GO:0008519","ammonium transmembrane transporter activ...",4,1,0.04,0.0393,"MF"
 31,"GO:0016020","membrane",1185,18,9.57,4.1e-05,"CC"

H. amplexicaulis F1-F3 Abajo

1,"GO:0007018","microtubule-based movement",42,17,2.96,9.5e-10,"BP"
 2,"GO:0006979","response to oxidative stress",88,24,6.21,4.2e-09,"BP"
 3,"GO:0006260","DNA replication",40,19,2.82,1e-06,"BP"
 4,"GO:0007076","mitotic chromosome condensation",4,4,0.28,2.4e-05,"BP"
 5,"GO:0006073","cellular glucan metabolic process",52,11,3.67,0.00081,"BP"
 6,"GO:0032502","developmental process",42,11,2.96,0.00147,"BP"
 7,"GO:0055114","oxidation-reduction process",728,70,51.36,0.00342,"BP"
 8,"GO:0042545","cell wall modification",10,4,0.71,0.00364,"BP"
 9,"GO:0006261","DNA-dependent DNA replication",6,6,0.42,0.00488,"BP"
 10,"GO:0007049","cell cycle",69,15,4.87,0.00491,"BP"
 11,"GO:0006269","DNA replication, synthesis of RNA primer",2,2,0.14,0.00497,"BP"
 12,"GO:0009956","radial pattern formation",2,2,0.14,0.00497,"BP"
 13,"GO:0006270","DNA replication initiation",2,2,0.14,0.00497,"BP"
 14,"GO:0006275","regulation of DNA replication",2,2,0.14,0.00497,"BP"
 15,"GO:0008356","asymmetric cell division",2,2,0.14,0.00497,"BP"
 16,"GO:0008017","microtubule binding",54,20,3.77,1.9e-10,"MF"
 17,"GO:0003777","microtubule motor activity",42,17,2.94,8.8e-10,"MF"
 18,"GO:0004601","peroxidase activity",89,24,6.22,5e-09,"MF"
 19,"GO:0005507","copper ion binding",31,11,2.17,4.1e-06,"MF"
 20,"GO:0016762","xyloglucan:xyloglucosyl transferase acti...",19,8,1.33,2.1e-05,"MF"
 21,"GO:0020037","heme binding",234,33,16.36,7.7e-05,"MF"
 22,"GO:0004553","hydrolase activity, hydrolyzing O-glycos...",224,30,15.66,0.00041,"MF"
 23,"GO:0004190","aspartic-type endopeptidase activity",63,12,4.4,0.0012,"MF"
 24,"GO:0005089","Rho guanyl-nucleotide exchange factor ac...",8,4,0.56,0.00132,"MF"
 25,"GO:0140097","catalytic activity, acting on DNA",73,13,5.1,0.00146,"MF"
 26,"GO:0003677","DNA binding",668,65,46.7,0.00338,"MF"
 27,"GO:0030599","pectinesterase activity",10,4,0.7,0.00353,"MF"
 28,"GO:0003896","DNA primase activity",2,2,0.14,0.00488,"MF"
 29,"GO:0009055","electron transfer activity",86,13,6.01,0.0064,"MF"
 30,"GO:0008236","serine-type peptidase activity",96,14,6.71,0.00657,"MF"
 31,"GO:0005871","kinesin complex",42,17,2.73,1.7e-10,"CC"
 32,"GO:0005618","cell wall",30,12,1.95,1.2e-07,"CC"
 33,"GO:0048046","apoplast",19,8,1.24,1.1e-05,"CC"
 34,"GO:0000786","nucleosome",21,7,1.37,0.00023,"CC"
 35,"GO:0000796","condensin complex",2,2,0.13,0.00421,"CC"
 36,"GO:0042555","MCM complex",2,2,0.13,0.00421,"CC"
 37,"GO:0005634","nucleus",448,42,29.15,0.00557,"CC"

38,"GO:0005694","chromosome",58,15,3.77,0.01837,"CC"

39,"GO:0044459","plasma membrane part",19,4,1.24,0.03099,"CC"

H. amplexicaulis F1-F3 Arriba

1,"GO:0015979","photosynthesis",49,26,2.27,3.2e-13,"BP"

2,"GO:0009765","photosynthesis, light harvesting",14,10,0.65,3.4e-11,"BP"

3,"GO:0055114","oxidation-reduction process",728,59,33.79,8.2e-06,"BP"

4,"GO:0006855","drug transmembrane transport",35,6,1.62,0.0049,"BP"

5,"GO:0015977","carbon fixation",4,2,0.19,0.0121,"BP"

6,"GO:0006091","generation of precursor metabolites and ...",63,16,2.92,0.0213,"BP"

7,"GO:0098655","cation transmembrane transport",62,7,2.88,0.0238,"BP"

8,"GO:0015985","energy coupled proton transport, down el...",14,3,0.65,0.0246,"BP"

9,"GO:0015986","ATP synthesis coupled proton transport",14,3,0.65,0.0246,"BP"

10,"GO:0017144","drug metabolic process",108,10,5.01,0.0276,"BP"

11,"GO:0006754","ATP biosynthetic process",51,6,2.37,0.0295,"BP"

12,"GO:0006812","cation transport",159,13,7.38,0.0328,"BP"

13,"GO:0009201","ribonucleoside triphosphate biosynthetic...",53,6,2.46,0.0349,"BP"

14,"GO:0009206","purine ribonucleoside triphosphate biosy...",53,6,2.46,0.0349,"BP"

15,"GO:0009142","nucleoside triphosphate biosynthetic pro...",53,6,2.46,0.0349,"BP"

16,"GO:0016705","oxidoreductase activity, acting on paire...",171,23,7.14,6e-07,"MF"

17,"GO:0005506","iron ion binding",172,23,7.18,6.6e-07,"MF"

18,"GO:0020037","heme binding",234,27,9.77,1.4e-06,"MF"

19,"GO:0010333","terpene synthase activity",16,6,0.67,2.9e-05,"MF"

20,"GO:0008171","O-methyltransferase activity",21,6,0.88,0.00016,"MF"

21,"GO:0042132","fructose 1,6-bisphosphate 1-phosphatase ...",5,3,0.21,0.00068,"MF"

22,"GO:0004185","serine-type carboxypeptidase activity",28,6,1.17,0.00088,"MF"

23,"GO:0016491","oxidoreductase activity",869,70,36.3,0.00093,"MF"

24,"GO:0016655","oxidoreductase activity, acting on NAD(P...)",6,3,0.25,0.00131,"MF"

25,"GO:0003824","catalytic activity",4391,237,183.41,0.00175,"MF"

26,"GO:0030246","carbohydrate binding",96,11,4.01,0.0021,"MF"

27,"GO:0015297","antiporter activity",50,7,2.09,0.00443,"MF"

28,"GO:0015238","drug transmembrane transporter activity",38,6,1.59,0.00452,"MF"

29,"GO:0016844","strictosidine synthase activity",3,2,0.13,0.00507,"MF"

30,"GO:0016830","carbon-carbon lyase activity",45,8,1.88,0.00646,"MF"

31,"GO:0016020","membrane",1180,78,42.43,9.6e-14,"CC"

32,"GO:0009522","photosystem I",9,8,0.32,1.3e-06,"CC"

33,"GO:0009538","photosystem I reaction center",5,4,0.18,7.6e-06,"CC"

34,"GO:0009523","photosystem II",19,7,0.68,0.00015,"CC"

35,"GO:0009654","photosystem II oxygen evolving complex",15,4,0.54,0.00157,"CC"

36,"GO:0042651","thylakoid membrane",18,6,0.65,0.00344,"CC"

37,"GO:0019898","extrinsic component of membrane",17,3,0.61,0.02118,"CC"

H. amplexicaulis F3-F5 Abajo

1,"GO:0006979","response to oxidative stress",88,17,2.77,1.2e-09,"BP"

2, "GO:0055114", "oxidation-reduction process", 728, 47, 22.93, 7.1e-07, "BP"
3, "GO:0030244", "cellulose biosynthetic process", 23, 6, 0.72, 5.8e-05, "BP"
4, "GO:0042545", "cell wall modification", 10, 4, 0.32, 0.00017, "BP"
5, "GO:0005975", "carbohydrate metabolic process", 357, 26, 11.25, 0.00286, "BP"
6, "GO:0006265", "DNA topological change", 8, 2, 0.25, 0.02439, "BP"
7, "GO:0005985", "sucrose metabolic process", 9, 2, 0.28, 0.03072, "BP"
8, "GO:0009791", "post-embryonic development", 1, 1, 0.03, 0.0315, "BP"
9, "GO:0048580", "regulation of post-embryonic development", 1, 1, 0.03, 0.0315, "BP"
10, "GO:0048581", "negative regulation of post-embryonic de...", 1, 1, 0.03, 0.0315, "BP"
11, "GO:0051017", "actin filament bundle assembly", 1, 1, 0.03, 0.0315, "BP"
12, "GO:0009908", "flower development", 1, 1, 0.03, 0.0315, "BP"
13, "GO:0009909", "regulation of flower development", 1, 1, 0.03, 0.0315, "BP"
14, "GO:2000026", "regulation of multicellular organismal d...", 1, 1, 0.03, 0.0315, "BP"
15, "GO:0003006", "developmental process involved in reprod...", 1, 1, 0.03, 0.0315, "BP"
16, "GO:0020037", "heme binding", 234, 29, 7.37, 1.8e-10, "MF"
17, "GO:0004601", "peroxidase activity", 89, 17, 2.8, 1.7e-09, "MF"
18, "GO:0030599", "pectinesterase activity", 10, 4, 0.31, 0.00017, "MF"
19, "GO:0004553", "hydrolase activity, hydrolyzing O-glycos...", 224, 18, 7.05, 0.00024, "MF"
20, "GO:0005507", "copper ion binding", 31, 6, 0.98, 0.00035, "MF"
21, "GO:0003824", "catalytic activity", 4391, 193, 138.24, 4e-04, "MF"
22, "GO:0016760", "cellulose synthase (UDP-forming) activit...", 21, 5, 0.66, 4e-04, "MF"
23, "GO:0004190", "aspartic-type endopeptidase activity", 63, 8, 1.98, 0.00075, "MF"
24, "GO:0016788", "hydrolase activity, acting on ester bond...", 307, 23, 9.67, 0.00217, "MF"
25, "GO:0009055", "electron transfer activity", 86, 8, 2.71, 0.00551, "MF"
26, "GO:0008725", "DNA-3-methyladenine glycosylase activity", 4, 2, 0.13, 0.00568, "MF"
27, "GO:0080019", "fatty-acyl-CoA reductase (alcohol-formin...", 4, 2, 0.13, 0.00568, "MF"
28, "GO:0016705", "oxidoreductase activity, acting on paire...", 171, 12, 5.38, 0.00764, "MF"
29, "GO:0005506", "iron ion binding", 172, 12, 5.41, 0.00799, "MF"
30, "GO:0016899", "oxidoreductase activity, acting on the C...", 6, 2, 0.19, 0.01362, "MF"
31, "GO:0005618", "cell wall", 30, 6, 0.85, 0.00014, "CC"
32, "GO:0005856", "cytoskeleton", 98, 7, 2.77, 0.01885, "CC"
33, "GO:0016020", "membrane", 1180, 42, 33.34, 0.02031, "CC"
34, "GO:0044430", "cytoskeletal part", 89, 6, 2.51, 0.03743, "CC"
35, "GO:0005576", "extracellular region", 29, 3, 0.82, 0.04664, "CC"
36, "GO:0015630", "microtubule cytoskeleton", 71, 5, 2.01, 0.04786, "CC"

H. amplexicaulis F3-F5 Arriba

1, "GO:0006468", "protein phosphorylation", 797, 59, 32.85, 2.8e-06, "BP"
2, "GO:0051704", "multi-organism process", 51, 7, 2.1, 0.0046, "BP"
3, "GO:0009065", "glutamine family amino acid catabolic pr...", 3, 2, 0.12, 0.0049, "BP"
4, "GO:0055085", "transmembrane transport", 418, 28, 17.23, 0.0069, "BP"
5, "GO:0009875", "pollen-pistil interaction", 48, 6, 1.98, 0.0133, "BP"
6, "GO:0008037", "cell recognition", 48, 6, 1.98, 0.0133, "BP"
7, "GO:0048544", "recognition of pollen", 48, 6, 1.98, 0.0133, "BP"

8,"GO:0044706","multi-multicellular organism process",48,6,1.98,0.0133,"BP"
9,"GO:0009856","pollination",48,6,1.98,0.0133,"BP"
10,"GO:0000003","reproduction",52,6,2.14,0.0192,"BP"
11,"GO:0022414","reproductive process",52,6,2.14,0.0192,"BP"
12,"GO:0055114","oxidation-reduction process",728,41,30.01,0.0216,"BP"
13,"GO:0050832","defense response to fungus",1,1,0.04,0.0412,"BP"
14,"GO:0009767","photosynthetic electron transport chain",1,1,0.04,0.0412,"BP"
15,"GO:0009772","photosynthetic electron transport in pho...",1,1,0.04,0.0412,"BP"
16,"GO:0004672","protein kinase activity",795,58,28.38,6.7e-08,"MF"
17,"GO:0030247","polysaccharide binding",41,10,1.46,1.2e-06,"MF"
18,"GO:0016758","transferase activity, transferring hexos...",296,28,10.57,2.1e-06,"MF"
19,"GO:0016705","oxidoreductase activity, acting on paire...",171,20,6.1,2.6e-06,"MF"
20,"GO:0005506","iron ion binding",172,19,6.14,1.1e-05,"MF"
21,"GO:0020037","heme binding",234,21,8.35,8.9e-05,"MF"
22,"GO:0005524","ATP binding",1001,56,35.73,0.00034,"MF"
23,"GO:0008171","O-methyltransferase activity",21,5,0.75,0.00071,"MF"
24,"GO:0010333","terpene synthase activity",16,4,0.57,0.00206,"MF"
25,"GO:0005509","calcium ion binding",78,8,2.78,0.00643,"MF"
26,"GO:0004611","phosphoenolpyruvate carboxykinase activi...",5,2,0.18,0.01182,"MF"
27,"GO:0033897","ribonuclease T2 activity",5,2,0.18,0.01182,"MF"
28,"GO:0016892","endoribonuclease activity, producing 3'-...",5,2,0.18,0.01182,"MF"
29,"GO:0016894","endonuclease activity, active with eithe...",5,2,0.18,0.01182,"MF"
30,"GO:0004497","monooxygenase activity",14,3,0.5,0.01223,"MF"
31,"GO:0016021","integral component of membrane",525,24,11.01,3.4e-05,"CC"
32,"GO:0016020","membrane",1180,40,24.75,0.0041,"CC"

H. amplexicaulis F5-F7 Abajo

1,"GO:0030244","cellulose biosynthetic process",23,12,1.14,1.5e-10,"BP"
2,"GO:0005975","carbohydrate metabolic process",357,50,17.71,3.4e-07,"BP"
3,"GO:0055114","oxidation-reduction process",728,61,36.11,1.7e-05,"BP"
4,"GO:0006979","response to oxidative stress",88,13,4.36,0.00035,"BP"
5,"GO:0042545","cell wall modification",10,4,0.5,0.00098,"BP"
6,"GO:0006633","fatty acid biosynthetic process",47,8,2.33,0.00193,"BP"
7,"GO:0009832","plant-type cell wall biogenesis",6,3,0.3,0.00216,"BP"
8,"GO:0000226","microtubule cytoskeleton organization",23,5,1.14,0.00467,"BP"
9,"GO:0007017","microtubule-based process",81,13,4.02,0.00678,"BP"
10,"GO:0000910","cytokinesis",9,3,0.45,0.00811,"BP"
11,"GO:0044264","cellular polysaccharide metabolic proces...",53,17,2.63,0.0128,"BP"
12,"GO:0048367","shoot system development",4,2,0.2,0.01376,"BP"
13,"GO:0044036","cell wall macromolecule metabolic proces...",11,3,0.55,0.01481,"BP"
14,"GO:0007018","microtubule-based movement",42,6,2.08,0.01644,"BP"
15,"GO:0006928","movement of cell or subcellular componen...",42,6,2.08,0.01644,"BP"
16,"GO:0016760","cellulose synthase (UDP-forming) activit...",21,12,1.2,1.2e-11,"MF"
17,"GO:0020037","heme binding",234,36,11.2,3.2e-10,"MF"

18,"GO:0004553","hydrolase activity, hydrolyzing O-glycos...",224,35,10.72,3.7e-10,"MF"
 19,"GO:0005506","iron ion binding",172,25,8.23,5.5e-07,"MF"
 20,"GO:0016705","oxidoreductase activity, acting on paire...",171,23,8.18,6.1e-06,"MF"
 21,"GO:0008017","microtubule binding",54,11,2.58,3.9e-05,"MF"
 22,"GO:0010333","terpene synthase activity",16,6,0.77,6.1e-05,"MF"
 23,"GO:0016747","transferase activity, transferring acyl ...",146,19,6.99,6.3e-05,"MF"
 24,"GO:0003824","catalytic activity",4391,288,210.08,6.7e-05,"MF"
 25,"GO:0004601","peroxidase activity",89,14,4.26,7.5e-05,"MF"
 26,"GO:0030599","pectinesterase activity",10,4,0.48,0.00086,"MF"
 27,"GO:0004857","enzyme inhibitor activity",35,7,1.67,0.00114,"MF"
 28,"GO:0016757","transferase activity, transferring glyco...",357,42,17.08,0.00589,"MF"
 29,"GO:0016758","transferase activity, transferring hexos...",296,34,14.16,0.0089,"MF"
 30,"GO:0015020","glucuronosyltransferase activity",10,3,0.48,0.01014,"MF"
 31,"GO:0005618","cell wall",30,7,1.23,0.00015,"CC"
 32,"GO:0005576","extracellular region",29,5,1.19,0.00572,"CC"
 33,"GO:0005871","kinesin complex",42,6,1.73,0.00651,"CC"
 34,"GO:0016020","membrane",1180,60,48.49,0.01061,"CC"
 35,"GO:0015630","microtubule cytoskeleton",71,10,2.92,0.02483,"CC"
 36,"GO:0031225","anchored component of membrane",7,2,0.29,0.03067,"CC"
 37,"GO:0048046","apoplast",19,3,0.78,0.04042,"CC"
 38,"GO:0008352","katanin complex",1,1,0.04,0.0411,"CC"

H. amplexicaulis F5-F7 Arriba

1,"GO:0006468","protein phosphorylation",797,16,7.48,0.002,"BP"
 2,"GO:0009733","response to auxin",20,2,0.19,0.015,"BP"
 3,"GO:0009690","cytokinin metabolic process",5,1,0.05,0.046,"BP"
 4,"GO:0034754","cellular hormone metabolic process",5,1,0.05,0.046,"BP"
 16,"GO:0004672","protein kinase activity",795,15,6.31,0.0011,"MF"
 17,"GO:0016705","oxidoreductase activity, acting on paire...",171,6,1.36,0.0023,"MF"
 18,"GO:0005506","iron ion binding",172,6,1.36,0.0023,"MF"
 19,"GO:0020037","heme binding",234,7,1.86,0.0024,"MF"
 20,"GO:0032559","adenyl ribonucleotide binding",1146,17,9.09,0.0066,"MF"
 21,"GO:0030246","carbohydrate binding",96,4,0.76,0.0069,"MF"
 22,"GO:0046429","4-hydroxy-3-methylbut-2-en-1-yl diphosph...",1,1,0.01,0.0079,"MF"
 23,"GO:0015036","disulfide oxidoreductase activity",58,3,0.46,0.0108,"MF"
 24,"GO:0043168","anion binding",1454,19,11.53,0.0152,"MF"
 25,"GO:0016667","oxidoreductase activity, acting on a sul...",68,3,0.54,0.0165,"MF"
 26,"GO:0005452","inorganic anion exchanger activity",3,1,0.02,0.0236,"MF"
 27,"GO:0043531","ADP binding",146,4,1.16,0.0283,"MF"
 28,"GO:0016668","oxidoreductase activity, acting on a sul...",4,1,0.03,0.0314,"MF"
 29,"GO:0047134","protein-disulfide reductase activity",4,1,0.03,0.0314,"MF"
 30,"GO:0061733","peptide-lysine-N-acetyltransferase activ...",5,1,0.04,0.039,"MF"

Hoja de datos suplementaria 2: Términos de GO con los valores p más bajos para OED entre especies. Se obtuvieron los 15 términos con valores más bajos para cada categoría: procesos biológicos (BP), función molecular (MF) y componente celular (CC). Sólo se muestran los términos que superaron el valor de corte (valor $p < 0,05$). Para cada término se indica: Identificador, nombre, N° de términos significativos, N° de términos esperados, valor p y categoría del término, respectivamente.

