



Centro de Investigación Científica de Yucatán, A.C.

Posgrado en Ciencias Biológicas

“Estudio de la formación de fibras perivasculares en  
*Agave fourcroydes* mediante un enfoque  
transcriptómico”

Tesis que presenta

José Roberto Cruz Balam

En opción al título de

MAESTRO EN CIENCIAS

(Ciencias Biológicas: Bioquímica y Biología Molecular)

Mérida, Yucatán, México

2024

*CENTRO DE INVESTIGACIÓN CIENTÍFICA DE YUCATÁN, A. C.*  
*POSGRADO EN CIENCIAS BIOLÓGICAS*



**RECONOCIMIENTO**

Por medio de la presente, hago constar que el trabajo de tesis de José Roberto Cruz Balam titulado **Estudio de la formación de fibras perivasculares en *Agave fourcroydes* mediante un enfoque transcriptómico**, fue realizado en la unidad de Biología Integrativa, en la línea de investigación de Biología celular y del desarrollo, en el laboratorio 10 del Centro de Investigación Científica de Yucatán, A.C. bajo la dirección del Dr. Fulgencio Alatorre Cobos, dentro de la opción de Bioquímica y Biología Molecular perteneciente al Programa de Posgrado en Ciencias Biológicas de este Centro.

Atentamente



---

Dr. Cecilia Hernández Zepeda  
Directora de Docencia

Mérida, Yucatán, México, a 04 de julio de 2024

## DECLARACIÓN DE PROPIEDAD

Declaro que la información contenida en las secciones de: Materiales y Métodos, Resultados y Discusión de este documento, proviene de las actividades de investigación realizadas durante el período que se me asignó para desarrollar mi trabajo de tesis, en las Unidades y Laboratorios del Centro de Investigación Científica de Yucatán, A.C., y que, a razón de lo anterior, y en contraprestación de los servicios educativos o de apoyo que me fueron brindados, dicha información, en términos de la Ley Federal del Derecho de Autor y la Ley Federal de Protección a la Propiedad Industrial, le pertenece patrimonialmente a dicho Centro de Investigación. Por otra parte, en virtud de lo ya manifestado, reconozco que de igual manera los productos intelectuales o desarrollos tecnológicos que deriven o pudieran derivar de lo correspondiente a dicha información, le pertenecen patrimonialmente al Centro de Investigación Científica de Yucatán, A.C., y se registrarán, en todo caso, por lo dispuesto por la Ley Federal del Derecho de Autor y la Ley Federal de Protección a la Propiedad Industrial, en el tenor de lo expuesto en la presente Declaración.

Firma:  \_\_\_\_\_

Ing. José Roberto Cruz Balam

## **AGRADECIMIENTOS**

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada.

Al Instituto de Ecología A.C. (INECOL), por el asesoramiento para el uso del equipo de súper cómputo "Huitzilín", recurso básico para el análisis computacional de este trabajo.

Al Dr. Fulgencio Alatorre Torres, expreso mi más profundo agradecimiento por la invaluable oportunidad de formar parte de su grupo de investigación. Agradezco profundamente su excepcional mentoría, paciencia infinita y la transmisión de conocimientos durante el desarrollo de este proyecto.

Al Dr. Enrique Laclette, expreso mi agradecimiento por su invaluable apoyo y guía durante el desarrollo de este proyecto. Agradezco su infinita paciencia al responder mis dudas y brindarme la dirección necesaria. Su valiosa colaboración ha sido fundamental para el éxito de este proyecto.

Al Dr. Felipe Sánchez por sus comentarios y observaciones en cada tutoría.

A la estimada M.C. Elidé Berzunza, le expreso mi más profundo agradecimiento por sus valiosos consejos, su apoyo incondicional en el desarrollo de este proyecto y la cálida amistad que me ha brindado.

Al M.C. Marcos Couch por su invaluable paciencia, su generosa disposición para compartir sus amplios conocimientos y la cálida amistad que me ha brindado. Su apoyo ha sido fundamental para mi crecimiento personal y profesional.

A mis compañeros: Samuel, Jorge, Diego, Lucía, Dafne, Hugo, les expreso mi más profundo agradecimiento por su invaluable amistad y apoyo durante mi estancia en el CICY.

A Alejandra Arceo, por su invaluable apoyo durante mi estancia en el CICY.

A mis compañeros Diobel, Laurení y Alberto, gracias por su amistad y apoyo.

## **DEDICATORIAS:**

A mi familia.

Mis padres, Alicia del Socorro Balam Canche e Isaías Pech Ceh, por su apoyo constante y por siempre creer en mí.

A todas esas personas que me han brindado su apoyo durante estos dos años, y sin duda forman parte de este éxito.

Gracias.

## ÍNDICE

INTRODUCCIÓN: .....	1
CAPÍTULO I :REVISIÓN DE LITERATURA.....	2
1.1 IMPORTANCIA ACTUAL DE LAS FIBRAS NATURALES .....	2
1.2 ESCLERÉNQUIMA Y CLASIFICACIÓN ANATÓMICA DE FIBRAS .....	3
1.3 ETAPAS DE FORMACIÓN DE LAS FIBRAS .....	5
1.4 ANATOMÍA Y PATRONES DE DISTRIBUCIÓN DE FIBRAS.....	8
1.5 COMPOSICIÓN DE LAS FIBRAS VEGETALES .....	9
1.6 LIGNINA.....	12
1.7 SÍNTESIS DE LOS MONÓMEROS DE LIGNINA .....	13
1.8 MECANISMOS PARA EL TRANSPORTE DE MONOLIGNOLES A TRAVÉS DE MEMBRANAS .....	15
1.9 POLIMERIZACIÓN DE MONOLIGNOLES EN LA PARED CELULAR SECUNDARIA.....	16
1.10 TEORÍAS SOBRE LA LIGNIFICACIÓN CELULAR.....	17
1.11 JUSTIFICACIÓN.....	19
1.12 HIPÓTESIS .....	20
1.13 OBJETIVOS.....	21
1.14 ESTRATEGIA EXPERIMENTAL .....	22
CAPÍTULO II MATERIALES Y MÉTODOS.....	23
2.1 DETERMINAR MEDIANTE ANÁLISIS HISTOLÓGICOS LAS ETAPAS CLAVES EN LA FORMACIÓN DE LAS FIBRAS PERIVASCULARES .....	23
2.2 GENERAR UN TRANSCRIPTOMA DE NOVO A PARTIR DE PLANTAS EN DESARROLLO DE <i>A. FOURCROYDES</i> QUE REPRESENTAN EL PROCESO DE FORMACIÓN DE LAS FIBRAS PERIVASCULARES. .	26

2.3 DEFINIR PATRONES DE EXPRESIÓN TRANSCRIPCIONAL RELACIONADAS CON LA BIOSÍNTESIS DE LIGNINA EN TEJIDOS ADYACENTES A LAS FIBRAS PERIVASCUALRES EN SUS DIFERENTES ETAPAS DE DESARROLLO. ....	30
CAPÍTULO III RESULTADOS .....	33
3.1 DETERMINAR MEDIANTE ANÁLISIS HISTOLÓGICOS LAS ETAPAS CLAVES EN LA FORMACIÓN DE LAS FIBRAS PERIVASCULARES .....	33
3.2 GENERAR UN TRANSCRIPTOMA DE <i>NOVO</i> A PARTIR DE PLANTAS EN DESARROLLO DE <i>A. FOURCROYDES</i> QUE REPRESENTAN EL PROCESO DE FORMACIÓN DE LAS FIBRAS PERIVASCULARES. .	35
3.3 DEFINIR PATRONES DE EXPRESIÓN TRANSCRIPCIONAL RELACIONADAS CON LA BIOSÍNTESIS DE LIGNINA EN TEJIDOS ADYACENTES A LAS FIBRAS PERIVASCUALRES EN SUS DIFERENTES ETAPAS DE DESARROLLO. ....	49
CAPITULO IV: DISCUSIÓN GENERAL.....	68
5.1 DISCUSIÓN.....	68
CAPITULO V: CONCLUSIONES Y PERSPECTIVAS .....	71
5.1 CONCLUSIONES .....	71
5.2 PERSPECTIVAS .....	72
BIBLIOGRAFÍA .....	73

