



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS
Instituto de Investigación en Ciencias Básicas y Aplicadas
Área Terminal en Bioquímica y Biología Molecular

PARTICIPACIÓN DE LOS GENES *PLETHORA* EN EL
DESARROLLO DE LA RAÍZ PRIMARIA DE LAS
CACTÁCEAS.

T E S I S

QUE PRESENTA:

RAMSÉS URIEL ALBARRÁN HERNÁNDEZ

Para obtener el grado de:
LICENCIADO EN CIENCIAS

Director de la Tesis:
Dr. Gustavo Rodríguez Alonso

Sinodales:
Dra. Carmen Nina Pastor Colón
Dra. Verónica Narváez Padilla
Dra. Svetlana Shishkova
Dra. Verónica Lira Ruan

Junio, 2020.

Cuernavaca, Morelos, México.



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS



Control Escolar de Licenciatura

VOTOS DE APROBATORIOS

Secretaria ejecutiva del Instituto de Investigación en Ciencias Básicas Aplicadas de la Universidad Autónoma del Estado de Morelos.
P r e s e n t e .

Por medio de la presente le informamos que después de revisar la versión escrita de la tesis que realizó el C. **RAMSES URIEL ALBARRÁN HERNÁNDEZ** con número de matrícula **20154000419** cuyo título es:

“Participación de los genes *PLETHORA* en el desarrollo de la raíz primaria de las cactáceas”.

Consideramos que **SI** reúne los méritos que son necesarios para continuar los trámites para obtener el título de **Licenciado Ciencias Área Terminal de Bioquímica y Biología Molecular.**

Cuernavaca, Mor a 18 de diciembre del 2020

Atentamente
Por una universidad culta

Se adiciona página con la e-firma UAEM de los siguientes:

DRA. CARMEN NINA PASTOR COLÓN
DRA. VERÓNICA MERCEDES NARVAEZ PADILLA
DR. GUSTAVO RODRÍGUEZ ALONSO
DRA. SVETLANA SHISHKOVA
DRA. VERÓNICA LIRA RUAN

PRESIDENTE
SECRETARIO
VOCAL
PRIMER SUPLENTE
SEGUNDO SUPLENTE



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

Sello electrónico

CARMEN NINA PASTOR COLON | Fecha:2020-12-29 09:22:17 | Firmante

t5RI4NYIELGBVPhCxZS7xmanFGboVgm9/H5l1ZjoMgMxqj2M3KctCgxsamY28NUd2aT5YOTkD/WaXrneRgvsCtn02/nJXmHj/AzJKD3SYROg5uWr1k5WHs8TTZPMQIZOBuhK6QyqJiSEY5QPghtAAsnL1pVzkZGVggjLa+JDqMAduvSrpAme3pEwlgXSDarkvmlG521mXg057FDgsme/NMZcw9Ri2Vt4z0vec5p+ghX8T52vJid9FwSfQVQz6wH7FiGnUpnZDH00YXNy0f6vB5CcBjP0Cy/EZAXd3LU42ZifnbnHtZsBNMFJx1bUv/F7etibuF/B51OQXy6hb4qdhg==

VERONICA MERCEDES NARVAEZ PADILLA | Fecha:2020-12-29 18:43:10 | Firmante

w0ohVqN4xXP28bO19Tn8DZVNHpZN39JDvNTm3Klbrm8sh5Y4hNlyZ6gBOa8GTqtnFD1V7swBRNzdCPI2qm0Vx96ULqHKgnijGcNcSIMGwslq5p0psG16jxAqmCXC9xCNjjOqUQ/Pe7jn+CUkyEFUEDoOdczHrmAt8GAbgSEazFPwV5T2/bAKl0vXkCEZa6jw54XlciMKtJT6XNUz42F7TCS4G76Y3/VqaLYRpmYQGSn+7uhPyrLHLLQwwl6aSN8GZqU4ilBSdtctfHEUGVQF8RfPepTLUVIA8J8sPmn6Fzhi1Yw2xQyLp7RgRtKGXBeLnWk23Jbqm+v97bQTLc+SyTw==

VERONICA LIRA RUAN | Fecha:2021-01-02 09:41:49 | Firmante

c8QqwOib8aWgSoFI25MsJ3xzaxA1qBtaiuDfXtSAPmflMSB/7FTBWCUsh805ySiKtJCdy7lrBpFJHZyU/ZXhaer1glsQ7Kblmzdl9wSACq1HyNQPGMrpfj0jByLcAN4H/xEN2YknVBdURc4pnp0gLe4MIKTqeppjSqG0pG8nYQ91Xxzt/x6FmLnEQdSMdn/DQkGDc+22Q1SHDfJfW+52cR+S2boZjjsA2A4X4HWny/gXHjBrnOdVA9p6ZPy5jkRwmHeKWsUg3+M7q+Q3fFE0LE7HKEebSH77qall5yHz0yRELOtnmSIQ3Q4s0UAZciljnJmL6S0Wiuy/2zhKrdSA==

GUSTAVO RODRÍGUEZ ALONSO | Fecha:2021-01-14 15:53:25 | Firmante

nLous5Cu/UtpmFKYYGPIJ/LPkWIEEky5BYjd8MP9j/tBsUCyit8yvTQUdHnjpzWtkYXOlpC+/tn/pueDMkeFoUuiCW5j+QAKaWTa2loa5oNHOWDDIs0QH1a0saMck7R4VMA7F6FdlmnVvnogcZpc8Nd1VDLIZtk3Xy+B95MMd5dPU1wPmlk4NmiTJCJhHzDZfloojiT04E3Ww8tby0E6wY26sR2BtHMk2CT8vVN6l+s8+nBOF8nlkCAM92c/sgR1PUQbx5Kh3MP4BcVh0nc/d8EwbGHlnPo3scDzm/yplu15DzsZk83MhxPul602PJHd4EkkDJppiPyimfHvTw==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o escaneando el código QR ingresando la siguiente clave:



tIBxov

<https://efirma.uaem.mx/noRepudio/dmJA2asNPiZnVgbrnGUxpcSTJx8CfWHP>



La presente tesis fue realizada en el departamento de Biología Molecular de Plantas del Instituto de Biotecnología-UNAM, bajo la asesoría del Dr. Gustavo Rodríguez Alonso.

Este proyecto fue apoyado por: CONACyT mediante el estímulo de “Ayudante de SNIII” por el Dr. Joseph Doubrovsky Jankovsky de agosto de 2019 a julio de 2020; el Instituto de Biotecnología mediante la “Ayuda de alimentación y hospedaje” los meses de abril, mayo y octubre de 2019 así como por el apoyo económico para asistir al XVII congreso nacional de Bioquímica y Biología Molecular de plantas en el 2019.

AGRADECIMIENTOS.

Este proyecto presentó dos facetas: un crecimiento personal y otro académico. En primera instancia fue personal debido a mi incapacidad mental para ser constante en ciertas tareas, o concretar retos que se complicaban y a la mitad de mi formación mis propias metas y objetivos se venían abajo. Sin embargo, tomé las enseñanzas de mi madre Olga H. A., quien siempre me impulsó a darlo todo, aunque eso implique caerse algunas veces; de mi padre Ranulfo A. M., quien me ayudó a superarme a mí mismo en cada ocasión, que no hay obstáculo imposible de vencer; y obtuve la fortaleza de mi hermano Sinhue R., quien alimenta mi fuerza de voluntad, que los límites no existen; así como de mi hermano Arath G., quien me da razones para seguir creciendo, que hay alguien viendo cuales caminos estoy tomando. De esta manera agradezco a mi familia que cada día fue una inspiración para continuar, a cumplir las expectativas que todos se crean de mí.

Por otro lado, mi salud mental también se vio beneficiada por mis amigos más cercanos quienes durante largas noches de debates y conversaciones en medio de lágrimas de tristeza y felicidad, llenas de retroalimentación y en muchas ocasiones a base de regaños me ayudaron a esclarecer mis pensamientos, a no dejarme vencer y son quienes también me dieron motivos para terminar este proyecto. Así mismo, quiero agradecer todo el apoyo brindado por Antonio L., Giselle F., Fernando B., Evelyn C., Armando P., Adrián R., Max V., Kevin L., Sofia E., Mauricio D., Ginna M., Tania B., Vinicio N., Rebeca D., Karen S., entre otros.

Fue también un crecimiento académico pues, además de las amistades que formé en el laboratorio del departamento de biología molecular de plantas a cargo del Dr. Joseph Dubrovsky, el aprendizaje adquirido en todo el tiempo que estuve con ellos fue en gran parte a que nunca me hicieron sentir abandonado. Es decir, en ninguna ocasión se rindieron conmigo, siempre salía de ese sitio con algún conocimiento nuevo y es debido a que todos se esforzaron porque así fuera. A pesar de todo, un soporte muy importante para mí fue el Dr. Gustavo Rodríguez quien, además de ser el asesor de esta tesis, siempre me apoyó y me escuchó cuando nadie más lo hacía y sin importar las circunstancias siempre esperó lo mejor de mí, a la Dra. Svetlana quien también fue clave para culminar este proyecto y quien me brindó todo el apoyo moral y académico a su disposición, y por último pero no con menos importancia a la Dra. Nina Pastor quien me impulsó cada vez que llegaba a su oficina con nada en el corazón. De la misma manera quiero agradecer a mis compañeros de laboratorio Héctor M., Selene N., Santiago G., Ramces G. y Jazmín R. quienes fueron parte importante en mi formación.

Finalmente, quiero agradecer la ayuda de Sophia E., Pablo V. y Marcela, por el apoyo técnico prestado y su amistad. Al equipo de football “Ciencias Físicas” y al equipo de volleyball “Mafia Rusa” por todas las tardes de juegos y buenos momentos que me hicieron olvidar las amarguras de la vida durante media hora.

"Nada en la vida es para ser temido, es sólo para ser comprendido. Ahora es el momento de entender más, de modo que podamos temer menos".

Maria Salomea Skłodowska-Curie

ÍNDICE

ÍNDICE DE TABLAS, FIGURAS Y ANEXOS.....	VI
LISTA DE ABREVIATURAS.....	VIII
1 INTRODUCCIÓN.....	1
1.1 Estructura de la raíz.....	1
1.2 Anatomía de la raíz.....	1
1.3 Crecimiento determinado e indeterminado de la raíz.....	3
1.4 Crecimiento determinado de la raíz primaria de las cactáceas.....	4
1.5 Vías de regulación del RAM.....	5
1.5.1 Las proteínas AP2.....	6
2 ANTECEDENTES DIRECTOS Y JUSTIFICACIÓN.....	10
3 OBJETIVOS.....	12
3.1 General.....	12
3.2 Particulares.....	12
4 MATERIALES Y MÉTODOS.....	13
4.1 Identificación de los genes PLT en las plantas.....	13
4.2 Identificación de los genes PLT en plantas vasculares sin semilla.....	13
4.3 Verificación de las secuencias y alineamiento múltiple.....	14
4.4 Selección de modelo evolutivo.....	14
4.5 Análisis filogenético.....	14
4.6 Análisis de los dominios AP2 de las proteínas PLT.....	15
4.7 Obtención de material biológico.....	15
4.8 Aclaración de raíces y análisis de la raíz primaria mediante microscopía DIC.....	16
5 RESULTADOS Y DISCUSIÓN.....	17
5.1 Identificación de genes PLT de las plantas.....	17
5.2 Predicción de modelo evolutivo.....	19
5.3 Análisis filogenético.....	20
5.4 Análisis filogenético en plantas vasculares sin semilla.....	24
5.5 Análisis filogenético en cactáceas.....	25
5.6 Análisis de las proteínas PLT.....	26
5.7 Análisis del dominio AP2 de las proteínas PLT de <i>P. pringlei</i>	31
5.8 Aclaración y análisis de la raíz primaria mediante microscopía DIC.....	34

5.9	Obtención de material biológico para cada etapa de desarrollo de la raíz primaria de <i>P. pringlei</i>	35
5.10	Extracción de ARN total.....	36
6	CONCLUSIONES	38
7	REFERENCIAS.....	40
8	ANEXOS	43

ÍNDICE DE FIGURAS, TABLAS Y ANEXOS

Figura 1. Tipos de raíces que conforman al sistema radical de una planta.

Figura 2. Estructura y anatomía de la raíz primaria de *Arabidopsis thaliana*.

Figura 3. Etapas del desarrollo de la raíz primaria de *Pachycereus pringlei*.

Figura 4. Dominio de expresión de los genes *PLT*.

Figura 5. Mutantes de los genes *PLT* a 10 d.d.g.

Figura 6. Los genes *PLT* son necesarios para la morfogénesis de primordios de raíces laterales en el tipo silvestre.

Figura 7. Identificación de proteínas PLT de especies seleccionadas.

Figura 8. Filogenia molecular de las proteínas PLT de las embriofitas.

Figura 9. Identificación de los genes *PLT* en plantas vasculares sin semilla.

Figura 10. Identificación de los genes PLT en cactáceas. El clado de las proteínas PLT está señalado en azul y el clado de las proteínas AIL en negro.

Figura 11. Matriz de identidad entre las proteínas PLT de *A. thaliana* y los transcritos traducidos de *P. pringlei*.

Figura 12. Los dominios AP2 están altamente conservados en todas las especies de plantas.

Figura 13. Matriz de identidad entre los dominios de las proteínas PLT de *A. thaliana* y los transcritos traducidos de *P. pringlei*.

Figura 14. Los dominios AP2 de las proteínas PLT *A. thaliana* y *P. pringlei* están altamente conservados.

Figura 15. Abundancia de los transcritos PLT de algunas especies de plantas monocotiledóneas y dicotiledóneas.

Figura 16. Alineamiento de los CBD y las PLT de *P. pringlei* y *A. thaliana*.

Figura 17. Motivos consenso de PLT5 y PLT2 (SELEX).

Figura 18. Estructura de AtERF96 unido a un motivo de ADN.

Figura 19. Etapas de desarrollo de la raíz primaria de la cactácea *P. pringlei*.

Figura 20. Ensayo de extracción de mRNA en la raíz primaria de *P. pringlei*.

Tabla 1. Especies de plantas en las cuales se identificaron secuencias de proteínas AIL.

Tabla 2. Predicción del modelo evolutivo de los órdenes Caryophyllales y Brassicales.

Tabla 3. Colecta de tejido para cada réplica biológica.

ANEXOS

Tabla 1. Predicción del modelo evolutivo para el orden Caryophyllales.

Tabla 2. Predicción del modelo evolutivo para la cactácea *P. pringlei*.

Tabla 3. Predicción del modelo evolutivo para el orden Brassicales.

Tabla 4. Predicción del modelo evolutivo para las secuencias AIL de los órdenes Brassicales y Caryophyllales.

Tabla 5. Porcentajes de identidad y similitud entre las proteínas PLT y sus dominios de *A. thaliana* (AraTha) y *P. pringlei* (PacPri).

Figura 1. Filogenia de las proteínas AIL en las embriofitas.

Figura 2. Fracción del alineamiento de las secuencias de los dominios AP2 de *A. thaliana* y *P. pringlei*.

Figura 3. Alineamiento entre la secuencia consenso del cristal AtERF96 y la proteína CBD.

LISTA DE ABREVIATURAS

- AIN: *AINTEGUMENTA*.
- AP2: *APETALA2*.
- PLT: *PLETHORA*.
- ARFS: *AUXIN RESPONSE FACTORS* / factores de respuesta a auxinas.
- ERFs: *ETHYLENE RESPONSE FACTORS* / factores de respuesta al etileno.
- EREBPs: *ETHYLENE RESPONSE ELEMENTS BINDING PROTEINS* / proteínas de unión a elementos de respuesta a etileno.
- CEI: cortex-endodermis initial cells / células iniciales del córtex y endodermis.
- CEID: cortex-endodermis initial daughters cells / células hijas de las iniciales del córtex y endodermis.
- LR: lateral root / raíz lateral.
- LRP: lateral root primordium / primordio de raíz lateral.
- RAM: root apical meristem / meristemo apical de la raíz.
- SAM: shoot apical meristem / meristemo apical del brote.
- WT: wild type / tipo silvestre.
- d.d.g.: días después de la germinación.
- N-term: extremo amino terminal.
- C-term: extremo carboxilo terminal.
- MTBS: microtubule buffer stabilizing / amortiguador estabilizador de microtúbulos.
- RPKM: reads per kilobase per million mapped reads / lecturas por kilobase por millón de lecturas mapeadas.
- RPM: revoluciones por minuto.
- RT-qPCR: reverse transcription-quantitative polymerase chain reaction / reacción en cadena de la polimerasa, cuantitativa y en tiempo real.
- BIC: Bayesian information criterion / criterio de información bayesiana.
- AIC: Akaike information criterion / criterio de información Akaike.
- SELEX: Systematic evolution of ligands by exponential enrichment / Evolución sistemática de ligandos mediante enriquecimiento exponencial.

1 INTRODUCCIÓN

Una de las estructuras que permitió a las plantas colonizar a todos los nichos ecológicos terrestres fue la raíz. Las raíces están presentes en la mayoría de las plantas vasculares, generalmente tienen gravitropismo positivo, y en el ápice una estructura llamada cofia que protege a la raíz durante su crecimiento a través del sustrato. Además, la raíz permite la interacción de la planta con bacterias y hongos presentes en el suelo [1].

1.1 Estructura de la raíz.

De acuerdo con su origen, la raíz puede clasificarse en tres tipos: raíz primaria, raíz lateral (Fig. 1a) y raíz adventicia (Fig. 1b). Se conoce como raíz primaria a aquella que se origina a partir de la radícula, la cual es una raíz embrionaria, mientras que las raíces laterales se originan a partir de una raíz parental. Por otra parte, las raíces adventicias se forman a partir de un órgano distinto a la raíz, como pueden ser el tallo o las hojas, tal como se muestra en la figura 1b. En su conjunto, los distintos tipos de raíz forman al sistema radical (Fig. 1) [2].

1.2 Anatomía de la raíz

En la raíz de la mayoría de las plantas, tales como en *Arabidopsis thaliana*, se pueden distinguir longitudinalmente tres zonas de desarrollo, las cuales son: i) la zona meristemática, en donde la mayoría de las células se dividen constantemente; ii) la zona de elongación, donde las células cesan su actividad mitótica y se expanden longitudinalmente; y iii) la zona de diferenciación, la cual se caracteriza por presentar células completamente elongadas y vacuolizadas. En esta zona algunas células epidérmicas prolongan sus membranas citoplasmáticas creando pelos radicales, el cual es un fenotipo fácilmente observable y que indica diferenciación celular (Fig. 2a) [3].