F1 Abajo *A. lanata* vs *H. amplexicaulis*

1, "GO:0055114", "oxidation-reduction process", 495, 118, 80.06, 1.8e-06, "BP"
 2, "GO:0006468", "protein phosphorylation", 560, 123, 90.58, 7.5e-05, "BP"
 3, "GO:0030244", "cellulose biosynthetic process", 8, 6, 1.29, 0.00037, "BP"
 4, "GO:0048544", "recognition of pollen", 23, 11, 3.72, 0.00038, "BP"
 5, "GO:0042546", "cell wall biogenesis", 11, 7, 1.78, 0.00051, "BP"
 6, "GO:0005975", "carbohydrate metabolic process", 250, 63, 40.44, 0.0012, "BP"
 7, "GO:0006979", "response to oxidative stress", 60, 19, 9.7, 0.00206, "BP"
 8, "GO:0006355", "regulation of transcription, DNA-templat...", 473, 99, 76.51, 0.00234, "BP"
 9, "GO:0071554", "cell wall organization or biogenesis", 30, 15, 4.85, 0.00616, "BP"
 10, "GO:0071805", "potassium ion transmembrane transport", 12, 6, 1.94, 0.00674, "BP"
 11, "GO:0020037", "heme binding", 159, 57, 25.81, 8.7e-10, "MF"
 12, "GO:0005506", "iron ion binding", 118, 42, 19.15, 1.9e-07, "MF"
 13, "GO:0016705", "oxidoreductase activity, acting on paire...", 118, 38, 19.15, 1.1e-05, "MF"
 14, "GO:0004553", "hydrolase activity, hydrolyzing O-glycos...", 163, 48, 26.46, 1.3e-05, "MF"
 15, "GO:0016747", "transferase activity, transferring acyl ...", 104, 34, 16.88, 2.3e-05, "MF"
 16, "GO:0016760", "cellulose synthase (UDP-forming) activit...", 7, 6, 1.14, 0.00011, "MF"
 17, "GO:0004190", "aspartic-type endopeptidase activity", 52, 19, 8.44, 3e-04, "MF"
 18, "GO:0004672", "protein kinase activity", 558, 120, 90.57, 0.00037, "MF"
 19, "GO:0004601", "peroxidase activity", 58, 20, 9.41, 5e-04, "MF"
 20, "GO:0010333", "terpene synthase activity", 11, 6, 1.79, 0.00398, "MF"
 21, "GO:0016020", "membrane", 875, 140, 94.61, 1.7e-10, "CC"
 22, "GO:0005618", "cell wall", 21, 8, 2.27, 0.00095, "CC"
 23, "GO:0009521", "photosystem", 21, 7, 2.27, 0.0048, "CC"
 24, "GO:0031224", "intrinsic component of membrane", 400, 57, 43.25, 0.00896, "CC"
 25, "GO:0005576", "extracellular region", 18, 6, 1.95, 0.00901, "CC"
 26, "GO:0031012", "extracellular matrix", 2, 2, 0.22, 0.01164, "CC"
 27, "GO:0048046", "apoplast", 10, 4, 1.08, 0.01647, "CC"
 28, "GO:0031225", "anchored component of membrane", 6, 3, 0.65, 0.01944, "CC"
 29, "GO:0016021", "integral component of membrane", 394, 54, 42.6, 0.02454, "CC"
 30, "GO:0009522", "photosystem I", 7, 3, 0.76, 0.03135, "CC"

F1 Arriba *A. lanata* vs *H. amplexicaulis*

1, "GO:0006260", "DNA replication", 38, 11, 2.85, 6.8e-05, "BP"
 2, "GO:0007275", "multicellular organism development", 25, 7, 1.87, 0.0018, "BP"
 3, "GO:0034622", "cellular protein-containing complex asse...", 40, 9, 3, 0.0023, "BP"

4,"GO:0006281","DNA repair",92,15,6.89,0.0031,"BP"
5,"GO:0006333","chromatin assembly or disassembly",17,5,1.27,0.0067,"BP"
6,"GO:0071103","DNA conformation change",26,6,1.95,0.0108,"BP"
7,"GO:0007015","actin filament organization",19,5,1.42,0.0111,"BP"
8,"GO:0006323","DNA packaging",19,5,1.42,0.0111,"BP"
9,"GO:0051276","chromosome organization",68,14,5.1,0.0114,"BP"
10,"GO:0034314","Arp2/3 complex-mediated actin nucleation",7,3,0.52,0.0116,"BP"
11,"GO:0043531","ADP binding",86,22,6.94,7.1e-07,"MF"
12,"GO:0003899","DNA-directed 5'-3' RNA polymerase activi...",30,9,2.42,0.00041,"MF"
13,"GO:0008137","NADH dehydrogenase (ubiquinone) activity",4,3,0.32,0.00196,"MF"
14,"GO:0140097","catalytic activity, acting on DNA",63,12,5.08,0.00395,"MF"
15,"GO:0003676","nucleic acid binding",950,97,76.63,0.00597,"MF"
16,"GO:0032559","adenyl ribonucleotide binding",863,101,69.61,0.0067,"MF"
17,"GO:0004386","helicase activity",45,9,3.63,0.00863,"MF"
18,"GO:0010333","terpene synthase activity",11,4,0.89,0.00873,"MF"
19,"GO:0003678","DNA helicase activity",31,7,2.5,0.01017,"MF"
20,"GO:0005524","ATP binding",777,79,62.67,0.01464,"MF"
21,"GO:0005634","nucleus",379,53,29.94,2.6e-06,"CC"
22,"GO:0044427","chromosomal part",41,9,3.24,0.0036,"CC"
23,"GO:0043232","intracellular non-membrane-bounded organ...",253,35,19.98,0.0065,"CC"
24,"GO:0005885","Arp2/3 protein complex",7,3,0.55,0.0133,"CC"
25,"GO:0000786","nucleosome",15,4,1.18,0.0256,"CC"
26,"GO:0032993","protein-DNA complex",15,4,1.18,0.0256,"CC"
27,"GO:0000793","condensed chromosome",9,3,0.71,0.0284,"CC"
28,"GO:0106068","SUMO ligase complex",4,2,0.32,0.0334,"CC"
29,"GO:0030915","Smc5-Smc6 complex",4,2,0.32,0.0334,"CC"
30,"GO:0005685","U1 snRNP",4,2,0.32,0.0334,"CC"

F1 Abajo - A. lanata vs R. pilosa

1,"GO:0055114","oxidation-reduction process",495,73,40.97,2.2e-07,"BP"
2,"GO:0005975","carbohydrate metabolic process",250,32,20.69,0.0077,"BP"
3,"GO:0006979","response to oxidative stress",60,11,4.97,0.0091,"BP"
4,"GO:0006073","cellular glucan metabolic process",24,6,1.99,0.0115,"BP"
5,"GO:0044042","glucan metabolic process",24,6,1.99,0.0115,"BP"
6,"GO:0044264","cellular polysaccharide metabolic proces...",25,6,2.07,0.0141,"BP"
7,"GO:0071554","cell wall organization or biogenesis",30,6,2.48,0.0334,"BP"
8,"GO:0070726","cell wall assembly",4,2,0.33,0.0366,"BP"
9,"GO:0009664","plant-type cell wall organization",4,2,0.33,0.0366,"BP"
10,"GO:0071668","plant-type cell wall assembly",4,2,0.33,0.0366,"BP"
11,"GO:0020037","heme binding",159,32,12.93,1e-06,"MF"
12,"GO:0043531","ADP binding",86,20,6.99,1.2e-05,"MF"
13,"GO:0005506","iron ion binding",118,24,9.6,2e-05,"MF"
14,"GO:0016705","oxidoreductase activity, acting on paire...",118,22,9.6,0.00017,"MF"
15,"GO:0004553","hydrolase activity, hydrolyzing O-glycos...",163,27,13.25,0.00026,"MF"

16,"GO:0004185","serine-type carboxypeptidase activity",20,7,1.63,0.00068,"MF"
 17,"GO:0004601","peroxidase activity",58,12,4.72,0.00204,"MF"
 18,"GO:0016491","oxidoreductase activity",595,86,48.38,0.00586,"MF"
 19,"GO:0016762","xyloglucan:xyloglucosyl transferase acti...",10,4,0.81,0.0061,"MF"
 20,"GO:0004097","catechol oxidase activity",2,2,0.16,0.0066,"MF"
 21,"GO:0016020","membrane",875,75,56.37,0.0012,"CC"
 22,"GO:0048046","apoplast",10,4,0.64,0.0025,"CC"
 23,"GO:0005664","nuclear origin of replication recognitio...",2,2,0.13,0.0041,"CC"
 24,"GO:0031225","anchored component of membrane",6,3,0.39,0.0045,"CC"
 25,"GO:0005618","cell wall",21,5,1.35,0.009,"CC"
 26,"GO:0009523","photosystem II",14,4,0.9,0.0099,"CC"

F1 Arriba - A. lanata vs R. pilosa

1,"GO:0006468","protein phosphorylation",560,67,43.6,0.00013,"BP"
 2,"GO:0016036","cellular response to phosphate starvatio...",2,2,0.16,0.00604,"BP"
 3,"GO:0051704","multi-organism process",26,6,2.02,0.0129,"BP"
 4,"GO:0032200","telomere organization",3,2,0.23,0.0172,"BP"
 5,"GO:0060249","anatomical structure homeostasis",3,2,0.23,0.0172,"BP"
 6,"GO:0006108","malate metabolic process",3,2,0.23,0.0172,"BP"
 7,"GO:0000723","telomere maintenance",3,2,0.23,0.0172,"BP"
 8,"GO:0006952","defense response",14,4,1.09,0.01931,"BP"
 9,"GO:0071555","cell wall organization",14,4,1.09,0.01931,"BP"
 10,"GO:0045229","external encapsulating structure organiz...",14,4,1.09,0.01931,"BP"
 11,"GO:0030247","polysaccharide binding",25,9,1.88,4.8e-05,"MF"
 12,"GO:0004672","protein kinase activity",558,67,41.9,5e-05,"MF"
 13,"GO:0016705","oxidoreductase activity, acting on paire...",118,21,8.86,0.00015,"MF"
 14,"GO:0043531","ADP binding",86,17,6.46,0.00018,"MF"
 15,"GO:0016758","transferase activity, transferring hexos...",205,29,15.4,0.00062,"MF"
 16,"GO:0005506","iron ion binding",118,17,8.86,0.00664,"MF"
 17,"GO:0032559","adenyl ribonucleotide binding",863,90,64.81,0.01293,"MF"
 18,"GO:0030554","adenyl nucleotide binding",865,90,64.96,0.0137,"MF"
 19,"GO:0008144","drug binding",822,78,61.73,0.01424,"MF"
 20,"GO:0046906","tetrapyrrole binding",159,20,11.94,0.01506,"MF"
 21,"GO:0045261","proton-transporting ATP synthase complex...",5,2,0.26,0.024,"CC"
 22,"GO:0031225","anchored component of membrane",6,2,0.31,0.035,"CC"
 23,"GO:0016020","membrane",875,54,45.59,0.046,"CC"

F1 Abajo - R. pilosa vs H. amplexicaulis

1,"GO:0006468","protein phosphorylation",560,141,84.44,1.3e-11,"BP"
 2,"GO:0048544","recognition of pollen",23,14,3.47,6e-07,"BP"
 3,"GO:0006355","regulation of transcription, DNA-templat...",473,97,71.32,0.00047,"BP"
 4,"GO:0009733","response to auxin",15,8,2.26,0.00062,"BP"
 5,"GO:0042546","cell wall biogenesis",11,6,1.66,0.00269,"BP"
 6,"GO:0016567","protein ubiquitination",52,15,7.84,0.00789,"BP"

7,"GO:0044036","cell wall macromolecule metabolic proces...",10,5,1.51,0.01001,"BP"
8,"GO:0071554","cell wall organization or biogenesis",30,13,4.52,0.01601,"BP"
9,"GO:0006508","proteolysis",237,48,35.74,0.01655,"BP"
10,"GO:0072488","ammonium transmembrane transport",2,2,0.3,0.02271,"BP"
11,"GO:0004672","protein kinase activity",558,139,81.34,8.9e-12,"MF"
12,"GO:0020037","heme binding",159,48,23.18,2.6e-07,"MF"
13,"GO:0003700","DNA-binding transcription factor activit...",241,63,35.13,1.2e-06,"MF"
14,"GO:0004190","aspartic-type endopeptidase activity",52,21,7.58,4.6e-06,"MF"
15,"GO:0005506","iron ion binding",118,36,17.2,6.4e-06,"MF"
16,"GO:0030247","polysaccharide binding",25,13,3.64,1.2e-05,"MF"
17,"GO:0016705","oxidoreductase activity, acting on paire...",118,34,17.2,4.3e-05,"MF"
18,"GO:0008725","DNA-3-methyladenine glycosylase activity",4,4,0.58,0.00045,"MF"
19,"GO:0016747","transferase activity, transferring acyl ...",104,27,15.16,0.00153,"MF"
20,"GO:0010333","terpene synthase activity",11,6,1.6,0.00226,"MF"
21,"GO:0016020","membrane",875,115,81.37,1.1e-06,"CC"
22,"GO:0016021","integral component of membrane",394,51,36.64,0.0041,"CC"
23,"GO:0031012","extracellular matrix",2,2,0.19,0.0086,"CC"
24,"GO:0005618","cell wall",21,5,1.95,0.039,"CC"
25,"GO:0030312","external encapsulating structure",21,5,1.95,0.039,"CC"

F1 Arriba – R. pilosa vs H. amplexicaulis

1,"GO:0006260","DNA replication",38,8,2.61,0.0036,"BP"
2,"GO:0006560","proline metabolic process",2,2,0.14,0.0047,"BP"
3,"GO:0006412","translation",188,22,12.91,0.0089,"BP"
4,"GO:0006281","DNA repair",92,13,6.32,0.0094,"BP"
5,"GO:0071840","cellular component organization or bioge...",236,26,16.21,0.0101,"BP"
6,"GO:0006325","chromatin organization",46,8,3.16,0.0119,"BP"
7,"GO:0009147","pyrimidine nucleoside triphosphate metab...",3,2,0.21,0.0135,"BP"
8,"GO:0009065","glutamine family amino acid catabolic pr...",3,2,0.21,0.0135,"BP"
9,"GO:0051276","chromosome organization",68,10,4.67,0.0166,"BP"
10,"GO:0006996","organelle organization",147,17,10.1,0.0225,"BP"
11,"GO:0043531","ADP binding",86,22,6.23,1.1e-07,"MF"
12,"GO:0008168","methyltransferase activity",113,18,8.19,0.0012,"MF"
13,"GO:0008137","NADH dehydrogenase (ubiquinone) activity",4,3,0.29,0.0014,"MF"
14,"GO:0003735","structural constituent of ribosome",121,18,8.77,0.0026,"MF"
15,"GO:0045300","acyl-[acyl-carrier-protein] desaturase a...",5,3,0.36,0.0034,"MF"
16,"GO:0016645","oxidoreductase activity, acting on the C...",10,4,0.72,0.004,"MF"
17,"GO:0046983","protein dimerization activity",85,12,6.16,0.0187,"MF"
18,"GO:0061733","peptide-lysine-N-acetyltransferase activ...",4,2,0.29,0.0285,"MF"
19,"GO:0034212","peptide N-acetyltransferase activity",4,2,0.29,0.0285,"MF"
20,"GO:0004402","histone acetyltransferase activity",4,2,0.29,0.0285,"MF"
21,"GO:0005840","ribosome",119,17,8.2,0.0023,"CC"
22,"GO:0005664","nuclear origin of replication recognitio...",2,2,0.14,0.0047,"CC"
23,"GO:0005759","mitochondrial matrix",6,3,0.41,0.0055,"CC"