## Índice de figuras:

Figura 1.1. Tipos de fibras vegetales. A) Fibras halladas en tallo de maíz. Se indican los diferentes tipos de fibras en función de posición dentro del órgano (tomado de Megías <i>et al.</i> , 2024). B) Fibras de floema primario (FF) en tallo de lino (tomado de Gorshkova <i>et al.</i> , 2023) C) Fibras corticales en raíz de arroz (tomado de <a href="https://labs.plb.ucdavis.edu">https://labs.plb.ucdavis.edu</a> ). D) Fibras perivasculares en hojas de <i>A. tequilana</i> (tomado de Maceda <i>et al.</i> , 2022). .....	4
Figura 1.2. Desarrollo de las fibras del esclerénquima: (a) Formación de fibras por fusión de células del parénquima y disolución de septos entre ellos; (b) Elongación coordinada de la parte superior de la fibra y crecimiento continuo de la planta; (c) Crecimiento intrusivo apical y (d) crecimiento intrusivo difusivo. Figura redibujada de Chernova & Gorshkova, (2007).( Tomado y modificado de referencia (2007) .....	6
Figura 1.3. Localización de las fibras estructurales y de cinta dentro de una hoja de agave. Figura modificada de Yan <i>et al.</i> , (2000) .....	9
Figura 1.4. a) Los monolignoles más comunes encontrados en los bloques de construcción de la lignina: 1: alcohol p-cumarílico, 2: alcohol coniferílico, 3: alcohol sinapílico, b) otros tipos de monolignoles (coniferil-derivados): 4: coniferaldehído, 5: alcohol (recolectada de Chávez-Sifontes & Domine, 2013).....	13
Figura 1.5. Ruta biosintética de la lignina (tomada de Chávez & Domine, 2013).....	14
Figura 1.6. Se muestra la estrategia experimental que se empleó en este trabajo. La caracterización de las fibras y el ensamblado de <i> novo</i> del transcriptoma de <i> agave fourcroydes</i> . .....	22
Figura 2.1. Colecta del material vegetal ) Hojas recolectadas de las tres etapas de desarrollo vegetativo. B) Zona utilizada para la extracción del ARN total. C) Promedio de las plantas que fueron muestreadas. D) Preparación de la muestra para la extracción.....	24
Figura 2.2. Flujo de trabajo del protocolo de extracción de ARN total. ....	27
Figura 3.1. Cinética temporal del desarrollo de las fibras de <i> Agave fourcroydes</i> secuenciadas por RNAseq en este estudio. A) Fibras extraídas. B) Fotografías de autofluorescencia de la clorofila c) Fotografías generadas por el microscopio electrónico de barrido. D) Teñido de fibras estructurales con azul de toluidina. E) Teñido de fibras estructurales con azul de toluidina. ....	34
Figura 3.2. Lignificación de fibras de <i> Agave fourcroydes</i> en las tres etapas de desarrollo secuenciadas por RNAseq. Arriba autofluorescencia de lignina revelada mediante microscopía láser confocal. Abajo, autofluorescencia de lignina y campo claro.....	35

Figura 3.3. Geles de agarosa al 1% teñidos con bromuro de etidio. A) Muestras que corresponden al tejido adyacente a las fibras B) Muestras de tejidos correspondientes las hojas completas de la zona basal. ....	38
Figura 3.4 Análisis de integridad mediante electroforesis de microfluidos (Bioanalyzer). FU: Intensidad de fluorescencia. RIN: Número de integridad. Los picos mas elevados, de derecha a izquierda, corresponden a los RNAs ribosomales 28S y 18S. ....	39
Figura 3.5. Distribución de la calidad de las lecturas para tejido de <i>A. fourcroydes</i> generado por la herramienta FastQC. A) Distribución de la calidad de las lecturas generadas por Illumina. B) Distribución de la calidad de las lecturas filtradas por Trimmomatic. El eje Y indica los puntajes de calidad Phred de las bases secuenciadas, mientras que el eje X representa las pares de base secuenciadas. ....	41
Figura 3.6. Análisis de componentes principales. se muestra la distribución de los datos en el plano definido por la primera componente principal (PC 1) en el eje horizontal, que explica el 28.20% de la varianza total de los datos. El eje vertical representa la segunda componente principal (PC 2), que captura una porción adicional de la varianza. ....	51
Figura 3.7. Expresión diferencial para los contrastes de Fibra versus hoja completa. A) Mapa de calor muestra los patrones de niveles de expresión de los DEG encontrados a lo largo de las 3 etapas de desarrollo analizadas. B) Diagrama de Venn de las secuencias putativas identificadas en los análisis de expresión diferencial, mostrando aquellos compartidos y únicos para etapa. C) Número de DEGs identificados en los diferentes contrastes realizados en el análisis de expresión diferencial. ....	53
Figura 3.8. Expresión diferencial para las contras 200vs40, 200vs15, 40vs15. A) Mapa de calor muestra la correlación entre los perfiles de expresión de los tejidos de <i>A. fourcroydes</i> . B) Diagrama de Venn de las secuencias putativas identificadas en la expresión diferencial. C) Número de DEGs identificados en los diferentes contras comparados. ....	59
Figura 3.9. Términos GO en la categoría de Proceso Biológicos para el transcriptoma a partir de la notación funcional del proteoma de <i>Arabidopsis thaliana</i> . ....	62
Figura 3.10 Actividad de peroxidasas (POX) y lacasas (LAC) en fibras de cinta y haces vasculares de <i>Agave fourcroydes</i> . La detección de la actividad de ambas enzimas se realizó usando el reactivo DAB en cortes transversales de hoja. ....	67

Índice de tablas:

Tabla 1.1. Composición química de fibras naturales (Morán, 2021).	10
Tabla 1.2. Composición química de fibras de agave (% del peso total seco)(Morán <i>et al.</i> , 2023).	11
Tabla 2.1. Características morfológicas de las plantas muestreadas.	23
Tabla 2.2. Lecturas generadas por biblioteca secuenciada.	27
Tabla 3.1. Resultados de pureza y concentración de RNA, así como el número de integridad de RNA obtenidos de las muestras de <i>Agave fourcroydes</i> .	36
Tabla 3.2. Lecturas crudas y filtradas con Trimmomatic.	42
Tabla 3.3. Estadísticas de los transcriptomas ensamblados usando generados por la herramienta Trinity.	43
Tabla 3.4. Comparación de las estadísticas de los ensamblados usando MIRA vs Trinity.	44
Tabla 3.5. Porcentajes de alineamiento de las bibliotecas secuenciadas.	45
Tabla 3.6. Resultado de la herramienta BUSCO.	46
Tabla 3.7. Secuencias conservadas en diferentes herramientas.	48
Tabla 3.8. Análisis BUSCO de las diferentes herramientas.	48
Tabla 3.9. Anotación del proteoma de <i>Agave fourcroydes</i> .	50
Tabla 3.10. Lista de los genes con la expresión diferencialmente más alta de las muestras correspondiente a 15 cm, 40 cm y 2 M.	55
Tabla 3.11. Lista de los genes con la expresión diferencialmente más alta de las muestras correspondiente a 200vs40, 200vs15 y 40vs15.	60
Tabla 3.12. Análisis de enriquecimiento de genes, con referencia a la expresión diferencial realizado Fibra entre Fibra	63
Tabla 3.13 Genes identificados del término GO:0009498, compartidos en todos los comparativos.	64

## Resumen:

El *Agave fourcroydes* o henequén es una planta endémica de la Península de Yucatán, caracterizada por sus hojas ricas en fibras de alta calidad. Estas han sido empleadas para la elaboración de múltiples productos, convirtiéndose en un símbolo de identidad y tradición. El cultivo de *A. fourcroydes* ha experimentado un declive significativo desde principios del siglo XIX. No obstante, en la última década, el interés por esta especie ha resurgido, impulsado por investigaciones que revelan su enorme potencial en diversos campos. La formación de fibras ha sido poco estudiada en agaves, enfocándose la mayoría de los estudios transcriptómicos en el metabolismo del azúcar y floración. Este estudio ha permitido el ensamblado *de novo* del transcriptoma de *A. fourcroydes*, con un total de 264,962 transcritos. El análisis de expresión diferencial entre dos tipos de tejidos secuenciados, fibra versus hoja completa, identificó 4,340 y 8,545 transcritos sobreexpresados y subexpresados en el comparativo de "15 cm", respectivamente; 4,000 y 7,269 transcritos sobreexpresados y subexpresados en el comparativo de "40 cm". Finalmente, el comparativo de "2M" identificó 5,994 y 6,252 transcritos sobreexpresados y subexpresados. Para el análisis de expresión diferencial contrastando fibra versus fibra, el comparativo "200vs40" identificó 650 y 394 transcritos sobreexpresados y sub expresados; en la comparación "200vs15" se identificaron 723 y 489 transcritos sobreexpresados y sub expresados. Finalmente, en el comparativo "40vs15" se identificaron 476 y 261 sobre expresados y sub expresados respectivamente. Dichos análisis de expresión génica lograron la identificación de genes como LAC3, NOV, PRX52, DegP7, incluso de factores de transcripción como APRR2, SPL1, GRL/NAP1 y miembros de la familia bHLH. Algunos de estos genes son considerados marcadores de la formación de fibras en otros cultivos, pero no se conocía de su función en henequén.