El crecimiento de las raíces que conforman al sistema radical se mantiene mediante la proliferación y elongación celular, lo cual ocurre en las zonas meristemática y de elongación, respectivamente. Las células del sistema radical se producen en el meristemo apical de la raíz, (RAM, por sus siglas en inglés; Fig. 2b), el cual está contenido en la zona meristemática. En contraste con la mayoría de las células del RAM, las células del centro quiescente (CQ) tienen una tasa de proliferación celular casi nula. El CQ actúa como centro organizador de la raíz y en *Arabidopsis thaliana* está conformado típicamente por cuatro células [4]. Las

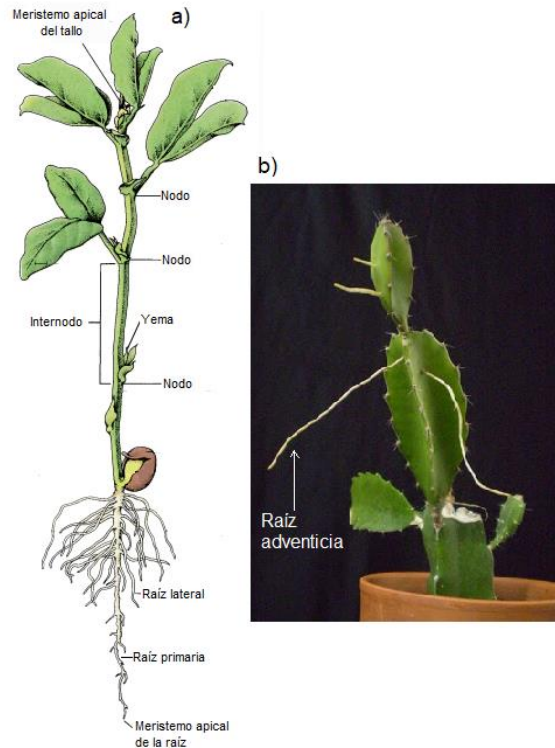


Figura 1. Tipos de raíces que conforman al sistema radical de una planta. a) Raíz primaria y raíces laterales. Modificado de Raven *et al.*, 2014 b) Raíz adventicia. (Tomada de <https://plantsarethe strangest people.blogspot.com/2008/01/unsung-hero-hylocereus-undatus.html>)

células adyacentes al CQ se conocen como células iniciales (Fig. 2b); éstas se consideran como células troncales y en el RAM de *A. thaliana* se reconocen cuatro tipos: iniciales de la epidermis y de la cofia lateral; iniciales del córtex y de la endodermis (CEI, por sus siglas en inglés); iniciales provasculares e iniciales de la columela (Fig. 2b) [3]. Las células iniciales, en conjunto con el CQ, forman al nicho de células troncales. Las células iniciales generan a los diferentes tipos celulares de la raíz mediante divisiones asimétricas, es decir, una de las dos células resultantes de la división permanece como célula inicial y la otra adquiere una identidad celular distinta. Los tipos celulares de la raíz se organizan concéntricamente alrededor del cilindro vascular y son: epidermis, córtex, endodermis y periciclo, el cual rodea al cilindro vascular [3].

En el RAM se generan nuevas células bajo estrictos patrones de división celular. Estas nuevas células serán desplazadas por otras hasta que entran en la zona de elongación, en donde se expanden longitudinalmente, y posteriormente entran a la zona de diferenciación, en donde exhibirán características particulares al diferenciarse, tal como la presencia de pelos radicales [4]. Debido a que la proliferación celular en el RAM de la mayoría de las plantas se mantiene por

largos períodos, se considera que la raíz puede crecer indefinidamente; es decir, se acepta que la raíz presenta crecimiento indeterminado.

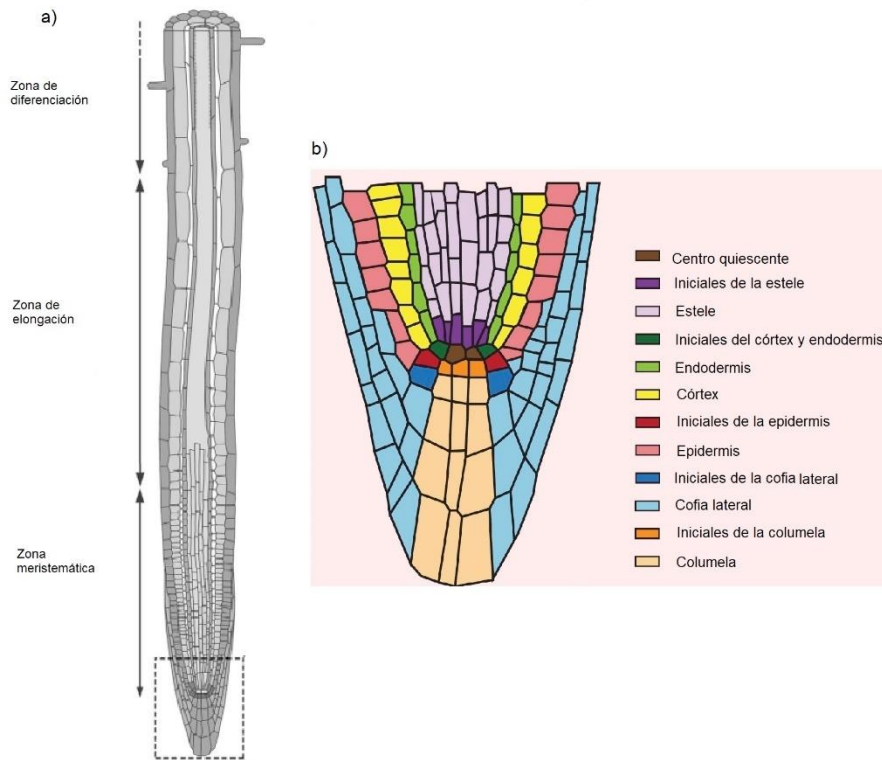


Figura 2. Estructura y anatomía de la raíz primaria de *Arabidopsis thaliana*. a) zonas de desarrollo de la raíz, el recuadro punteado señala al ápice de la raíz. Modificada de Peret, 2017: doi:10.6084/m9.figshare.5143987.v4. b) Meristemo apical de la raíz y los tipos celulares que lo conforman. Modificada de [5].

1.3 Crecimiento determinado e indeterminado de la raíz.

El crecimiento indeterminado ocurre en la raíz primaria de la mayoría de las especies de plantas debido a que el RAM permanece organizado y conserva sus funciones, al menos potencialmente. Por otro lado, el crecimiento del tipo determinado de la raíz ocurre cuando el meristemo se agota, es decir, cuando las células del RAM pierden su capacidad mitótica, por lo que éste deja de ser funcionalmente activo y el crecimiento de la raíz se detiene [6]. La presencia del RAM es la característica principal para distinguir el tipo de crecimiento determinado o indeterminado de la raíz. Algunas especies que habitan en zonas con inviernos muy fríos o temporadas de sequía, por ejemplo *Libocedrus decurrens*, tienen raíces cuyo RAM entra en una fase de latencia, pero puede reiniciar su actividad cuando las condiciones ambientales son adecuadas [7]. De esta manera, aunque el crecimiento de la raíz puede detenerse en condiciones

adversas, se considera que la raíz posee crecimiento indeterminado pues el RAM está presente y continúa siendo funcional [6].

1.4 Crecimiento determinado de la raíz primaria de las cactáceas.

En algunos casos, como en las cactáceas, el RAM está genéticamente programado para agotarse, es decir, a partir de ciertas etapas del desarrollo se abate la proliferación celular y ocurre la diferenciación de todas las células del ápice de la raíz [8]. El crecimiento determinado representa un programa especial en el desarrollo de la raíz primaria de las cactáceas, durante el cual las células de la zona meristemática se dividen pocas veces y posteriormente se diferencian [9]. Se ha sugerido que el crecimiento determinado proporciona una ventaja adaptativa a las cactáceas que habitan ambientes áridos y semi-áridos [10]; esto se ha propuesto debido a que algunas cactáceas presentan inducción temprana de raíces laterales, desarrollan un sistema radical muy ramificado, y todas las células epidérmicas generan pelos radicales lo cual favorece el establecimiento de la plántula [9].

El crecimiento determinado se ha documentado, por ejemplo, para las raíces proteoides en especies de la familia Proteaceae [11], las raíces adventicias de *Ficus pumila* [12], y las raíces laterales de *Zea mays* [13]. Sin embargo, el crecimiento determinado de la raíz primaria solo se ha observado en la familia Cactaceae. Las primeras especies de cactáceas en las cuales se observó el crecimiento determinado de la raíz primaria fueron *Ferocactus peninsulae*, *Stenocereus thurberi*, *S. gummosus* [9] y *Pachycereus pringlei* [14]. Posteriormente, este tipo de crecimiento se observó en la raíz primaria de todas las especies analizadas de la subfamilia Cactoideae y en varias especies de la subfamilia Opuntioideae [10].

El desarrollo de la raíz primaria de las cactáceas se puede dividir en tres etapas: a) inicial, cuando el RAM está presente en la raíz; b) intermedia, cuando el RAM es más pequeño pero sigue presente y los pelos radicales están cercanos al ápice de la raíz; y c) terminal, cuando el RAM está agotado, todas las células del ápice están diferenciadas y como consecuencia los pelos radicales cubren toda la raíz, incluyendo al ápice (Fig. 3) [15].

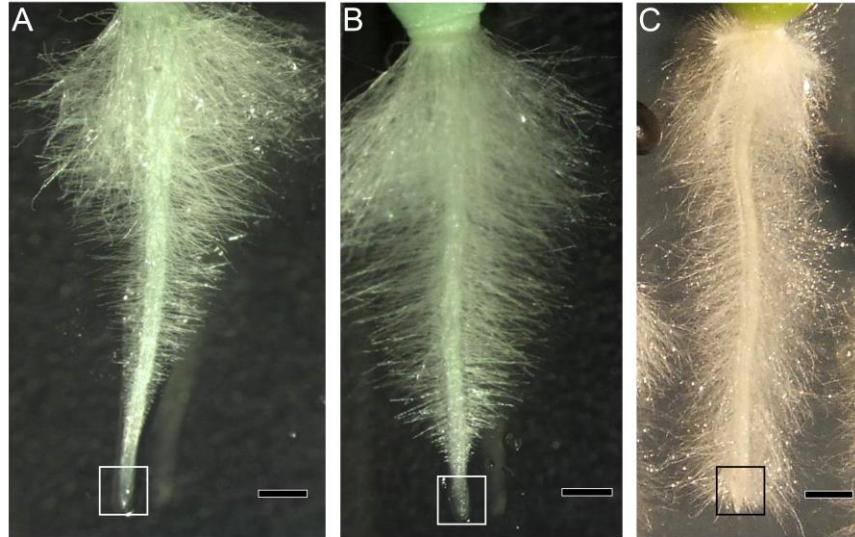


Figura 3. Etapas del desarrollo de la raíz primaria de *Pachycereus pringlei*, a) etapa inicial, b) etapa intermedia, c) etapa terminal. Los recuadros indican el ápice de la raíz. Barra y cuadro: 1mm. Modificada de Rodríguez-Alonso *et al.*, 2018

El tiempo durante el cual la raíz primaria de las cactáceas puede crecer, es decir, el tiempo en el cual se alcanza el agotamiento del RAM, es distinto para cada especie; por ejemplo la raíz primaria de *F. peninsulae* crece solo de 2 a 3 días después de la germinación (d.d.g.), mientras que la raíz primaria de *Pachycereus pringlei* crece durante 7 a 10 d.d.g [10]. En 1997, Dubrovsky [9] germinó semillas de los cactus *Stenocereus thurberi* y *Ferocactus peninsulae* en distintas condiciones, tales como germinar las semillas en papel filtro o en medio de cultivo sólido, con o sin la adición de sacarosa. En las dos especies analizadas se observó agotamiento del RAM en todas las condiciones probadas. Estos resultados sugieren que el agotamiento del RAM y, por lo tanto, el crecimiento determinado de la raíz primaria de las cactáceas se debe a un programa genético y no a las condiciones ambientales o de crecimiento.

1.5 Vías de regulación del RAM.

La proliferación de las células del RAM está controlada por distintas vías de regulación, las cuales se interconectan para formar una red regulatoria [16]. Una de estas vías, de gran relevancia para el mantenimiento del RAM, es la de los factores de transcripción PLETHORA (PLT), los cuales regulan la división y diferenciación celular. En la raíz, las proteínas PLT forman un gradiente de concentración, su acción depende de su abundancia y esta última define a las tres zonas de desarrollo de la raíz en *A. thaliana* [17]. Otra vía de regulación es la de SHORT ROOT-SCARECROW (SHR-SCR) la cual mantiene la identidad del CQ y participa en el establecimiento del patrón radial de la raíz [18]. Además, los

cofactores transcripcionales SCR y SHR controlan las divisiones asimétricas de las CEI y de sus células hijas (CEID, células hijas de las iniciales del córtex y endodermis) para dar lugar a los linajes celulares del córtex y la endodermis [16].

Las proteínas PLT y SCR sobrelapan parcialmente su dominio de expresión en el ápice de la raíz, particularmente en el CQ, en donde la expresión de un factor de transcripción adicional, WUSCHEL-RELATED HOMEODOMAIN 5 (WOX5), es alta y se ha mostrado que éste también se requiere para el control de la tasa de división del CQ y el mantenimiento de las células iniciales de la columela [16].

1.5.1 Las proteínas AP2

Las proteínas PLT forman parte de la familia AINTEGUMENTA-LIKE (AIL) y éstas a su vez de la superfamilia proteica APETALA 2/ETHYLENE RESPONSE FACTORS (AP2/ERF). La superfamilia AP2/ERF comprende 147 miembros en *A. thaliana*, 157 miembros en arroz (*Oryza sativa*) y 148 miembros en soya (*Glycine max*) [19]. Las proteínas AP2 están involucradas en distintos procesos celulares tales como la replicación y el metabolismo del ADN, así como en la regulación del ciclo celular [18, 14], han sufrido diversos eventos de duplicación e intercambio de dominios durante su evolución y forman parte de uno de los grupos más grandes de factores de transcripción específicos de plantas [19].

Todas las proteínas de la superfamilia AP2 contienen al menos un dominio AP2, el cual es de unión al ADN y está altamente conservado [19]. Los genes *AP2/ERF* se dividen en tres clases de acuerdo con el número de dominios AP2 y la existencia o no de dominios adicionales [22]. La primera de estas clases incluye a los genes que codifican para proteínas que contienen exclusivamente dos dominios AP2 (AP2), por ejemplo las proteínas AP2, AINTEGUMENTA (ANT), PLETHORA (PLT) y GLOSSY (GL), entre otras; la segunda clase, llamada ERF, incluye genes que codifican para proteínas con un solo dominio AP2; y la tercera clase incluye genes que codifican para proteínas que tienen dos dominios de unión al ADN: un dominio AP2 y un dominio B3 [22].

Las proteínas AP2, como las PLETHORA, se caracterizan por la presencia de una región enlazadora que tiene de 25 a 26 aminoácidos de longitud entre los dos dominios AP2 [20]. Aunque en general las proteínas PLT son altamente similares entre sí, las secuencias en las regiones C-term y N-term son variables [19].

1.5.1.1 Vía de regulación de las proteínas PLT

En *Arabidopsis* los genes *PLT* comprenden a un clado de seis factores de transcripción, los cuales son *PLT1/AIL3*, *PLT2/AIL4*, *PLT3/AIL6*, *BABY BOOM* o *BBM/PLT4/AIL2*, *PLT5/AIL5* y *PLT7/AIL7* [23].

Las proteínas PLT forman un gradiente de abundancia en el ápice de la raíz debido a su movimiento de corta distancia entre células, desde su lugar de traducción hacia las células contiguas a través del plasmodesmo [24]. Las proteínas PLT establecen una trayectoria de diferenciación celular gradual, desde el ápice hacia la base de la raíz, mediante dos mecanismos: a) por activación de proteínas efectoras para el crecimiento y la división celular; y b) por represión de la diferenciación celular [25]. Este gradiente de proteínas mantiene la identidad de las células troncales y regula la proliferación, expansión y diferenciación celular, definiendo así a las tres zonas del desarrollo de las raíces. La expresión de los genes *PLT* se restringe a los tejidos en desarrollo, en donde controlan, por ejemplo, la morfogénesis de los primordios de órganos laterales, tales como raíces laterales u hojas [21].

Los genes *PLT* están regulados por la fitohormona auxina a través de la acción, directa o indirecta, de AUXIN RESPONSE FACTORS (ARFs) [23]. Por ejemplo, la acumulación de auxinas en las células del periciclo especifica el sitio de iniciación de primordios de raíces laterales (LRP) mediante la inducción de las proteínas PLT3, PLT5 y PLT7; mientras que la misma hormona concentrada en la zona periférica del meristemo apical del brote (SAM, por sus siglas en inglés) especifica el sitio de iniciación de los primordios florales o de la hoja [26].

Los dominios de expresión de *PLT1*, *PLT2*, *PLT3* y *BBM* en el ápice de la raíz de *A. thaliana*, se traslapan parcialmente para mantener la división celular y reprimir la diferenciación de las células iniciales (Fig. 4). Además, se ha demostrado que la función de estos genes es parcialmente redundante [27]. En el nicho de células troncales la concentración de las proteínas PLT1 y PLT2 es muy alta, con un máximo de concentración en el CQ. La abundancia de las proteínas PLT1-4 se reduce progresivamente desde la zona meristemática hasta la zona de diferenciación, en donde no se detectan estas proteínas [17].

Las mutantes simples de *plt1* y *plt2* no muestran fenotipos muy distintos a las plantas del tipo silvestre (WT, por sus siglas en inglés; Fig. 5), aunque se observan alteraciones sutiles en la tasa de división celular y en el número de las células meristemáticas [23]. El RAM de la mutante doble *plt1plt2* es más pequeño

y desorganizado en comparación con el tipo silvestre, y su raíz primaria es más pequeña (Fig. 5) [17, 19].

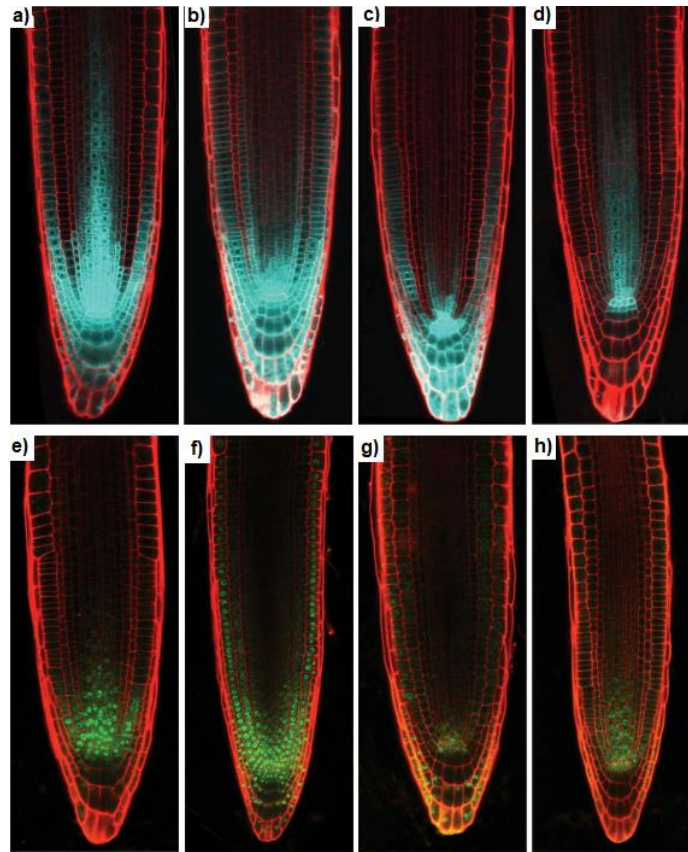


Figura 4. Dominio de expresión de los genes *PLT*. Proteína reportera CFP bajo el promotor de PLT1 (a), PLT2 (b), PLT3 (c) y BBM (d). Proteína reportera YFP fusionada al fragmento genómico de PLT1 (e), PLT2 (f), PLT3 (g) y BBM (h). Modificada de Galinha *et al.*, 2007.