24, "GO:0000793", "condensed chromosome", 9,3,0.62,0.0197, "CC"
 25, "GO:0043229", "intracellular organelle", 750,70,51.68,0.0229, "CC"
 26, "GO:0043226", "organelle", 751,70,51.75,0.0236, "CC"
 27, "GO:0043228", "non-membrane-bounded organelle", 253,33,17.43,0.0309, "CC"
 28, "GO:0043232", "intracellular non-membrane-bounded organ...", 253,33,17.43,0.0309, "CC"
 29, "GO:0005694", "chromosome", 50,9,3.45,0.0399, "CC"
 30, "GO:0044427", "chromosomal part", 41,8,2.83,0.0443, "CC"

F3 Abajo - A. lanata vs H. amplexicaulis

1, "GO:0048544", "recognition of pollen", 23,11,3.08,6.7e-05, "BP"
 2, "GO:0055114", "oxidation-reduction process", 495,95,66.33,7.9e-05, "BP"
 3, "GO:0006355", "regulation of transcription, DNA-templat...", 473,91,63.38,1e-04, "BP"
 4, "GO:0006468", "protein phosphorylation", 560,103,75.04,0.00022, "BP"
 5, "GO:0071554", "cell wall organization or biogenesis", 30,9,4.02,0.01374, "BP"
 6, "GO:0009308", "amine metabolic process", 21,7,2.81,0.01595, "BP"
 7, "GO:0042126", "nitrate metabolic process", 2,2,0.27,0.01793, "BP"
 8, "GO:0042128", "nitrate assimilation", 2,2,0.27,0.01793, "BP"
 9, "GO:2001057", "reactive nitrogen species metabolic proc...", 2,2,0.27,0.01793, "BP"
 10, "GO:0003006", "developmental process involved in reprod...", 2,2,0.27,0.01793, "BP"
 11, "GO:0003700", "DNA-binding transcription factor activit...", 241,59,31.78,9.1e-07, "MF"
 12, "GO:0004672", "protein kinase activity", 558,102,73.58,0.00021, "MF"
 13, "GO:0020037", "heme binding", 159,36,20.97,0.00064, "MF"
 14, "GO:0005506", "iron ion binding", 118,28,15.56,0.00117, "MF"
 15, "GO:0016705", "oxidoreductase activity, acting on paire...", 118,27,15.56,0.00249, "MF"
 16, "GO:0043565", "sequence-specific DNA binding", 149,32,19.65,0.00309, "MF"
 17, "GO:0043531", "ADP binding", 86,21,11.34,0.00321, "MF"
 18, "GO:0030247", "polysaccharide binding", 25,9,3.3,0.00327, "MF"
 19, "GO:0016701", "oxidoreductase activity, acting on singl...", 22,8,2.9,0.00513, "MF"
 20, "GO:0016702", "oxidoreductase activity, acting on singl...", 16,6,2.11,0.01278, "MF"
 21, "GO:0031224", "intrinsic component of membrane", 400,52,37.87,0.0051, "CC"
 22, "GO:0016021", "integral component of membrane", 394,50,37.3,0.0102, "CC"
 23, "GO:0005618", "cell wall", 21,6,1.99,0.0107, "CC"
 24, "GO:0030312", "external encapsulating structure", 21,6,1.99,0.0107, "CC"
 25, "GO:0005576", "extracellular region", 18,5,1.7,0.0222, "CC"
 26, "GO:0016020", "membrane", 875,102,82.84,0.0295, "CC"

F3 Arriba - A. lanata vs H. amplexicaulis

1, "GO:0006979", "response to oxidative stress", 60,20,6.32,1.4e-06, "BP"
 2, "GO:0006308", "DNA catabolic process", 3,3,0.32,0.0012, "BP"
 3, "GO:0006259", "DNA metabolic process", 139,27,14.65,0.0068, "BP"
 4, "GO:0046942", "carboxylic acid transport", 24,7,2.53,0.0097, "BP"
 5, "GO:0030071", "regulation of mitotic metaphase/anaphase...", 5,3,0.53,0.0099, "BP"
 6, "GO:0006928", "movement of cell or subcellular componen...", 37,9,3.9,0.0125, "BP"
 7, "GO:0007018", "microtubule-based movement", 37,9,3.9,0.0125, "BP"

8,"GO:0055114","oxidation-reduction process",495,67,52.16,0.015,"BP"
9,"GO:0034754","cellular hormone metabolic process",6,3,0.63,0.0182,"BP"
10,"GO:0009690","cytokinin metabolic process",6,3,0.63,0.0182,"BP"
11,"GO:0004601","peroxidase activity",58,20,6.53,2.4e-06,"MF"
12,"GO:0043531","ADP binding",86,25,9.68,4.8e-06,"MF"
13,"GO:0020037","heme binding",159,36,17.9,2.5e-05,"MF"
14,"GO:0033897","ribonuclease T2 activity",3,3,0.34,0.0014,"MF"
15,"GO:0016788","hydrolase activity, acting on ester bond...",240,43,27.02,0.0049,"MF"
16,"GO:0008184","glycogen phosphorylase activity",2,2,0.23,0.0127,"MF"
17,"GO:0004519","endonuclease activity",22,9,2.48,0.0146,"MF"
18,"GO:0046943","carboxylic acid transmembrane transporte...",14,5,1.58,0.015,"MF"
19,"GO:0015171","amino acid transmembrane transporter act...",14,5,1.58,0.015,"MF"
20,"GO:0005342","organic acid transmembrane transporter a...",14,5,1.58,0.015,"MF"
21,"GO:0005871","kinesin complex",37,9,3.61,0.0072,"CC"
22,"GO:0030312","external encapsulating structure",21,5,2.05,0.0465,"CC"
23,"GO:0005618","cell wall",21,5,2.05,0.0465,"CC"

F3 Abajo - A. lanata vs R. pilosa

1,"GO:0055114","oxidation-reduction process",495,104,59.02,5.8e-10,"BP"
2,"GO:0009308","amine metabolic process",21,8,2.5,0.0019,"BP"
3,"GO:0042545","cell wall modification",10,5,1.19,0.0036,"BP"
4,"GO:0048367","shoot system development",4,3,0.48,0.0061,"BP"
5,"GO:0008152","metabolic process",3461,447,412.69,0.0062,"BP"
6,"GO:1901607","alpha-amino acid biosynthetic process",31,9,3.7,0.0082,"BP"
7,"GO:0010817","regulation of hormone levels",8,4,0.95,0.0094,"BP"
8,"GO:0006568","tryptophan metabolic process",5,3,0.6,0.014,"BP"
9,"GO:0006586","indolalkylamine metabolic process",5,3,0.6,0.014,"BP"
10,"GO:0042430","indole-containing compound metabolic pro...",5,3,0.6,0.014,"BP"
11,"GO:0016491","oxidoreductase activity",595,119,70.29,1.4e-06,"MF"
12,"GO:0020037","heme binding",159,38,18.78,1.2e-05,"MF"
13,"GO:0016705","oxidoreductase activity, acting on paire...",118,30,13.94,2.9e-05,"MF"
14,"GO:0005506","iron ion binding",118,30,13.94,2.9e-05,"MF"
15,"GO:0016829","lyase activity",81,21,9.57,0.00033,"MF"
16,"GO:0050662","coenzyme binding",153,30,18.07,0.00325,"MF"
17,"GO:0030599","pectinesterase activity",10,5,1.18,0.00343,"MF"
18,"GO:0016746","transferase activity, transferring acyl ...",132,26,15.59,0.00556,"MF"
19,"GO:0050660","flavin adenine dinucleotide binding",38,10,4.49,0.01072,"MF"
20,"GO:0030145","manganese ion binding",5,3,0.59,0.01366,"MF"
21,"GO:0005618","cell wall",21,8,1.93,0.00031,"CC"
22,"GO:0005840","ribosome",119,21,10.93,0.0019,"CC"
23,"GO:0009521","photosystem",21,6,1.93,0.00926,"CC"
24,"GO:0015934","large ribosomal subunit",6,3,0.55,0.01237,"CC"
25,"GO:0044391","ribosomal subunit",7,3,0.64,0.0202,"CC"
26,"GO:0009523","photosystem II",14,4,1.29,0.03305,"CC"

F3 Arriba - A. lanata vs R. pilosa

1, "GO:0006260", "DNA replication", 38, 10, 3.06, 0.00062, "BP"
2, "GO:0006468", "protein phosphorylation", 560, 64, 45.1, 0.00168, "BP"
3, "GO:0000724", "double-strand break repair via homologou...", 2, 2, 0.16, 0.00647, "BP"
4, "GO:0015849", "organic acid transport", 24, 6, 1.93, 0.01012, "BP"
5, "GO:0046942", "carboxylic acid transport", 24, 6, 1.93, 0.01012, "BP"
6, "GO:0015711", "organic anion transport", 32, 7, 2.58, 0.01196, "BP"
7, "GO:0005976", "polysaccharide metabolic process", 32, 7, 2.58, 0.01196, "BP"
8, "GO:0003333", "amino acid transmembrane transport", 19, 5, 1.53, 0.01489, "BP"
9, "GO:0006865", "amino acid transport", 19, 5, 1.53, 0.01489, "BP"
10, "GO:0060249", "anatomical structure homeostasis", 3, 2, 0.24, 0.01837, "BP"
11, "GO:0005524", "ATP binding", 777, 89, 65.47, 0.0011, "MF"
12, "GO:0043531", "ADP binding", 86, 16, 7.25, 0.0019, "MF"
13, "GO:0004672", "protein kinase activity", 558, 65, 47.02, 0.0036, "MF"
14, "GO:0030246", "carbohydrate binding", 61, 12, 5.14, 0.0043, "MF"
15, "GO:0015171", "amino acid transmembrane transporter act...", 14, 5, 1.18, 0.0044, "MF"
16, "GO:0004553", "hydrolase activity, hydrolyzing O-glycos...", 163, 24, 13.73, 0.0047, "MF"
17, "GO:0030695", "GTPase regulator activity", 15, 5, 1.26, 0.0061, "MF"
18, "GO:0016799", "hydrolase activity, hydrolyzing N-glycos...", 10, 4, 0.84, 0.0069, "MF"
19, "GO:0009916", "alternative oxidase activity", 2, 2, 0.17, 0.0071, "MF"
20, "GO:0005085", "guanyl-nucleotide exchange factor activi...", 11, 4, 0.93, 0.0102, "MF"
21, "GO:0044427", "chromosomal part", 41, 8, 2.85, 0.0059, "CC"
22, "GO:0000793", "condensed chromosome", 9, 3, 0.63, 0.0202, "CC"
23, "GO:0106068", "SUMO ligase complex", 4, 2, 0.28, 0.0262, "CC"
24, "GO:0030915", "Smc5-Smc6 complex", 4, 2, 0.28, 0.0262, "CC"
25, "GO:0005875", "microtubule associated complex", 45, 7, 3.13, 0.0326, "CC"
26, "GO:0005634", "nucleus", 379, 35, 26.33, 0.0343, "CC"
27, "GO:0015630", "microtubule cytoskeleton", 56, 8, 3.89, 0.0363, "CC"
28, "GO:0005871", "kinesin complex", 37, 6, 2.57, 0.039, "CC"
29, "GO:0045261", "proton-transporting ATP synthase complex...", 5, 2, 0.35, 0.0417, "CC"
30, "GO:0005856", "cytoskeleton", 79, 10, 5.49, 0.0429, "CC"

F3 Abajo - R. pilosa vs H. amplexicaulis

1, "GO:0006468", "protein phosphorylation", 560, 86, 55.62, 9.5e-06, "BP"
2, "GO:0006355", "regulation of transcription, DNA-templat...", 473, 69, 46.98, 0.00042, "BP"
3, "GO:0048544", "recognition of pollen", 23, 8, 2.28, 0.00114, "BP"
4, "GO:0006508", "proteolysis", 237, 35, 23.54, 0.00973, "BP"
5, "GO:0006020", "inositol metabolic process", 2, 2, 0.2, 0.00985, "BP"
6, "GO:0034637", "cellular carbohydrate biosynthetic proce...", 26, 7, 2.58, 0.01122, "BP"
7, "GO:0016567", "protein ubiquitination", 52, 11, 5.17, 0.01152, "BP"
8, "GO:0032446", "protein modification by small protein co...", 54, 11, 5.36, 0.01524, "BP"
9, "GO:0009070", "serine family amino acid biosynthetic pr...", 3, 2, 0.3, 0.02759, "BP"
10, "GO:0006563", "L-serine metabolic process", 3, 2, 0.3, 0.02759, "BP"
11, "GO:0003700", "DNA-binding transcription factor activit...", 241, 46, 23.46, 4.6e-06, "MF"

12,"GO:0004672","protein kinase activity",558,85,54.32,1e-05,"MF"
 13,"GO:0030247","polysaccharide binding",25,9,2.43,0.00036,"MF"
 14,"GO:0005506","iron ion binding",118,22,11.49,0.00199,"MF"
 15,"GO:0004190","aspartic-type endopeptidase activity",52,12,5.06,0.0035,"MF"
 16,"GO:0005509","calcium ion binding",55,12,5.35,0.00568,"MF"
 17,"GO:0016705","oxidoreductase activity, acting on paire...",118,20,11.49,0.0094,"MF"
 18,"GO:0010333","terpene synthase activity",11,4,1.07,0.01681,"MF"
 19,"GO:0019787","ubiquitin-like protein transferase activ...",63,12,6.13,0.01687,"MF"
 20,"GO:0030246","carbohydrate binding",61,17,5.94,0.01879,"MF"
 21,"GO:0031012","extracellular matrix",2,2,0.15,0.0053,"CC"
 22,"GO:0031224","intrinsic component of membrane",400,40,29.36,0.0157,"CC"
 23,"GO:0016020","membrane",875,76,64.22,0.0201,"CC"
 24,"GO:0016021","integral component of membrane",394,39,28.92,0.0203,"CC"

F3 Arriba - R. pilosa vs H. amplexicaulis

1,"GO:0055114","oxidation-reduction process",495,77,44.52,3.6e-07,"BP"
 2,"GO:0006979","response to oxidative stress",60,19,5.4,5.6e-07,"BP"
 3,"GO:0009690","cytokinin metabolic process",6,4,0.54,0.00084,"BP"
 4,"GO:0006855","drug transmembrane transport",19,6,1.71,0.00501,"BP"
 5,"GO:0006596","polyamine biosynthetic process",5,3,0.45,0.00629,"BP"
 6,"GO:0097164","ammonium ion metabolic process",5,3,0.45,0.00629,"BP"
 7,"GO:0032200","telomere organization",3,2,0.27,0.02276,"BP"
 8,"GO:0060249","anatomical structure homeostasis",3,2,0.27,0.02276,"BP"
 9,"GO:0006308","DNA catabolic process",3,2,0.27,0.02276,"BP"
 10,"GO:0000723","telomere maintenance",3,2,0.27,0.02276,"BP"
 11,"GO:0020037","heme binding",159,40,15.5,9.7e-09,"MF"
 12,"GO:0004601","peroxidase activity",58,18,5.66,5.3e-06,"MF"
 13,"GO:0043531","ADP binding",86,19,8.39,0.00049,"MF"
 14,"GO:0005506","iron ion binding",118,23,11.51,0.00087,"MF"
 15,"GO:0016705","oxidoreductase activity, acting on paire...",118,23,11.51,0.00087,"MF"
 16,"GO:0016747","transferase activity, transferring acyl ...",104,21,10.14,0.00089,"MF"
 17,"GO:0019139","cytokinin dehydrogenase activity",6,4,0.59,0.00114,"MF"
 18,"GO:0009055","electron transfer activity",65,15,6.34,0.00117,"MF"
 19,"GO:0050660","flavin adenine dinucleotide binding",38,9,3.71,0.00924,"MF"
 20,"GO:0015238","drug transmembrane transporter activity",22,6,2.15,0.01605,"MF"
 21,"GO:0005618","cell wall",21,6,1.47,0.0024,"CC"
 22,"GO:0005759","mitochondrial matrix",6,3,0.42,0.0057,"CC"
 23,"GO:0048046","apoplast",10,3,0.7,0.028,"CC"

F5 Abajo – A. lanata vs H. amplexicaulis

1,"GO:0006468","protein phosphorylation",560,107,65.52,2.9e-08,"BP"
 2,"GO:0048544","recognition of pollen",23,10,2.69,0.00012,"BP"
 3,"GO:0006355","regulation of transcription, DNA-templat...",473,75,55.34,0.00254,"BP"
 4,"GO:0005991","trehalose metabolic process",10,5,1.17,0.00328,"BP"

5,"GO:0005992","trehalose biosynthetic process",9,4,1.05,0.01441,"BP"
6,"GO:0046351","disaccharide biosynthetic process",10,4,1.17,0.02183,"BP"
7,"GO:0034637","cellular carbohydrate biosynthetic proce...",26,7,3.04,0.026,"BP"
11,"GO:0043531","ADP binding",86,30,10.72,5.6e-08,"MF"
12,"GO:0004672","protein kinase activity",558,108,69.57,7.1e-07,"MF"
13,"GO:0030247","polysaccharide binding",25,10,3.12,0.00048,"MF"
14,"GO:0016891","endoribonuclease activity, producing 5'-...",10,5,1.25,0.00437,"MF"
15,"GO:0005516","calmodulin binding",19,7,2.37,0.00592,"MF"
16,"GO:0046983","protein dimerization activity",85,19,10.6,0.00738,"MF"
17,"GO:0003700","DNA-binding transcription factor activit...",241,43,30.05,0.00867,"MF"
18,"GO:0008144","drug binding",822,124,102.48,0.00961,"MF"
19,"GO:0032559","adenyl ribonucleotide binding",863,145,107.59,0.01014,"MF"
20,"GO:0030554","adenyl nucleotide binding",865,145,107.84,0.01097,"MF"
21,"GO:0016459","myosin complex",5,3,0.45,0.0063,"CC"
22,"GO:0071944","cell periphery",61,11,5.5,0.0177,"CC"
23,"GO:0017053","transcriptional repressor complex",3,2,0.27,0.0228,"CC"