## Abstract

*Agave fourcroydes* is an endemic plant of the Yucatan, characterized by its leaves rich in high-quality fibers. These have been used to produce multiple products, becoming a symbol of identity and tradition. Despite its historical importance, the cultivation of *A. fourcroydes* has experienced a significant decline since the beginning of the 19th century. However, in the last decade, interest in this species has resurged, driven by research that reveals its enormous potential in various fields. Although some previous studies reported agave leaf transcriptomics, these researches focused on sugar metabolism and flowering. This study has allowed the *de novo* assembly of the *A. fourcroydes* transcriptome, with 264,962 transcripts. The differential expression analysis using the contrast fiber versus whole leaf identified 4,340 and 8,545 overexpressed and underexpressed transcripts in the "15 cm" comparison, respectively; and 4,000 and 7,269 overexpressed and underexpressed transcripts in the "40 cm" comparison. Finally, the "2 M" comparison identified 5,994 and 6,252 overexpressed and underexpressed transcripts. For the differential expression analysis contrasting fiber versus fiber, the "200vs40" comparison identified 650 and 394 overexpressed and underexpressed transcripts; in the "200vs15" comparison, 723 and 489 overexpressed and underexpressed transcripts were identified. Finally, in the "40vs15" comparison, 476 and 261 overexpressed and underexpressed transcripts were identified, respectively. These gene expression analyses achieved the identification of genes such as *LAC3*, *NOV*, *PRX52*, and *DegP7*, as well as transcription factors such as *APRR2*, *SPL1*, and *GRL/NAP1*, and members of the bHLH family. Some DEGs are considered gene markers of fiber formation in other crops, but their functions in agave henequen was not reported yet

### INTRODUCCIÓN:

El “Henequén” es una planta originada a partir del *Agave angustifolia*, domesticada en la península de Yucatán por los mayas desde la época prehispánica. Se utilizaba para la producción de fibras de gran calidad. Estas fibras jugaban un rol importante dentro de las necesidades básicas de la cultura maya, fungiendo como materia prima para la manufactura de diversos materiales básicos de uso doméstico, herramientas de casa, pesca, etc. El cultivo alcanzó su edad dorada en el siglo XIX (Dolores, 1988; Castillo Mora *et al.*, 2010). Con la llegada de las fibras sintéticas, su cultivo disminuyó drásticamente, hasta llegar actualmente a cerca de 20,000 hectáreas cultivables, obteniendo una producción aproximada de 27,000 toneladas por año (Nava *et al.*, 2015)

Las fibras se componen de células largas y delgadas de esclerenquima, caracterizadas por poseer una pared secundaria lignificada; esta proporciona soporte, dureza y rigidez a las hojas de esta planta. Las fibras de *Agave fourcroydes* se componen, en promedio, de cerca de un 60 % de celulosa y, en segundo lugar, lignina, aunque dependiendo de la especie podemos encontrar taninos, gomas, pectinas y otros polisacáridos (Mónica *et al.*, 2006).

La importancia de este cultivo radica en el potencial que tiene para convertirse en materia prima de múltiples industrias, por ejemplo, en la producción de bioenergía (Morán *et al.*, 2020). Además, el estudio de sus componentes activos ha aumentado por parte de la industria farmacéutica y agrícola (Espinosa *et al.*, 2021). No obstante, cerca del 60 % de las investigaciones actuales se centra en los componentes estructurales de sus fibras, destacando el potencial que aún poseen para el aprovechamiento de la biomasa lignocelulósica (Trujillo *et al.*, 2023). Esto debido a que aún se desconocen por completo los mecanismos que regulan el proceso de formación de las fibras y la biosíntesis de lignina, fundamentales para el desarrollo de nuevas variedades con fibras diseñadas y para el correcto manejo de todos los componentes de la planta (Raya *et al.*, 2021).

## CAPÍTULO I :REVISIÓN DE LITERATURA

### 1.1 Importancia actual de las fibras naturales

En última década, la conciencia por el cuidado al ambiente se incrementado mundialmente derivado de problemas graves como la escasez de agua y el calentamiento global; esto ha conducido al desarrollo de materiales eco-amigables que sustituyan a los sintéticos. En este escenario, hay un interes renovado por las fibras vegetales, las cuales pueden ser extraídas de fuentes renovables a bajo costo (Kurien *et al.*, 2023; Yano & Fu, 2023; Tavares *et al.*, 2020).

Las fibras vegetales tienen un gran número de beneficios que incluyen su baja densidad, alta resistencia mecánica, dureza, bajo costo de producción y pueden ser obtenidas de cultivos de rápido crecimiento y alta capacidad para fijar CO<sub>2</sub> atmosférico (Kurien *et al.*, 2023). Las fibras ha sido aprovechadas desde los pueblos prehispánicos hasta la actualidad para la confección de diversos artículos como textiles, papel, sogas, entre otros. Su incorporación en diversas industrias ha fungido como catalizador para poder desarrollar nuevos materiales, y ésto lo podemos observar en la elaboración de artesanías, calzado, muebles. Otro ejemplo de la versatilidad en el uso de fibras naturales está presente en el diseño y construcción de biomateriales o "materiales verdes", pues al mezclar fibras con otros polímeros o elementos de construcción (plásticos, cementos) se aumenta significativamente la resistencia de las estructuras elaboradas (Kurien *et al.*, 2023; Tavares *et al.*, 2020; Macías, 2006). Tomando ventaja de la composición química de las fibras, recientemente se está obteniendo nanocelulosa para su uso médico en implantes biológicos donde ayudan al soporte e inmovilización de células. La nanoestructura y las similitudes entre la celulosa de fibras naturales y el colágeno hacen mas atractivo el uso de la nanocelulosa de fibras que aquella proveniente de bacterias (Khalil *et al.*, 2014). También las propiedades físicas de las fibras (porosidad, morfología, capacidad de absorción) están siendo consideradas el diseño de nuevas presentaciones para fármacos como amoxicilina, diclofenaco y tigeciclina (Tavares *et al.*, 2020).

Todos los usos tradicionales y nuevos de las fibras naturales han impactado en su valor de mercado. Aparte del algodón, el cual el domina el mercado mundial de las fibras naturales, otros cultivos de fibras naturales también representan un ingreso importante para las economías mundiales. Por ejemplo, el 2017 el valor de mercado del cañamo en China se estimó en US 1.7 billones (Yano & Fu, 2023). En México, la demanda fibras naturales en 2022 fue de US\$38.1M. Ese mismo año, Yucatán se transformó en la cuarta entidad federativa con más compras

---

internacionales, con un total de US\$4.47M (Data México, 2023). También en 2022, el valor de mercado mundial de las fibras naturales se calculó en US 4.90 billones, con un crecimiento anual de 7.3% entre 2023 y 2032 (The Brainy Insights, 2024).

Durante el siglo XIX y principios del XX, el henequén o *Agave fourcroydes* Lem, también conocido como el “oro verde”, transformó radicalmente la economía y sociedad yucateca. La producción de henequén colocó a México como el mayor exportador mundial de fibras duras, representando más del 90 % de la producción global (Balam *et al.*, 2006). La era del henequén moldeó profundamente la estructura social y el panorama físico de Yucatán. En este estado, se estableció un sistema de trabajo casi feudal, donde miles de trabajadores mayas laboraban en las haciendas henequeneras (Katz, 1962) . Simultáneamente, la riqueza generó un desarrollo urbano sin precedentes. La ciudad de Mérida se transformó en una ciudad cosmopolita con mansiones, teatros y bulevares al estilo parisino. Simultáneamente, impulsó la construcción de infraestructuras modernas como ferrocarriles, puertos y telégrafos, integrando a Yucatán en la economía internacional (Fraga, 2004).

Actualmente la relevancia económica de henequén ha disminuido significativamente debido a la competencia de fibras sintéticas y otras naturales, sin embargo este cultivo mantiene un papel importante en la identidad cultural yucateca. También, gracias al renovado interés mundial por las fibras naturales, las fibras del henequén han comenzado a diversificar sus usos tradicionales; por ejemplo, están implementado en industrias de la construcción como aditivo en concreto y en textiles de alta gama, gracias a sus propiedades mecánicas y baja huella de carbono (Cazaurang *et al.*, 1991). Además, investigaciones recientes han demostrado su potencial como materia prima para la generación de biocombustibles y como fuente de compuestos bioactivos, abriendo nuevas oportunidades para este cultivo tradicional (Núñez *et al.*, 2011).