El máximo de expresión del gen *PLT3* se encuentra en las células iniciales de la columela, en las células provasculares y en el CQ, mientras que el dominio de expresión de *BBM* se encuentra en las células iniciales y el CQ (Fig. 4) [17]. Las mutantes simples de *plt3^{-/-}* presentan raíces y un RAM más corto en comparación con las plantas WT; las mutantes dobles *plt3^{-/-}bbm^{-/-}* tienen un RAM más pequeño y raíces más cortas que las mutantes simples de estos genes; en contraste, la mutante triple *plt1^{-/-}plt2^{-/-}plt3^{-/-}* no desarrolla una raíz post-embriónica (Fig. 5) y la segregante de *plt1^{-/-}plt2^{+/-}bbm-1^{-/-}* también carece de raíz [17].

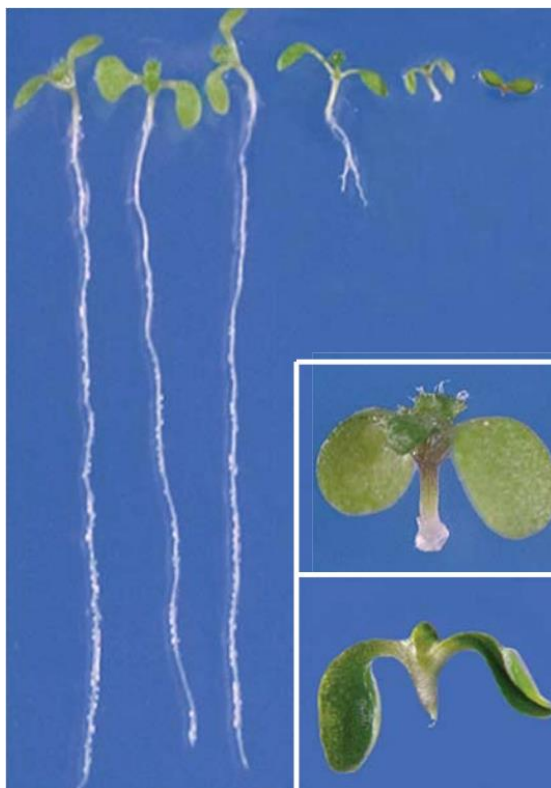


Figura 5. Plántulas de 10 d.d.g. de *A. thaliana*. De izquierda a derecha se muestra al WT, *plt3*^{-/-}, *bbm-1*^{-/-}, *plt1*^{-/-}*plt2*^{-/-}, *plt1*^{-/-}*plt2*^{-/-}*plt3*^{-/-} y a la segregantes de *plt1*^{-/-}*plt2*^{+/-}*bbm-1*^{-/-}. Los recuadros muestran una magnificación de las mutantes *plt1*^{-/-}*plt2*^{-/-}*plt3*^{-/-} (arriba) y la segregante de *plt1*^{-/-}*plt2*^{+/-}*bbm-1*^{-/-} (abajo). Modificada de Galinha *et al.*, 2007.

Los genes *PLT3*, *PLT5* y *PLT7* regulan de manera parcialmente redundante las divisiones celulares durante la morfogénesis temprana de los LRPs (Fig. 6) [28]. En las mutantes triples *plt3*^{-/-}*plt5*^{-/-}*plt7*^{-/-}, los LRPs en la etapa más temprana del desarrollo son morfológicamente indistinguibles respecto al WT (Fig. 6A, F). Sin embargo, en etapas posteriores, la mayoría de las células de la triple mutante no se dividen correctamente (Fig. 6G-I). [28].

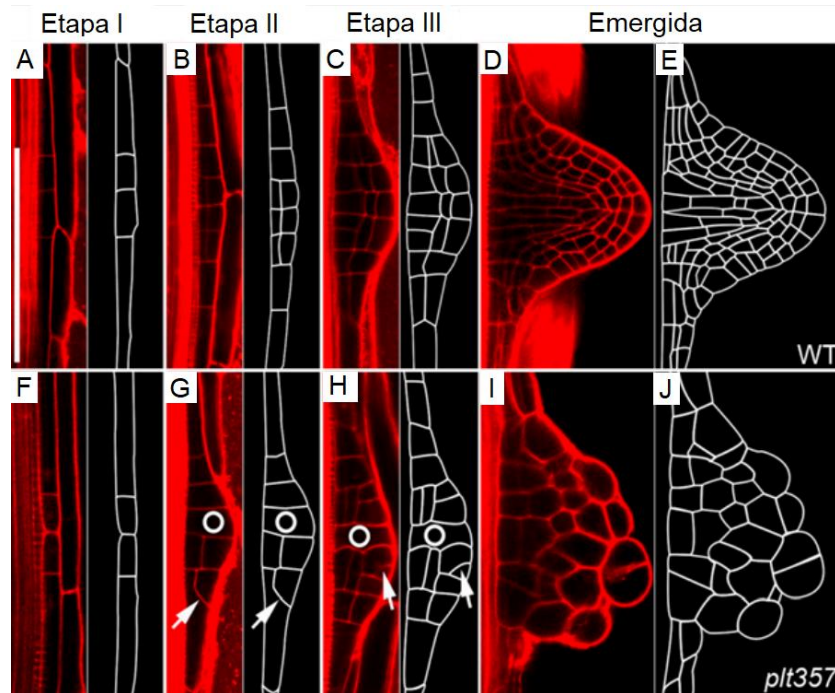


Figura 6. Los genes *PLT* son necesarios para la morfogénesis del LRP en *A. thaliana* (A-E). La morfogénesis de los LRP se ve afectada en la mutante *plt3^{-/-}plt5^{-/-}plt7^{-/-}* (F-J). Imágenes de microscopía confocal de LRPs teñidos con yoduro de propidio, los círculos indican ausencia de división celular y las flechas divisiones celulares anormales. Modificada de Du y Scheres, 2017.

2 ANTECEDENTES DIRECTOS Y JUSTIFICACIÓN

En *A. thaliana* los genes *PLT* son reguladores muy importantes para el mantenimiento y establecimiento del RAM en la raíz [17, 21] y para el desarrollo de este órgano. Se ha reportado que las mutantes triples de *plt1plt2plt3* carecen de desarrollo post-embrionario de la raíz primaria [17]. Sin embargo, a pesar de su relevancia en el desarrollo del sistema radical, el papel de los genes *PLT* solo se ha estudiado exhaustivamente en *A. thaliana* y no en otras especies de plantas.

En la planta modelo *A. thaliana* los genes *PLT* tienen su máximo de expresión en la zona meristemática, tanto a nivel de proteína como de transcrito (Fig. 4 a-h); su nivel de expresión se reduce en la zona de elongación y no se detecta su expresión en la zona de diferenciación [15, 22]. Los datos obtenidos a partir del transcriptoma del ápice de la raíz primaria de *P. pringlei* sugieren que en las cactáceas los genes *PLT* se expresan en las etapas inicial, intermedia y en células diferenciadas de la etapa terminal del crecimiento, en contraste con la nula expresión de estos genes en la zona de diferenciación de la raíz de *A. thaliana* [15]. *P. pringlei* resulta un buen modelo de estudio para el agotamiento del RAM, debido a que presenta crecimiento determinado de la raíz primaria de manera

natural. Dada la relevancia de los genes *PLT* en el desarrollo de la raíz, resulta de interés estudiar la participación de estos genes en la regulación del crecimiento determinado de la raíz primaria. Actualmente se cuenta con un transcriptoma ensamblado *de novo* del ápice de la raíz primaria de *P. pringlei* así como un borrador de genoma para esta especie [15]. El acceso a genomas preliminares de cuatro cactáceas columnares, así como a los genomas de otras embriofitas depositados en bases públicas de datos, permitirá la construcción de una filogenia molecular para los genes *PLT*. Esta información será de utilidad para comprender cómo evolucionaron y se diversificaron los factores de transcripción *PLT* de las plantas.

En su trabajo de maestría, López-Valle (2015) realizó un árbol filogenético preliminar, a partir del cual se observó que las proteínas *PLT* se agrupan en dos clados. Dicho análisis se realizó con el método de Neighbor Joining para agrupar a las secuencias de *P. pringlei* y *A. thaliana* de acuerdo con su similitud. En este proyecto se plantea construir un árbol filogenético con el método de Maximum Likelihood, realizando previamente un análisis de selección de modelo evolutivo para encontrar a aquel que se ajuste mejor a las secuencias de algunos géneros de especies de embriofitas cuyo genoma está secuenciado. Con base en la topología del árbol filogenético que se obtuvo previamente, López-Valle sugirió que el ancestro común de ambos clados sería similar a *PLT5*, sin embargo, estos análisis se llevaron a cabo con un número reducido de secuencias proteicas. En el presente trabajo, esta hipótesis podrá ser corroborada o rechazada pues la disponibilidad de un mayor número de genomas secuenciados, mejor ensamblados y anotados nos permitirá obtener una filogenia molecular con mayor resolución y soporte estadístico.

Adicionalmente, Huang y Schiefelbein [29] obtuvieron el transcriptoma de cada zona de desarrollo de la raíz primaria de siete especies de plantas: *A. thaliana*, *Cucumis sativus*, *Glycine max*, *Oryza sativa*, *Selaginella moellendorffi*, *Solanum lycopersicum*, y *Zea mays* [29]. A pesar de las diferencias entre cada especie, se demostró que en general existen programas transcripcionales conservados en el desarrollo en la raíz de las plantas vasculares. Sin embargo, Huang y Schiefelbein no analizaron la abundancia de los transcritos *PLT*, por lo tanto, en este trabajo también se propone analizar la expresión virtual de los transcritos de los genes *PLT* en las angiospermas con transcriptomas secuenciados de la raíz.

3 OBJETIVOS

3.1 General

- Estudiar la evolución de las proteínas PLT de las embriofitas y su posible participación en el crecimiento determinado de la raíz primaria de la cactácea *P. pringlei*.

3.2 Particulares

- Identificar a las proteínas AP2 y PLT en las embriofitas.
- Construir una filogenia molecular de las proteínas PLT de las embriofitas.
- Identificar a las proteínas PLT en plantas vasculares sin semilla.
- Identificar los posibles ortólogos a los genes *PLT* de *A. thaliana* en *P. pringlei* y otras cactáceas.
- Analizar la expresión de los genes *PLT* de cactáceas y otras angiospermas.

4 MATERIALES Y MÉTODOS

4.1 Identificación de los genes *PLT* en las plantas

Se realizó la búsqueda por homología de secuencia de los genes *PLT* de angiospermas y gimnospermas mediante el algoritmo tBLASTn con una matriz de sustitución BLOSUM62, implementado en la base de datos Phytozome¹. Se utilizó como sonda a la secuencia de aminoácidos de la proteína PLT1 (AT3G20840) de *A. thaliana*. La búsqueda de genes *PLT* en el genoma de especies de la familia Solanaceae se realizó en la base de datos SolGenomics Network². Adicionalmente se utilizó la base de datos del National Center for Biotechnology Information³ (NCBI) para especies adicionales de plantas.

Para identificar a los transcritos *PLT* de *P. pringlei* se realizó una búsqueda mediante tBLASTn de las secuencias de las proteínas *PLT* de *A. thaliana* contra el transcriptoma de *P. pringlei* utilizando el programa de CLC Genomics Workbench (versión 9.5.4) con un costo de gaps de 5 por apertura y 2 por extensión. Esta búsqueda se realizó en el transcriptoma de la raíz primaria de la cactácea *P. pringlei* así como en los borradores de genomas de las cactáceas *Carnegiea gigantea*, *Lophocereus schottii*, *Pereskia humboldtii* y en *Stenocereus thurberi*.

Se consideraron como secuencias de proteínas *PLT* aquellas que tuvieron un e-value máximo de 1×10^{-10} y que se pudieron traducir en una secuencia peptídica con longitud mayor a 100 aminoácidos.

4.2 Identificación de los genes *PLT* en plantas vasculares sin semilla

Se realizó un tBLASTn con una matriz de sustitución BLOSUM62 en los genomas de los helechos *Azolla filiculoides* y *Salvinia cucullata* en la base de datos FernBase⁴ y se seleccionaron las secuencias con un e-value menor a 1×10^{-10} .

También se analizó el transcriptoma de *Isoetes tegetiformans* e *Isoetes spp.* [30], el cual se obtuvo del servidor de DRASearch⁵ con acceso preliminar del trabajo de Hetherington, *et. al* [30]. Se realizó un BLAST con las secuencias de proteínas *PLT* de *A. thaliana* mediante el programa CLC bajo los parámetros por

¹ <https://phytozome.jgi.doe.gov/pz/portal.html#>

² <https://solgenomics.net/>

³ <https://www.ncbi.nlm.nih.gov/>

⁴ <https://www.fernbase.org/>

⁵ <https://trace.ddbj.nig.ac.jp/DRASearch/study?acc=SRP135936>

default y se seleccionaron aquellas secuencias que cumplieran con los criterios antes utilizados.

4.3 Verificación de las secuencias y alineamiento múltiple.

Cada una de las secuencias se validaron manualmente al verificar la presencia de los dos dominios AP2 (IPR036955) que caracterizan a las proteínas PLT. Se utilizó en el servidor InterPro del European Bioinformatics Institute, EBI¹. Las secuencias que incluyeron dominios adicionales, más de dos dominios AP2, o aquellas que tuvieron un solo dominio AP2, se descartaron. Las secuencias verificadas se alinearon con el algoritmo Clustal omega².

4.4 Selección de modelo evolutivo.

Se realizó un alineamiento múltiple de las secuencias de *P. pringlei* y otro para los siguientes grupos de secuencias: Brassicales, Caryophyllales, Brassicales+Caryophyllales. El alineamiento múltiple de cada grupo se utilizó para encontrar al modelo evolutivo al que mejor se ajustan datos. La selección de modelo evolutivo se realizó al utilizar el programa Molecular Evolutionary Genetic Analysis (MEGA V10.0.5) mediante el método de Maximum Likelihood (ML) con un umbral de 65% de delección parcial de gaps.

4.5 Análisis filogenético.

A partir del alineamiento múltiple de todas las secuencias de proteínas AP2 se construyó una filogenia con el método Neighbor-joining (NJ) al utilizar el programa MEGA y el modelo evolutivo que se obtuvo del análisis anterior. La robustez estadística de la topología filogenética se evaluó mediante un análisis de Bootstrap de 1,000 réplicas, con una distribución gamma de 1 y delección parcial de gaps del 65%. Se tomaron las secuencias que se agrupaban en el mismo nodo monofilético con las proteínas PLT de *A. thaliana* y se construyó un árbol filogenético adicional a partir de estas secuencias mediante el método ML y el modelo evolutivo obtenido anteriormente.

¹ <http://www.ebi.ac.uk/interpro/>

² <https://www.ebi.ac.uk/Tools/msa/clustalo/>

4.6 Análisis de los dominios AP2 de las proteínas PLT.

La visualización de la conservación de los dominios AP2 de las secuencias que codifican para proteínas PLT se realizó en la plataforma WEBLOGO¹ con los parámetros por default y 25 símbolos por línea. Se identificó el inicio y fin de los dominios AP2 para cada proteína PLT de *A. thaliana* y para cada secuencia PLT ortóloga putativa de *P. pringlei* en el servidor InterPro. Se tomaron las secuencias que solo codificaban para los dos dominios AP2 y la región interdominio de cada proteína PLT para ambas especies de plantas y se realizó una matriz de identidad y similitud.

4.7 Obtención de material biológico.

Las semillas de *P. pringlei* (S. Watson) Britton & Rose que se utilizaron se obtuvieron a partir de frutos colectados en el Desierto de Sonora, en Sonora, México; en las localidades de Guaymas (27.9288, -110.8888) y Bahía de Kino (28.831476, -111.803929) en junio de 2017. Las semillas se esterilizaron al colocarlas en un infusor y sumergir a éste en solución superkill, la cual contiene 60% de cloro comercial (hipoclorito de sodio 4-6% e hidróxido de sodio 0.5-0.10%) y Tritón x-100 0.8%. Las semillas se colocaron dentro de un infusor y se pusieron en agitación durante 20 minutos. Posteriormente, las semillas se enjuagaron con 4 lavados (con agua destilada y esterilizada durante 10 minutos cada uno) en agitación a 160 RPM. Las semillas esterilizadas se sembraron en el medio de cultivo Linsmaier & Skoog (LS, PhytoTechnology Laboratories) 0.2X, sin sacarosa, el cual se gelificó con agar 0.8% (Bacto™ agar; Becton, Dickinson and company). Las cajas Petri con semillas sembradas se colocaron en posición vertical y se incubaron a 28 °C con un fotoperíodo 12 h/12 h de luz y oscuridad, en una cámara de crecimiento Precision Scientific.

Se colectaron ápices de la raíz primaria de *P. pringlei* en tres etapas de desarrollo: a) inicial, cuando el RAM aún está presente en la raíz; b) intermedia, cuando el RAM sigue presente en la raíz, aunque es más pequeño y los pelos radicales se acercan al ápice de la raíz; y c) terminal; cuando el RAM se ha agotado y en consecuencia todas las células del ápice se diferenciaron; en esta etapa se observan pelos radicales en todas las células de la epidermis, incluyendo a las células del ápice. La muestra para la etapa inicial incluyó ápices de ~0.5 mm (~2-3 días después de germinación, d.d.g) y 200-250 ápices de raíz por réplica; la muestra para la etapa intermedia incluyó ápices de ~0.7 mm (~5-7 d.d.g) con 300-350 ápices por muestra, mientras que la muestra para la etapa terminal incluye

¹ <https://weblogo.berkeley.edu/logo.cgi>

ápices de ~1 mm (~7-10 d.d.g) con 400-450 ápices por réplica. Se colectaron tres réplicas biológicas por etapa.