F5 Arriba – A. lanata vs H. amplexicaulis

1,"GO:0055114","oxidation-reduction process",495,96,63.23,5.9e-06,"BP"
2,"GO:0006979","response to oxidative stress",60,21,7.66,7.4e-06,"BP"
3,"GO:0006108","malate metabolic process",3,3,0.38,0.0021,"BP"
4,"GO:0046942","carboxylic acid transport",24,8,3.07,0.0076,"BP"
5,"GO:0009733","response to auxin",15,6,1.92,0.0076,"BP"
6,"GO:0006259","DNA metabolic process",139,27,17.76,0.0155,"BP"
7,"GO:0015977","carbon fixation",2,2,0.26,0.0163,"BP"
8,"GO:0015980","energy derivation by oxidation of organi...",9,4,1.15,0.0196,"BP"
9,"GO:0045333","cellular respiration",9,4,1.15,0.0196,"BP"
10,"GO:0006865","amino acid transport",19,6,2.43,0.0265,"BP"
11,"GO:0020037","heme binding",159,45,20.4,1.1e-07,"MF"
12,"GO:0004601","peroxidase activity",58,20,7.44,1.8e-05,"MF"
13,"GO:0005506","iron ion binding",118,28,15.14,0.00075,"MF"
14,"GO:0016705","oxidoreductase activity, acting on paire...",118,27,15.14,0.00166,"MF"
15,"GO:0033897","ribonuclease T2 activity",3,3,0.38,0.0021,"MF"
16,"GO:0015171","amino acid transmembrane transporter act...",14,6,1.8,0.00526,"MF"
17,"GO:0016788","hydrolase activity, acting on ester bond...",240,46,30.79,0.00995,"MF"
18,"GO:0016655","oxidoreductase activity, acting on NAD(P...",8,4,1.03,0.01227,"MF"
19,"GO:0008171","O-methyltransferase activity",12,5,1.54,0.01249,"MF"
20,"GO:0008964","phosphoenolpyruvate carboxylase activity",2,2,0.26,0.01644,"MF"
21,"GO:0000786","nucleosome",15,6,1.76,0.0048,"CC"

F5 Abajo – A. lanata vs R. pilosa

1,"GO:0055114","oxidation-reduction process",495,70,48.5,0.00064,"BP"
2,"GO:1901607","alpha-amino acid biosynthetic process",31,9,3.04,0.00217,"BP"

3,"GO:0007018","microtubule-based movement",37,9,3.63,0.00787,"BP"
4,"GO:0006568","tryptophan metabolic process",5,3,0.49,0.00804,"BP"
5,"GO:0071555","cell wall organization",14,5,1.37,0.00836,"BP"
6,"GO:0048366","leaf development",2,2,0.2,0.00958,"BP"
7,"GO:0042545","cell wall modification",10,4,0.98,0.01182,"BP"
8,"GO:0009084","glutamine family amino acid biosynthetic...",6,3,0.59,0.01491,"BP"
9,"GO:0008152","metabolic process",3461,363,339.13,0.01684,"BP"
10,"GO:0009067","aspartate family amino acid biosynthetic...",7,3,0.69,0.02422,"BP"
11,"GO:0043531","ADP binding",86,23,8.84,1.1e-05,"MF"
12,"GO:0050662","coenzyme binding",153,27,15.72,0.0033,"MF"
13,"GO:0003824","catalytic activity",3286,375,337.63,0.0047,"MF"
14,"GO:0003774","motor activity",42,10,4.32,0.0086,"MF"
15,"GO:0016741","transferase activity, transferring one-c...",119,21,12.23,0.0091,"MF"
16,"GO:0030145","manganese ion binding",5,3,0.51,0.0092,"MF"
17,"GO:0008168","methyltransferase activity",113,20,11.61,0.0104,"MF"
18,"GO:0004097","catechol oxidase activity",2,2,0.21,0.0105,"MF"
19,"GO:0008839","4-hydroxy-tetrahydrodipicolinate reducta...",2,2,0.21,0.0105,"MF"
20,"GO:0003942","N-acetyl-gamma-glutamyl-phosphate reduct...",2,2,0.21,0.0105,"MF"
21,"GO:0005871","kinesin complex",37,9,3.09,0.0025,"CC"
22,"GO:0005618","cell wall",21,6,1.75,0.0058,"CC"
23,"GO:0032040","small-subunit processome",7,3,0.58,0.0155,"CC"
24,"GO:0030684","preribosome",9,3,0.75,0.0329,"CC"

F5 Arriba – A. lanata vs R. pilosa

1,"GO:0009733","response to auxin",15,7,1.72,0.00072,"BP"
2,"GO:0046942","carboxylic acid transport",24,8,2.76,0.00398,"BP"
3,"GO:0006570","tyrosine metabolic process",4,3,0.46,0.00553,"BP"
4,"GO:0048523","negative regulation of cellular process",19,6,2.18,0.0164,"BP"
5,"GO:0006865","amino acid transport",19,6,2.18,0.0164,"BP"
6,"GO:0003333","amino acid transmembrane transport",19,6,2.18,0.0164,"BP"
7,"GO:0006508","proteolysis",237,38,27.25,0.01916,"BP"
8,"GO:0055085","transmembrane transport",294,45,33.81,0.02463,"BP"
9,"GO:1905039","carboxylic acid transmembrane transport",21,6,2.41,0.0269,"BP"
10,"GO:1903825","organic acid transmembrane transport",21,6,2.41,0.0269,"BP"
11,"GO:0016758","transferase activity, transferring hexos...",205,41,23.14,0.00016,"MF"
12,"GO:0020037","heme binding",159,32,17.95,0.00072,"MF"
13,"GO:0015171","amino acid transmembrane transporter act...",14,6,1.58,0.00274,"MF"
14,"GO:0005506","iron ion binding",118,24,13.32,0.00277,"MF"
15,"GO:0016705","oxidoreductase activity, acting on paire...",118,24,13.32,0.00277,"MF"
16,"GO:0003700","DNA-binding transcription factor activit...",241,41,27.21,0.00431,"MF"
17,"GO:0030247","polysaccharide binding",25,8,2.82,0.00475,"MF"
18,"GO:0005215","transporter activity",323,54,36.46,0.01099,"MF"
19,"GO:0004809","tRNA (guanine-N2-)-methyltransferase act...",2,2,0.23,0.01273,"MF"

20,"GO:0009916","alternative oxidase activity",2,2,0.23,0.01273,"MF"
21,"GO:0016020","membrane",875,103,87.25,0.0079,"CC"
22,"GO:0005576","extracellular region",18,5,1.79,0.0272,"CC"

F5 Abajo – R. pilosa vs H. amplexicaulis

1,"GO:0006468","protein phosphorylation",560,109,63.27,8.2e-10,"BP"
2,"GO:0048544","recognition of pollen",23,10,2.6,9e-05,"BP"
3,"GO:0006355","regulation of transcription, DNA-templat...",473,75,53.44,0.00094,"BP"
4,"GO:0072488","ammonium transmembrane transport",2,2,0.23,0.01274,"BP"
5,"GO:0015696","ammonium transport",2,2,0.23,0.01274,"BP"
6,"GO:0009733","response to auxin",15,5,1.69,0.02067,"BP"
7,"GO:0009725","response to hormone",32,8,3.62,0.02269,"BP"
8,"GO:0009719","response to endogenous stimulus",32,8,3.62,0.02269,"BP"
9,"GO:0016567","protein ubiquitination",52,11,5.87,0.02793,"BP"
10,"GO:0070647","protein modification by small protein co...",60,12,6.78,0.03335,"BP"
11,"GO:0004672","protein kinase activity",558,109,62.36,7.8e-10,"MF"
12,"GO:0030247","polysaccharide binding",25,15,2.79,5.1e-09,"MF"
13,"GO:0003700","DNA-binding transcription factor activit...",241,51,26.93,3.5e-06,"MF"
14,"GO:0005524","ATP binding",777,113,86.83,0.0012,"MF"
15,"GO:0010333","terpene synthase activity",11,5,1.23,0.0045,"MF"
16,"GO:0016682","oxidoreductase activity, acting on diphe...",4,3,0.45,0.0051,"MF"
17,"GO:0008519","ammonium transmembrane transporter activ...",2,2,0.22,0.0125,"MF"
18,"GO:0009916","alternative oxidase activity",2,2,0.22,0.0125,"MF"
19,"GO:0043565","sequence-specific DNA binding",149,26,16.65,0.0134,"MF"
20,"GO:0005509","calcium ion binding",55,12,6.15,0.0164,"MF"
21,"GO:0016020","membrane",875,86,69.61,0.0027,"CC"

F5 Arriba – R. pilosa vs H. amplexicaulis

1,"GO:0006979","response to oxidative stress",60,20,6.31,1.4e-06,"BP"
2,"GO:0055114","oxidation-reduction process",495,81,52.05,1.4e-05,"BP"
3,"GO:0006412","translation",188,38,19.77,4.2e-05,"BP"
4,"GO:0046131","pyrimidine ribonucleoside metabolic proc...",3,3,0.32,0.0012,"BP"
5,"GO:0009148","pyrimidine nucleoside triphosphate biosy...",2,2,0.21,0.011,"BP"
6,"GO:0046051","UTP metabolic process",2,2,0.21,0.011,"BP"
7,"GO:0006228","UTP biosynthetic process",2,2,0.21,0.011,"BP"
8,"GO:0006241","CTP biosynthetic process",2,2,0.21,0.011,"BP"
9,"GO:0046036","CTP metabolic process",2,2,0.21,0.011,"BP"
10,"GO:0009220","pyrimidine ribonucleotide biosynthetic p...",2,2,0.21,0.011,"BP"
11,"GO:0020037","heme binding",159,41,17.38,7.9e-08,"MF"
12,"GO:0003735","structural constituent of ribosome",121,31,13.22,3.6e-06,"MF"
13,"GO:0004601","peroxidase activity",58,19,6.34,6.4e-06,"MF"
14,"GO:0005506","iron ion binding",118,26,12.9,0.00033,"MF"
15,"GO:0016705","oxidoreductase activity, acting on paire...",118,24,12.9,0.00179,"MF"

16,"GO:0004550","nucleoside diphosphate kinase activity",2,2,0.22,0.01193,"MF"
 17,"GO:0030983","mismatched DNA binding",10,4,1.09,0.01732,"MF"
 18,"GO:0016741","transferase activity, transferring one-c...",119,21,13.01,0.01748,"MF"
 19,"GO:0043531","ADP binding",86,16,9.4,0.02243,"MF"
 20,"GO:0016798","hydrolase activity, acting on glycosyl b...",177,28,19.34,0.02742,"MF"
 21,"GO:0005840","ribosome",119,30,11.33,1.7e-07,"CC"
 22,"GO:1990204","oxidoreductase complex",14,5,1.33,0.0073,"CC"
 23,"GO:0005759","mitochondrial matrix",6,3,0.57,0.0137,"CC"
 24,"GO:0009654","photosystem II oxygen evolving complex",11,4,1.05,0.0154,"CC"
 25,"GO:0042651","thylakoid membrane",13,4,1.24,0.0286,"CC"
 26,"GO:0019898","extrinsic component of membrane",13,4,1.24,0.0286,"CC"
 27,"GO:0009523","photosystem II",14,4,1.33,0.0372,"CC"
 28,"GO:0009521","photosystem",21,5,2,0.0427,"CC"

F7 Abajo – A. lanata vs H. amplexicaulis

1,"GO:0006468","protein phosphorylation",560,107,78.55,0.00022,"BP"
 2,"GO:0048544","recognition of pollen",23,9,3.23,0.00258,"BP"
 3,"GO:0044264","cellular polysaccharide metabolic proces...",25,9,3.51,0.00497,"BP"
 4,"GO:0006073","cellular glucan metabolic process",24,8,3.37,0.0133,"BP"
 5,"GO:0044042","glucan metabolic process",24,8,3.37,0.0133,"BP"
 6,"GO:0015748","organophosphate ester transport",8,4,1.12,0.01681,"BP"
 7,"GO:0015914","phospholipid transport",8,4,1.12,0.01681,"BP"
 8,"GO:0034637","cellular carbohydrate biosynthetic proce...",26,8,3.65,0.02191,"BP"
 9,"GO:0015969","guanosine tetraphosphate metabolic proce...",5,3,0.7,0.02205,"BP"
 10,"GO:0034035","purine ribonucleoside bisphosphate metab...",5,3,0.7,0.02205,"BP"
 11,"GO:0043531","ADP binding",86,32,12.9,3e-07,"MF"
 12,"GO:0005524","ATP binding",777,153,116.58,8.9e-05,"MF"
 13,"GO:0030247","polysaccharide binding",25,12,3.75,9.5e-05,"MF"
 14,"GO:0004672","protein kinase activity",558,108,83.72,0.002,"MF"
 15,"GO:0003774","motor activity",42,14,6.3,0.0023,"MF"
 16,"GO:0042626","ATPase activity, coupled to transmembran...",32,11,4.8,0.0051,"MF"
 17,"GO:0016891","endoribonuclease activity, producing 5'-...",10,5,1.5,0.0098,"MF"
 18,"GO:0008017","microtubule binding",49,14,7.35,0.0107,"MF"
 19,"GO:0003917","DNA topoisomerase type I activity",4,3,0.6,0.012,"MF"
 20,"GO:0015631","tubulin binding",50,14,7.5,0.0129,"MF"
 21,"GO:0016459","myosin complex",5,4,0.61,0.00098,"CC"
 22,"GO:0071944","cell periphery",61,16,7.45,0.00181,"CC"
 23,"GO:0005871","kinesin complex",37,10,4.52,0.0105,"CC"
 24,"GO:0005875","microtubule associated complex",45,11,5.5,0.01614,"CC"
 25,"GO:0030684","preribosome",9,4,1.1,0.01658,"CC"
 26,"GO:0048046","apoplast",10,4,1.22,0.02503,"CC"
 27,"GO:0015630","microtubule cytoskeleton",56,12,6.84,0.03336,"CC"
 28,"GO:0030312","external encapsulating structure",21,6,2.56,0.03452,"CC"
 29,"GO:0005618","cell wall",21,6,2.56,0.03452,"CC"

30,"GO:0017053","transcriptional repressor complex",3,2,0.37,0.04097,"CC"

F7 Arriba – A. lanata vs H. amplexicaulis

1,"GO:0015979","photosynthesis",35,16,5.16,0.00039,"BP"
2,"GO:0003333","amino acid transmembrane transport",19,9,2.8,0.00072,"BP"
3,"GO:0009765","photosynthesis, light harvesting",10,5,1.47,0.00909,"BP"
4,"GO:0050801","ion homeostasis",4,3,0.59,0.01136,"BP"
5,"GO:0006873","cellular ion homeostasis",4,3,0.59,0.01136,"BP"
6,"GO:0006950","response to stress",210,43,30.96,0.01308,"BP"
7,"GO:0009733","response to auxin",15,6,2.21,0.01533,"BP"
8,"GO:0015977","carbon fixation",2,2,0.29,0.02171,"BP"
9,"GO:0006333","chromatin assembly or disassembly",17,6,2.51,0.02924,"BP"
10,"GO:1901136","carbohydrate derivative catabolic proces...",13,5,1.92,0.03172,"BP"
11,"GO:0015171","amino acid transmembrane transporter act...",14,9,1.96,2.1e-05,"MF"
12,"GO:0020037","heme binding",159,42,22.29,2.2e-05,"MF"
13,"GO:0016758","transferase activity, transferring hexos...",205,48,28.74,0.00016,"MF"
14,"GO:0005506","iron ion binding",118,31,16.55,0.00029,"MF"
15,"GO:0016705","oxidoreductase activity, acting on paire...",118,29,16.55,0.00142,"MF"
16,"GO:0008171","O-methyltransferase activity",12,6,1.68,0.00324,"MF"
17,"GO:0016655","oxidoreductase activity, acting on NAD(P...",8,4,1.12,0.01681,"MF"
18,"GO:0008964","phosphoenolpyruvate carboxylase activity",2,2,0.28,0.01964,"MF"
19,"GO:0005088","Ras guanyl-nucleotide exchange factor ac...",6,3,0.84,0.03954,"MF"
20,"GO:0005089","Rho guanyl-nucleotide exchange factor ac...",6,3,0.84,0.03954,"MF"
21,"GO:0000786","nucleosome",15,8,2.14,4e-04,"CC"
22,"GO:0009538","photosystem I reaction center",5,4,0.71,0.0018,"CC"
23,"GO:0009521","photosystem",21,11,3,0.0038,"CC"
24,"GO:0009522","photosystem I",7,6,1,0.0198,"CC"
25,"GO:0042651","thylakoid membrane",13,5,1.86,0.0277,"CC"
26,"GO:0009523","photosystem II",14,5,2,0.0382,"CC"

F7 Abajo – A. lanata vs R. pilosa

1,"GO:0007018","microtubule-based movement",37,16,5.31,1.8e-05,"BP"
2,"GO:1901607","alpha-amino acid biosynthetic process",31,13,4.45,0.00017,"BP"
3,"GO:0006073","cellular glucan metabolic process",24,11,3.45,0.00021,"BP"
4,"GO:0055114","oxidation-reduction process",495,95,71.09,0.00103,"BP"
5,"GO:0009073","aromatic amino acid family biosynthetic ...",18,8,2.59,0.002,"BP"
6,"GO:0009066","aspartate family amino acid metabolic pr...",8,5,1.15,0.00231,"BP"
7,"GO:0006979","response to oxidative stress",60,16,8.62,0.00868,"BP"
8,"GO:0000105","histidine biosynthetic process",7,4,1.01,0.01029,"BP"
9,"GO:0052803","imidazole-containing compound metabolic ...",7,4,1.01,0.01029,"BP"
10,"GO:0006547","histidine metabolic process",7,4,1.01,0.01029,"BP"
11,"GO:0008017","microtubule binding",49,21,7.26,1.9e-06,"MF"
12,"GO:0043531","ADP binding",86,30,12.75,2.6e-06,"MF"
13,"GO:0003777","microtubule motor activity",37,16,5.48,2.8e-05,"MF"

14, "GO:0003824", "catalytic activity", 3286, 561, 487.09, 0.00011, "MF"
 15, "GO:0016762", "xyloglucan:xyloglucosyl transferase acti...", 10, 6, 1.48, 0.00128, "MF"
 16, "GO:0005524", "ATP binding", 777, 144, 115.18, 0.00139, "MF"
 17, "GO:0008168", "methyltransferase activity", 113, 29, 16.75, 0.00166, "MF"
 18, "GO:0003774", "motor activity", 42, 20, 6.23, 0.00202, "MF"
 19, "GO:0047262", "polygalacturonate 4-alpha-galacturonosyl...", 15, 7, 2.22, 0.00332, "MF"
 20, "GO:0004601", "peroxidase activity", 58, 17, 8.6, 0.00341, "MF"
 21, "GO:0005871", "kinesin complex", 37, 16, 4.39, 1.2e-06, "CC"
 22, "GO:0005618", "cell wall", 21, 9, 2.49, 0.00032, "CC"
 23, "GO:0048046", "apoplast", 10, 6, 1.19, 0.00036, "CC"
 24, "GO:0016459", "myosin complex", 5, 4, 0.59, 0.00088, "CC"
 25, "GO:0032040", "small-subunit processome", 7, 4, 0.83, 0.00507, "CC"
 26, "GO:0017053", "transcriptional repressor complex", 3, 2, 0.36, 0.03883, "CC"