## **1.2 Esclerénquima y clasificación anatómica de fibras**

El esclerénquima es un tipo de célula caracterizado por proporcionar rigidez y resistencia a las estructuras vegetales de todas las plantas superiores (Evert, 2006; Esau, 1953). Estas poseen paredes secundarias engrosadas, lignificadas y carentes de citoplasma (Yokoyama *et al.*, 2014). Una de las particularidades de la pared secundaria de estas células son las punteaduras pareadas que encuentran por toda la célula; estas puntuaduras pueden ser simples o múltiples, lo que permite la comunicación entre las células vecinas. A diferencia de otro tipo de células, no poseen una forma determinada, por lo cual podemos encontrarlas desde alargadas hasta cortas y estrelladas (Esau, 1953). Otra particularidad de las células esclerenquimáticas es la carencia

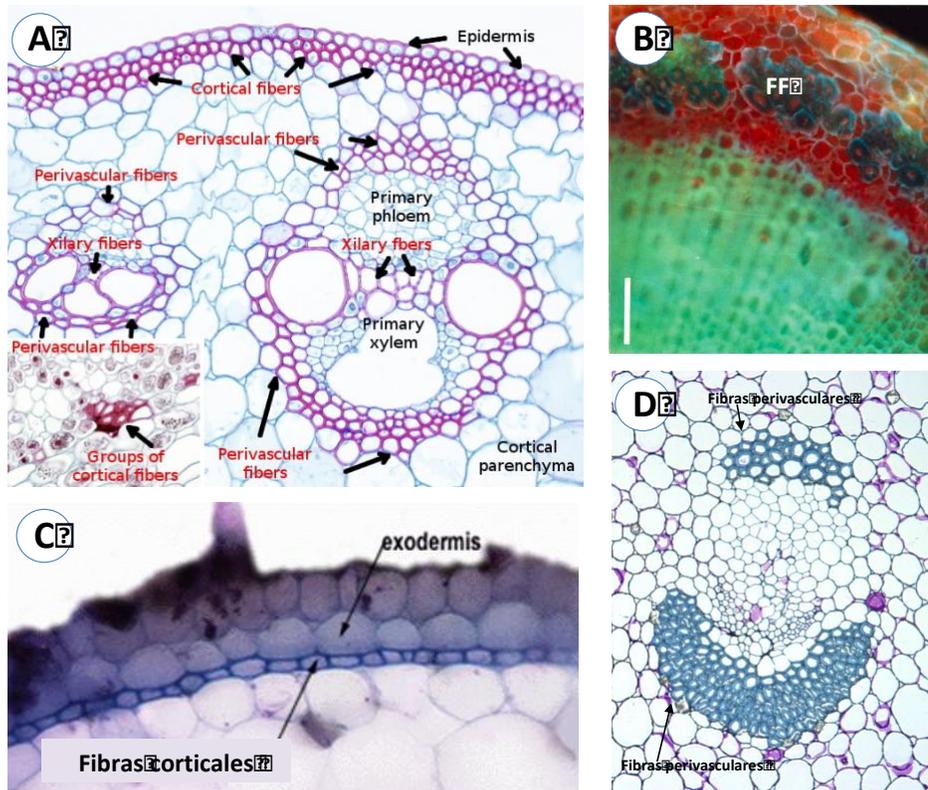
---

de espacios intercelulares, ocasionando que sus paredes se depositen de forma uniforme en todo el contorno de la célula.

Las células esclerenquimáticas se puede clasificar en dos tipos principales: fibras y esclereidas (Esau, 1953). Las esclereidas no poseen una forma determinada, por lo cual podemos encontrarlas desde alargadas hasta cortas y esféricas. Se utilizan para rellenar espacios y proporcionan apoyo mecánico (Evert, 2006; Martín, 1973). Por otro lado, las fibras esclerenquimáticas se caracterizan por tener paredes celulares secundarias lignificadas que proporciona fuerza y elasticidad a las estructuras vegetales, confiriendo gran resistencia a tallos y hojas; este tipo de fibras se pueden hallar en en el cáñamo, el lino, el yute, bamboo entre otras más. Las fibras esclerenquimáticas suelen encontrarse en grupos o paquetes, esta configuración proporciona resistencia y soporte a las plantas, permitiendo a las plantas permanecer erectas (Kurien *et al.*, 2023; Jarvis, 2012; Esau, 1953).

Las fibras esclerenquimáticas se pueden clasificar como fibras xilares y fibras extraxilares, en función si están asociadas al xilema o no, respectivamente (Esau, 1953) (figura 1.1). Las fibras extraxilares se distinguen por tener una pared secundaria gruesa, llegando a ocupar hasta el 90 % del área de la célula. En algunas especies vegetales, este tipo de fibras pueden presentar una estructura multilameada, por ejemplo en bamboo. Hay diferentes tipos de fibras extraxilares en función de su posición dentro de los tejidos vegetales y origen ontogénico. Las fibras extraxilares asociadas o inmersas en el floema primario o secundario, llamadas fibras de floema, se originan del protofloema. En el caso de las fibras de lino (*Linum usitatissimum* L.) son fibras asociadas a floema primario (figura 1.1B), y en el caso de vid (*Vitis vinifera*) son fibras asociadas a floema secundario. Otro tipo de fibras extraxilares son las fibras corticales, las cuales a diferencia de las de floema se originan en el cortex y se hallan justo debajo de la epidermis. Este tipo de fibras se hallan generalmente en las monocotiledóneas, como maíz y arroz (figura 1.1C) (Zhong *et al.*, 2011), y están apiladas en dos o tres capas de células; son células alargadas y delgadas, con extremos afilados; su función principal es proporcionar soporte y rigidez a la planta. Las fibras perivasculares también son una clase de fibras extraxilares y como su nombre lo indica se encuentran alrededor de los haces vasculares (figura 1.1D); son largas y delgadas, pero no tan afiladas como las fibras corticales; su función principal es proteger los haces vasculares del daño (Megías *et al.*, 2024; Esau, 1953).

Figura 1.1. Tipos de fibras vegetales. A) Fibras halladas en tallo de maíz. Se indican los diferentes tipos de fibras en función de posición dentro del órgano



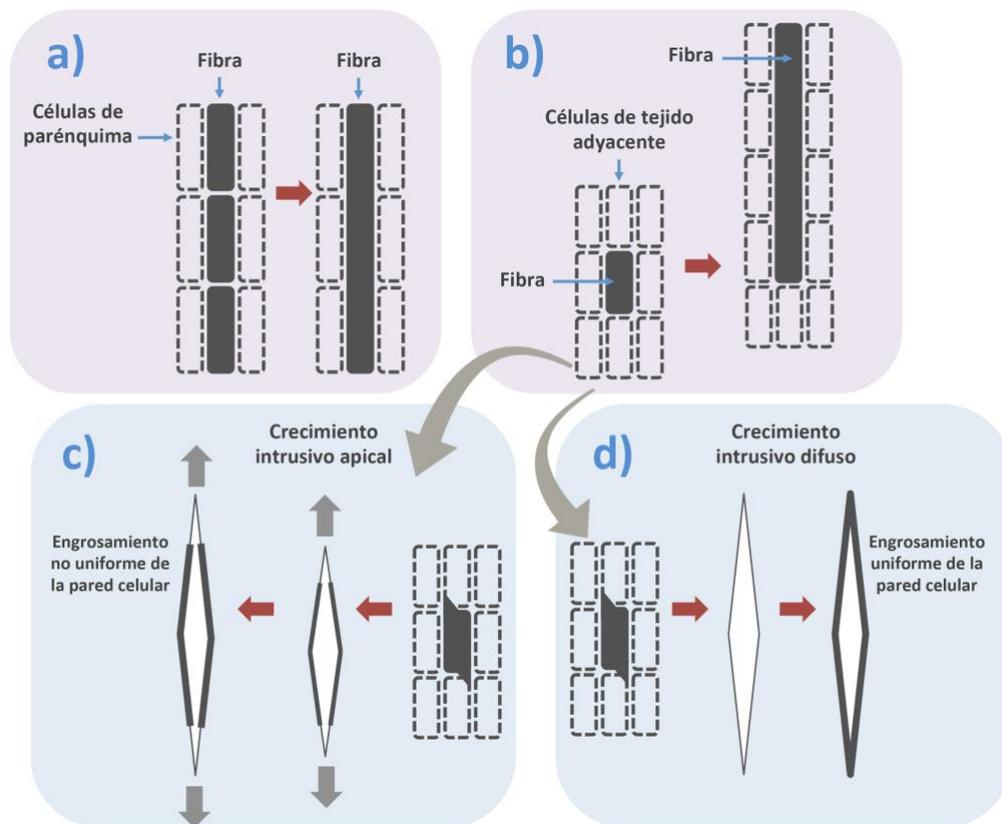
(tomado de Megías *et al.*, 2024). B) Fibras de floema primario (FF) en tallo de lino (tomado de Gorshkova *et al.*, 2023) C) Fibras corticales en raíz de arroz (tomado de <https://labs.plb.ucdavis.edu>). D) Fibras perivasculares en hojas de A. tequilana (tomado de Maceda *et al.*, 2022).

### 1.3 Etapas de formación de las fibras

Durante el desarrollo de las fibras se pueden distinguir cuatro etapas: iniciación, elongación, intrusión y senescencia de la fibra. Cada una de estas etapas se describen a continuación.

Iniciación de la fibra. El estudio de las fibras en su etapa inicial ha sido un reto, principalmente por la dificultad que implica la identificación en sus primeras etapas de desarrollo. Se supone que este proceso es el resultado de la fusión de las células del parénquima (figura 2.2A), iniciando con la disolución de la pared celular entre ellas, aunque en estudios realizados en *Linum usitatissimum* apoya la teoría de la polinuclearidad (Barros *et al.*, 2015). Otra complicación surge para el estudio de las fibras en sus etapas tempranas en que solo podemos encontrarlas en sitios determinados dentro de las plantas, con un origen a partir de la actividad de los meristemos primarios y secundarios. Muchos de los factores que regulan la formación inicial de fibras son desconocidos. Sin embargo, el interés por las fibras ha aumentado significativamente, por lo cual

se han realizado estudios enfocados en la regulación hormonal, han determinado la participación del ácido giberélico sobre la longitud de la pared celular, este mismo estudio también determinó el efecto de la sinergia existente entre las auxinas, ácido giberélico y las citoquininas (Smith *et al.*, 2013; Chernova & Gorshkova, 2007). Uno de los principales errores detectados al tratar de identificar la etapa de iniciación, son la identificación errónea de las fibras, debido a que se presentan en diversas configuraciones o se encuentran dañadas por diversos factores, incluso se ha documentado la existencia de células que logran mantener paredes delgadas por mucho tiempo. Esto ocasiona que las células utilizadas en los ensayos no son representativas (Gorshkova *et al.*, 2012).