4.8 Aclaración de raíces y análisis de la raíz primaria mediante microscopía DIC.

Se aclaró la raíz primaria de *P. pringlei* en las tres etapas de desarrollo. Para la etapa inicial se utilizaron 3 raíces, las cuales tenían una edad de 3 d.d.g.; para la etapa intermedia se utilizaron 4 raíces con 7 d.d.g; mientras que para la etapa terminal se utilizaron 4 raíces 9 d.d.g. Adicionalmente, se realizó el aclaramiento de 3 raíces de plantas con un mes de edad.

Las raíces se cortaron y se colocaron en una caja de ELISA de 16 pozos, colocando una raíz por pozo. Se añadió solución de fijación (Tritón X-100 0.1%, paraformaldehído 4% y MTBS 0.2 X) hasta cubrir la raíz por completo. Posteriormente, la caja se cubrió con papel aluminio y se incubó a 4 °C durante toda la noche. Se retiró la solución de fijación y se añadió la solución de cambio (glicerol 30% y dimetilsulfoxido o DMSO 2%), para incubar a 4 °C durante 2 a 3 días. Finalmente, se retiró la solución de cambio y se añadió la solución de aclaración, la cual está compuesta por 17 g de NaI y 0.04 g de Na₂S₂O₃ disueltos en glicerol al 65% con DMSO 2% en un volumen final de 25 ml, para incubarse durante 3 días a 4 °C.

Se montaron de 2 a 3 raíces sobre un portaobjetos (26.4 x 76.2 mm) al cual se le colocó un marco de parafilm (26.4 x 50 mm). Se agregó solución de aclaración a la preparación hasta llenar el marco de parafilm y un cubreobjetos (22 x 40.1 mm). Las aclaraciones de las raíces por cada etapa de desarrollo, así como de raíces de 30 d.d.g se analizaron en un microscopio Olympus BX 53. Para visualizar las raíces se utilizó la técnica de contraste de interferencia diferencial (DIC, por sus siglas en inglés) o Nomarski. Una vez ajustada la imagen, se tomaron fotos con la cámara Photometrics Cool SNAP, la cual está integrada al microscopio. Para observar y ajustar las imágenes se utilizó el programa QED Capture v2.003.

5 RESULTADOS Y DISCUSIÓN

5.1 Identificación de genes *PLT* de las plantas

Los genes *PLT* pertenecen a la subfamilia proteica *AINTEGUMENTA-LIKE* (*AIL*), de la superfamilia AP2/ERF y codifican para factores de transcripción con dos dominios AP2. Esta característica fue aprovechada para identificar a los ortólogos putativos de PLTs en las embriofitas. Para identificar a las secuencias ortólogas de los genes *PLT* en el genoma secuenciado de distintas plantas se utilizó como sonda a la secuencia de la proteína PLT1 de *A. thaliana*. Se realizó una búsqueda mediante tBLASTn en las bases de datos SolGenomics para las solanáceas; la colección de nucleótidos (nr/nt) de la base de datos del servidor del NCBI para las gimnospermas y Caryophyllales; FernBase para dos especies de helechos; y Phytozome para el resto de las especies.

Se consideró que las secuencias obtenidas podrían ser proteínas AP2 putativas cuando su longitud fue mayor a 100 aminoácidos, puesto que las proteínas PLT de *A. thaliana* tienen una longitud aproximadamente de 560 aminoácidos y tan solo ambos dominios AP2, en conjunto con su región interdominio, tienen una longitud de 160 aminoácidos aproximadamente. Además, se observó empíricamente que aquellas secuencias con un e-value mayor a 1×10^{-10} tenían una longitud menor a 100 aminoácidos, por lo cual se filtraron a las secuencias conservando sólo a aquellas que tuvieron un e-value menor a 1×10^{-10} , con el objetivo de disminuir la probabilidad de falsos positivos, es decir, para evitar que la secuencia resultante del tBLASTn contra el genoma de la planta sea similar a la sonda sólo por azar.

Se realizó el análisis de dominios de cada secuencia obtenida con la herramienta InterPro del EMBL-EBI. Esta herramienta reconoce firmas específicas de dominios proteicos y se utilizó para comprobar que las secuencias sólo incluyen dos dominios AP2. Después de este análisis se obtuvieron 1,161 secuencias que codifican para proteínas de la subfamilia AIL. Estas secuencias provienen del genoma de 61 especies, las cuales se agrupan en 54 géneros, 32 familias y 26 órdenes taxonómicos, los cuales abarcan desde las briofitas hasta las angiospermas (Tabla 1).

Tabla 1. Especies de plantas en las cuales se identificaron secuencias de proteínas ALL.

Orden	Familia	Genero	Especie	# secuencias
Amborellales	Amborellaceae	<i>Amborella</i>	<i>Amborella trichopoda</i>	12
Apiales	Apiaceae	<i>Daucus</i>	<i>Daucus carota</i>	10
Asparagales	Asparagaceae	<i>Asparagus</i>	<i>Asparagus officinalis</i>	22
Brassicales	Brassicaceae	<i>Arabidopsis</i>	<i>Arabidopsis halleri</i>	13
			<i>Arabidopsis thaliana</i>	13
			<i>Arabidopsis lyrata</i>	14
		<i>Brassica</i>	<i>Brassica rapa</i>	16
		<i>Boechera</i>	<i>Boechera stricta</i>	14
		<i>Capsella</i>	<i>Capsella grandiflora</i>	14
			<i>Capsella rubella</i>	22
		<i>Thellungiella</i>	<i>Thellungiella halophila</i>	3
	Caricaceae	<i>Carica</i>	<i>Carica papaya</i>	13
Caryophyllales	Chenopodiaceae	<i>Chenopodium</i>	<i>Chenopodium quinoa</i>	16
		<i>Beta</i>	<i>Beta vulgaris</i>	17
	Amaranthaceae	<i>Spinacia</i>	<i>Spinacia oleracea</i>	15
		<i>Amaranthus</i>	<i>A. hypochondriacus</i>	10
		<i>Pachycereus</i>	<i>Pachycereus Pringlei</i>	5
	Cactaceae	<i>Stenocereus</i>	<i>Stenocereus thurberi</i>	2
		<i>Pereskia</i>	<i>Pereskia humdoldtii</i>	1
		<i>Lophocereus</i>	<i>Lophocereus schottii</i>	3
<i>Carnegiea</i>	<i>Carnegiea gigantea</i>	2		
Cucurbitales	Cucurbitaceae	<i>Cucumis</i>	<i>Cucumis sativus</i>	29
Fabales	Fabaceae	<i>Trifolium</i>	<i>Trifolium pratense</i>	20
		<i>Phaseolus</i>	<i>Phaseolus vulgaris</i>	18
		<i>Medicago</i>	<i>Medicago truncatula</i>	26
		<i>Glycine</i>	<i>Glycine max</i>	45
		<i>Physcomitrella</i>	<i>Physcomitrella patens</i>	17
Gentianales	Rubiaceae	<i>Coffea</i>	<i>Coffea arabica</i>	23
Ginkgoales	Ginkgoaceae	<i>Ginkgo</i>	<i>Ginkgo biloba</i>	3
Isoetales	Isoetaceae	<i>Isoetes</i>	<i>Isoetes spp.</i>	2
			<i>Isoetes tegetiformans</i>	2
Lamiales	Phrymaceae	<i>Mimulus</i>	<i>Mimulus guttatus</i>	12
Malpighiales	Salicaceae	<i>Populus</i>	<i>Populus trichocarpa</i>	32
	Linaceae	<i>Linum</i>	<i>Linum usitatissimum</i>	30
	Euphorbiaceae	<i>Manihot</i>	<i>Manihot esculenta</i>	28
		<i>Ricinus</i>	<i>Ricinus communis</i>	14
Malvales	Malvaceae	<i>Theobroma</i>	<i>Theobroma cacao</i>	17

		<i>Gossypium</i>	<i>Gossypium hirsutum</i>	44
			<i>Gossypium raimondii</i>	31
Marchantiales	Marchantiaceae	<i>Marchantia</i>	<i>Marchantia polymorpha</i>	16
Myrtales	Myrtaceae	<i>Eucalyptus</i>	<i>Eucalyptus grandis</i>	19
Pinales	Pinaceae	<i>Picea</i>	<i>Picea glauca</i>	1
			<i>Picea abies</i>	2
Poales	Poaceae	<i>Oropetium</i>	<i>Oropetium thomaeum</i>	17
		<i>Panicum</i>	<i>Panicum hallii</i>	25
			<i>Panicum virgatum</i>	42
		<i>Brachypodium</i>	<i>Brachypodium distachyon</i>	15
			<i>Brachypodium stacei</i>	24
		<i>Zea</i>	<i>Zea mays</i>	36
		<i>Oryza</i>	<i>Oryza sativa</i>	21
		<i>Sorghum</i>	<i>Sorghum bicolor</i>	22
	Bromeliaceae	<i>Ananas</i>	<i>Ananas comosus</i>	19
Ranunculales	Ranunculaceae	<i>Aquilegia</i>	<i>Aquilegia coerulea</i>	17
Rosales	Rosaceae	<i>Prunus</i>	<i>Prunus persica</i>	17
		<i>Fragaria</i>	<i>Fragaria vesca</i>	16
		<i>Malus</i>	<i>Malus domestica</i>	34
Salviniales	Salviniaceae	<i>Salvinia</i>	<i>Salvinia cucullata</i>	6
		<i>Azolla</i>	<i>Azolla filiculoides</i>	8
Sapindales	Rutaceae	<i>Citrus</i>	<i>Citrus sinensis</i>	28
			<i>Citrus clementina</i>	19
Saxifragales	Crassulaceae	<i>Kalanchoe</i>	<i>Kalanchoe fedtschenkoi</i>	14
Selaginellales	Selaginellaceae	<i>Selaginella</i>	<i>Selaginella moellendorffii</i>	19
Solanales	Solanaceae	<i>Solanum</i>	<i>Solanum melongena</i>	14
			<i>Solanum tuberosum</i>	18
			<i>Solanum lycopersicum</i>	20
		<i>Capsicum</i>	<i>Capsicum annuum</i>	20
Sphagnales	Sphagnaceae	<i>Sphagnum</i>	<i>Sphagnum fallax</i>	9
Vitales	Vitaceae	<i>Vitis</i>	<i>Vitis vinifera</i>	16
Zingiberales	Musaceae	<i>Musa</i>	<i>Musa acuminata</i>	23

* Los datos se obtuvieron del transcriptoma ensamblado por Rodríguez-Alonso *et al.*, 2018

** Los resultados de estas especies provienen de la sección 5.4

5.2 Predicción de modelo evolutivo.

Para identificar la tasa de sustitución de aminoácidos durante la evolución de las proteínas AIL, se realizó la predicción del modelo evolutivo. Para este análisis se obtuvo, por separado, un alineamiento múltiple con las secuencias del orden Caryophyllales (*Beta vulgaris*, *Chenopodium quinoa*, *Pachycereus pringlei* y *Spinacia oleracea*), debido a que en este orden se encuentra la especie de

estudio, *P. pringlei*; y un alineamiento adicional para Brassicales (*Arabidopsis halleri*, *A. thaliana*, *A. lyrata*, *Boechera stricta*, *Brassica rapa*, *Capsella grandiflora*, *Capsella rubella*, *Carica papaya*, y *Thellungiella halophila*), debido a que en este orden se encuentra la especie *A. thaliana*, la cual es la planta modelo mejor estudiada y será la especie contra la que se compararán los resultados obtenidos. Se realizó un alineamiento múltiple de las secuencias ortólogas putativas de las proteínas AIL de ambos órdenes (Brassicales y Caryophyllales) y se repitió el análisis (Tabla 2 y Anexos: Tabla 1 y 2) para comprobar que el modelo evolutivo es el mismo para todas las especies. Los resultados de estos tres análisis muestran que las proteínas PLT se ajustan al modelo de sustitución Jones-Taylor-Thornton (JTT), el cual se eligió debido a que los valores más altos del criterio de información bayesiana (BIC, por sus siglas en inglés) y el criterio de información Akaike (AIC, por sus siglas en inglés) así como el logaritmo natural de verosimilitud más negativo corresponden a este modelo (Anexo, Tabla 2) [31].

5.3 Análisis filogenético.

Para identificar a las proteínas PLT entre las 1,161 secuencias AIL que se obtuvieron, se construyó un árbol filogenético mediante el método *Neighbor-Joining* (NJ) al utilizar el modelo JTT. Este análisis se realizó anteriormente en el laboratorio [32], sin embargo, ahora contamos con 1,161 secuencias de proteínas AIL en comparación con las 309 obtenidas en dicho trabajo. Dado que el conjunto de secuencias a analizar en este trabajo fue, aproximadamente, cuatro veces más grande, inicialmente se utilizó el algoritmo de NJ debido que es más práctico computacionalmente al procesar un número grande de secuencias. La construcción de la filogenia se realizó con 500 réplicas Bootstrap y una delección parcial de gaps del 65%, es decir, solo se analizaron aquellas columnas del alineamiento que tuvieran al menos el 65% de sitios ocupados por un aminoácido.

A partir del árbol que se obtuvo (Anexos: Figura 1) se extrajeron todas las secuencias que derivan del nodo monofilético que incluye a los PLT de *A. thaliana*. A partir de este análisis, las secuencias obtenidas se dividieron en 299 secuencias putativas de proteínas PLT y 862 secuencias de proteínas AIL (Anexos: Figura 1). Se observó que las proteínas PLT están presentes en todas las especies de plantas analizadas con excepción de la hepática *Marchantia polymorpha*, y de dos musgos, o briofitas en sentido estricto (*Sphagnum fallax* y *Physcomitrella patens*). (Fig. 7). Es necesario resaltar que estas tres especies pertenecen al clado Bryophyta en sentido amplio (*sensu lato*), es decir, son plantas no vasculares y sin raíces. Además de las hepáticas y musgos, el clado Bryophyta también incluye a los antoceros. Posterior a este análisis y hacia el final de este proyecto se liberó el transcriptoma y genoma de *Anthoceros angustus* (número de acceso:

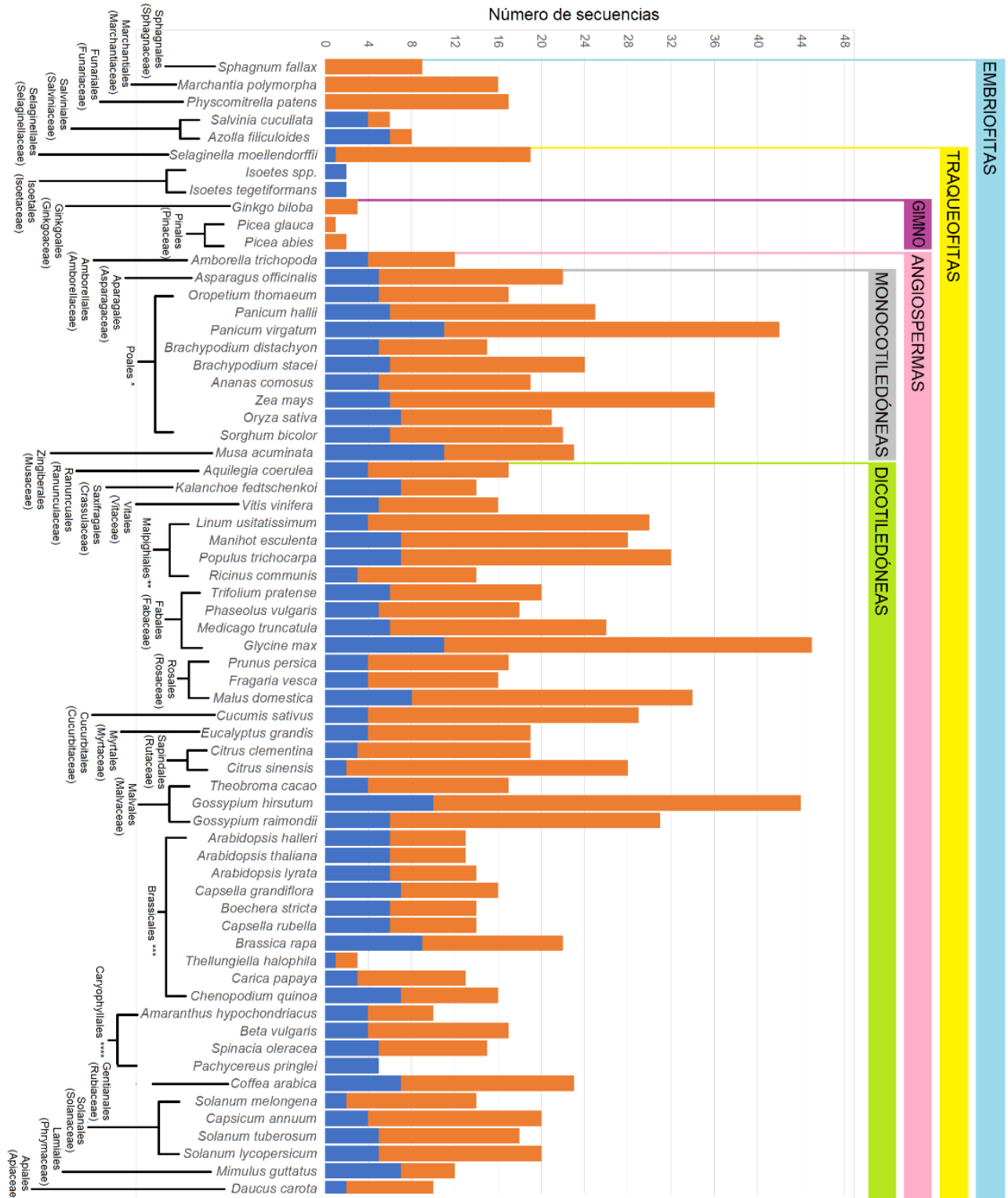


Figura 7. Proteínas PLT y AIL en las especies de plantas analizadas. En las plantas cuyo genoma está secuenciado, se identificó a los genes que codifican para proteínas con 2 dominios AP2, la cual es una característica de las proteínas PLT. Para el orden Poales* se muestran especies de las familias Poaceae y Bromeliaceae; el orden Malpighiales** a las familias Linaceae, Euphorbiaceae y Salicaceae; el orden Brassicales*** a las familias Brassicaceae y Caricaceae; y el orden Caryophyllales**** a las familias Amaranthaceae y Cactaceae. ▼secuencias traducidas a partir de transcritos de *P. pringlei*

PRJNA543716), en un análisis realizado por Zhang, J. *et al.*, 2020, por lo que se realizó la búsqueda de proteínas AP2 en esta especie [33]. La búsqueda se realizó mediante tBLASTn en el programa CLC (Genomics Workbench V 9.5.4) con los parámetros por default. No se encontraron proteínas AIL que cumplieran con los requisitos establecidos tales como un e-value mayor a 1×10^{-10} . Sin embargo, Zhang, J. *et al.* [33] sí reportan proteínas AP2, esto puede explicarse a partir de la diferencia en los umbrales de búsqueda y reconocimiento de proteínas.