F7 Arriba – A. lanata vs R. pilosa

1, "GO:0003333", "amino acid transmembrane transport", 19, 7, 2.33, 0.0053, "BP"
 2, "GO:0006570", "tyrosine metabolic process", 4, 3, 0.49, 0.0066, "BP"
 3, "GO:0016036", "cellular response to phosphate starvatio...", 2, 2, 0.24, 0.015, "BP"
 4, "GO:0009267", "cellular response to starvation", 2, 2, 0.24, 0.015, "BP"
 5, "GO:0015977", "carbon fixation", 2, 2, 0.24, 0.015, "BP"
 6, "GO:0042594", "response to starvation", 2, 2, 0.24, 0.015, "BP"
 7, "GO:0018904", "ether metabolic process", 18, 6, 2.2, 0.0166, "BP"
 8, "GO:0006662", "glycerol ether metabolic process", 18, 6, 2.2, 0.0166, "BP"
 9, "GO:0032501", "multicellular organismal process", 48, 11, 5.87, 0.0272, "BP"
 10, "GO:0009733", "response to auxin", 15, 5, 1.84, 0.0284, "BP"
 11, "GO:0005506", "iron ion binding", 118, 26, 14.42, 0.0018, "MF"
 12, "GO:0016705", "oxidoreductase activity, acting on paire...", 118, 26, 14.42, 0.0018, "MF"
 13, "GO:0015171", "amino acid transmembrane transporter act...", 14, 6, 1.71, 0.0041, "MF"
 14, "GO:0020037", "heme binding", 159, 31, 19.43, 0.0051, "MF"
 15, "GO:0015035", "protein disulfide oxidoreductase activit...", 40, 11, 4.89, 0.0069, "MF"
 16, "GO:0004809", "tRNA (guanine-N2-)-methyltransferase act...", 2, 2, 0.24, 0.0149, "MF"
 17, "GO:0008964", "phosphoenolpyruvate carboxylase activity", 2, 2, 0.24, 0.0149, "MF"
 18, "GO:0003700", "DNA-binding transcription factor activit...", 241, 41, 29.45, 0.0162, "MF"
 19, "GO:0140110", "transcription regulator activity", 253, 42, 30.92, 0.0222, "MF"
 20, "GO:0017048", "Rho GTPase binding", 6, 3, 0.73, 0.0274, "MF"
 21, "GO:0000786", "nucleosome", 15, 6, 1.62, 0.0032, "CC"
 22, "GO:0044421", "extracellular region part", 4, 3, 0.43, 0.0046, "CC"
 23, "GO:0005615", "extracellular space", 2, 2, 0.22, 0.0116, "CC"
 24, "GO:0009523", "photosystem II", 14, 5, 1.51, 0.0124, "CC"
 25, "GO:0005634", "nucleus", 379, 52, 40.98, 0.027, "CC"

F7 Abajo – R. pilosa vs H. amplexicaulis

1, "GO:0006468", "protein phosphorylation", 560, 84, 53.37, 5.9e-06, "BP"
 2, "GO:0048544", "recognition of pollen", 23, 8, 2.19, 0.00086, "BP"

3,"GO:0006355","regulation of transcription, DNA-templat...",473,63,45.08,0.00275,"BP"
4,"GO:0000070","mitotic sister chromatid segregation",8,4,0.76,0.00417,"BP"
5,"GO:0005992","trehalose biosynthetic process",9,4,0.86,0.00695,"BP"
6,"GO:0016567","protein ubiquitination",52,10,4.96,0.02278,"BP"
7,"GO:0006260","DNA replication",38,8,3.62,0.02413,"BP"
8,"GO:0070647","protein modification by small protein co...",60,11,5.72,0.02427,"BP"
9,"GO:0009070","serine family amino acid biosynthetic pr...",3,2,0.29,0.02547,"BP"
10,"GO:0006563","L-serine metabolic process",3,2,0.29,0.02547,"BP"
11,"GO:0004672","protein kinase activity",558,84,54.14,1.6e-05,"MF"
12,"GO:0030247","polysaccharide binding",25,10,2.43,5.8e-05,"MF"
13,"GO:0003700","DNA-binding transcription factor activit...",241,40,23.38,0.00044,"MF"
14,"GO:0010333","terpene synthase activity",11,5,1.07,0.00238,"MF"
15,"GO:0005524","ATP binding",777,98,75.39,0.00267,"MF"
16,"GO:0043531","ADP binding",86,16,8.34,0.00773,"MF"
17,"GO:0005506","iron ion binding",118,20,11.45,0.00907,"MF"
18,"GO:0009916","alternative oxidase activity",2,2,0.19,0.0094,"MF"
19,"GO:0005509","calcium ion binding",55,11,5.34,0.01494,"MF"
20,"GO:0003677","DNA binding",540,67,52.39,0.01811,"MF"
21,"GO:0016020","membrane",875,72,61.76,0.036,"CC"

F7 Arriba – R. pilosa vs H. amplexicaulis

1,"GO:0055114","oxidation-reduction process",495,91,59.13,6.2e-06,"BP"
2,"GO:0006979","response to oxidative stress",60,19,7.17,3.9e-05,"BP"
3,"GO:0071555","cell wall organization",14,6,1.67,0.0036,"BP"
4,"GO:0006525","arginine metabolic process",5,3,0.6,0.0141,"BP"
5,"GO:0008643","carbohydrate transport",5,3,0.6,0.0141,"BP"
6,"GO:0006952","defense response",14,5,1.67,0.019,"BP"
7,"GO:0042545","cell wall modification",10,4,1.19,0.0234,"BP"
8,"GO:0008152","metabolic process",3461,439,413.46,0.0235,"BP"
9,"GO:0006633","fatty acid biosynthetic process",32,8,3.82,0.0307,"BP"
10,"GO:0009308","amine metabolic process",21,6,2.51,0.0318,"BP"
11,"GO:0020037","heme binding",159,44,19.51,8.7e-08,"MF"
12,"GO:0005506","iron ion binding",118,30,14.48,5.9e-05,"MF"
13,"GO:0004601","peroxidase activity",58,18,7.12,0.00012,"MF"
14,"GO:0003824","catalytic activity",3286,476,403.22,0.00021,"MF"
15,"GO:0005507","copper ion binding",18,8,2.21,7e-04,"MF"
16,"GO:0004553","hydrolase activity, hydrolyzing O-glycos...",163,37,20,0.00097,"MF"
17,"GO:0016705","oxidoreductase activity, acting on paire...",118,26,14.48,0.00188,"MF"
18,"GO:0008168","methyltransferase activity",113,25,13.87,0.00215,"MF"
19,"GO:0016747","transferase activity, transferring acyl ...",104,23,12.76,0.00319,"MF"
20,"GO:0004650","polygalacturonase activity",7,4,0.86,0.0058,"MF"
21,"GO:0005618","cell wall",21,7,2,0.0023,"CC"
22,"GO:0031225","anchored component of membrane",6,3,0.57,0.0137,"CC"
23,"GO:0044430","cytoskeletal part",71,13,6.76,0.014,"CC"

24,"GO:0009654","photosystem II oxygen evolving complex",11,4,1.05,0.0154,"CC"
 25,"GO:0005871","kinesin complex",37,8,3.52,0.0199,"CC"
 26,"GO:0005875","microtubule associated complex",45,9,4.29,0.0227,"CC"
 27,"GO:0031226","intrinsic component of plasma membrane",3,2,0.29,0.0254,"CC"
 28,"GO:0042651","thylakoid membrane",13,4,1.24,0.0286,"CC"
 29,"GO:0019898","extrinsic component of membrane",13,4,1.24,0.0286,"CC"
 30,"GO:0043228","non-membrane-bounded organelle",253,33,24.1,0.0293,"CC"

Hoja de datos suplementaria 3: Lista de ortogrupos asociadas a procesos del desarrollo.

Fotorrespiración

Ortogrupo	Ortólogo anotado	Anotación
OG0000295	<i>Sobic.001G065600.1.p,Sobic.006G220500.1.p,Sobic.002G374700.1.p,Sobic.006G220600.3.p</i>	Aldolase-type TIM barrel family protein
OG0000378	<i>Sobic.004G247000.1.p,Sobic.001G116400.2.p,Sobic.001G451500.1.p</i>	glutamine synthase clone R1,glutamine synthase clone F11
OG0000783	<i>Sobic.005G042000.1.p,Sevir.3G319300.1.p</i>	ribulose bisphosphate carboxylase small chain 1A
OG0000964	<i>Sobic.008G039900.1.p</i>	glycine decarboxylase P-protein 2
OG0001962	<i>Sobic.007G226800.1.p,Sobic.002G233700.1.p</i>	dicarboxylate transport 2.1
OG0002384	<i>Sobic.005G113300.1.p,Sobic.008G144800.1.p</i>	serine hydroxymethyltransferase 4
OG0003519	<i>Sobic.007G213700.1.p</i>	alanine:glyoxylate aminotransferase
OG0003992	<i>Sobic.004G055700.2.p</i>	Single hybrid motif superfamily protein
OG0004041	<i>Sobic.003G373600.1.p,GRMZM2G113959_P01</i>	serine hydroxymethyltransferase 7,serine hydroxymethyltransferase 6
OG0004061	<i>Sobic.003G152900.1.p</i>	mitochondrial lipoamide dehydrogenase 1
OG0004294	<i>Sobic.001G097100.3.p</i>	serine transhydroxymethyltransferase 1
OG0004978	<i>GRMZM2G024823_P02,Sobic.001G288900.1.p</i>	Mitochondrial substrate carrier family protein
OG0005385	<i>Sobic.006G220800.1.p</i>	Glycine cleavage T-protein family P-loop containing nucleoside
OG0005556	<i>Sobic.009G004100.1.p</i>	triphosphate hydrolases superfamily protein
OG0006771	<i>Sobic.004G001300.1.p</i>	hydroxypyruvate reductase
OG0007574	<i>Sobic.006G130300.1.p</i>	2-phosphoglycolate phosphatase 1
OG0007822	<i>Sobic.008G062500.1.p</i>	Single hybrid motif superfamily protein
OG0008057	<i>Sobic.001G283000.1.p</i>	membrane protein,putative
OG0010248	<i>Sobic.001G007000.1.p</i>	2-phosphoglycolate phosphatase 2
OG0012065	<i>Sobic.008G112300.1.p</i>	dicarboxylate transporter 1
OG0012541	<i>Sobic.002G006600.1.p</i>	glutamate:glyoxylate aminotransferase
OG0000516	<i>Sobic.004G011566.1.p,Sobic.010G274500.2.p</i>	catalase 1,catalase 2
OG0011137	<i>Sobic.010G024600.1.p</i>	Glycine cleavage T-protein family
OG0001049	<i>Sobic.003G258800.1.p</i>	NADH-dependent glutamate synthase 1

Ciclo de CBB

Ortogrupo	Ortólogo anotado	Anotación
OG0000622	<i>Sobic.010G188300.1.p,Sobic.003G393900.1.p</i>	Aldolase superfamily protein
OG0000741	<i>Sobic.009G183700.1.p,Sevir.5G354000.2.p,Sobic.010G221800.1.p</i>	Phosphoglycerate kinase family protein,phosphoglycerate kinase
OG0000759	<i>Sobic.003G172800.2.p</i>	ribulose-bisphosphate carboxylases
OG0000783	<i>Sobic.005G042000.1.p,Sevir.3G319300.1.p</i>	ribulose bisphosphate carboxylase small chain 1A
OG0000861	<i>Sobic.008G053200.1.p,Sobic.005G056400.1.p</i>	fructose-bisphosphate aldolase 2
OG0001568	<i>Sobic.005G231600.1.p</i>	rubisco activase
OG0001652	<i>Sobic.001G003100.1.p,Sobic.009G098100.1.p</i>	chaperonin-60alpha
OG0002203	<i>Sobic.004G002200.1.p,Sobic.010G011200.1.p</i>	TCP-1/cpn60 chaperonin family protein,chaperonin 60 beta
OG0002237	<i>Sobic.006G034300.1.p,Sobic.010G023700.1.p</i>	Transketolase
OG0003654	<i>Sobic.004G272100.1.p</i>	phosphoribulokinase
OG0004555	<i>Sobic.006G105900.1.p</i>	glyceraldehyde 3-phosphate dehydrogenase A subunit 2
OG0005782	<i>Sobic.001G425400.1.p</i>	high cyclic electron flow 1
OG0006166	<i>Sobic.003G359100.1.p</i>	sedoheptulose-bisphosphatase
OG0007543	<i>Sobic.003G096000.2.p</i>	Aldolase superfamily protein
OG0007618	<i>GRMZM2G480850_P01</i>	CP12 domain-containing protein 1
OG0008019	<i>Sobic.002G277100.1.p</i>	triosephosphate isomerase
OG0010329	<i>Sobic.001G236800.1.p</i>	phosphoglycerate kinase 1
OG0010691	<i>Sobic.001G069000.1.p</i>	NagB/RpiA/CoA transferase-like superfamily protein
OG0012267	<i>Sobic.006G248700.2.p</i>	P-loop containing nucleoside triphosphate hydrolases superfamily protein
OG0012671	<i>Sobic.002G051100.1.p</i>	Ribose 5-phosphate isomerase,type A protein
OG0014845	<i>Sobic.001G519800.1.p</i>	glyceraldehyde-3-phosphate dehydrogenase B subunit

Fotosíntesis

Ortogrupo	Ortólogo anotado	Anotación
OG0000081	<i>Sobic.002G339200.2.p,Sobic.002G289100.1.p,Sobic.002G288300.1.p,Sobic.009G234600.1.p,GRMZM2G105518_P01</i>	light-harvesting chlorophyll B-binding protein 3,light-harvesting chlorophyll-protein complex II subunit B1,photosystem II light harvesting complex gene B1B2,photosystem II light harvesting complex gene 2.1
OG0000278	<i>Sobic.010G160700.1.p,Sobic.003G301800.1.p,Sobic.004G106900.1.p</i>	phosphoenolpyruvate carboxylase 2,phosphoenolpyruvate carboxylase 3
OG0000783	<i>Sobic.005G042000.1.p,Sevir.3G319300.1.p</i>	ribulose bisphosphate carboxylase small chain 1A
OG0000861	<i>Sobic.008G053200.1.p,Sobic.005G056400.1.p</i>	fructose-bisphosphate aldolase 2
OG0001093	<i>Sobic.002G052100.1.p,Sobic.002G052000.2.p</i>	Chlorophyll A-B binding family protein
OG0001136	<i>Sobic.007G065500.1.p,Sobic.002G322000.2.p</i>	glucose-6-phosphate/phosphate translocator 2
OG0001317	<i>Sobic.003G367500.1.p,Sobic.009G152700.1.p</i>	Inositol monophosphatase family protein
OG0001553	<i>Sobic.003G234500.1.p</i>	carbonic anhydrase 2
OG0001568	<i>Sobic.005G231600.1.p</i>	rubisco activase
OG0001652	<i>Sobic.001G003100.1.p,Sobic.009G098100.1.p</i>	chaperonin-60alpha

OG0001703	<i>Sobic.002G026300.1.p,Sobic.002G160300.1.p</i>	<i>Phosphofructokinase family protein</i>
OG0001724	<i>Sobic.002G330500.1.p,Sevir.2G354200.1.p,Sobic.009G137700.1.p</i>	<i>isopentenyl pyrophosphate:dimethylallyl pyrophosphate isomerase 2,isopentenyl diphosphate isomerase 1</i>
OG0001725	<i>Sobic.003G055300.1.p,Sobic.002G414400.1.p</i>	<i>photosystem II subunit X</i>
OG0001814	<i>Sobic.004G238500.1.p,Sobic.003G122100.1.p</i>	<i>Pyridine nucleotide-disulphide oxidoreductase family protein</i>
OG0002054	<i>Sobic.010G005500.1.p</i>	<i>ferredoxin-NADP(+)-oxidoreductase 2</i>
OG0002068	<i>Sobic.001G005200.2.p</i>	<i>PGR5-LIKE A</i>
OG0002204	<i>Sobic.009G132900.1.p</i>	<i>pyruvate orthophosphate dikinase</i>
OG0002462	<i>Sevir.2G035400.1.p,Sobic.002G033400.1.p</i>	<i>root FNR 2</i>
OG0002475	<i>Sobic.008G088821.1.p,Sobic.009G246400.1.p</i>	<i>solaneyl diphosphate synthase 1,solaneyl diphosphate synthase 2</i>
OG0002605	<i>Sobic.007G071500.1.p</i>	<i>photosystem II subunit R</i>
OG0003214	<i>Sobic.008G051000.1.p</i>	<i>magnesium chelatase i2</i>
OG0003469	<i>Sobic.008G063500.1.p</i>	<i>photosystem I reaction center subunit PSI-N,chloroplast,putative / PSI-N,putative (PSAN)</i>
OG0003574	<i>Sobic.004G056400.1.p</i>	<i>glyceraldehyde-3-phosphate dehydrogenase of plastid 2</i>
OG0003592	<i>Sobic.003G313000.1.p</i>	<i>photosystem II reaction center W</i>
OG0003643	<i>Sobic.001G046900.1.p</i>	<i>photosystem I subunit D-2</i>
OG0003792	<i>Sobic.008G004800.1.p</i>	<i>ferritin 4</i>
OG0003966	<i>Sobic.002G329600.1.p</i>	<i>photosystem II subunit Q-2</i>
OG0003992	<i>Sobic.004G055700.2.p</i>	<i>Single hybrid motif superfamily protein</i>
OG0004318	<i>Sobic.002G338000.1.p</i>	<i>light harvesting complex photosystem II</i>
OG0004496	<i>Sobic.010G184600.1.p</i>	<i>Plant haem oxygenase (decyclizing) family protein</i>
OG0004520	<i>Sobic.001G079500.1.p</i>	<i>salt tolerance zinc finger</i>
OG0004555	<i>Sobic.006G105900.1.p</i>	<i>glyceraldehyde 3-phosphate dehydrogenase A subunit 2</i>
OG0004571	<i>Sobic.007G023300.1.p</i>	<i>photosystem II BY</i>
OG0004614	<i>Sobic.004G261800.1.p</i>	<i>Phosphofructokinase family protein</i>
OG0004657	<i>Sobic.009G143100.1.p</i>	<i>photosynthetic electron transfer C</i>
OG0004918	<i>Sobic.004G193400.1.p</i>	<i>Photosystem II 5 kD protein</i>
OG0005003	<i>Sobic.003G370000.1.p</i>	<i>Chlorophyll A-B binding family protein</i>
OG0005295	<i>Sobic.002G352100.1.p</i>	<i>photosystem I light harvesting complex gene 2</i>
OG0005335	<i>GRMZM2G451224_P01,Sobic.009G229700.1.p</i>	<i>photosystem I subunit H2</i>
OG0005399	<i>Sobic.006G264900.1.p</i>	
OG0005402	<i>Sobic.010G229400.1.p</i>	<i>geranyl diphosphate synthase 1</i>
OG0005462	<i>Sobic.004G023400.1.p</i>	<i>low psii accumulation2</i>
OG0005771	<i>Sobic.007G136900.1.p</i>	<i>light-harvesting chlorophyll-protein complex I subunit A4</i>
OG0005782	<i>Sobic.001G425400.1.p</i>	<i>high cyclic electron flow 1</i>
OG0005886	<i>Sobic.001G417200.1.p</i>	<i>ATPase,F0 complex,subunit B/B',bacterial/chloroplast</i>
OG0005896	<i>Sobic.002G242000.1.p</i>	<i>photosystem I subunit G</i>
OG0005907	<i>Sobic.010G101600.1.p</i>	<i>Phosphofructokinase family protein</i>
OG0006162	<i>Sobic.001G394400.1.p</i>	<i>phytochrome B</i>
OG0006224	<i>Sobic.002G047400.1.p</i>	<i>protein containing PDZ domain,a K-box domain,and a TPR region</i>