**Figura 1.2.** Desarrollo de las fibras del esclerénquima: (a) Formación de fibras por fusión de células del parénquima y disolución de septos entre ellos; (b) Elongación coordinada de la parte superior de la fibra y crecimiento continuo de la planta; (c) Crecimiento intrusivo apical y (d) crecimiento intrusivo difuso. Figura redibujada de Chernova & Gorshkova, (2007). (Tomado y modificado de referencia (2007))

---

Elongación de la fibra. Una vez iniciado el proceso de formación de la fibra, continua con el crecimiento de la misma, donde la elongación de la fibra es el resultado de la coordinación de los tejidos circundantes del entrenudo, gracias a este alargamiento las células obtienen longitudes que varían desde docenas o cientos de milímetros. Sin embargo, el crecimiento intrusivo ocurre cuando la célula que está dando origen a la fibra crece más que los tejidos circundantes y se introduce entre las células (características clave de la biogénesis de la fibra) (Chernova & Gorshkova, 2007). Este crecimiento puede ser apical o difuso; el crecimiento apical ocurre en los extremos de la célula, mientras que el crecimiento difuso ocurre por toda la superficie de la célula, como se ilustra en la figura 2. 2 (Yokoyama *et al.*, 2014). Este mecanismo se encuentra altamente regulado dentro de la planta, lo que lo hace vulnerable a diversos tipos de estrés, afectando tanto al crecimiento coordinado como al intrusivo, de formas que aún no se comprende por completo. En el 2006, Chervona propone que parte del mecanismo está sujeta a la regulación hormonal, ya que se ha visto que el crecimiento de la fibra se ve afectado en presencia de hormonas exógenas, principalmente auxinas y ácido giberélico. En algunas especies, al suministrar de manera combinadas o individual estas hormonas, se acelera el crecimiento del entrenudo del tejido. El potasio también es muy importante, pues los cambios en su concentración afectan el diámetro de los vasos, aunque no se observan cambios en el diámetro de las fibras, sugiriendo una relación en el creciente de las fibras vegetales con el transporte de iones (Snegireva *et al.*, 2010; Chernova & Gorshkova, 2007).

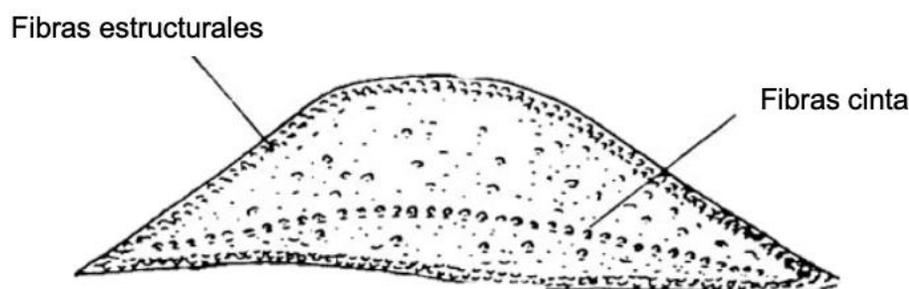
Intrusión de la fibra. El proceso de intrusión se inicia en la punta de la fibra, la cual secreta una enzima que debilita la lámina media entre las células adyacentes logrando colocarse entre las células vecinas sin dañarlas (Lev, 2001; Wenham & Cusick, 1975). Durante este proceso, aparece una estructura en forma de rodilla, la cual desaparece una vez finalizado el crecimiento de la fibra. La duración del crecimiento inclusivo dependerá de múltiples factores, debido a que este proceso cambia dependiendo del tipo de fibra y la especie en cuestión. Se ha demostrado que el engrosamiento de la pared celular secundaria es uniforme a lo largo de toda la fibra, sugiriendo la existencia de un mecanismo de regulación encargado de formación la pared celular secundaria en la superficie interna de la fibra que aún no se comprende por completo (Gorshkova *et al.*, 2012; Snegireva *et al.*, 2010; Chernova & Gorshkova, 2007). Al igual que en la etapa anterior, el estudio de las fibras se complica debido a que las células están incrustadas entre las células vecinas del parénquima y tienen paredes celulares delgadas y frágiles, comprometiendo integridad del tejido para su extracción (Gorshkova *et al.*, 2012).

---

Senescencia de la fibra. La senescencia es la última etapa de la formación de las fibras vegetales. Da inicio con la reducción del tamaño del citoplasma y las vacuolas, también con la rotura del tonoplasto en las fibras y la presencia de vacuolas autolíticas en el citoplasma. Posteriormente, se produce la descomposición del tonoplasto, una vacuolización intensa y un aclaramiento del citoplasma; el aparato de Golgi se reduce y el sistema tilacoidal de los cloroplastos se degrada. Sin embargo, el núcleo y las mitocondrias no presentan cambios significativos en esta etapa. También se han podido observar grandes gránulos de almidón en esta etapa. En algunas especies se presentan protoplastos conservados y mitocondrias con estructura nativa en las fibras hasta la etapa de madurez completa, lo que sugiere que algunas células siguen siendo viables incluso a esta edad (Gorshkova *et al.*, 2012; Lev, 2010; Chernova & Gorshkova, 2007) aunque se hipotetizaba que las células que componen las fibras deben de morir para poder dar inicio a la lignificación de sus paredes secundarias (Evert, 2006)

#### **1.4 Anatomía y patrones de distribución de fibras**

Las células constan de cuatro componentes principales: lumen, pared celular primaria; secundaria y terciaria. Las paredes celulares constan de varias capas que se mantienen unidas mediante sustancias pegajosas y cerosas conocidas como lignina, pectina y hemicelulosas (Morán, 2021; Cazaurang *et al.*, 1991). Dependiendo de la especie, la proporción de estos compuestos puede variar. Por ejemplo, las fibras del henequén consta de una estructura compuesta por delgadas fibras filiformes; las células que componen las fibrillas constan un lumen con un tamaño promedio de 3 a 12  $\mu\text{m}$  con esquinas y bordes redondeados. Este lumen es de un tamaño significativo si se compara con el de la pared secundaria. Entre las células hay una región denominada laminilla media constituida principalmente por lignina y pectina; usualmente tiene formas hexagonal y pentagonal; ellas están acomodadas en varias capas compuestas principalmente por celulosa orientadas en direcciones angulares que forman una hélice plana de forma transversal (Lev, 2010; Chernova & Gorshkova, 2007). Otra particularidad de las fibras de henequén es que tiene secciones decrecientes en su diámetro, manteniendo el tramo medio constante. Este comportamiento podría atribuirse a las diferencias en el contenido de lignina y hemicelulosa a lo largo de la fibra o un daño ocasionado al calcular el área de la sección transversal (Cazaurang *et al.*, 1991).



**Figura 1.3.** Localización de las fibras estructurales y de cinta dentro de una hoja de agave. Figura modificada de Yan *et al.*, (2000)

Dentro de las hojas de agave existe tres tipos básicos de fibras: estructurales, fibras de cinta y de xilema. En la Figura 1.3, podemos observar el patrón de distribución de éstas en un corte transversal de la hoja. Las fibras estructurales dan rigidez a la hoja y se encuentran en la periferia de la hoja; son fáciles de identificar porque se presentan generalmente en forma de "herradura". Las fibras de cinta crecen en asociación con los tejidos conductores de la planta y generalmente se encuentran en el medio de la hoja, distribuidas desde la base hasta la punta, poseen una excelente resistencia mecánica. Finalmente, las fibras del xilema crecen opuestas a las fibras de cinta y están conectadas a través de los tejidos conductores, compuestas por células de paredes delgadas, que a diferencia de las anteriores, se pierden durante el proceso de extracción de fibra debido a su fragilidad (Morán, 2021; Silva *et al.*, 2008).