Nuestros resultados sugieren que las proteínas PLT evolucionaron a partir de las proteínas AIL después de la divergencia entre las briofitas y las plantas vasculares, lo cual ocurrió hace ~360 M.A., en el período devónico temprano [1, 2]. Otro grupo de plantas en los cuales no fue posible identificar proteínas PLT fue el de las gimnospermas *Ginkgo biloba*, *Picea glauca*, y *Picea abies* (Fig. 7). Esto podría ser explicado al considerar que la búsqueda de proteínas PLT de las gimnospermas se realizó en borradores de genoma, fragmentados y de baja cobertura, por lo que la ausencia de los genes *PLT* podría deberse a que los *loci* correspondientes no se ensamblaron correctamente en las versiones disponibles del genoma.

En *Arabidopsis* los genes *PLT* comprenden a *PLT1-PLT5* y *PLT7*. Para identificar a los clados de los posibles ortólogos de estas proteínas en todas las especies de plantas analizadas se realizó un árbol filogenético con el algoritmo Maximum-Likelihood (ML) debido a que este método filogenético es más exhaustivo que el método NJ y más robusto. Para la construcción de la filogenia mediante el método ML se tomaron las 299 secuencias de proteínas PLT putativas identificadas anteriormente (Fig. 8).

En la filogenia obtenida se puede observar que las proteínas PLT se agrupan en cuatro clados, los cuales corresponden a PLT1/PLT2, PLT3/PLT7, PLT4 y PLT5 de *A. thaliana* (Fig. 8). Además, se observa un grupo de secuencias provenientes de plantas monocotiledóneas que no se agrupan en los clados antes mencionados. Esto sugiere que existió una duplicación génica después de la divergencia entre plantas mono y dicotiledóneas, lo cual ocurrió hace 170 M.A., aproximadamente [34]¹.

En la topología del árbol se identificó a un posible ortólogo de *P. pringlei* para cada proteína PLT de *A. thaliana*, con excepción de PLT5. Se sabe que las proteínas PLT1-4 están involucradas en el mantenimiento del RAM en *A. thaliana* [17]. En *P. pringlei* encontramos un ortólogo para cada una de estas proteínas

¹ Consultado en <http://www.mobot.org/MOBOT/Research/APweb/>.

dentro de cada clado identificado en la filogenia. Esta observación es congruente con lo reportado por diversos autores, incluyendo a Huang y Schiefelbein en el 2015, quienes mostraron que, en general, los programas genéticos del desarrollo de la raíz están conservados en las traqueofitas [29]; por Ferrari *et. al.* en el 2019, quienes observaron procesos conservados en el mantenimiento de la raíz en la planta basal *S. moellendorffi* [35]; y por Hetherington *et. al.*, en el 2019, quienes secuenciaron el transcriptoma de las dos especies de *Isoetes* que fueron utilizadas en este proyecto [30].

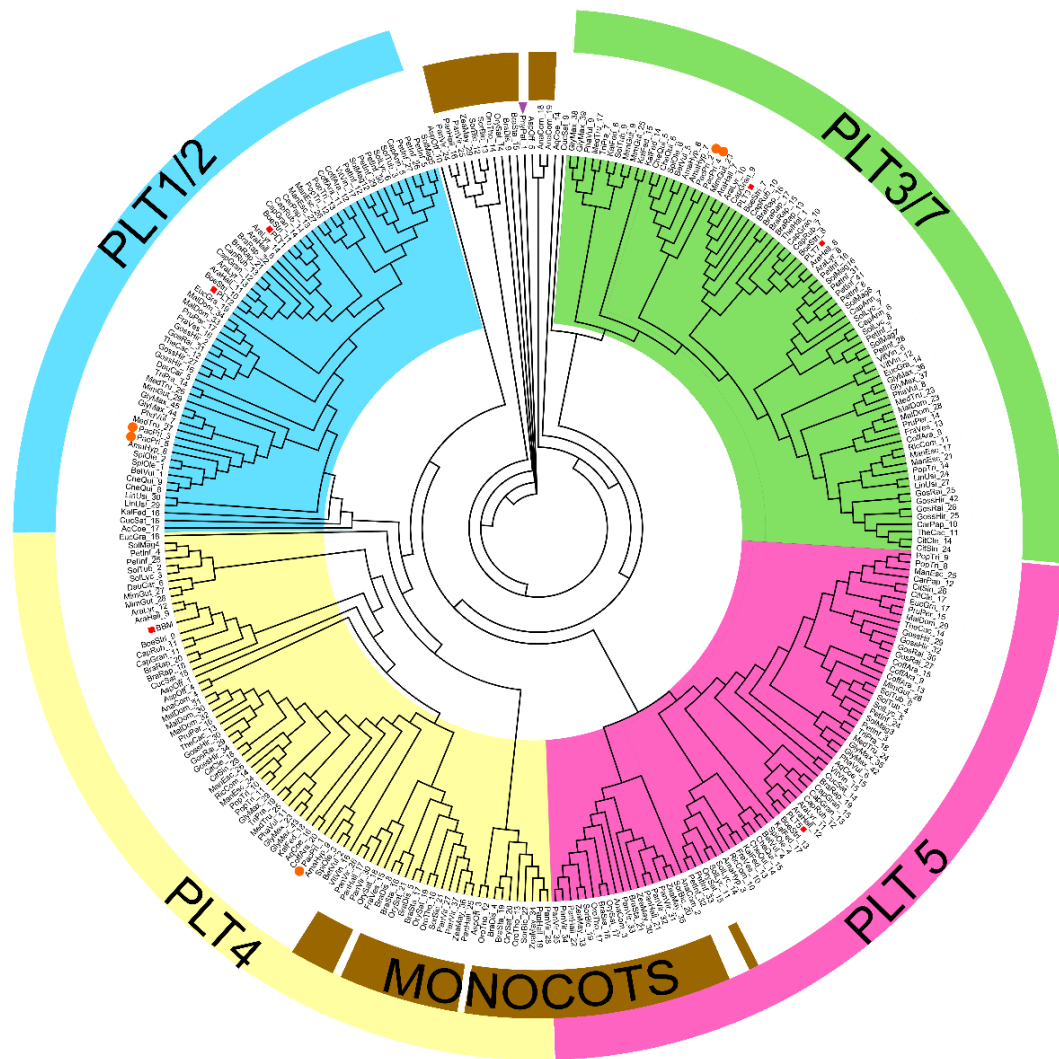


Figura 8. Filogenia molecular de las proteínas PLT de las embriofitas. El árbol filogenético se obtuvo con el método de ML con 1,000 réplicas de bootstrap bajo el modelo JTT. Se muestran los clados de cada una de las proteínas PLT. Proteínas PLT de *P. pringlei* (●); *A. thaliana* (■); y Monocotiledóneas (color café). Árbol enraizado con la secuencia de una proteína AIL de *Physcomitrella patens* (▼).

Las traqueofitas, es decir, las plantas vasculares, comprenden a dos grandes grupos: las Euphyllophyta y las Lycophyta. En estos grupos la raíz evolucionó de manera independiente. En la actualidad, las raíces de las Lycophyta (Lycopodiales, Selaginellales, Isoetales) muestran caracteres ancestrales, tales como ramificación de la raíz por bifurcación del RAM en Isoetales y Selaginellales; o la ausencia de endodermis en la raíz, como sucede en *Lycopodium*. Huang y Schiefelbein [29] demostraron que la lycophyta *S. moellendorffii* expresa un programa molecular similar al del desarrollo de la raíz de las angiospermas. Una observación similar fue reportada por Hetherington *et al.*, para los Isoetales [30]. Es decir, existe convergencia a nivel molecular y de órgano a pesar del origen independiente de la raíz en las Euphyllophyta y Lycophyta. Para estudiar la posible conservación de las proteínas PLT en este programa convergente del desarrollo de la raíz, realizamos la búsqueda e identificación de proteínas PLT en plantas vasculares sin semilla.

5.4 Análisis filogenético en plantas vasculares sin semilla

Las plantas vasculares sin semilla están compuestas por las Lycophytas y Monilophytas. Para las Monilophytas existen genomas disponibles de los helechos *Azolla filiculoides* y *Salvinia cucullata*; para las Lycophytas solo se dispone del genoma de *Sellaginella moellendorffii* y del transcriptoma de los Isoetales, como *Isoetes tegetiformans* e *Isoetes spp.*, a los cuales tuvimos acceso mediante comunicación personal con el Dr. Alexander Hetherington [30].

Se identificaron a las proteínas AIL de las plantas vasculares sin semilla mediante un tBLASTn y las secuencias se filtraron conforme lo descrito en los métodos. Se obtuvieron ocho secuencias AIL en *A. filiculoides*, seis en *S. cucullata* y dos secuencias para cada una de las Isoetales. Estas secuencias se alinearon con la secuencia PLT identificada en el genoma de *S. moellendorffii*, las seis secuencias de proteínas PLT y seis secuencias de proteínas AIL de *A. thaliana*; así como una secuencia AIL del musgo *P. patens* como grupo externo. Con el alineamiento múltiple se construyó una filogenia con el método ML y el modelo JTT con 1,000 réplicas de Bootstrap (Fig. 9).

En la filogenia se observa que existen proteínas PLT en todas las especies analizadas de plantas vasculares sin semilla. La topología obtenida muestra que las proteínas PLT1-4 y PLT7 de *A. thaliana* se agrupan en un clado mientras que la proteína PLT5 agrupa con las secuencias de Lycophyta (Fig. 9). Esto podría sugerir que PLT5 es la secuencia ancestral a partir de la cual las demás proteínas PLT divergieron. Sin embargo, este resultado debe corroborarse mediante análisis

diferencias de las demás especies analizadas en este proyecto, tales como: presencia de hojas suculentas, un tallo leñoso y la presencia de raíces con crecimiento indeterminado.

Se realizó una filogenia bajo el método ML con 1,000 réplicas de Bootstrap y con el método evolutivo JTT. El análisis incluyó 6 secuencias AIL de: 2 de *S. thurberi* y de *C. gigantea*; 1 de *P. humboldtii*; 3 de *Lophocereus schottii* [36] y de *A. thaliana*. Así mismo, 5 secuencias PLT putativas de *P. pringlei* [15]; 6 secuencias PLT de *A. thaliana*; y 1 secuencia AIL del musgo *P. patens* para enraizar la topología. La filogenia (Fig. 10) muestra dos clados: el clado de las proteínas PLT, señalado en azul, y el clado de las proteínas AIL, en negro. Se puede observar que todas las secuencias analizadas de especies de la familia Cactaceae codifican para proteínas PLT, con excepción de 1 secuencia AIL de las especies *P. humboldtii*, *C. gigantea* y *S. thurberi*. A pesar de ello, las secuencias PLT putativas agrupan en un clado con un soporte de rama alto.

Se observa que, similar a la topología de los genes *PLT* para las embriofitas (Fig. 8) y las plantas vasculares sin semilla (Fig. 9), los genes *PLT1* y *PLT2* de *P. pringlei* agrupan juntos, así como *PLT3* y *PLT7*, mientras que *PLT5* no agrupa con alguna otra secuencia. Esta observación refuerza la hipótesis de que *PLT5* podría ser la secuencia ancestral a partir de la cual las demás proteínas *PLT* divergieron.

5.6 Análisis de las proteínas *PLT*.

Las proteínas *PLT* son factores de transcripción que reconocen al ADN mediante los dominios AP2, además de interactuar con otras proteínas tales como *WOX5* [37] y *TCP* [16, 38]; es decir, su función depende de su estructura terciaria y, por lo tanto, de sus estructuras secundaria y primaria. Tradicionalmente se consideraba que las proteínas *PLT* solo contienen dos dominios AP2, sin embargo, recientemente se observó la presencia de dominios intrínsecamente desordenados parecidos a priones en *PLT3*, los cuales son necesarios para su interacción con *WOX5* [37]. Para estudiar la conservación de las proteínas *PLT* en *A. thaliana* y *P. pringlei* se realizaron alineamientos pareados mediante una matriz de sustitución BLOSUM62 entre cada una de las secuencias de proteínas *PLT* de ambas especies. Los resultados de este análisis se usaron para construir una matriz de identidad y similitud (Fig. 10, Anexos: Tabla 5).

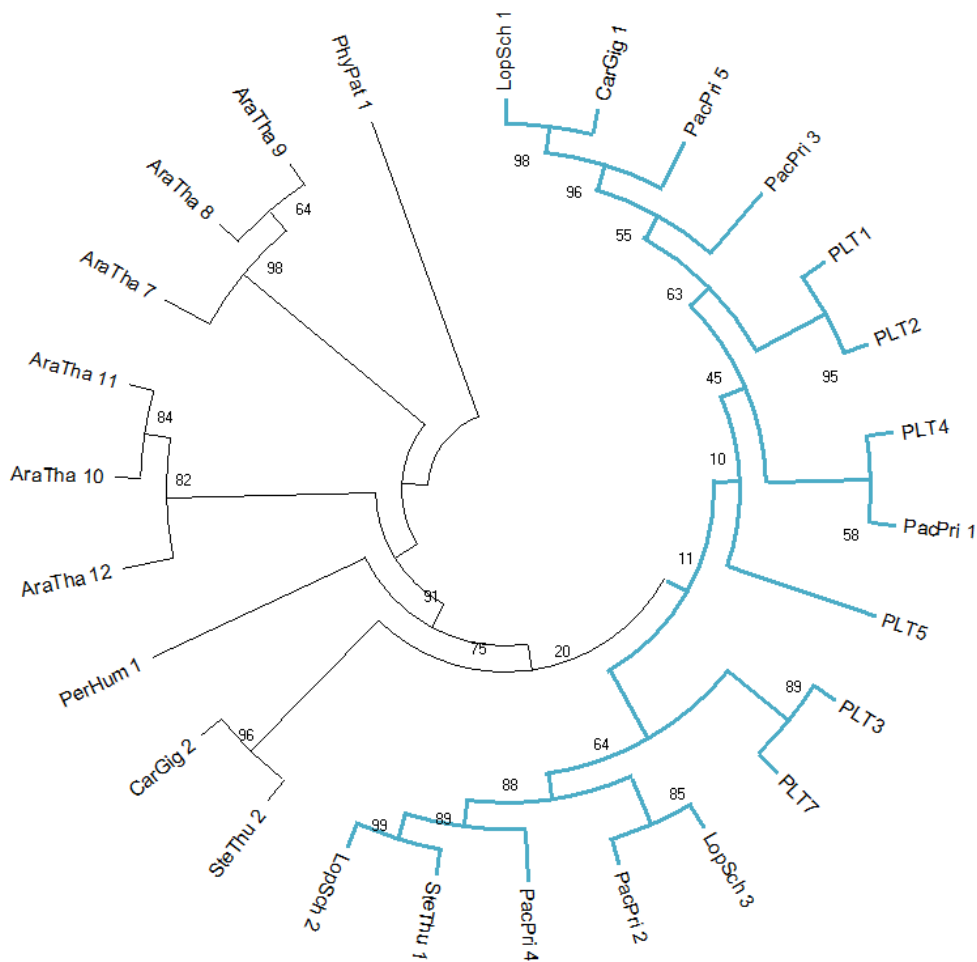


Figura 10. Topología filogenética de *PLT* en cactáceas. El clado de las proteínas *PLT* está señalado en azul y el clado de las proteínas *AIL* en negro. La topología se llevó a cabo mediante 1000 réplicas Bootstrap y el método ML.

En la matriz (Fig. 11) se observa que, en *A. thaliana* los valores de identidad van del 34 al 74.5% mientras que los de similitud van del 36.5% al 80%. Por otro lado, en *P. pringlei* los valores de identidad van del 30.9% al 63.5% mientras que los de similitud van del 37.9% al 69.1%. Si comparamos los resultados entre las dos especies, los valores van de 29 - 69.5%, y de 30.3 - 70.9% de identidad y similitud, respectivamente. Es decir, las proteínas *PLT* de *A. thaliana* son más parecidas entre sí que a las secuencias de proteínas *PLT* putativas de *P. pringlei*.

Estos resultados podrían sugerir que las proteínas *PLT* entre ambas especies, al ser poco similares, reconocen diferentes blancos en el ADN. Se ha propuesto que la región interdominio es importante para actividades transcripcionales específicas, tales como la interacción con otras proteínas o para facilitar la localización nuclear [39]. Adicionalmente, en los dominios AP2 hay una

secuencia conservada que forma una α -hélice anfipática en el dominio de unión al ADN [20]. Por esta razón, se analizó la conservación de aminoácidos de los dominios AP2 de las proteínas PLT de embriofitas (Fig. 12). Se observó que, precisamente, los dominios AP2 son altamente conservados y su variación es mínima, al igual que la región interdominio entre los dos dominios AP2.

		<i>PLT de A. thaliana</i>					<i>PLT de P. pringlei</i>					
		1	2	3	4	5	7	1	2	3	4	5
PLT de <i>A. thaliana</i>	1		80	44.9	51.6	50.6	44.2	49.3	42.9	53.6	37.7	59.5
	2	74.5		43.7	51.1	52.3	46.4	47.6	44.4	57.1	46.8	70.9
	3	36.6	35.2		48.7	44.1	66.6	47	55.3	37.6	53.8	43.4
	4	37.9	39.9	39.4		44.8	44.3	51.5	45.2	43.9	43.5	50.2
	5	39.5	39.7	35.9	35.6		36.5	43.7	46.8	41.1	46.4	38.5
	7	35.3	36.1	60	34.8	46.7		42.5	60.6	40.4	53.9	45.9
	PLT de <i>P. pringlei</i>	1	40.1	37.8	36.1	42.2	36	33.8		37.9	46	41.7
2		35	35.4	48.3	37.4	38.3	50.4	30.9		45	69.1	48.5
3		45.2	47	29	33.7	32.6	32.1	36	36.4		46.7	66.2
4		48.4	36.8	44.1	34.9	37	45.3	34	63.5	36.4		49
5		69.5	61.7	34.6	39.9	30.4	35.7	38.6	38.7	59.4	39.2	

Figura 11. Matriz de identidad y similitud entre las proteínas PLT de *A. thaliana* y los transcritos traducidos de *P. pringlei*. Los valores sobre la diagonal corresponden a la similitud entre las secuencias, mientras que los valores bajo la diagonal corresponden a la identidad entre las secuencias.

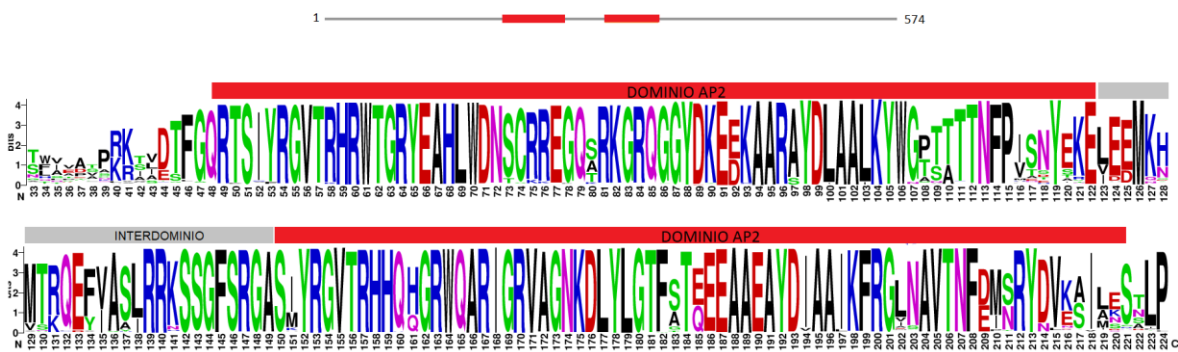


Figura 12. Los dominios AP2 están altamente conservados en todas las especies de plantas analizadas.