OG0006250	GRMZM2G043109_P01	dicarboxylate diiron protein, putative (Crd1)
OG0006426	Sobic.010G241000.1.p	photosystem II subunit O-2
OG0006664	Sobic.001G179100.1.p	post-illumination chlorophyll fluorescence increase
OG0006730	Sobic.002G033800.1.p	one helix protein
OG0006771	Sobic.004G001300.1.p	hydroxypyruvate reductase
OG0006822	Sobic.006G016900.1.p	cysteine synthase C1
OG0006924	Sobic.004G126000.1.p	Pyridine nucleotide-disulphide oxidoreductase family protein
OG0006949	Sobic.004G227200.1.p	ferredoxin/thioredoxin reductase subunit A (variable subunit) 2
OG0006989	Sobic.001G289500.1.p	pfkB-like carbohydrate kinase family protein
OG0007358	Sobic.001G041800.1.p	ALBINA 1
OG0007437	Sobic.002G119500.3.p	photosystem I subunit E-2
OG0007507	Sobic.004G336200.1.p	CLP protease proteolytic subunit 6
OG0007618	GRMZM2G480850_P01	CP12 domain-containing protein 1
OG0007652	Sobic.004G191200.1.p	PsbQ-like 2
OG0007822	Sobic.008G062500.1.p	Single hybrid motif superfamily protein
OG0007870	Sobic.001G087100.1.p	phytochrome C
OG0008135	Sobic.003G110100.1.p	O-acetylserine (thiol) lyase B
OG0008169	Sobic.004G235200.1.p	ATP synthase delta-subunit gene
OG0008260	Sobic.004G151600.1.p	FAD/NAD(P)-binding oxidoreductase
OG0008816	Sobic.007G027000.1.p	
OG0008830	Sobic.008G179900.1.p	O-acetylserine (thiol) lyase (OAS-TL) isoform A1
OG0009019	Sobic.009G225600.1.p	2Fe-2S ferredoxin-like superfamily protein
OG0009148	Sobic.001G078000.1.p	cytochrome b6f complex subunit (petM), putative
OG0009162	Sobic.001G070500.1.p	photosystem I subunit F
OG0009555	Sobic.003G203000.1.p	one-helix protein 2
OG0009577	Sobic.002G316700.1.p	ATPase, F1 complex, gamma subunit protein
OG0009639	Sobic.003G221700.1.p	Mog1/PsbP/DUF1795-like photosystem II reaction center PsbP family protein
OG0009902	Sobic.009G137000.1.p	FAD/NAD(P)-binding oxidoreductase family protein
OG0010329	Sobic.001G236800.1.p	phosphoglycerate kinase 1
OG0010635	Sobic.009G219100.1.p	STT7 homolog STN7
OG0010691	Sobic.001G069000.1.p	NagB/RpiA/CoA transferase-like superfamily protein
OG0010890	Sobic.010G277600.1.p	photosystem II stability/assembly factor, chloroplast (HCF136)
OG0011043	Sobic.003G433900.1.p	NAD(P)H:plastoquinone dehydrogenase complex subunit O
OG0011222	Sobic.010G027200.1.p	CLP protease proteolytic subunit 2
OG0011330	Sobic.007G212400.2.p	Mog1/PsbP/DUF1795-like photosystem II reaction center PsbP family protein
OG0011719	Sobic.005G087000.1.p	light harvesting complex of photosystem II 5
OG0011754	Sobic.006G214700.1.p	Arabidopsis thaliana protein of unknown function (DUF794)
OG0011853	Sobic.006G162200.1.p	subunit NDH-M of NAD(P)H:plastoquinone dehydrogenase complex
OG0011875	Sobic.006G084500.1.p	Photosystem II reaction center PsbP family protein
OG0011917	Sobic.002G198600.1.p	inorganic carbon transport protein-related

OG0011936	<i>Sobic.009G130800.1.p</i>	<i>Mog1/PsbP/DUF1795-like photosystem II reaction center PsbP family protein</i>
OG0012017	<i>Sobic.001G347800.1.p</i>	<i>heme oxygenase 2</i>
OG0012212	<i>Sobic.002G159100.1.p</i>	<i>Chlorophyll A-B binding family protein</i>
OG0012267	<i>Sobic.006G248700.2.p</i>	<i>P-loop containing nucleoside triphosphate hydrolases superfamily protein</i>
OG0012298	<i>Sobic.010G189300.1.p</i>	<i>photosystem I light harvesting complex gene 3</i>
OG0012301	<i>Sobic.003G328200.1.p</i>	<i>Mog1/PsbP/DUF1795-like photosystem II reaction center PsbP family protein</i>
OG0012377	<i>Sobic.002G215000.2.p</i>	<i>photosystem I light harvesting complex gene 6</i>
OG0012444	<i>Sobic.002G030200.1.p</i>	<i>photosystem II subunit P-1</i>
OG0012496	<i>Sobic.002G305700.1.p</i>	<i>NDH-dependent cyclic electron flow 1</i>
OG0012671	<i>Sobic.002G051100.1.p</i>	<i>Ribose 5-phosphate isomerase, type A protein</i>
OG0013713	<i>Sobic.006G073500.1.p</i>	<i>photosystem I subunit O</i>
OG0013743	<i>Sobic.001G415400.1.p</i>	<i>PsbP-like protein 2</i>
OG0014227	<i>Sobic.004G056900.1.p</i>	<i>photosystem I light harvesting complex gene 1</i>
OG0014374	<i>Sobic.003G052500.1.p</i>	<i>photosystem I subunit I</i>
OG0015229	<i>GRMZM2G384988_P01</i>	<i>ATP synthase subunit beta</i>
OG0016280	<i>Sobic.003G378600.1.p</i>	<i>oxidoreductases, acting on NADH or NADPH, quinone or similar compound as acceptor</i>

Ciclo C4

Ortogrupo	Ortólogo anotado	Anotación
OG0000342	<i>Sobic.009G152600.1.p, Sobic.006G276700.2.p, Sobic.004G268000.1.p</i>	<i>pyrophosphorylase 3, pyrophosphorylase 4</i>
OG0000495	<i>Sobic.009G108700.1.p, Sobic.009G069600.1.p</i>	<i>NADP-malic enzyme 4, NADP-malic enzyme 3</i>
OG0001149	<i>Sobic.002G375800.1.p</i>	<i>alanine aminotransferase 2</i>
OG0001553	<i>Sobic.003G234500.1.p</i>	<i>carbonic anhydrase 2</i>
OG0002204	<i>Sobic.009G132900.1.p</i>	<i>pyruvate orthophosphate dikinase</i>
OG0002332	<i>Sobic.002G159900.2.p</i>	<i>Glucose-6-phosphate/phosphate translocator-related</i>
OG0002544	<i>Sobic.003G431900.1.p, Sobic.009G124000.1.p</i>	<i>Protein of unknown function (DUF3411)</i>
OG0002658	<i>Sobic.003G002300.1.p</i>	<i>Glucose-6-phosphate/phosphate translocator-related</i>
OG0002688	<i>Sobic.003G050800.1.p, Sobic.009G062500.1.p</i>	<i>Glucose-6-phosphate/phosphate translocator-related</i>
OG0003087	<i>Sobic.009G240700.1.p</i>	<i>Lactate/malate dehydrogenase family protein</i>
OG0003499	<i>Sobic.004G219900.1.p</i>	<i>phosphoenolpyruvate carboxylase kinase 1</i>
OG0003989	<i>Sobic.007G166200.1.p</i>	<i>lactate/malate dehydrogenase family protein</i>
OG0004015	<i>Sobic.002G324400.1.p</i>	<i>PPDK regulatory protein</i>
OG0004441	<i>Sobic.003G236800.1.p</i>	<i>Sodium Bile acid symporter family</i>
OG0005162	<i>Sobic.007G137600.1.p</i>	<i>malate dehydrogenase</i>
OG0005953	<i>Sobic.001G432800.1.p</i>	<i>phosphoenolpyruvate carboxykinase 1</i>
OG0006181	<i>Sobic.007G113600.2.p</i>	<i>adenylate kinase family protein</i>
OG0006537	<i>Sobic.004G165000.1.p</i>	<i>Sodium Bile acid symporter family</i>
OG0009009	<i>Sobic.002G309400.1.p</i>	<i>NAD-dependent malic enzyme 1</i>

OG0011845	<i>Sobic.001G201700.1.p</i>	NAD-dependent malic enzyme 2
OG0012060	<i>Sobic.002G385700.2.p</i>	malate dehydrogenase
OG0012294	<i>Sobic.004G331700.1.p</i>	aspartate aminotransferase 5
OG0012980	<i>Sobic.004G311100.1.p</i>	pyrophosphorylase 6
OG0015176	<i>Sobic.004G338000.1.p</i>	phosphoenolpyruvate carboxylase kinase 1
OG0000278	<i>Sobic.010G160700.1.p,Sobic.003G301800.1.p,Sobic.004G106900.1.p</i>	phosphoenolpyruvate carboxylase 2,phosphoenolpyruvate carboxylase 3
OG0014652	<i>Sobic.003G100600.1.p</i>	phosphoenolpyruvate carboxylase 4

Desarrollo de HV

Ortogrupo	Ortólogo anotado	Anotación
OG0000210	<i>GRMZM2G157596_P01,Sobic.001G040800.1.p,GRMZM2G357399_P01</i>	ADP-ribosylation factor A1E,ADP-ribosylation factor A1F
OG0000398	<i>Sobic.010G051000.1.p,Sobic.004G240900.1.p,Sobic.005G025900.1.p</i>	Auxin efflux carrier family protein
OG0001588	<i>Sobic.009G207000.1.p,Sobic.004G122000.2.p</i>	Protein of unknown function (DUF502),like COV 1
OG0001885	<i>Sobic.005G017500.2.p,Sobic.008G023401.1.p</i>	GRAS family transcription factor
OG0001978	<i>Sobic.008G168100.1.p,Sobic.001G157400.1.p</i>	Homeobox-leucine zipper family protein / lipid-binding START domain-containing protein
OG0002523	<i>Sobic.010G229000.1.p,Sobic.004G051900.1.p</i>	auxin response factor 6,auxin response factor 8
OG0002834	<i>Sobic.002G260700.1.p,Sobic.007G189100.1.p</i>	ARF GTPase-activating protein
OG0003287	<i>Sobic.002G249200.1.p</i>	Auxin efflux carrier family protein
OG0003523	<i>Sobic.001G537300.2.p</i>	Homeobox-leucine zipper family protein / lipid-binding START domain-containing protein
OG0003915	<i>Sobic.002G357800.1.p</i>	GRAS family transcription factor
OG0007241	<i>Sobic.003G336100.2.p</i>	Auxin efflux carrier family protein
OG0008202	<i>Sobic.002G108800.1.p</i>	Auxin efflux carrier family protein
OG0010840	<i>Sobic.003G028300.1.p</i>	Homeobox-leucine zipper family protein / lipid-binding START domain-containing protein
OG0011147	<i>Sobic.004G210900.1.p</i>	Auxin efflux carrier family protein
OG0012068	<i>Sobic.002G342800.1.p</i>	GRAS family transcription factor
OG0012184	<i>Sobic.004G058700.1.p</i>	Homeodomain-like superfamily protein
OG0014246	<i>Sobic.006G255300.1.p</i>	Transcriptional factor B3 family protein / auxin-responsive factor AUX/IAA-related
OG0014482	<i>Sobic.007G038000.1.p</i>	Leucine-rich repeat protein kinase family protein
OG0015235	<i>Sobic.006G182000.1.p</i>	transcription factor-related
OG0018305	<i>Sobic.003G235800.1.p</i>	Auxin efflux carrier family protein

Expansión de la pared celular

Ortogrupo	Ortólogo anotado	Anotación
OG0000032	<i>Sobic.007G166500.1.p,Sobic.003G059900.1.p,Sobic.003G111800.1.p,Sobic.003G338801.1.p,Sobic.004G238801.1.p,Sobic.001G033300.1.p,Sobic.004G120</i>	expansin A16,expansin A17,expansin A1,expansin A4,expansin 11

OG0000053	000.1.p,Sobic.006G031900.1.p,Sobic.001G499900.1.p,Sobic.001G499800.1.p Sobic.001G301000.1.p,Sobic.001G542100.1.p,GRM ZM2G703152_P02,Sobic.001G155600.3.p,Sobic.00 4G294500.1.p,Sobic.001G301300.1.p,Sobic.001G15 5700.5.p	expansin B2,expansin B4
OG0000659	Sobic.009G002600.1.p,Sobic.009G002500.1.p,Sobic .009G002400.2.p	polygalacturonase inhibiting protein 1
OG0000722	Sobic.010G246600.1.p,Sobic.010G246400.1.p	Xyloglucan endotransglucosylase/hydrolase family protein
OG0000729	Sobic.002G370300.1.p,Sobic.006G172000.1.p,Sobic .003G292100.1.p,Sobic.009G203200.1.p	Pectin lyase-like superfamily protein
OG0000748	Sobic.003G094800.1.p,Sobic.003G384800.3.p	Pectinacylesterase family protein
OG0000825	Sobic.009G194200.1.p,Sobic.001G075600.1.p	Cellulose-synthase-like C12
OG0000891	GRMZM2G128876_P02,Sobic.007G090436.1.p	xyloglucan endotransglucosylase/hydrolase 25,Xyloglucan endotransglucosylase/hydrolase family protein
OG0000901	Sobic.005G176100.1.p,Sobic.003G087900.1.p	Mannose-6-phosphate isomerase,type I
OG0001549	Sevir.1G122100.1.p	fucosyltransferase 1
OG0001800	GRMZM2G349428_P01,Sobic.003G445000.1.p,Sevi r.5G475800.1.p,Sobic.009G013700.1.p	Pectinacylesterase family protein
OG0001858	Sobic.001G284600.1.p	endoxyloglucan transferase A3
OG0002073	Sobic.001G089100.3.p	Glycosyl hydrolase family protein
OG0002195	Sobic.003G148400.1.p,Sobic.002G416000.1.p	pectin methylesterase 2
OG0002217	Sobic.002G237900.1.p	cellulose synthase like E1
OG0002300	Sobic.001G414650.1.p,Sobic.003G348800.1.p	UDP-glucuronic acid decarboxylase 1
OG0002366	Sobic.003G304600.1.p,Sobic.001G173300.1.p	reversibly glycosylated polypeptide 1,reversibly glycosylated polypeptide 2
OG0002682	Sobic.001G026800.1.p,Sobic.002G090400.1.p	Pectin lyase-like superfamily protein
OG0002884	Sobic.009G060400.1.p,Sobic.003G052400.1.p	Protein of unknown function (DUF1421)
OG0003056	Sobic.001G063500.1.p	FASCICLIN-like arabinogalactan protein 17 precursor
OG0003091	Sobic.001G086000.1.p	COBRA-like extracellular glycosyl-phosphatidyl inositol-anchored protein family
OG0003234	Sobic.002G049000.1.p,GRMZM2G170934_P01	FRIGIDA-like protein
OG0003338	Sobic.009G243500.1.p,Sobic.003G232600.1.p	Pectin lyase-like superfamily protein
OG0003588	Sobic.010G197300.1.p,Sobic.004G075900.1.p	Nucleotide-diphospho-sugar transferases superfamily protein,cellulose synthase-like A02
OG0003789	Sobic.003G142100.1.p,Sobic.010G152900.1.p	beta-xylosidase 2,beta-D-xylosidase 4
OG0003825	Sobic.003G223100.1.p	Pectin lyase-like superfamily protein
OG0004375	Sobic.006G014400.1.p	Pectin lyase-like superfamily protein
OG0004508	Sobic.006G097900.1.p	fucosyltransferase 1
OG0004519	Sobic.001G027800.2.p	Glycosyl hydrolase superfamily protein
OG0004544	Sobic.003G082700.1.p	glycosyl hydrolase family 10 protein / carbohydrate- binding domain-containing protein
OG0005055	Sobic.005G110436.1.p	Glycosyl hydrolase family protein
OG0005161	Sobic.009G055900.1.p	FASCICLIN-like arabinogalactan 2
OG0005363	Sobic.003G041300.1.p	FRIGIDA-like protein
OG0005510	Sobic.009G111000.2.p	Plant invertase/pectin methylesterase inhibitor superfamily
OG0005762	Sobic.006G168900.1.p	hydroxyproline-rich glycoprotein family protein
OG0006148	Sobic.005G211500.1.p	beta-xylosidase 2