### 1.5 Composición de las fibras vegetales

Las fibras vegetales son estructuras compuestas principalmente por celulosa, lignina y hemicelulosa (Yokoyama *et al.*, 2014). La celulosa es el principal componente, confiriendo resistencia y rigidez a las células de esclerénquima. Por otro lado, la lignina es la encargada de proporcionar estabilidad y resistencia ante la degradación. Mientras que la hemicelulosa tiene la función de mantener las estructuras vegetales. Sin embargo, existen otros compuestos que forman parte de la estructura de la fibras, pero estos dependen en gran medida de la especie a la cual pertenece la fibra. El porcentaje de cada uno de estos componentes varía según la especie, localización y tipo de tejido vegetal en cuestión. Aunque en general, se ha demostrado que las fibras con un mayor contenido de celulosa son más fuertes y resistentes (Chokshi *et al.*, 2022). Sin embargo, podemos encontrar otros compuestos que forman parte de su composición, por ejemplo pectina, ceras y cenizas; la concentración de estos compuestos es

considerablemente menor a comparación de los otros compuestos mencionados anteriormente, y dependen en gran medida de cada especie. A continuación se menciona los rangos en el que es posible encontrarlos: pectinas oscilan entre 0.45 % a 10 % y ceras van entre 0.09% hasta un máximo del 4 %. Por último, el contenido de ceniza oscila 0.60% a 8.00%. En la tabla 1.1 podemos comparar la concentración de los principales compuestos que integran de algunas de las fibras vegetales más conocidas y aprovechadas a nivel mundial (Chokshi *et al.*, 2022). Resaltan especies como el algodón, ortiga y piña con fibras con niveles de celulosa por arriba del 80 %, en contraste con el capoc por abajo del 20 %. En cambio, algodón y ramio poseen fibras con los contenidos de lignina mas bajos reportados, 0.75 y 0.8 %, respectivamente, que posiblemente expliquen la suavidad al tacto.

**Tabla 1.1.** Composición química de fibras naturales (Morán, 2021).

Cultivo	Celulosa (%)	Hemicelulosa (%)	Lignina (%)	Pectina (%)	Ceras (%)
Abacá	62.5	21	12	0.8	3
Algodón	89	4	0.75	6	0.6
Bambú	34.5	20.5	26	-	-
Cáñamo	81	20	4	0.9	0.8
Capoc	13.16	-	-	-	-
Coco	46	0.3	45	4	-
Curauá	73.6	5	7.5	-	-
Esparto	45.4	38.5	14.9	-	2
Henequén	60	28	8	-	0.5
Isora	74	-	23	-	1.09
Ortiga	86	10	-	-	4
Phormim	67	30	11	-	-
Piassava	28.6	25.8	45	-	-

Piña	80.5	17.5	8.3	4	-
Plátano	62.5	12.5	7.5	4	-
Ramio	72	14	0.8	1.95	-
Sisal	60	11.5	8	1.2	-

La tabla 1.2 compara las diferencias entre las concentraciones de los principales componentes de las fibras de algunas especies de agaves. En ella se indica que *A. angustifolia* y *A. cantala* destacan por tener las concentraciones más altas de hemicelulosa de todas las especies examinadas, en contraste con *A. fourcroydes* y *A. lechuguilla*, que tienen los porcentajes más bajos, con un promedio de 3 - 7%. Las concentraciones de celulosa varían drásticamente entre las especies comparadas. Los agaves utilizados principalmente para la obtención de fibras duras como *A. fourcroydes*, *A. sisalana* y *A. lechuguilla* destacan por poseer las concentraciones más altas de celulosa con valores de 77.6 %, 71% y 79.8%, respectivamente. En cambio, *A. atrovirens* y *A. salmiana* muestran valores similares en un rango de 24 - 36 %. En cuanto a lignina, *A. angustifolia* fue el agave con la mayor concentración, con un 20.69 %, seguido por *A. lechuguilla*, *A. fourcroydes* y H11648, con 15.3 %, 13.1 % y 15.2 %, respectivamente. Los estudios en *A. tequilana* y *A. americana* fueron los únicos trabajos donde se mencionaron concentraciones de ceras, con 0.19 % y 0.26 %; en el resto de los reportes no se indican dichas concentraciones.

**Tabla 1.2.** Composición química de fibras de agave (% del peso total seco)(Morán *et al.*, 2023).

Especie/Híbrido	Celulosa	Hemicelulosa	Lignina	Ceras	Referencias
<i>A. americana</i>	68.42	15.67	4.85	0.26	Mylsamy y Rajendran 2010
<i>A. angustifolia</i>	48.04	34.08	20.69	NR	Hidalgo-Reyes et al., 2015
<i>A. atrovirens</i>	24-36	10	2	NR	Pérez-Pimienta et al., 2017
<i>A. cantala</i>	48.97	34.41	11.41	NR	Yudhanto et al., 2021
<i>A. fourcroydes</i>	77.6	5-7	13.1	NR	Escamilla-Treviño, 2012

<i>A. lechuguilla</i>	79.8	3-6	15.3	NR	Vieira et al., 2002
<i>A. salmiana</i>	32	11	10	NR	Pérez-Pimienta et al., 2017
<i>A. sisalana</i>	71	11.7	8.7	NR	Sahu y Gupta, 2017
<i>A. tequilana</i>	62.41	18.28	11.57	0.19	Binoj y Binin, 2019
H11648	60.20	16.03	15.20	NR	Magaton, 2015

NR= No reportado

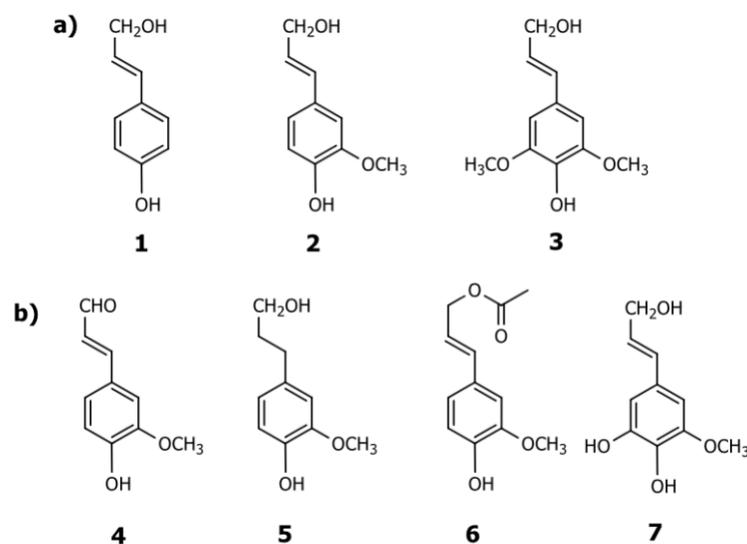
## 1.6 Lignina

La lignina es uno de los biopolímeros más abundantes encontrados en el planeta, no obstante ésto su estructura y su mecanismo de regulación no se comprenden por completo. La lignina se puede hallar depositada en la pared celular de las plantas superiores y su composición química consta de unidades de fenilpropanos unidas entre sí por enlaces covalentes con las fibras de celulosa y hemicelulosa para proporcionar rigidez y resistencia a la pared celular secundaria (Chernova & Gorshkova, 2007; Vázquez, 1999). En las fibras, la lignina representa una proporción importante y es de gran relevancia económica en aquellos cultivos destinados a bioenergía, pues su eliminación puede ser laboriosa y encarecer el proceso.

La lignificación es el proceso por el cual se sintetizan y depositan lignina usando diversos alcoholes fenólicos, conocidos como monolignoles los cuales se polimerizan y depositan en las paredes celulares secundarias de ciertas células vegetales, como las células del xilema y fibras. Este proceso implica la síntesis, transporte y la polimerización oxidativa de los monolignoles para formar la macromolécula aromática tridimensional compleja y racemizada de la lignina (Ralph *et al.*, 2004; Boerjan *et al.*, 2003). El proceso involucra una serie de pasos altamente regulados y coordinados que requieren la participación de múltiples enzimas y proteínas en diferentes compartimentos celulares para cada una de las diferentes fases desde la biosíntesis hasta la deposición de las subunidades polimerizadas. En las siguientes secciones se describirán cada una de ellas.