Para entender la conservación de las proteínas PLT en el desarrollo la raíz de *P. pringlei* se elaboró una matriz de identidad y similitud con las secuencias de los dominios AP2 y la región interdominio de las proteínas PLT de *A. thaliana* y *P.*

pringlei (Fig. 13). En la matriz de la figura 13 se observa que los dominios AP2 de las proteínas PLT están altamente conservados entre las dos especies analizadas. Por un lado, los dominios AP2 de las proteínas PLT de *A. thaliana* tienen valores de identidad que van del 74.6% al 96.1% y de similitud que van del 81.2% al 98.3%. Por otro lado, los dominios AP2 de las proteínas putativas de *P. pringlei* tienen valores de identidad que van 84.7% al 96.5% y similitud que van del 92.7% al 100%. En contraste, si comparamos las dos especies, los valores de identidad van del 75.1% al 93.8% y de similitud que van del 82.2% al 99.4%. Estos resultados muestran que, aunque las regiones C-term y N-term de las proteínas PLT son variables, los dominios AP2 están altamente conservados entre *A. thaliana* y *P. pringlei*, por lo que las proteínas PLT podrían reconocer los mismos blancos en el ADN. Para explorar esta hipótesis, se analizó la conservación de cada aminoácido que compone a los dominios AP2 de *A. thaliana* y *P. pringlei*.

		PLT de <i>A. thaliana</i>					PLT de <i>P. pringlei</i>					
		1	2	3	4	5	7	1	2	3	4	5
PLT de <i>A. thaliana</i>	1		98.3	81.5	96.7	95	91.7	97.8	92.3	93.9	92.8	97.8
	2	96.1		82.7	98.3	96.6	93.2	99.4	93.8	95.5	94.4	99.4
	3	75.5	76.6		81.2	84.4	87.4	82.2	86.9	85.8	87.4	82.7
	4	87.8	90.4	74.6		94.9	92.1	99.4	91.6	94.9	92.7	98.9
	5	90	91.5	77.6	88.7		95.3	96	96	97.7	96.5	96
	7	82.8	84.2	83.7	84.2	87.8		92.7	98.2	97.6	98.2	93.8
	PLT de <i>P. pringlei</i>	1	89.4	92.1	75.1	93.8	91	84.7		92.7	96	93.2
2		83.4	86	82.7	84.3	89.6	93.6	86.5		97.1	98.8	93.3
3		88.3	89.8	78.4	87	91.9	88.2	90.4	90.6		97.6	96
4		81.7	85.3	81.1	84.2	87.8	93.5	85.9	96.5	88.8		94.8
5		92.2	93.8	75.1	91	90.4	84.2	94.4	86.5	94.9	84.7	

Figura 13. Matriz de identidad y similitud entre los dominios de las proteínas PLT de *A. thaliana* y los transcritos traducidos de *P. pringlei* muestra un alto grado de similitud e identidad.

Se realizó un alineamiento múltiple entre las secuencias de los dominios AP2 de las especies *A. thaliana* y *P. pringlei*; encontramos una inserción de 21 aminoácidos en el segundo dominio AP2 de las secuencias de la proteína PLT3 de *A. thaliana*, sin embargo, a pesar de esta inserción, comprobamos que existe una alta conservación entre las secuencias de los dominios (Anexo: Fig. 2). Estos resultados sugieren que, si las proteínas PLT de *P. pringlei* y de *A. thaliana* están altamente conservadas, entonces podrían reconocer a los mismos blancos en el ADN en ambas especies, por lo que las diferencias en el patrón de desarrollo de la raíz entre estas *A. thaliana* y *P. pringlei* podrían ser explicados por dos hipótesis

alternativas: que las proteínas PLT de estas especies son reguladas por distintos reguladores o que las proteínas PLT interaccionan con proteínas diferentes.

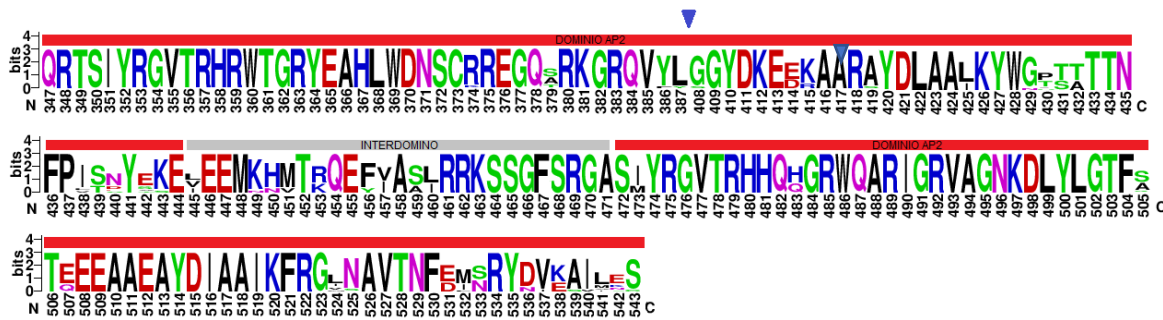


Figura 14. Los dominios AP2 de las proteínas PLT *A. thaliana* y *P. pringlei* están altamente conservados. El triángulo indica una inserción de 21 aminoácidos en PLT3 de *A. thaliana*.

Para indagar si las diferencias del patrón de desarrollo en *A. thaliana* y *P. pringlei* se deben a que las proteínas PLT se regulan de forma distinta se analizó la abundancia de los transcritos de PLT en 4 plantas. Se utilizaron los datos de dos especies de plantas monocotiledóneas (*Oryza sativa* y *Zea mays*) y dos dicotiledóneas (*Solanum lycopersicum* y *Glycine max*) obtenidos a partir del trabajo de Huang y Schiefelbein [29] para comparar los niveles de abundancia de transcrito de PLT reportadas para *A. thaliana* y los datos del transcriptoma de *P. pringlei* ensamblado por Rodríguez-Alonso *et al.* [15] (Fig. 15).

Al analizar la abundancia de los transcritos de PLT en *P. pringlei* se observó que los genes *PLT* se expresan en todas las etapas del desarrollo de la raíz e incluso en células diferenciadas en la etapa terminal y, además, presentan niveles de abundancia más altos que en comparación con *A. thaliana* en esta etapa de desarrollo. Esto contrasta con lo reportado para *A. thaliana* en los trabajos de Aida *et al.* [23], y de Galinha *et al.* [17], en donde se demostró que la pérdida de función de los genes *PLT* conduce al crecimiento determinado por el agotamiento del RAM.

Cuando observamos los niveles de abundancia de los transcritos de las proteínas PLT en otras especies observamos que PLT3 y PLT5 en dicotiledóneas siguen el patrón de expresión observado para *A. thaliana*. Para las monocotiledóneas *O. sativa* y *Z. mays* se encontró que las proteínas PLT que no se agrupan en los clados clásicos, se transcriben mayoritariamente en el ápice de la raíz que en las zonas de elongación o de diferenciación.

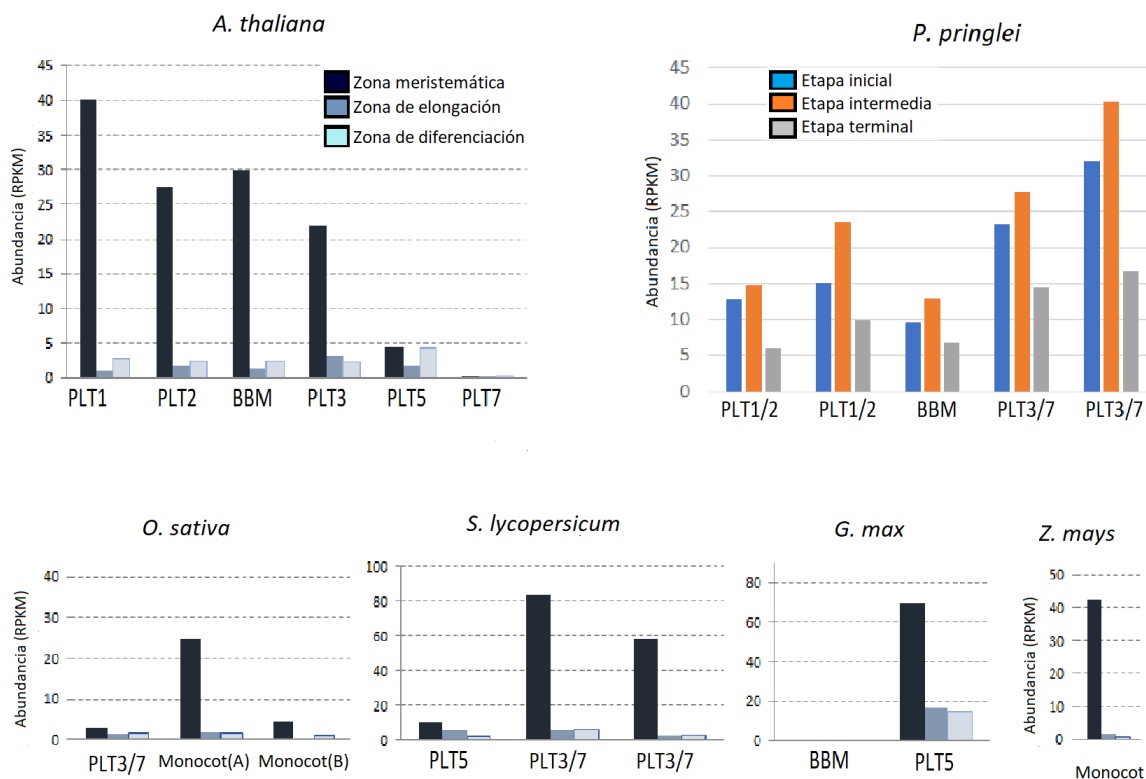


Figura 15. Abundancia de los transcritos PLT de algunas especies de plantas monocotiledóneas y dicotiledóneas. Los niveles de transcritos indican que, a diferencia de *A. thaliana*, *P. pringlei* mantiene la expresión de los genes PLT incluso en la etapa terminal, la cual incluye solo a células diferenciadas.

5.7 Análisis del dominio AP2 de las proteínas PLT de *P. pringlei*.

Las proteínas PLT pertenecen a la subfamilia AP2/ERF la cual, a su vez, pertenece a la superfamilia APETALA2 [39], una de las familias más grandes y conservadas de las plantas. Las proteínas AP2/ERF se caracterizan por poseer, al menos, un dominio AP2 que comprende entre 60 a 70 aminoácidos y están clasificadas en cuatro subfamilias; las AP2, las relacionadas a ABI3/VP1 (RAV), las proteínas de unión a elementos de respuesta a deshidratación (DREB), y los factores de respuesta a etileno (ERF) [40].

Las proteínas EREBPS reconocen al ADN mediante el dominio AP2, el cual se unen específicamente a la caja GCC, que es una secuencia conservada de 7 nucleótidos (AGCCGCC). Entre los EREBP hay una región de, aproximadamente, 60 aminoácidos altamente conservada; esta región se refiere a al dominio de unión a la caja GCC (GBD, GCC-box domain), altamente similar a nivel de aminoácidos al dominio AP2 de las proteínas PLT. La interacción entre EREBP y

la caja CGG se ha estudiado mediante cristalografía y se sabe que esta interacción está mediada principalmente por argininas [41].

Se realizó un alineamiento entre las GBD y el dominio AP2 de las proteínas PLT de *Arabidopsis* y las PLT putativas de *Pachycereus*, como se observa en la figura 16. El alineamiento muestra secuencias de algunos EREBPS y ERF analizados por Allen *et. al.* [41], y secuencias de algunas proteínas PLT de *A. thaliana* y *P. pringlei*. Con el objetivo de encontrar el sitio de unión al ADN se observó que la R162 es reemplazada por una histidina (H) en *P. pringlei* y en *A. thaliana*. Debido a que los dos aminoácidos son polares con carga y que la Histidina presenta dos grupos amino disponibles, esto sugiere que, aunque existe una sustitución de aminoácido, las interacciones con los grupos amino de la Timina y la Guanina podrían también ocurrir en el dominio AP2 de las proteínas PLT de *P. pringlei* y de *A. thaliana*.

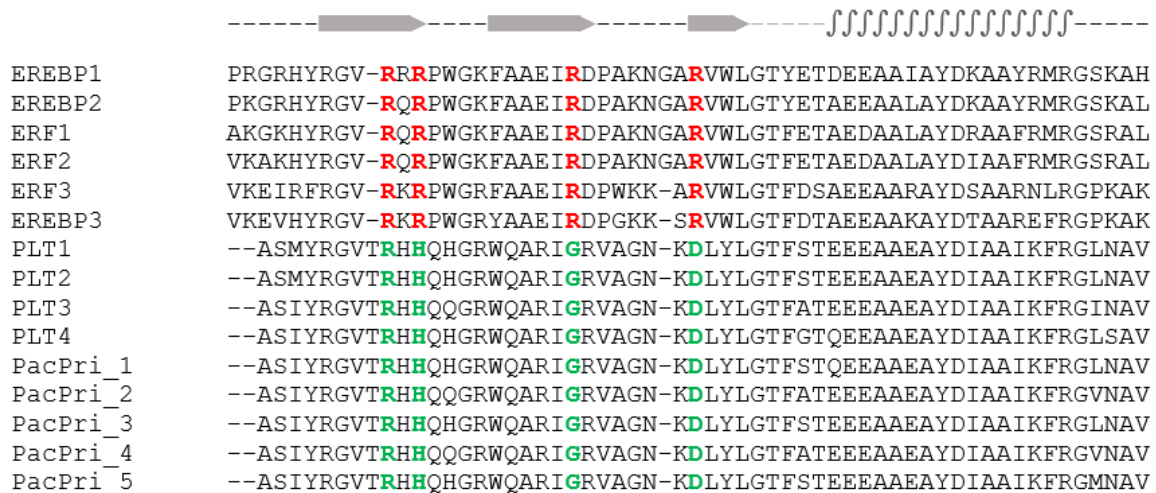


Figura 16. Alineamiento de 3 secuencias de EREBP, 3 de ERF (obtenidos del trabajo de Allen, *et al.* en 1998) con 4 PLT de *A. thaliana* (PLT1-4) y 5 PLT de *P. pringlei* (PacPri_1-5). En rojo se muestran las posiciones en los EREBP y ERF que interactúan con el DNA, mientras que se muestran en verde las posiciones de los AP2/PLT. Además, están marcadas las hebras β con un bloque y una punta, así como las hélices con el símbolo \uparrow . El alineamiento se obtuvo en Clustal y la predicción de estructura secundaria en Jpred ⁴¹.

Por otro lado, Santuari *et. al.* [21] identificaron mediante SELEX (Systematic evolution of ligands by exponential enrichment, por sus siglas en inglés) los motivos de ADN que son reconocidos por PLT2 y PLT5 (Fig. 17).

¹ <http://www.compbio.dundee.ac.uk/jpred4/index.html>

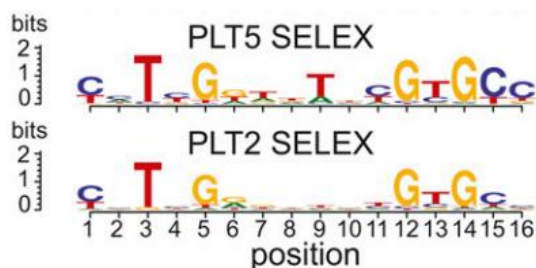


Figura 17. Motivos consenso de PLT5 y PLT2 (SELEX). Santuari, *et. al.*, 2016.

Con la finalidad de conocer el sitio de unión de las proteínas PLT putativas de *P. pringlei* se realizó una búsqueda en el banco de datos de proteínas (PDB; protein data bank) con la secuencia de la proteína PLT1 de *P. pringlei*. Para este análisis obtuvimos la estructura del cristal de AtERF96¹ (Fig. 18) el cual es un ERF de *A. thaliana*.

Se obtuvo la secuencia de aminoácidos del cristal AtERF96 y realizamos un alineamiento (Anexos: Figura 3) contra la secuencia consenso de los CBD del artículo de Allen *et al.* De la misma manera que en la figura 16, localizamos las argininas que podrían interactuar con el ADN y podemos observar que en el CBD están todas conservadas. En las proteínas PLT estas posiciones están sustituidas por una H, por lo cual los EREBPs y las PLT reconocen cajas distintas. Sin embargo, las sustituciones que se observan en *P. pringlei* y en *A. thaliana* son las mismas, por lo que este análisis refuerza la hipótesis de que las proteínas PLT de *P. pringlei* están reconociendo los mismos blancos que en *A. thaliana*.

En conjunto, los resultados del análisis del sitio de unión de las proteínas PLT y de las proteínas ERF sugieren que la interacción entre los aminoácidos y sus sitios blanco en el ADN parece estar conservada; es decir, los resultados muestran que algunos de los nucleótidos del sitio de unión de AtERF96 (TAGCCGCCAGC) que interactúan con la proteína están también conservados en la secuencia consenso de PLT5 (CCTCGGTTTTCTGCC) [21].

¹ DOI: 10.2210/pdb5WX9/pdb

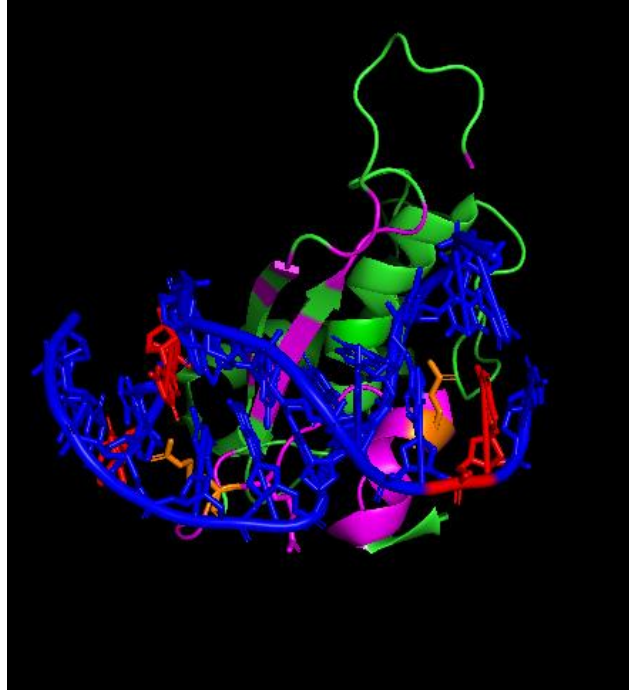


Figura 18. Estructura de AtERF96 (DOI: 10.2210/pdb5WX9/pdb) unido a un motivo de ADN. En azul se muestran las hebras de ADN, en verde la proteína y marcados en rojo las bases nitrogenadas que interaccionan con los aminoácidos marcados en naranja. Esta imagen fue obtenida por PyMOL.