OG0006323	<i>Sobic.004G227900.1.p</i>	<i>expansin B3</i>
OG0007028	<i>Sobic.009G157300.1.p</i>	<i>Glycosyl hydrolase family protein</i>
OG0007656	<i>Sobic.002G302000.1.p</i>	<i>xyloglucan endotransglucosylase/hydrolase 32</i>
OG0007869	<i>Sobic.004G308200.1.p</i>	<i>fucosyltransferase 1</i>
OG0008003	<i>Sobic.001G297000.1.p</i>	<i>hydroxyproline-rich glycoprotein family protein</i>
OG0008059	<i>Sobic.003G017400.1.p</i>	<i>phosphoglucosamine mutase-related</i>
OG0008215	<i>Sobic.003G252700.1.p</i>	<i>Protein of unknown function (DUF1421)</i>
OG0008259	<i>Sobic.004G129400.1.p,Sobic.003G005400.1.p</i>	<i>Plant invertase/pectin methylesterase inhibitor superfamily,pectin methylesterase PCR fragment F</i>
OG0008713	<i>Sobic.006G191700.1.p</i>	<i>expansin A13</i>
OG0008762	<i>AC200493.4_FGP001</i>	<i>Glycosyl hydrolase family protein</i>
OG0008878	<i>Sobic.004G151400.1.p</i>	<i>Nucleotide-diphospho-sugar transferases superfamily protein</i>
OG0009225	<i>Sobic.001G330500.1.p</i>	<i>polygalacturonase inhibiting protein 1</i>
OG0009344	<i>Sobic.001G050500.1.p</i>	<i>Galactosyltransferase family protein</i>
OG0010160	<i>Sobic.002G385800.1.p</i>	<i>Nucleotide-diphospho-sugar transferases superfamily protein</i>
OG0010467	<i>Sobic.002G214900.1.p</i>	<i>Pectin lyase-like superfamily protein</i>
OG0010544	<i>Sobic.002G271200.1.p</i>	<i>UDP-D-glucose/UDP-D-galactose 4-epimerase 1</i>
OG0010643	<i>Sobic.003G442300.1.p</i>	<i>UDP-D-apiose/UDP-D-xylose synthase 2</i>
OG0010669	<i>Sobic.003G294900.2.p</i>	
OG0010753	<i>Sobic.002G068800.1.p</i>	<i>Galactosyltransferase family protein</i>
OG0011075	<i>Sobic.002G154600.1.p</i>	<i>FASCICLIN-like arabinogalactan 2</i>
OG0011236	<i>Sobic.003G299000.1.p</i>	<i>Protein of unknown function (DUF1421)</i>
OG0011772	<i>Sobic.003G331000.1.p</i>	<i>Exostosin family protein</i>
OG0011773	<i>Sobic.004G047400.2.p</i>	<i>Sugar isomerase (SIS) family protein</i>
OG0011888	<i>Sobic.009G026101.1.p</i>	<i>Nucleotide-diphospho-sugar transferases superfamily protein</i>
OG0011955	<i>Sobic.001G436100.1.p</i>	<i>UDP-D-glucuronate 4-epimerase 2</i>
OG0012524	<i>Sobic.001G409100.1.p</i>	<i>Nucleotide-diphospho-sugar transferases superfamily protein</i>
OG0012527	<i>Sobic.007G137400.1.p</i>	<i>Nucleotide-diphospho-sugar transferases superfamily protein</i>
OG0013222	<i>Sobic.002G343600.1.p</i>	<i>Leucine-rich repeat (LRR) family protein</i>
OG0013231	<i>Sobic.004G237800.1.p</i>	<i>Nucleotide-diphospho-sugar transferases superfamily protein</i>
OG0013630	<i>Sobic.006G186300.2.p</i>	<i>FASCICLIN-like arabinogalactan protein 8</i>
OG0013734	<i>Sobic.004G296100.1.p</i>	<i>hydroxyproline-rich glycoprotein family protein</i>
OG0013751	<i>Sobic.004G233700.1.p</i>	<i>Glycosyl hydrolase family protein</i>
OG0013888	<i>Sobic.004G308400.1.p</i>	<i>fucosyltransferase 1</i>
OG0013985	<i>Sobic.009G231100.1.p</i>	<i>Nucleotide-diphospho-sugar transferases superfamily protein</i>
OG0014053	<i>Sobic.003G247000.1.p</i>	<i>Glycosyl hydrolase superfamily protein</i>
OG0014339	<i>Sobic.003G376900.1.p</i>	<i>Plant invertase/pectin methylesterase inhibitor superfamily</i>
OG0014434	<i>Sobic.004G113900.1.p</i>	<i>Pectin lyase-like superfamily protein</i>
OG0014668	<i>Sobic.010G071300.1.p</i>	<i>Plant invertase/pectin methylesterase inhibitor superfamily</i>
OG0014819	<i>Sobic.006G235600.1.p</i>	<i>beta-D-xylosidase 4</i>

OG0016022	Sobic.009G070500.1.p	Leucine-rich repeat (LRR) family protein
OG0016240	Sobic.003G321200.1.p	methylesterase PCR A
OG0016301	Sobic.002G333800.1.p	cellulose-synthase like D2
OG0016544	Sobic.004G254300.1.p	FASCICLIN-like arabinogalactan 2
OG0016597	Sobic.002G430700.1.p	Nucleotide-diphospho-sugar transferases superfamily protein
OG0016856	Sobic.006G204500.1.p	Pectinacetylerase family protein
OG0016873	Sobic.004G269400.1.p	Pectinacetylerase family protein
OG0017049	Sobic.004G308600.1.p	fucosyltransferase 1
OG0017075	Sobic.004G238700.1.p	Nucleotide-diphospho-sugar transferases superfamily protein
OG0017208	Sobic.007G146200.1.p	Plant invertase/pectin methylesterase inhibitor superfamily
OG0017357	Sobic.001G179400.1.p	xyloglucan endotransglucosylase/hydrolase 9
OG0017394	Sobic.010G092400.1.p	Nucleotide-diphospho-sugar transferases superfamily protein
OG0018076	Sobic.001G521850.1.p	FASCICLIN-like arabinogalactan-protein 10
OG0018417	Sobic.010G246500.1.p	Xyloglucan endotransglucosylase/hydrolase family protein
OG0018418	Sobic.002G324100.1.p	xyloglucan endotransglucosylase/hydrolase 8
OG0018439	Sobic.005G110460.1.p	Glycosyl hydrolase family protein

Biosíntesis de suberina

Ortogrupo	Ortólogo anotado	Anotación
OG0000296	Sobic.002G134100.1.p,Sobic.007G170100.2.p,Sobic.007G170800.1.p,Sobic.008G076300.1.p,Sobic.004G060700.1.p	fatty acid reductase 1,fatty acid reductase 4
OG0000610	Sobic.010G055300.2.p,Sobic.008G015000.1.p	cytochrome P450,family 94,subfamily D,polypeptide 2
OG0004343	Sobic.001G450100.1.p	cytochrome P450,family 94,subfamily D,polypeptide 2
OG0004648	Sobic.006G177000.1.p	cytochrome P450,family 86,subfamily A,polypeptide 4
OG0005527	Sobic.009G022000.1.p	ABC-2 type transporter family protein
OG0006046	Sobic.010G202600.1.p	RING/U-box superfamily protein
OG0006384	Sobic.003G387300.1.p	ABC-2 type transporter family protein
OG0008734	Sobic.003G348100.1.p	ABC-2 type transporter family protein
OG0013822	Sobic.003G220000.1.p	ABC-2 type transporter family protein
OG0014877	Sobic.009G122300.1.p	ABC-2 type transporter family protein
OG0016849	Sobic.001G413700.1.p	ABC-2 type transporter family protein

Biosíntesis de lignina

Ortogrupo	Ortólogo anotado	Anotación
OG0000109	Sobic.002G416700.1.p,Sobic.001G328100.1.p,Sobic.005G011500.1.p,Sobic.002G357000.1.p	Peroxidase superfamily protein

OG0000128	<i>Sobic.005G045600.1.p,Sobic.009G197400.1.p,Sobic.001G246650.1.p,Sobic.009G197600.4.p,Sobic.009G197000.1.p,GRMZM2G017557_P01</i>	O-methyltransferase family protein
OG0000132	<i>Sobic.005G127000.1.p,GRMZM2G085661_P01,Sobic.003G089600.1.p,Sobic.007G003500.1.p,Sobic.007G003900.1.p</i>	cytochrome P450,family 71,subfamily A,polypeptide 22,subfamily B,polypeptide 37,polypeptide 25,polypeptide 34
OG0000135	<i>Sobic.006G227400.1.p,Sobic.006G226400.1.p,Sobic.006G227800.2.p</i>	NAD(P)-binding Rossmann-fold superfamily protein
OG0000152	<i>Sobic.002G216500.1.p,Sobic.001G364400.1.p,Sobic.002G044200.1.p,Sobic.007G044900.1.p</i>	Cytochrome P450 superfamily protein
OG0000186	<i>Sobic.009G162800.1.p,Sobic.003G352700.1.p,Sobic.003G353200.1.p,Sobic.001G422300.1.p</i>	laccase 17,laccase 2
OG0000231	<i>Sobic.007G076000.2.p,Sobic.002G195700.1.p,Sobic.002G195600.1.p,Sevir.2G207100.1.p</i>	cinnamyl alcohol dehydrogenase 9,elicitor-activated gene 3-1
OG0000248	<i>Sobic.004G068600.2.p,Sobic.004G068900.1.p,Sevir.1G058501.1.p,Sobic.004G068800.1.p,Sevir.7G289901.1.p</i>	cytochrome P450,family 71,subfamily B,polypeptide 2,family 76,subfamily C,polypeptide 4,polypeptide 13,polypeptide 34
OG0000279	<i>Sobic.002G205500.1.p,Sobic.002G075500.1.p,Sobic.001G045700.1.p</i>	Cellulose synthase family protein
OG0000308	<i>Sobic.004G069700.2.p,GRMZM2G397402_P01,Sevir.2G102300.2.p</i>	cytochrome P450,family 71,subfamily B,polypeptide 37,subfamily A,polypeptide 13,polypeptide 24
OG0000319	<i>GRMZM2G536099_P02,Sobic.002G250100.2.p,Sobic.002G250000.4.p</i>	NAD(P)-binding Rossmann-fold superfamily protein
OG0000386	<i>Sevir.6G183100.1.p,Sobic.007G152500.1.p</i>	cytochrome P450,family 76,subfamily C,polypeptide 4,polypeptide 1
OG0000389	<i>Sobic.007G099400.1.p,Sobic.002G079500.1.p,Sobic.008G061401.1.p</i>	O-methyltransferase family protein
OG0000438	<i>Sobic.001G012000.1.p,Sevir.4G248800.1.p,Sobic.001G018600.1.p,Sobic.008G058500.1.p</i>	cytochrome P450,family 71,subfamily B,polypeptide 34
OG0000474	<i>Sobic.002G040400.1.p,Sobic.002G040500.1.p,Sobic.008G136500.1.p</i>	Cytochrome P450 superfamily protein
OG0000475	<i>Sobic.006G145700.1.p,Sobic.006G146000.1.p</i>	beta glucosidase 46,beta-glucosidase 47
OG0000482	<i>Sobic.004G006400.1.p,Sobic.003G076700.1.p,Sevir.9G394900.1.p</i>	Cytochrome P450 superfamily protein,cytochrome P450,family 89,subfamily A,polypeptide 5,family 87,polypeptide 6
OG0000508	<i>Sobic.008G109800.2.p,GRMZM2G090980_P01,Sobic.010G072200.2.p</i>	cinnamyl-alcohol dehydrogenase
OG0000514	<i>Sobic.001G192200.1.p,Sobic.005G149600.1.p,Sobic.001G044100.1.p</i>	cytochrome P450,family 76,subfamily C,polypeptide 4,polypeptide 1
OG0000526	<i>Sobic.001G176000.1.p</i>	cytochrome P450,family 71,subfamily B,polypeptide 2
OG0000572	<i>Sobic.003G010000.2.p,Sobic.003G009700.1.p,Sobic.003G010300.1.p,Sobic.003G009900.1.p</i>	cytochrome P450,family 71,subfamily A,polypeptide 25,polypeptide 22
OG0000586	<i>Sobic.001G235500.1.p,Sobic.002G065700.1.p</i>	cytochrome P450,family 71,subfamily B,polypeptide 37,polypeptide 34
OG0000628	<i>GRMZM2G172826_P01,Sobic.002G190300.1.p</i>	cytochrome P450,family 71,subfamily A,polypeptide 22,subfamily B,polypeptide 2
OG0000633	<i>Sevir.1G168100.1.p,Sobic.002G108400.1.p</i>	cytochrome P450,family 76,subfamily C,polypeptide 3,polypeptide 4
OG0000639	<i>Sobic.001G338900.1.p</i>	cytochrome P450,family 71 subfamily B,polypeptide 7
OG0000770	<i>Sobic.010G161900.1.p,Sobic.010G161800.1.p</i>	Peroxidase superfamily protein
OG0000830	<i>Sobic.009G033600.1.p,Sobic.003G151600.1.p,Sevir.3G054100.1.p</i>	S-adenosylmethionine synthetase 1,S-adenosylmethionine synthetase family protein
OG0000888	<i>Sobic.009G162500.1.p,Sobic.004G200800.1.p</i>	Cytochrome P450 superfamily protein
OG0000951	<i>Sobic.004G106100.1.p,Sobic.010G161700.1.p,Sobic.004G105700.1.p</i>	Peroxidase superfamily protein
OG0001188	<i>Sobic.001G195200.1.p</i>	cytochrome P450,family 87,subfamily A,polypeptide 6
OG0001258	<i>Sobic.003G121000.1.p,Sobic.003G121100.1.p</i>	Peroxidase superfamily protein
OG0001393	<i>Sobic.006G122000.1.p,Sobic.006G121600.1.p</i>	cytochrome P450,family 71,subfamily A,polypeptide 25,polypeptide 26

OG0001709	<i>Sobic.007G063300.1.p,Sobic.010G195800.1.p</i>	NAD(P)-binding Rossmann-fold superfamily protein
OG0001730	<i>Sobic.002G126600.1.p,Sobic.003G337400.1.p</i>	cinnamate-4-hydroxylase
OG0001787	<i>Sobic.001G137900.1.p</i>	cytochrome P450,family 87,subfamily A,polypeptide 6
OG0001843	<i>Sobic.009G196800.1.p</i>	cytochrome P450,family 71,subfamily A,polypeptide 25
OG0001894	<i>Sobic.004G235900.1.p</i>	Laccase/Diphenol oxidase family protein
OG0001907	<i>Sobic.005G224300.1.p</i>	O-methyltransferase family protein
OG0002036	<i>Sobic.001G369500.1.p</i>	cytochrome P450,family 71,subfamily B,polypeptide 36
OG0002059	<i>Sobic.001G196300.1.p</i>	ferulic acid 5-hydroxylase 1
OG0002118	<i>Sobic.002G272700.1.p,Sobic.007G179750.1.p</i>	cytochrome p450 78a9,cytochrome P450,family 78,subfamily A,polypeptide 6
OG0002217	<i>Sobic.002G237900.1.p</i>	cellulose synthase like E1
OG0002317	<i>Sobic.010G203500.1.p</i>	cytochrome P450,family 71 subfamily B,polypeptide 7
OG0002572	<i>Sevir.7G287000.1.p,Sobic.008G107200.1.p</i>	cytochrome P450,family 71,subfamily B,polypeptide 36,polypeptide 37
OG0002657	<i>Sobic.003G233200.2.p,Sobic.009G243400.1.p</i>	NAD(P)-binding Rossmann-fold superfamily protein
OG0002830	<i>Sobic.001G336200.1.p,Sobic.002G367600.1.p</i>	cytochrome P450,family 78,subfamily A,polypeptide 5,polypeptide 10
OG0002887	<i>Sobic.008G014000.1.p</i>	O-methyltransferase 1
OG0003560	<i>Sobic.002G183700.1.p</i>	Cupredoxin superfamily protein
OG0003565	<i>Sobic.008G021200.1.p</i>	vacuolar ATP synthase subunit C (VATC) / V-ATPase C subunit / vacuolar proton pump C subunit (DET3)
OG0003743	<i>Sobic.001G037900.1.p,Sobic.001G037800.2.p</i>	NAD(P)-binding Rossmann-fold superfamily protein
OG0003762	<i>Sobic.003G327800.1.p</i>	cytochrome P450,family 98,subfamily A,polypeptide 3
OG0003866	<i>GRMZM2G050076_P04</i>	NAD(P)-binding Rossmann-fold superfamily protein
OG0004167	<i>Sobic.004G273800.1.p</i>	myb domain protein 63
OG0004226	<i>Sobic.010G171800.1.p</i>	Plant L-ascorbate oxidase
OG0004242	<i>GRMZM2G015232_P01</i>	cytochrome P450,family 71,subfamily B,polypeptide 34
OG0004487	<i>GRMZM2G158342_P01,Sobic.007G178500.1.p</i>	cytochrome P450,family 706,subfamily A,polypeptide 4,polypeptide 6
OG0004502	<i>Sobic.006G014700.1.p</i>	cinnamyl alcohol dehydrogenase 9
OG0005320	<i>Sevir.9G007100.1.p</i>	cytochrome P450,family 87,subfamily A,polypeptide 6
OG0005873	<i>Sobic.007G059100.1.p,Sobic.001G456650.1.p</i>	O-methyltransferase 1
OG0005960	<i>GRMZM2G177424_P01</i>	O-methyltransferase family protein
OG0006433	<i>Sobic.006G146200.1.p</i>	beta glucosidase 46
OG0006698	<i>Sobic.004G071000.1.p</i>	GroES-like zinc-binding alcohol dehydrogenase family protein
OG0007037	<i>Sobic.001G270200.1.p</i>	cytochrome P450,family 87,subfamily A,polypeptide 6
OG0007305	<i>Sobic.002G146000.1.p</i>	cinnamoyl coa reductase 1
OG0007508	<i>Sobic.006G199800.1.p</i>	myb domain protein 63
OG0007666	<i>Sobic.001G360700.1.p</i>	Cytochrome P450 superfamily protein
OG0008165	<i>Sobic.003G341500.1.p</i>	laccase 5
OG0008973	<i>Sobic.007G047300.1.p</i>	O-methyltransferase 1
OG0009302	<i>Sobic.002G000400.1.p</i>	cytochrome P450,family 93,subfamily D,polypeptide 1
OG0009906	<i>Sobic.003G357600.1.p</i>	laccase 12
OG0011554	<i>Sobic.004G007500.1.p</i>	hydroperoxide lyase 1
OG0012091	<i>Sobic.004G065600.1.p</i>	cinnamoyl coa reductase 1