## 1.7 Síntesis de los monómeros de lignina

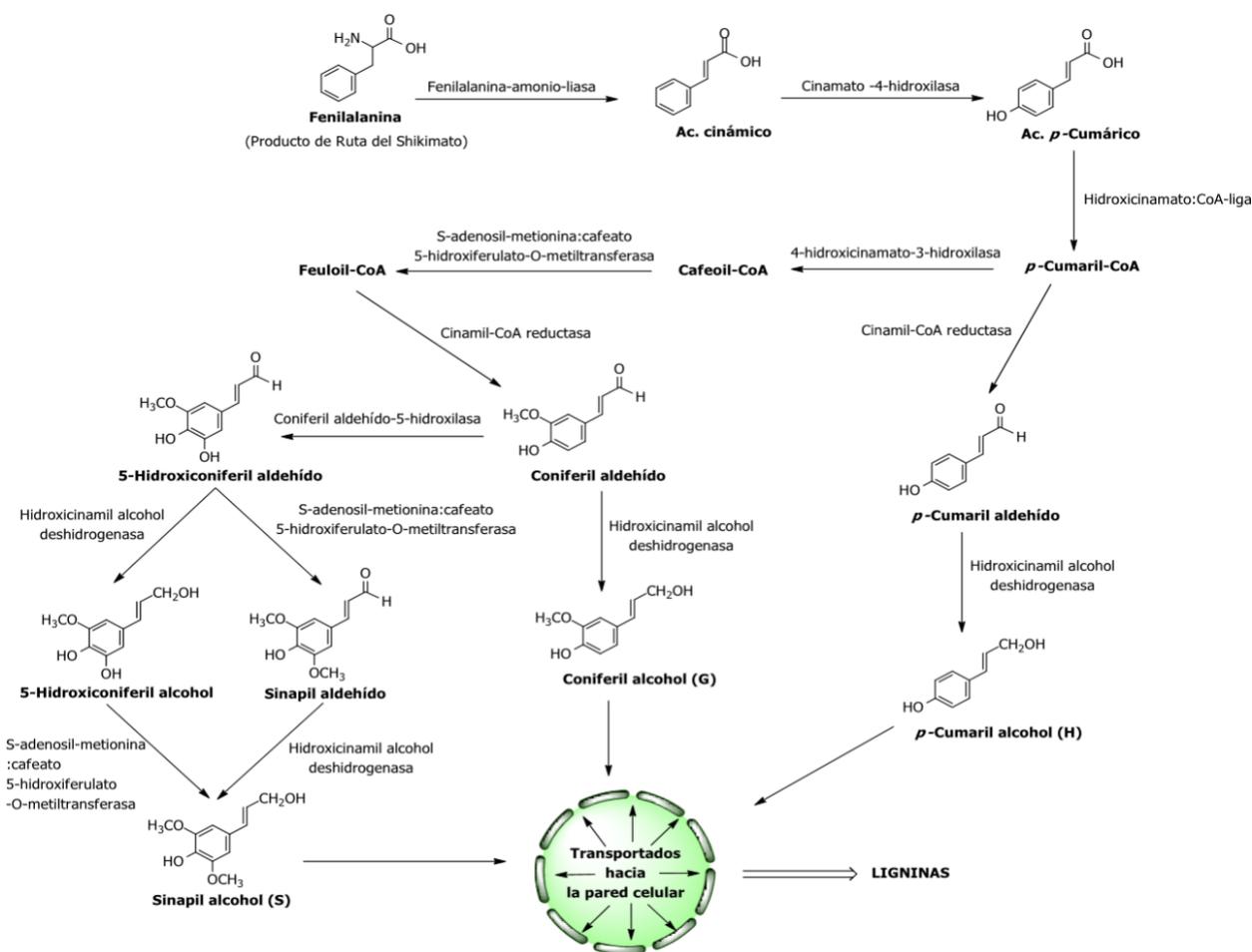
Los monómeros o monolignoles que conforman a la lignina son sintetizados a partir de la fenilalanina, la cual es producida plástidos a partir de la vía biosintética del shikimato (Maceda *et al.*, 2021; Barros *et al.*, 2015; Micic *et al.*, 2003). En la Figura 1.4 podemos visualizar estos tres componentes: unidades H- (hidroxifenilo), G- (guaicil) y S- (siringilo), también conocidas como lignina H, G y S, respectivamente.



**Figura 1.4.** a) Los monolignoles más comunes encontrados en los bloques de construcción de la lignina: 1: alcohol p-cumarílico, 2: alcohol coniferílico, 3: alcohol sinapílico, b) otros tipos de monolignoles (coniferil-derivados): 4: coniferaldehído, 5: alcohol (recolectada de Chávez-Sifontes & Domine, 2013)

La biosíntesis de lignina puede iniciar con la transformación de la fenilalanina en ácido cinámico, paso mediado por la enzima fenilalanina amonio liasa (PAL), que luego se convierte en ácido p-cumárico. En la primera ruta de biosíntesis, el p-cumaril CoA se transforma en L-tirosina con la ayuda de la enzima tirosina amonia liasa (TAL). Mientras que la segunda ruta es más compleja y produce múltiples moléculas como el alcohol p-cumarílico (Lignina H), el ácido ferúlico, el ácido sináptico, el alcohol catequilo (Lignina C), el alcohol coniferílico (Lignina G) y el alcohol sinapílico (Lignina S) (Barros *et al.*, 2015; Adler, 1977). Estos monolignoles que formarán parte del polímero de la lignina, aunque cabe destacar que las ligninas de la gimnospermas están compuestas casi

por completo de unidades G y con cantidades menores de unidades H, mientras que las dicotiledóneas usualmente se componen de unidades G y S (Vanholme *et al.*, 2010). En las plantas monocotiledóneas, la ruta de biosíntesis de lignina puede tener origen distinto pues estas plantas cuentan con la capacidad de utilizar la tirosina como un precursor adicional (Barros *et al.*, 2015).



**Figura 1.5.** Ruta biosintética de la lignina (tomada de Chávez & Domine, 2013).

La síntesis de monómeros de lignina tiene lugar tanto en el citoplasma como en la superficie externa del retículo endoplasmático, creando una compleja relación entre estos dos sitios (figura 1.5); por ejemplo, la fenilalanina amonía liasa (PAL) es citoplasmática en tanto que cinamato 4-hidrolasa (C4H) está localizada en el retículo endoplasmático. Además, se ha encontrado que la p-hidroxocinamoil-CoA (HCT) y la 4-(hidroxi)cinamoil CoA ligasa (4CL) están parcialmente asociados con el retículo endoplasmático tras la expresión de p-cumarato 3-hidrolasa (C3H) (Boerjan *et al.*, 2003; Whetten & Sederoff, 1995). Otras enzimas como las oxidorreductasas del

---

citocromo P450, que son responsables de la hidroxilación del anillo aromático, se encuentra en la superficie externa del retículo endoplasmático. Las demás enzimas se producen en áreas subcelulares especializadas en la interfaz entre el citoplasma y la superficie del retículo endoplásmico (Barros *et al.*, 2015).

### **1.8 Mecanismos para el transporte de monolignoles a través de membranas**

Los monómeros una vez sintetizados deben ser transportados a la pared celular por las fenoloxidasas donde empieza la adición de nuevos monómeros, los cuales son activados en sus extremos y van creciendo en conjunto al adicionarse las diferentes subunidades (Smith *et al.*, 2013). El proceso de transporte de monolignoles hasta la pared celular no está completamente descrito hasta ahora, pero se sugieren tres mecanismos involucrados en función: transportadores de membranas o difusión pasiva, exocitosis asociada a vesículas y transporte activo dependiente de ATP utilizando transportadores tipo ABC (Maceda *et al.*, 2021).

En los estudios sobre el mecanismo de transporte dependiente de ATP que utiliza transportadores tipo ABC, la membrana plasmática ha demostrado una selectividad por el transporte del alcohol hydroxycinnamyl y algunos aldehídos, y una incapacidad de aceptar compuestos como el ácido ferúlico o monolignoles glucosidados (Sibout & Höfte, 2012). Experimentos realizados por Alejandro *et al.* (2012) ha identificado exitosamente al gen AtABC29 en *Arabidopsis thaliana*, el cual posee la capacidad de reducir la toxicidad del alcohol p-coumaryl, pero no la del alcohol conoferyl, sugiriendo que este transportador es específico para el bombeo del alcohol p-coumaryl fuera de las células. También se ha reportado que la acumulación del p-coumaryl en el citosol causa la inhibición de la síntesis de monolignoles. Por otro lado, mediante simulaciones dinámicas moleculares realizadas por Vermaas *et al.*, (2019) se ha determinado que la permeación pasiva de membranas es capaz de soportar hasta 69 monolignoles diferentes relacionados con la lignina, a través de un modelo representativo de *Z. mays*.

En estudios tempranos realizados por Fujita & Harada (1968) y Pickett (1979) se observó un tráfico que va desde la vesícula hasta el citosol y plasmalema, al etiquetar el xilema con [3 H]-Phe, -tirosina y -cinnámico. Los investigadores encontraron estas etiquetas radiomarcadas asociadas al retículo endoplasmático rugoso, el aparato Golgi, y con algunas vesículas fusionadas con la membrana plasmática o agregadas en el citoplasma cerca de las bandas de microtubos de pared. Estos hallazgos sugieren que el aparato de Golgi está involucrado en la síntesis y transporte de monolignoles a la pared celular. Sin embargo, estudios autoradiográficos

---

posteriores sugieren que es poco probable que la exocitosis mediada por aparato de golgi desempeñe un papel importante en el transporte de monolignoles (Liu *et al.*, 2011).

### 1.9 Polimerización de monolignoles en la pared celular secundaria

Al finalizar el transporte de los monolignoles hacia la pared secundaria, el proceso de polimerización comienza con la oxidación de los monómeros. Estos monómeros de lignina se unen entre sí formando enlaces covalentes de dos tipos: condensados y no condensados. Los condensados C-C son del tipo 5-5',  $\beta$ -5',  $\beta$ - $\beta$  y  $\beta$ -1'; estos enlaces se caracterizan por ser más fuertes y por proporcionar mayor rigidez al monómero de lignina, mientras que los enlaces no condensados C-O- C como  $\beta$ -O-4' son más flexibles y permiten a la lignina adaptarse a los cambios en la forma de las células vegetales(Maceda *et al.*, 2021).