5.8 Aclaración y análisis de la raíz primaria mediante microscopía DIC.

El análisis del transcriptoma de *P. pringlei* mostró que los transcritos de los genes *PLT* se expresan incluso en células diferenciadas (Fig. 15). Un posible escenario para la presencia de transcritos *PLT* en el ápice es que en la etapa terminal de raíz exista iniciación de LRP puesto que las proteínas *PLT3*, *PLT5* y *PLT7* son requeridas para este proceso (Fig. 6). Para explorar si existen primordios de raíces laterales cercanos al ápice, se aclararon 4 raíces de la etapa inicial; 3 raíces de la etapa intermedia; y 4 raíces de la etapa terminal; adicionalmente se aclararon 15 raíces de 30 d.d.g para observar si en etapas posteriores existe formación de primordios cercanos al ápice (Fig. 19). Se observó que las células que se colectaron en la etapa inicial son, principalmente, células meristemáticas; en la etapa intermedia son células meristemáticas y elongadas; y en la etapa terminal todas son células diferenciadas (Fig. 3). No se observaron

primordios de raíces laterales en las etapas inicial, intermedia, ni terminal, y tampoco fue posible identificarlos en las raíces de 30 d.d.g (Fig. 19D). Esto descarta que la expresión de las proteínas PLT en el ápice de la raíz primaria de *P. pringlei* esté asociada a la formación de LRP, es decir, la expresión de las proteínas PLT en el ápice de la raíz de *P. pringlei* en la etapa intermedia y terminal no se debe a la iniciación de primordios de raíces laterales.

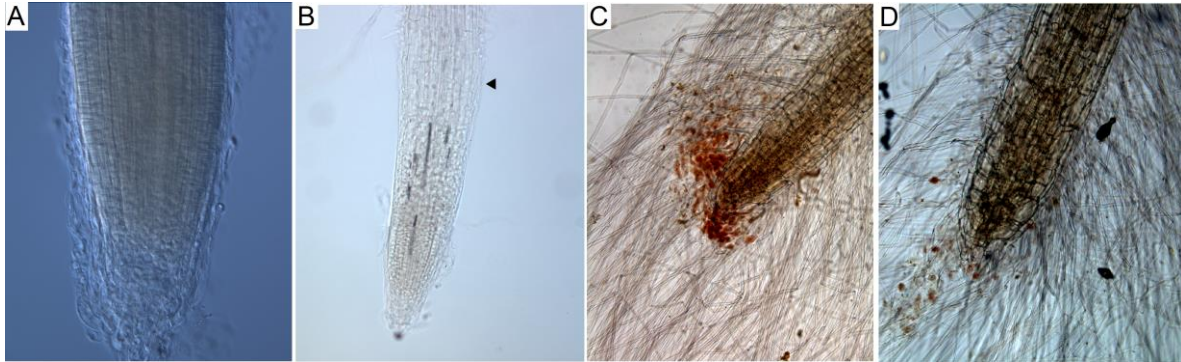


Figura 19. A) Raíz primaria de *P. pringlei* en etapa inicial, 20X. B) raíz primaria de *P. pringlei* en etapa intermedia, 10X C) raíz primaria de *P. pringlei* en etapa terminal, 10X. D) raíz primaria de *P. pringlei* 30 d.d.g., 10X.

5.9 Obtención de material biológico para cada etapa de desarrollo de la raíz primaria de *P. pringlei*.

La raíz primaria de la cactácea *P. pringlei* se puede dividir en tres etapas de desarrollo: a) inicial, cuando el RAM está presente en la raíz; b) intermedia, cuando el RAM es más pequeño, pero sigue presente y los pelos radicales están cercanos al ápice de la raíz; y c) terminal, cuando el RAM se ha agotado, todas las células del ápice están diferenciadas y como consecuencia los pelos radicales cubren toda la raíz, incluyendo al ápice. Para analizar la participación de los genes PLT se colectaron ápices de las tres etapas de desarrollo de la raíz primaria de la cactácea *P. pringlei*. Se colectaron tres réplicas biológicas para cada etapa de desarrollo de la raíz primaria de *P. pringlei*; en la etapa inicial, se colectaron ápices de 0.5 mm; para la etapa intermedia, ápices de 0.7 mm; y para la etapa terminal, ápices de 1 mm (Tabla 3). Estas muestras se utilizaron para la extracción de ARN, síntesis de ADNc y los ensayos de RT-qPCR; los cuales están en progreso.

Tabla 3. Número de ápices colectados para cada réplica biológica.

Réplica	Inicial	Intermedia	Terminal
1	351	300	400
2	212	310	440
3	245	370	410
Total	808	980	1,250

5.10 Extracción de ARN total

[Los resultados reportados en esta sección corresponden a experimentos preliminares, así como para conocer los protocolos de extracción. Por cuestiones de la cuarentena impuesta debido al COVID-19 no se pudieron culminar estos análisis con las muestras colectadas para dicho propósito].

Para comprobar la expresión virtual de los transcritos de los genes *PLT* se medirá la abundancia de transcrito de los genes *PLT* en cada una de las etapas de desarrollo de la raíz primaria de *P. pringlei*.

Debido a que la tercera réplica biológica no está completa (Tabla 3), las pruebas de extracción de RNA se han realizado con muestras de raíz total para estandarizar el protocolo de extracción. Se realizaron dos ensayos de extracción mediante el método de TRizol. Los resultados de la extracción de RNA (Fig. 20) muestran que la relación de los valores de absorbencia a 260/280 y 260/230 nm fueron de 2.01 y 0.95, respectivamente. La razón de absorbencia entre las longitudes de onda 260/280 nm indica la pureza de los ácidos nucleicos respecto a las proteínas, la cual debe aproximarse a un valor de 2.0. Por otro lado, la razón de absorbencia entre 260/230 también mide la pureza de los ácidos nucleicos, pero con respecto a polifenoles o polisacáridos. En una muestra limpia, se espera que este valor se encuentre entre 2.0 y 2.2. Los resultados obtenidos indican que la extracción no presenta contaminación por proteínas, como se observa en los valores obtenidos a 260/280. Por otro lado, observamos que la extracción presenta contaminación por polifenoles o polisacáridos puesto que los valores de absorbencia en las longitudes de onda 260 y 230 son similares. Esto indica contaminación por polisacáridos, lo cual es esperado debido a que la muestra analizada contiene, principalmente, células diferenciadas. Se ha demostrado empíricamente que la presencia de polisacáridos podría interferir en las reacciones enzimáticas de polimerización, por lo cual es necesario estandarizar la técnica de extracción de RNA para obtener una muestra con mayor pureza en la absorbencia 260/230.

Sample ID	User ID	Date	Time	ng/ul	A260	A280	260/280	260/230	Constant	Cursor Pos.	Cursor abs.	340 raw
Pp_P1	Default	27/09/2019	01:49 p.m.	964.59	24.115	12.020	2.01	0.81	40.00	230	29.802	0.883
Pp_P2	Default	27/09/2019	01:51 p.m.	941.27	23.532	11.893	1.98	0.96	40.00	230	24.485	1.106

Figura 20. Ensayo de extracción de mRNA en la raíz primaria de *P. pringlei*.

6 CONCLUSIONES

Al realizar la búsqueda de proteínas PLT en las plantas terrestres se identificaron 1,161 secuencias que codifican para proteínas AIL, de las cuales 299 corresponden a secuencias posibles ortólogos de las proteínas PLT. Estas secuencias provienen del genoma de 61 especies, las cuales se agrupan en 54 géneros, 32 familias y 26 órdenes taxonómicos (Tabla 1, Fig. 8). Los resultados muestran que las AP2 surgieron en la diversificación de bryophyta y lycophyta.

Las proteínas PLT están presentes en todas las especies de plantas analizadas, con excepción de las briofitas. Esto sugiere que las proteínas PLT evolucionaron a partir de las proteínas AIL después de la divergencia entre las briofitas y las plantas vasculares, hace 360 M.A aproximadamente (Fig. 9).

En la filogenia de las proteínas PLT (Fig.8) se puede observar que éstas se agrupan en cuatro clados, los cuales corresponden al de PL1/PLT2, PLT3/PLT7, BBM y PLT5 de *A. thaliana*. Además, se observa un grupo de secuencias provenientes de plantas monocotiledóneas que no agrupan en los clados antes mencionados. Esto sugiere que existió una duplicación génica después de la divergencia entre plantas mono y dicotiledóneas. Por otro lado, se identificó a un posible ortólogo de *P. pringlei* para cada proteína PLT de *A. thaliana* con excepción de PLT5.

En el árbol filogenético de las proteínas PLT de las plantas vasculares sin semilla se observa que las proteínas PLT1-4 y PLT7 de *A. thaliana* se agrupan en un clado mientras que la proteína PLT5 agrupa con las secuencias de las Lycophyta. Esto podría sugerir que PLT5 es la secuencia ancestral a partir de la cual divergió el resto de las PLTs. En resumen, los análisis filogenéticos muestran que las proteínas AP2 están conservadas en todas las especies de plantas, y que la aparición de las proteínas PLT, posiblemente a partir de una secuencia similar a PLT5, coincide con la aparición de las raíces y del sistema vascular.

Los análisis de identidad y similitud de los dominios AP2 sugieren que las proteínas PLT de *P. pringlei* podrían reconocer los mismos blancos en el ADN que en *A. thaliana*, por lo cual, las diferencias en el patrón de desarrollo entre estas dos especies podrían explicarse por dos razones: 1) las proteínas PLT están siendo reguladas por distintas vías de regulación, o 2) las proteínas PLT interaccionan con proteínas diferentes. Debido a que no se han cristalizado las proteínas PLT no es posible comprobar su interacción con otras proteínas, sin embargo, es posible obtener su estructura por homología con otros dominios AP2,

por lo que es necesario realizar estos análisis para descartar alguna de estas hipótesis.

7 REFERENCIAS

- [1] F. Augstein and A. Carlsbecker, "Getting to the roots: a developmental genetic view of root anatomy and function from arabidopsis to lycophytes," *Front. Plant Sci.*, vol. 9, no. September, pp. 1–15, 2018, doi: 10.3389/fpls.2018.01410.
- [2] N. Chaffey, "Raven biology of plants, 8th edn," *Ann. Bot.*, vol. 113, no. 7, pp. vii–vii, Jun. 2014, doi: 10.1093/aob/mcu090.
- [3] I. Verstraeten and I. De Smet, "Primary Root," *eLS*, pp. 1–8, 2017, doi: 10.1002/9780470015902.a0002059.pub2.
- [4] B. Scheres, P. Benfey, and L. Dolan, "Root Development," *Arab. B.*, vol. 1, p. e0101, 2002, doi: 10.1199/tab.0101.
- [5] T. Greb and J. U. Lohmann, "Minireview Plant Stem Cells Minireview," *Curr. Biol.*, vol. 26, no. 17, pp. R816–R821, 2016, doi: 10.1016/j.cub.2016.07.070.
- [6] S. Shishkova, T. L. Rost, and J. G. Dubrovsky, "Determinate root growth and meristem maintenance in angiosperms," *Annals of Botany*. 2008, doi: 10.1093/aob/mcm251.
- [7] H. Wilcox, "Growth Studies of the Root of Incense Cedar, *Libocedrus decurrens*. II. Morphological Features of the Root System and Growth Behavior," *Am. J. Bot.*, vol. 49, no. 3, p. 237, 2006, doi: 10.2307/2439545.
- [8] S. Shishkova *et al.*, "Determinate primary root growth as an adaptation to aridity in Cactaceae: Towards an understanding of the evolution and genetic control of the trait," *Ann. Bot.*, vol. 112, no. 2, pp. 239–252, 2013, doi: 10.1093/aob/mct100.
- [9] J. G. Dubrovsky, "Determinate primary-root growth in seedlings of Sonoran Desert Cactaceae; its organization, cellular basis, and ecological significance," *Planta*, vol. 203, no. 1, pp. 85–92, 1997, doi: 10.1007/s00050168.
- [10] S. Shishkova *et al.*, "Determinate primary root growth as an adaptation to aridity in Cactaceae: Towards an understanding of the evolution and genetic control of the trait," *Ann. Bot.*, 2013, doi: 10.1093/aob/mct100.
- [11] B. Dinkelaker, C. Hengeler, and H. Marschner, "Distribution and Function of Proteoid Roots and other Root Clusters," *Bot. Acta*, vol. 108, no. 3, pp. 183–200, Jun. 1995, doi: 10.1111/j.1438-8677.1995.tb00850.x.
- [12] E. P. Groot, E. J. Sweeney, and T. L. Rost, "Development of the adhesive pad on climbing fig (*Ficus pumila*) stems from clusters of adventitious roots," *Plant Soil*, vol. 248, no. 1/2, pp. 85–96, Jan. 2003, doi: 10.1023/A:1022342025447.
- [13] G. T. VARNEY and M. E. McCULLY, "The branch roots of *Zea*. II. Developmental loss of the apical meristem in field-grown roots," *New Phytol.*, vol. 118, no. 4, pp. 535–546, Aug. 1991, doi: 10.1111/j.1469-8137.1991.tb00993.x.
- [14] J. G. Dubrovsky and L. F. Gomez-Lomeli, "Water deficit accelerates determinate developmental program of the primary root and does not affect lateral root initiation in a Sonoran Desert cactus (*Pachycereus pringlei*, Cactaceae)," *Am. J. Bot.*, vol. 90, no. 6, pp. 823–831, Jun. 2003, doi: 10.3732/ajb.90.6.823.
- [15] G. Rodriguez-Alonso *et al.*, "Transcriptomics insights into the genetic regulation of root apical meristem exhaustion and determinate primary root growth in *Pachycereus pringlei* (Cactaceae)," *Sci. Rep.*, vol. 8, no. 1, pp. 1–11, 2018, doi: 10.1038/s41598-018-26897-1.

- [16] A. Shimotohno, R. Heidstra, I. Blilou, and B. Scheres, "Root stem cell niche organizer specification by molecular convergence of PLETHORA and SCARECROW transcription factor modules," *Genes Dev.*, vol. 32, no. 15–16, pp. 1085–1100, 2018, doi: 10.1101/gad.314096.118.
- [17] C. Galinha *et al.*, "PLETHORA proteins as dose-dependent master regulators of Arabidopsis root development," *Nature*, vol. 449, no. 7165, pp. 1053–1057, 2007, doi: 10.1038/nature06206.
- [18] L. Di Laurenzio *et al.*, "The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the Arabidopsis root," *Cell*, vol. 86, no. 3, pp. 423–433, 1996, doi: 10.1016/S0092-8674(00)80115-4.
- [19] S. El Ouakfaoui *et al.*, "Control of somatic embryogenesis and embryo development by AP2 transcription factors," *Plant Mol. Biol.*, vol. 74, no. 4, pp. 313–326, 2010, doi: 10.1007/s11103-010-9674-8.
- [20] J. K. Okamoto, B. Caster, R. Villarroel, M. Van Montagu, and K. D. Jofuku, "The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in Arabidopsis.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 94, no. 13, pp. 7076–81, 1997, [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/9192694><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC21287>.
- [21] L. Santuari *et al.*, "The PLETHORA Gene Regulatory Network Guides Growth and Cell Differentiation in Arabidopsis Roots.," *Plant Cell*, vol. 28, no. 12, pp. 2937–2951, 2016, doi: 10.1105/tpc.16.00656.
- [22] S. Kim, P. S. Soltis, K. Wall, and D. E. Soltis, "Phylogeny and domain evolution in the APETALA2-like gene family," *Mol. Biol. Evol.*, vol. 23, no. 1, pp. 107–120, 2006, doi: 10.1093/molbev/msj014.
- [23] M. Aida *et al.*, "The PLETHORA genes mediate patterning of the Arabidopsis root stem cell niche," *Cell*, vol. 119, no. 1, pp. 119–120, 2004, doi: 10.1016/j.cell.2004.09.018.
- [24] A. P. Mähönen *et al.*, "PLETHORA gradient formation mechanism separates auxin responses," *Nature*, vol. 515, no. 7525, pp. 125–129, 2014, doi: 10.1038/nature13663.
- [25] B. Scheres and B. A. Krizek, "Coordination of growth in root and shoot apices by AIL/PLT transcription factors," *Curr. Opin. Plant Biol.*, vol. 41, pp. 95–101, 2018, doi: 10.1016/j.pbi.2017.10.002.
- [26] B. A. Krizek, "Auxin regulation of Arabidopsis flower development involves members of the AINTEGUMENTA-LIKE/PLETHORA (AIL/PLT) family," *J. Exp. Bot.*, vol. 62, no. 10, pp. 3311–3319, 2011, doi: 10.1093/jxb/err127.
- [27] K. Birnbaum *et al.*, "A Gene Expression Map of the Arabidopsis Root," *Science (80-.)*, vol. 302, no. 5652, pp. 1956–1960, 2003, doi: 10.1126/science.1090022.
- [28] Y. Du and B. Scheres, "PLETHORA transcription factors orchestrate de novo organ patterning during Arabidopsis lateral root outgrowth," *Proc. Natl. Acad. Sci.*, vol. 114, no. 44, pp. 11709–11714, 2017, doi: 10.1073/pnas.1714410114.
- [29] L. Huang and J. Schiefelbein, "Conserved Gene Expression Programs in Developing Roots from Diverse Plants," *Plant Cell*, vol. 27, no. 8, pp. 2119–2132, 2015, doi: 10.1105/tpc.15.00328.