OG0012424	<i>Sobic.003G410500.1.p</i>	Aldolase-type TIM barrel family protein
OG0012503	<i>Sobic.007G043200.1.p</i>	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
OG0012895	<i>Sobic.007G050600.1.p</i>	cellulose synthase-like D3
OG0013717	<i>Sobic.001G254800.1.p</i>	cytochrome P450,family 78,subfamily A,polypeptide 7
OG0014122	<i>Sobic.002G220500.1.p</i>	cytochrome P450,family 76,subfamily C,polypeptide 4
OG0014919	<i>Sobic.008G014200.1.p</i>	cytochrome P450,family 87,subfamily A,polypeptide 9
OG0016137	<i>Sobic.004G105200.1.p</i>	Peroxidase superfamily protein
OG0016301	<i>Sobic.002G333800.1.p</i>	cellulose-synthase like D2
OG0016501	<i>Sobic.010G231600.1.p</i>	cytochrome P450,family 77,subfamily B,polypeptide 1
OG0016731	<i>Sobic.004G105100.1.p</i>	Peroxidase superfamily protein
OG0016893	<i>Sobic.001G199400.1.p</i>	Cupredoxin superfamily protein
OG0017142	<i>Sobic.003G183300.1.p</i>	Peroxidase superfamily protein
OG0017252	<i>Sobic.008G125700.1.p</i>	cellulose synthase-like D5
OG0017371	<i>Sobic.006G066400.1.p</i>	cytochrome P450,family 77,subfamily A,polypeptide 4
OG0018544	<i>Sobic.002G110100.1.p</i>	cytochrome P450,family 71,subfamily B,polypeptide 34

Biosíntesis de celulosa

Ortogrupo	Ortólogo anotado	Anotación
OG0000052	<i>Sobic.002G177600.1.p,Sevir.8G255001.1.p,GRM ZM2G354621_P02,GRMZM2G392176_P01,GRMZM2G332280_P01</i>	Leucine-rich repeat protein kinase family protein,Leucine-rich repeat transmembrane protein kinase protein
OG0000057	<i>Sobic.005G220740.1.p,Sevir.8G167900.1.p,Sevir.8G024912.1.p,GRMZM2G388776_P01,Sobic.003G265701.1.p</i>	Leucine-rich repeat protein kinase family protein,Leucine-rich receptor-like protein kinase family protein
OG0000100	<i>Sobic.003G067600.1.p,Sobic.010G274700.1.p,Sobic.003G069950.1.p,Sobic.003G070025.1.p,Sobic.002G075000.1.p</i>	Leucine-rich repeat protein kinase family protein
OG0000113	<i>Sobic.010G028501.1.p,Sobic.007G149700.1.p,Sobic.001G277500.2.p,Sobic.010G028600.1.p</i>	S-locus lectin protein kinase family protein,cysteine-rich RLK (RECEPTOR-like protein kinase) 23
OG0000124	<i>GRMZM2G350793_P01,Sobic.005G205100.1.p,Sevir.4G241800.1.p</i>	Leucine-rich repeat protein kinase family protein
OG0000153	<i>GRMZM2G155837_P02,Sevir.2G100500.1.p,Sobic.006G215400.1.p</i>	Leucine-rich repeat transmembrane protein kinase
OG0000211	<i>Sobic.005G038800.2.p,Sobic.004G214900.1.p,Sobic.009G040175.1.p</i>	Leucine-rich repeat protein kinase family protein,Leucine-rich receptor-like protein kinase family protein
OG0000213	<i>Sevir.8G052700.1.p,Sobic.005G057700.1.p,Sevir.8G052500.1.p,Sobic.005G057600.1.p</i>	Leucine-rich repeat transmembrane protein kinase,Leucine-rich repeat protein kinase family protein,EF-TU receptor
OG0000236	<i>Sobic.009G019000.1.p,Sevir.5G015601.1.p,GRMZM2G179270_P01,Sevir.2G325432.1.p,GRMZM5G823832_P01,Sevir.3G369420.1.p</i>	S-locus lectin protein kinase family protein
OG0000279	<i>Sobic.002G205500.1.p,Sobic.002G075500.1.p,Sobic.001G045700.1.p</i>	Cellulose synthase family protein
OG0000304	<i>Sevir.1G082000.1.p,Sobic.004G045000.1.p,Sobic.004G045100.1.p,Sobic.004G045500.1.p</i>	receptor like protein 2,Leucine-rich receptor-like protein kinase family protein
OG0000307	<i>Sobic.001G446400.1.p,Sobic.009G259500.1.p,Sobic.001G074000.1.p,Sobic.010G267700.1.p</i>	Leucine-rich receptor-like protein kinase family protein
OG0000314	<i>Sobic.010G241500.1.p,Sevir.1G083200.1.p</i>	Leucine-rich receptor-like protein kinase family protein
OG0000575	<i>GRMZM2G175848_P01</i>	cellulose synthase family protein

OG0000585	<i>Sobic.002G077400.1.p,Sevir.6G176650.1.p,Sevir.2G032000.1.p</i>	<i>Protein kinase family protein with leucine-rich repeat domain,Leucine-rich repeat receptor-like protein kinase family protein</i>
OG0000648	<i>Sobic.004G091300.1.p,Sevir.1G036307.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0000758	<i>Sobic.006G000600.1.p,Sobic.009G175600.1.p,Sobic.003G336700.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0000877	<i>Sobic.006G030600.2.p,Sobic.006G030300.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0000908	<i>Sobic.004G189900.1.p,Sobic.006G104500.1.p,Sobic.007G059600.1.p</i>	<i>somatic embryogenesis receptor-like kinase 1,somatic embryogenesis receptor-like kinase 2</i>
OG0000949	<i>Sobic.005G187700.1.p,GRMZM2G389948_P01</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0001085	<i>Sobic.010G067200.2.p,Sobic.010G067400.1.p,GRMZM2G428370_P01</i>	<i>Leucine-rich receptor-like protein kinase family protein</i>
OG0001090	<i>Sobic.006G228700.1.p</i>	<i>Malectin/receptor-like protein kinase family protein</i>
OG0001170	<i>Sobic.003G445900.1.p,Sevir.8G020475.1.p,Sobic.005G001100.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0001192	<i>Sobic.005G004900.1.p</i>	<i>TRICHOME BIREFRINGENCE-LIKE 34</i>
OG0001240	<i>Sobic.006G012900.1.p,Sobic.002G117500.1.p</i>	<i>UDP-Glycosyltransferase superfamily protein</i>
OG0001313	<i>Sobic.008G060700.1.p</i>	<i>Protein kinase family protein with leucine-rich repeat domain</i>
OG0001486	<i>GRMZM2G015406_P01,Sobic.005G131600.2.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0001664	<i>Sobic.001G406700.1.p,Sevir.9G444400.1.p,Sobic.002G429800.1.p</i>	<i>Plant protein of unknown function (DUF828)</i>
OG0001803	<i>Sobic.008G083300.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0001888	<i>Sobic.001G223550.1.p</i>	<i>Leucine-rich repeat transmembrane protein kinase</i>
OG0001890	<i>Sobic.006G080700.2.p</i>	<i>cellulose synthase-like B4</i>
OG0002061	<i>Sobic.002G176700.1.p,Sobic.002G176500.1.p</i>	<i>Leucine-rich repeat transmembrane protein kinase</i>
OG0002113	<i>Sobic.004G099700.1.p</i>	<i>Protein kinase family protein with leucine-rich repeat domain</i>
OG0002134	<i>AC217401.3_FGP003,Sobic.004G128200.3.p</i>	<i>leucine-rich repeat transmembrane protein kinase family protein</i>
OG0002217	<i>Sobic.002G237900.1.p</i>	<i>cellulose synthase like E1</i>
OG0002283	<i>Sobic.002G376100.1.p,Sobic.001G346100.1.p</i>	<i>STRUBBELIG-receptor family 6,STRUBBELIG-receptor family 7</i>
OG0002290	<i>Sobic.008G003600.1.p</i>	<i>Plant protein of unknown function (DUF828)</i>
OG0002295	<i>Sobic.002G334100.1.p</i>	<i>cellulose synthase-like D5</i>
OG0002366	<i>Sobic.003G304600.1.p,Sobic.001G173300.1.p</i>	<i>reversibly glycosylated polypeptide 1,reversibly glycosylated polypeptide 2</i>
OG0002379	<i>Sobic.009G177400.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0002383	<i>Sobic.001G480800.1.p,Sobic.001G261400.2.p</i>	<i>STRUBBELIG-receptor family 7,STRUBBELIG-receptor family 6</i>
OG0002550	<i>Sobic.001G038200.1.p,Sobic.001G038300.1.p</i>	<i>TRICHOME BIREFRINGENCE-LIKE 33</i>
OG0002635	<i>Sobic.005G034100.1.p</i>	<i>Plant protein of unknown function (DUF828)</i>
OG0002865	<i>Sobic.008G060000.1.p</i>	<i>leucine-rich repeat transmembrane protein kinase family protein</i>
OG0002917	<i>Sobic.005G182400.1.p</i>	<i>leucine-rich repeat transmembrane protein kinase family protein</i>
OG0003098	<i>Sobic.003G336900.1.p,Sobic.009G175100.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0003232	<i>Sobic.001G395700.1.p,Sobic.001G395600.2.p</i>	<i>TRICHOME BIREFRINGENCE-LIKE 6</i>
OG0003248	<i>Sobic.001G306800.1.p</i>	<i>Plant protein of unknown function (DUF828)</i>
OG0003254	<i>Sobic.008G186300.1.p</i>	<i>Leucine-rich receptor-like protein kinase family protein</i>
OG0003379	<i>Sobic.010G120800.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0003724	<i>Sobic.007G212000.1.p,Sobic.001G265700.1.p</i>	<i>PEP1 receptor 2</i>

OG0003797	<i>Sobic.006G262400.1.p,Sobic.006G262700.1.p</i>	<i>phytosulfokin receptor 1</i>
OG0004035	<i>Sobic.001G455000.1.p</i>	<i>Glycosyl hydrolases family 31 protein</i>
OG0004148	<i>Sobic.002G131400.1.p</i>	<i>mannosyltransferase family protein</i>
OG0004201	<i>Sobic.003G291600.1.p</i>	<i>Protein kinase family protein with leucine-rich repeat domain</i>
OG0004387	<i>Sobic.002G155400.1.p</i>	<i>BRI1 like</i>
OG0004421	<i>Sobic.010G199700.1.p</i>	<i>STRUBBELIG-receptor family 8</i>
OG0004537	<i>Sobic.006G134700.1.p</i>	<i>Six-hairpin glycosidases superfamily protein</i>
OG0004608	<i>Sobic.001G462200.1.p</i>	<i>ADP-glucose pyrophosphorylase family protein</i>
OG0004990	<i>Sobic.008G006800.1.p</i>	<i>inflorescence meristem receptor-like kinase 2</i>
OG0005111	<i>Sobic.005G111000.1.p</i>	<i>transmembrane kinase 1</i>
OG0005322	<i>Sobic.003G002100.1.p</i>	<i>HAESA-like 1</i>
OG0005329	<i>Sobic.008G186400.1.p</i>	<i>Leucine-rich repeat transmembrane protein kinase family protein</i>
OG0005350	<i>Sobic.009G105000.1.p</i>	<i>Plant protein of unknown function (DUF828)</i>
OG0005524	<i>Sobic.010G077000.1.p</i>	<i>Leucine-rich receptor-like protein kinase family protein</i>
OG0005534	<i>Sobic.008G074500.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0005797	<i>Sobic.004G248100.1.p</i>	<i>glycosyl hydrolase 9B1</i>
OG0006062	<i>Sobic.006G127400.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0006152	<i>Sobic.005G033000.1.p</i>	<i>ribose-5-phosphate isomerase 2</i>
OG0006447	<i>Sobic.002G075900.1.p</i>	<i>dolichyl-diphosphooligosaccharide-protein glycosyltransferase 48kDa subunit family protein</i>
OG0006462	<i>Sobic.007G144600.1.p</i>	<i>PEP1 receptor 1</i>
OG0006514	<i>Sobic.003G442500.1.p</i>	<i>cellulose synthase like G3</i>
OG0007756	<i>Sobic.009G203600.1.p</i>	<i>Protein kinase family protein with leucine-rich repeat domain</i>
OG0008097	<i>Sobic.006G116200.1.p</i>	<i>Leucine-rich repeat receptor-like protein kinase family protein</i>
OG0008792	<i>Sobic.001G116600.1.p,Sevir.9G117300.1.p</i>	<i>receptor-like kinase 1</i>
OG0009044	<i>Sobic.004G104800.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0009081	<i>Sobic.003G051100.1.p</i>	<i>NSP-interacting kinase 3</i>
OG0010400	<i>Sobic.003G020600.1.p</i>	<i>TRICHOME BIREFRINGENCE-LIKE 37</i>
OG0010888	<i>Sobic.007G113300.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0011085	<i>Sobic.008G190000.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0011209	<i>Sobic.003G277900.1.p</i>	<i>Leucine-rich receptor-like protein kinase family protein</i>
OG0011379	<i>Sobic.003G368400.1.p</i>	<i>UDP-Glycosyltransferase superfamily protein</i>
OG0011399	<i>Sobic.009G138200.1.p</i>	<i>BAK1-interacting receptor-like kinase 1</i>
OG0011420	<i>Sobic.005G084100.1.p</i>	<i>BAK1-interacting receptor-like kinase 1</i>
OG0011509	<i>Sobic.002G074000.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0011529	<i>Sobic.006G140000.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0011680	<i>Sobic.007G150000.1.p</i>	<i>TRICHOME BIREFRINGENCE-LIKE 16</i>
OG0011790	<i>Sobic.003G083500.1.p</i>	<i>leucine-rich repeat transmembrane protein kinase family protein</i>
OG0012011	<i>Sobic.002G239700.1.p</i>	<i>Leucine-rich receptor-like protein kinase family protein</i>
OG0012156	<i>Sobic.006G187000.1.p</i>	<i>Leucine-rich receptor-like protein kinase family protein</i>
OG0012669	<i>Sobic.004G222100.1.p</i>	<i>phytosulfokin receptor 1</i>

OG0012671	<i>Sobic.002G051100.1.p</i>	<i>Ribose 5-phosphate isomerase, type A protein</i>
OG0012895	<i>Sobic.007G050600.1.p</i>	<i>cellulose synthase-like D3</i>
OG0012923	<i>Sobic.003G375800.1.p</i>	<i>Leucine-rich receptor-like protein kinase family protein</i>
OG0012988	<i>Sobic.001G384500.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0013023	<i>Sobic.002G419200.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0013026	<i>Sobic.001G425700.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0013692	<i>Sobic.004G012200.1.p</i>	<i>phytosulfokin receptor 1</i>
OG0013758	<i>Sobic.002G009900.1.p</i>	<i>receptor-like protein kinase 2</i>
OG0013835	<i>Sobic.002G383100.1.p</i>	<i>Leucine-rich receptor-like protein kinase family protein</i>
OG0013913	<i>Sobic.010G020600.1.p</i>	<i>ERECTA-like 2</i>
OG0013915	<i>Sobic.009G061400.1.p</i>	<i>Leucine-rich repeat transmembrane protein kinase</i>
OG0014069	<i>Sobic.007G019600.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0014158	<i>Sobic.005G091000.1.p</i>	<i>Leucine-rich receptor-like protein kinase family protein</i>
OG0014317	<i>Sobic.001G274400.1.p</i>	<i>BRI1-like 2</i>
OG0014318	<i>Sobic.007G115400.1.p</i>	<i>HAESA-like 1</i>
OG0014378	<i>Sobic.009G079300.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0014482	<i>Sobic.007G038000.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0014542	<i>Sobic.001G113800.1.p</i>	<i>transmembrane kinase 1</i>
OG0014587	<i>Sobic.002G308200.1.p</i>	<i>Leucine-rich repeat transmembrane protein kinase</i>
OG0014599	<i>Sobic.002G363400.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0014613	<i>Sobic.001G407000.1.p</i>	<i>Plant protein of unknown function (DUF828)</i>
OG0014677	<i>Sobic.002G341200.1.p</i>	<i>STRUBBELIG-receptor family 3</i>
OG0015323	<i>Sobic.003G432000.1.p</i>	<i>Leucine-rich receptor-like protein kinase family protein</i>
OG0016004	<i>Sobic.002G032500.1.p</i>	<i>Leucine-rich receptor-like protein kinase family protein</i>
OG0016073	<i>Sobic.004G078500.1.p</i>	<i>STRUBBELIG-receptor family 8</i>
OG0016215	<i>Sobic.004G008600.1.p</i>	<i>Leucine-rich repeat transmembrane protein kinase family protein</i>
OG0016301	<i>Sobic.002G333800.1.p</i>	<i>cellulose-synthase like D2</i>
OG0016365	<i>Sobic.007G142900.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0016384	<i>Sobic.010G078300.1.p</i>	<i>Plant protein of unknown function (DUF828)</i>
OG0016488	<i>Sobic.003G174700.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0016530	<i>Sobic.001G038400.1.p</i>	<i>TRICHOME BIREFRINGENCE-LIKE 34</i>
OG0016682	<i>Sobic.002G006100.1.p</i>	<i>phytosylfokine-alpha receptor 2</i>
OG0016782	<i>Sobic.002G035600.1.p</i>	<i>Leucine-rich receptor-like protein kinase family protein</i>
OG0016983	<i>Sobic.010G024500.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0016994	<i>Sobic.001G381100.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0017109	<i>Sobic.006G217900.1.p</i>	<i>Leucine-rich receptor-like protein kinase family protein</i>
OG0017158	<i>Sobic.009G174100.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>
OG0017252	<i>Sobic.008G125700.1.p</i>	<i>cellulose synthase-like D5</i>
OG0017266	<i>Sobic.003G338200.1.p</i>	<i>Leucine-rich repeat protein kinase family protein</i>