- [30] L. Hetherington, A., Emms, D. & Dolan, “Gene expression data support the hypothesis that Isoetes rootlets are true roots and not modified leaves,” *bioRxiv*, pp. 10–14, 2019.
- [31] D. T. Jones, W. R. Taylor, and J. M. Thornton, “The rapid generation of mutation data matrices from protein sequences,” vol. 8, no. 3, pp. 275–282, 2018.
- [32] M. L. L. Valle, “Análisis molecular de los genes PpPLT en el desarrollo de la raíz primaria del cardón *Pachycereus pringlei*,” UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO, 2015.
- [33] A. Zwaenepoel *et al.*, “The hornwort genome and early land plant evolution,” *Nat. Plants*, vol. 6, no. February, 2020, doi: 10.1038/s41477-019-0588-4.
- [34] S. Magall, G. Sandra, and L. S. Luna, “A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity,” pp. 437–453, 2015.
- [35] A. C. Ferrari, D. Shivhare, B. O. Hansen, N. Winter, and A. Pasha, “Summary :,” *bioRxiv*, 2019.
- [36] D. Copetti *et al.*, “Extensive gene tree discordance and hemiplasy shaped the genomes of North American columnar cacti,” *Proc. Natl. Acad. Sci.*, vol. 114, no. 45, pp. 12003–12008, 2017, doi: 10.1073/pnas.1706367114.
- [37] A. Burkart, R., Strotmann, V., Kirschner, G. & Akinci, “PLETHORA and WOX5 interaction and subnuclear localisation regulates Arabidopsis root stem cell maintenance,” *bioRxiv*, pp. 1–37, 2019.
- [38] S. Li, “The Arabidopsis thaliana TCP transcription factors: A broadening horizon beyond development,” *Plant Signal. Behav.*, vol. 10, no. 7, pp. 1–12, 2015, doi: 10.1080/15592324.2015.1044192.
- [39] S. El Ouakfaoui *et al.*, “Control of somatic embryogenesis and embryo development by AP2 transcription factors,” *Plant Mol. Biol.*, vol. 74, no. 4, pp. 313–326, 2010, doi: 10.1007/s11103-010-9674-8.
- [40] C. Liu and T. Zhang, “Expansion and stress responses of the AP2/EREBP superfamily in cotton,” *BMC Genomics*, vol. 18, no. 1, pp. 1–16, 2017, doi: 10.1186/s12864-017-3517-9.
- [41] M. D. Allen, K. Yamasaki, M. Ohme-Takagi, M. Tateno, and M. Suzuki, “A novel mode of DNA recognition by a β -sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA,” *EMBO J.*, vol. 17, no. 18, pp. 5484–5496, 1998, doi: 10.1093/emboj/17.18.5484.

8 ANEXOS

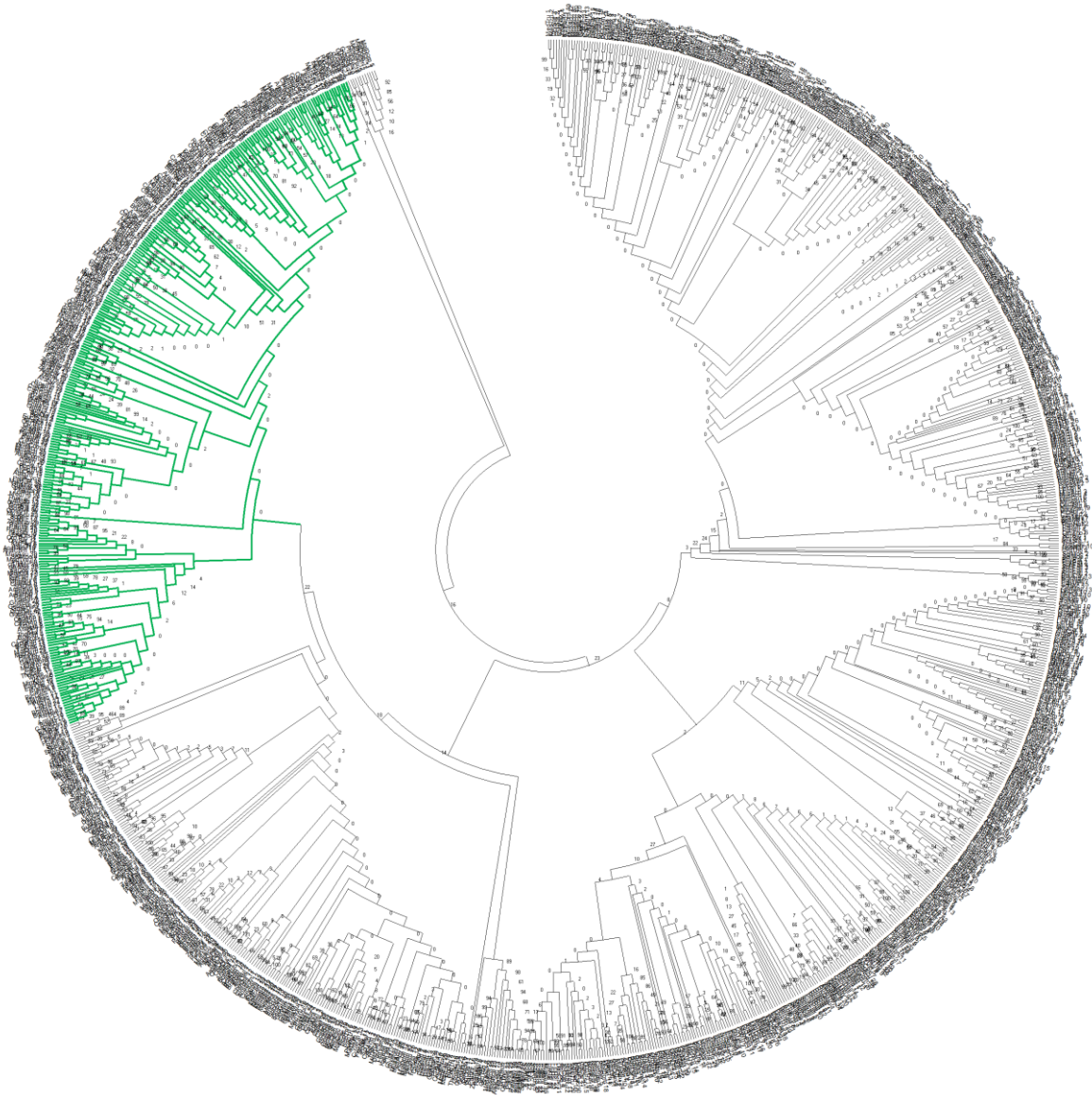


Figura 1. Filogenia de las proteínas AIL en las embriofitas. En verde está marcado el clado de los PLT.

Tabla 1. Predicción del modelo evolutivo del orden Caryophyllales. Incluye a las especies *Beta vulgaris*, *Chenopodium quinoa*, *Pachycereus pringlei* y *Spinacia oleracea*.

Caryophyllales					
Model	#Param	BIC	AICc	InL	Gamma
JTT+G+I+F	112	81079.56	80153.45	-39964.3	1.164944
WAG+G+I+F	112	81274.92	80348.82	-40062	1.20684
Dayhoff+G+I+F	112	81478.56	80552.45	-40163.8	1.101686
LG+G+I+F	112	81634.01	80707.91	-40241.5	1.05479
rtREV+G+I+F	112	81679.65	80753.54	-40264.3	1.09522
mtREV24+G+I+F	112	82074.97	81148.86	-40462	0.980974
cpREV+G+I+F	112	85088.24	84162.13	-41968.6	0.57384

Tabla 2. Predicción del modelo evolutivo de la cactácea *P. pringlei*.

Pachycereus Pringlei					
Model	#Param	BIC	AICc	InL	Gamma
JTT+G+F	27	12196.4	12034.32	-5989.91	1.494673
cpREV+G+F	27	12197.8	12035.73	-5990.61	0.592889
LG+G+I+F	28	12202.94	12034.88	-5989.17	1.864541
WAG+G+F	27	12208.83	12046.76	-5996.13	1.555985
Dayhoff+G+I+F	28	12211.37	12043.31	-5993.39	1.809162
rtREV+G+I+F	28	12216.41	12048.36	-5995.91	1.980436
mtREV24+G+I+F	28	12236.04	12067.99	-6005.72	1.77316

Tabla 3. Predicción del modelo evolutivo del orden Brassicales. Incluye a las especies *Arabidopsis halleri*, *A. thaliana*, *A. lyrata*, *Capsella grandiflora*, *Boechera stricta*, *Brassica rapa*, *Capsella rubella*, *Carica papaya* y *Thellungiella halophila*

Brassicales					
Model	#Param	BIC	AICc	InL	Gamma
JTT+G+I+F	162	43564	42197.68	-20936.1	1.06
WAG+G+I+F	162	43888.44	42522.12	-21098.3	1.06
Dayhoff+G+I+F	162	43946.49	42580.17	-21127.3	1
LG+G+F	161	44082.11	42724.21	-21200.3	0.76
rtREV+G+I+F	162	44236.5	42870.17	-21272.3	0.98
mtREV24+G+I+F	162	44239.36	42873.03	-21273.7	0.89
cpREV+G+I+F	162	46193.6	44827.27	-22250.9	0.36

Tabla 4. Predicción del modelo evolutivo para las secuencias AIL de los órdenes Brassicales y Caryophyllales.

Brassicales + Caryophyllales				
Modelo	BIC	AICc	lnL	Gamma
JTT+G+I+F	122608.1	120291.6	-59888.8	0.913808
WAG+G+I+F	123109.5	120793.1	-60139.5	0.982585
Dayhoff+G+I+F	123358.8	121042.3	-60264.1	0.876802
LG+G+I+F	123646.9	121330.5	-60408.2	0.87062
rtREV+G+I+F	123939.7	121623.2	-60554.6	0.912407
mtREV24+G+I+F	124325.4	122008.9	-60747.4	0.81881
cpREV+G+F	128113.7	125806.3	-62647.1	0.441455

Las proteínas PLT evolucionaron bajo el modelo de sustitución Jones-Taylor-Thornton (JTT); con los valores del criterio de información bayesiana (BIC, por sus siglas en inglés) y el criterio de información Akaike (AIC, por sus siglas en inglés) más altos y el logaritmo natural de verosimilitud más negativo resultante de los dos criterios anteriores.

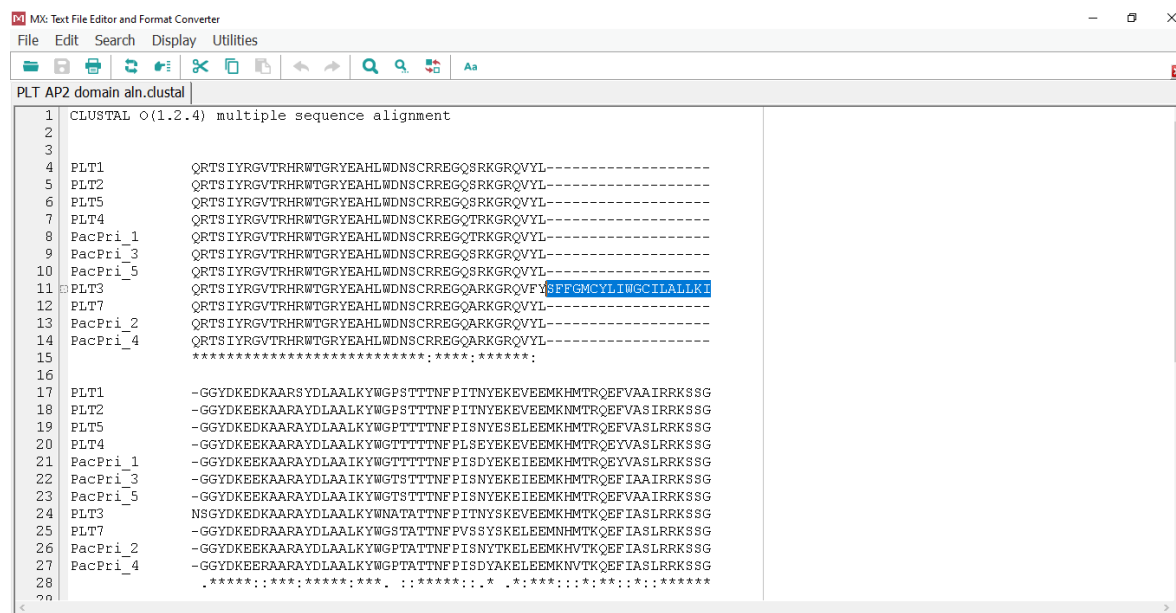


Figura 2. Fracción del alineamiento de las secuencias de los dominios AP2 de *A. thaliana* y *P. pringlei*. Se muestra solo la región del alineamiento que contiene la inserción de 21 aminoácidos en la proteína PLT3 de *A. thaliana*.

Tabla 5. Porcentajes de identidad y similitud entre las proteínas PLT y sus dominios de *A. thaliana* (AraTha) y *P. pringlei* (PacPri).

Secuencia completa			Secuencia de los dominios AP2		
AraTha	Identidad	Similitud	AraTha	Identidad	Similitud
PLT1 vs PLT2	74.5	80	PLT1 vs PLT2	96.1	98.3
PLT1 vs PLT3	36.6	44.9	PLT1 vs PLT3	75.5	81.5
PLT1 vs PLT4	37.9	51.6	PLT1 vs PLT4	87.8	96.7
PLT1 vs PLT5	39.5	50.6	PLT1 vs PLT5	90	95
PLT1 vs PLT7	35.3	44.2	PLT1 vs PLT7	82.8	91.7
PLT2 vs PLT3	35.2	43.7	PLT2 vs PLT3	76.6	82.7
PLT2 vs PLT4	39.9	51.1	PLT2 vs PLT4	90.4	98.3
PLT2 vs PLT5	39.7	52.3	PLT2 vs PLT5	91.5	96.6
PLT2 vs PLT7	36.1	46.4	PLT2 vs PLT7	84.2	93.2
PLT3 vs PLT4	39.4	48.7	PLT3 vs PLT4	74.6	81.2
PLT3 vs PLT5	34	44.1	PLT3 vs PLT5	77.6	84.4
PLT3 vs PLT7	60	66.6	PLT3 vs PLT7	83.7	87.4
PLT4 vs PLT5	35.6	44.8	PLT4 vs PLT5	88.7	94.9
PLT4 vs PLT7	34.8	44.3	PLT4 vs PLT7	84.2	92.1
PLT5 vs PLT7	35.9	46.7	PLT5 vs PLT7	87.8	95.3
PacPri			PacPri		
Pacpri_1 vs Pacpri_2	30.9	37.9	Pacpri_1 vs Pacpri_2	86.5	92.7
Pacpri_1 vs Pacpri_3	36	46	Pacpri_1 vs Pacpri_3	90.4	96
Pacpri_1 vs Pacpri_4	34	41.7	Pacpri_1 vs Pacpri_4	85.9	93.2
Pacpri_1 vs Pacpri_5	38.6	47.4	Pacpri_1 vs Pacpri_5	94.4	100
Pacpri_2 vs Pacpri_3	36.4	45	Pacpri_2 vs Pacpri_3	90.6	97.1
Pacpri_2 vs Pacpri_4	63.5	69.1	Pacpri_2 vs Pacpri_4	96.5	98.8
Pacpri_2 vs Pacpri_5	38.7	48.5	Pacpri_2 vs Pacpri_5	86.5	93.3
Pacpri_3 vs	36.4	46.7	Pacpri_3 vs	88.8	97.6

Pacpri_4			Pacpri_4		
Pacpri_3 vs Pacpri_5	59.4	66.2	Pacpri_3 vs Pacpri_5	94.9	96
Pacpri_4 vs Pacpri_5	39.2	49	Pacpri_4 vs Pacpri_5	84.7	93.8
AraTha vs PacPri			AraTha vs PacPri		
PLT1 vs Pacpri_1	40.1	49.3	PLT1 vs Pacpri_1	89.4	97.8
PLT1 vs Pacpri_2	35	42.9	PLT1 vs Pacpri_2	83.4	92.3
PLT1 vs Pacpri_3	45.2	53.6	PLT1 vs Pacpri_3	88.3	93.9
PLT1 vs Pacpri_4	37.7	48.4	PLT1 vs Pacpri_4	81.7	92.8
PLT1 vs Pacpri_5	59.5	69.5	PLT1 vs Pacpri_5	92.2	97.8
PLT2 vs Pacpri_1	37.8	47.6	PLT2 vs Pacpri_1	92.1	99.4
PLT2 vs Pacpri_2	35.4	44.4	PLT2 vs Pacpri_2	86	93.8
PLT2 vs Pacpri_3	47	57.1	PLT2 vs Pacpri_3	89.8	95.5
PLT2 vs Pacpri_4	36.8	46.8	PLT2 vs Pacpri_4	85.3	94.40%
PLT2 vs Pacpri_5	61.7	70.9	PLT2 vs Pacpri_5	93.8	99.4
PLT3 vs Pacpri_1	36.1	47	PLT3 vs Pacpri_1	75.1	82.2
PLT3 vs Pacpri_2	48.3	55.3	PLT3 vs Pacpri_2	82.7	86.9
PLT3 vs Pacpri_3	29	37.6	PLT3 vs Pacpri_3	78.4	85.8
PLT3 vs Pacpri_4	44.1	53.8	PLT3 vs Pacpri_4	81.1	87.4
PLT3 vs Pacpri_5	34.6	43.4	PLT3 vs Pacpri_5	75.1	82.7
PLT4 vs Pacpri_1	42.2	51.5	PLT4 vs Pacpri_1	93.8	99.4
PLT4 vs Pacpri_2	37.4	45.2	PLT4 vs Pacpri_2	84.3	91.6
PLT4 vs Pacpri_3	33.7	43.9	PLT4 vs Pacpri_3	87	94.9
PLT4 vs Pacpri_4	34.9	43.5	PLT4 vs Pacpri_4	84.2	92.7
PLT4 vs Pacpri_5	39.9	50.2	PLT4 vs Pacpri_5	91	98.9
PLT5 vs Pacpri_1	36	43.7	PLT5 vs Pacpri_1	91	96
PLT5 vs Pacpri_2	38.3	46.8	PLT5 vs Pacpri_2	89.6	96
PLT5 vs Pacpri_3	32.6	41.1	PLT5 vs Pacpri_3	91.9	97.7
PLT5 vs Pacpri_4	37	46.4	PLT5 vs Pacpri_4	87.8	96.5
PLT5 vs Pacpri_5	30.4	38.5	PLT5 vs Pacpri_5	90.4	96
PLT7 vs Pacpri_1	33.8	42.5	PLT7 vs Pacpri_1	84.7	92.7
PLT7 vs Pacpri_2	50.4	60.6	PLT7 vs Pacpri_2	93.6	98.2
PLT7 vs Pacpri_3	32.1	40.4	PLT7 vs Pacpri_3	88.2	97.6

PLT7 vs Pacpri_4	45.3	53.9	PLT7 vs Pacpri_4	93.5	98.2
PLT7 vs Pacpri_5	35.7	45.7	PLT7 vs Pacpri_5	84.2	93.8

Figura 3. Alineamiento entre la secuencia consenso del cristal AtERF96 y la proteína CBD.

```

CLUSTAL O(1.2.4) multiple sequence alignment

SeqPro      MDQGGRGVGAEHGKYRGVRRFPWGKYAAEIRD SRKHGERVWMLGTFDTAEAAARAYDQAAAY
CBD          -----KHYRGVRRFPWGKFAAEIRDPAKNGARVWMLGTFETAEDAALAYDRAAF
              :*****:*****:*****  *:* *****:***:** **:***
              .

SeqPro      SMRGQAAILNFPHEYNMGSGVSSSTAMAGSSSASASSSSRQVFEFYLDLDSVLEELLEEG
CBD          RMRGSRALLNFPLRV-----
              ***. *:***** .

SeqPro      EKPMKGKKK
CBD          -----

```