Los monómeros se acumulan y entrelazan deshidrogenizando los dímeros, de nuevo a radicales fenólicos, antes de que pueda acoplarse a otro monómero y de esta forma va creciendo el polímero. A este mecanismo se denomina acoplamiento terminal; en algunas especies puede ramificarse debido a la presencia de enlaces de tipo 5-5 o 4-O-5. La deshidrogenación de los monolignoles está estrictamente regulada por las lacasas dependientes de  $O_2$  y las peroxidasas dependientes de  $H_2O_2$ . Las peroxidasas pueden diferir en sus especificidades del sustrato; por ejemplo algunas son muy específicas para alcohol sinapílico, mientras que otros únicamente aceptan el alcohol coniferílico. Esta especificidad puede determinar la estructura final del polímero de la lignina mediante la expresión de isoformas específicas de las misma. Debido a que estas enzimas poseen actividades superpuestas, ha generado un enorme inconveniente al momento de estudiarlos (Weng & Chapple, 2010) Existen diversos modelos hipotéticos que tratan de explicar la interacciones de las lacasas y las peroxidasas durante la lignificación de la pared celular. Dentro de los modelos más destacados encontramos los siguientes: a) la diferenciación de la especificidad del sustrato, b) la intervención de la expresión espacio-temporal divergente y c) los complejos proteicos diferenciales. El primer modelo plantea que las lacasas son necesarias para la activación de los monómeros para producir los polímeros de lignina centrales, los cuales requieren de la acción de las peroxidasas para su ramificación o extensión. El segundo modelo propone que dichas enzimas exhiben una distribución y secreción selectiva en la matriz de las paredes celulares tanto primarias como secundarias, sugiriendo la existencia de mecanismos autónomos en las células, además de posibles interacciones cooperativas. El último modelo asocia el rol de las fenoloxidasas y otras proteínas, para formar complejos

---

proteicos que restringen la especificidad de sustrato o la localización en la pared celular, estos complejos proteicos podrían formarse en un momento o lugar específicos para permitir la lignificación (Barros *et al.*, 2015; Tobimatsu & Schuetz, 2019).

### 1.10 Teorías sobre la lignificación celular

Existen varias teorías y modelos propuestos que intentan explicar el proceso de lignificación, entre los modelos más aceptados se pueden encontrar tres: La lignificación cooperativa, semicooperativa y autónoma (Barros *et al.*, 2015)

La teoría de la lignificación cooperativa propone que la polimerización de los monolignoles y la deposición de lignina en la pared celular secundaria implica la participación coordinada de las células parénquimáticas adyacentes la fibra que se está formando, las cuales proporcionan las enzimas y proteínas necesarias en el proceso de lignificación (Liu, 2012; Barros *et al.*, 2015). Las enzimas localizadas en el citosol de las células vecinas sintetizan los monolignoles necesarios, los cuales son transportados al apoplasto por transportadores específicos; ahí, las lacasas y peroxidasas catalizan la deshidrogenización oxidativa de los monolignoles, generando radicales libres que se acoplan y polimerizan para formar la lignina, la cual se deposita finalmente en la pared celular secundaria del esclerénquima. Aún se debate si el inicio de este proceso se produce antes o después de la muerte de la célula a lignificar. El ejemplo mejor documentado de este tipo de lignificación lo podemos observar en los elementos traqueales (Barros *et al.*, 2015; Zhao, 2016).

A diferencia del modelo anterior, el modelo de la lignificación semicooperativa propone que tanto las células del esclerénquima como las células parenquimáticas vecinas participan activamente en la producción de monolignoles, lacasas, peroxidasas y las especies reactivas de oxígeno (Higuchi, 2006; Liu *et al.*, 2018). Por otro lado, la teoría de la lignificación autónoma sugiere que el proceso de lignificación ocurre mientras las células lignificantes aún están vivas y se detiene únicamente después de la muerte celular programada. Por lo tanto, cada célula es responsable de su propia lignificación. Este escenario ha sido documentado en fibras del xilema y en células esclerenquimáticas originadas a partir del parénquima (Barros *et al.*, 2015; Voxeur *et al.*, 2015)

Cabe recalcar que los mecanismos de lignificación varían dependiendo de la especie estudiada y de la función de la célula analizadas, por lo que la lignificación puede presentarse desde la autonomía celular hasta la cooperación total. Sin embargo, se requieren estudios más

---

exhaustivos para poder tener una comprensión más completa de la función biológica y la importancia de estos diferentes niveles de autonomía celular en la lignificación.

---

## 1.11 Justificación

La lignina es uno de los compuestos principales que conforman la estructura de las fibras, impregnándose en la matriz de celulosa y hemicelulosa de la pared celular secundaria para otorgar las propiedades de dureza y rigidez que caracteriza a muchos tipos de fibras. A diferencia del xilema donde hay grandes avances en el conocimiento sobre los circuitos genéticos que controlan su diferenciación y desarrollo, en el caso de fibras este conocimiento es aún limitado. Biológicamente, las fibras constituyen uno de los principales tejidos de soporte y protección, junto con el xilema; económicamente, las fibras tienen un alto valor de mercado, el cual se ha ido incrementando por el cambio mundial hacia las economías verdes con menor impacto ambiental. Por tanto, hay necesidad de abonar en el conocimiento de las rutas biosintéticas y mecanismos genéticos que controlan la formación de fibras, especialmente en aquellos cultivos de importancia económica.

El henequén o *Agave fourcroydes* es una especie endémica de la península de Yucatán; hasta antes del uso masivo de fibras artificiales, este cultivo fue la principal fuente de fibras duras a nivel mundial. Aunque actualmente *A. sisalana* supera la producción de *A. fourcroydes*, éste último es todavía importante en países como México, y sigue siendo una fuente importante de germoplasma para programas de mejoramiento de agave con fines de obtención de fibra retomados recientemente por Brasil, el principal mundial productor de fibras de agave (Raya *et al.*, 2023). Aunque se han reportado la secuenciación de genomas algunas especies de agave como *A. tequilana* y *A. fourcroydes* (Peña *et al.*, 2023), los mecanismos involucrados en la formación de fibras en agaves han sido explorados hasta ahora usando únicamente enfoques transcriptómicos. El análisis de transcriptomas de hoja de henequén y *agave sisalana* mostró expresión diferencial para genes tipo calosa sintasa, sugiriendo un papel estructural del compuesto calosa para compensar los bajos niveles de lignina que se han reportado en las fibras de agave (Raya *et al.*, 2023). Sorprendentemente, en estas dos especies de agave, muchos de los genes de la ruta de fenilpropanoides no fueron hallados con niveles de expresión significativamente altos (Raya *et al.*, 2023), dejando preguntas abiertas sobre los mecanismos genéticos que controlan la biosíntesis del componente lignina en fibras.

En henequén se ha demostrado que existen diferencias significativas en las concentraciones de lignina, celulosa y hemicelulosa durante la formación de las fibras de henequén; esto coincide con cambios en los niveles de expresión de algunos genes tipo *CESA*, involucrados con la biosíntesis de celulosa (Morán *et al.*, 2021; García *et al.*, 2022). Similarmente a lo hallado en henequén (Raya *et al.*, 2023), un análisis expresión de genes biosintéticos relacionados con celulosa y lignina en

---

transcriptomas de hoja de *A. tequilana* no mostró incrementos significativos en genes *CAD* y *CESA* (Maceda *et al.* 2022); sin embargo, el análisis de expresión de *CAD5* y *CESA7* usando qRT-PCR reveló que dichos genes están mayormente en tejido parenquimatoso adyacente a la fibra, sugiriendo mecanismos de formación del esclerénquima similar al reportado en xilema (Maceda *et al.*, 2022). Los bajos niveles de expresión para *CAD3/CAD7* hallados en transcriptomas de hoja completa en *A. angustifolia*, y validados posteriormente por qRT-PCRs, fueron explicado en términos de una posible expresión tejido-específica de dichos genes (Huang *et al.*, 2022).

Recientemente, un estudio de transcriptómica comparativa entre diferentes tejidos de *A. angustifolia* reveló que genes tipo *CESA* son expresados preferentemente en hojas jóvenes (Valadez *et al.*, 2023); ello sugiere un metabolismo mas activo relacionado a celulosa en etapas tempranas en el desarrollo de la hoja de agaves.

La evidencia hasta ahora recopilada en diferentes especies de agaves productores de fibra ha dado luz sobre algunos eventos moleculares involucrados en la formación de este tejido. Aproximaciones mediante RNAseq y qRT-PCR en *A. tequilana* y *A. angustifolia* sugieren patrones de expresión tejido-específico que involucraría mecanismos de lignificación cooperativa y semi-cooperativa operando durante la formación de fibras perivasculares. Un análisis a nivel transcriptómico tejido-temporal durante la formación de la fibra pudiera revelar no solo genes biosintéticos desregulados involucrados con los componentes estructurales de la fibra, sino también genes efectores y factores de transcripción controlando la formación, desarrollo y senescencia de la fibra.

### 1.12 Hipótesis

La lignificación cooperativa es un proceso en el cual las células no lignificantes proveen de los precursores necesarios a las células lignificantes para la formación de las fibras; en ella, participan múltiples enzimas, las cuales interactúan de forma sinérgica y coordinada. En *A. fourcroydes*, la formación de fibras perivasculares incluye un mecanismo de lignificación cooperativa o semicooperativa. Usando una estrategia de RNA-seq será posible identificar genes expresados diferencialmente, los cuales participan en dicho proceso de lignificación.

---

## 1.13 Objetivos

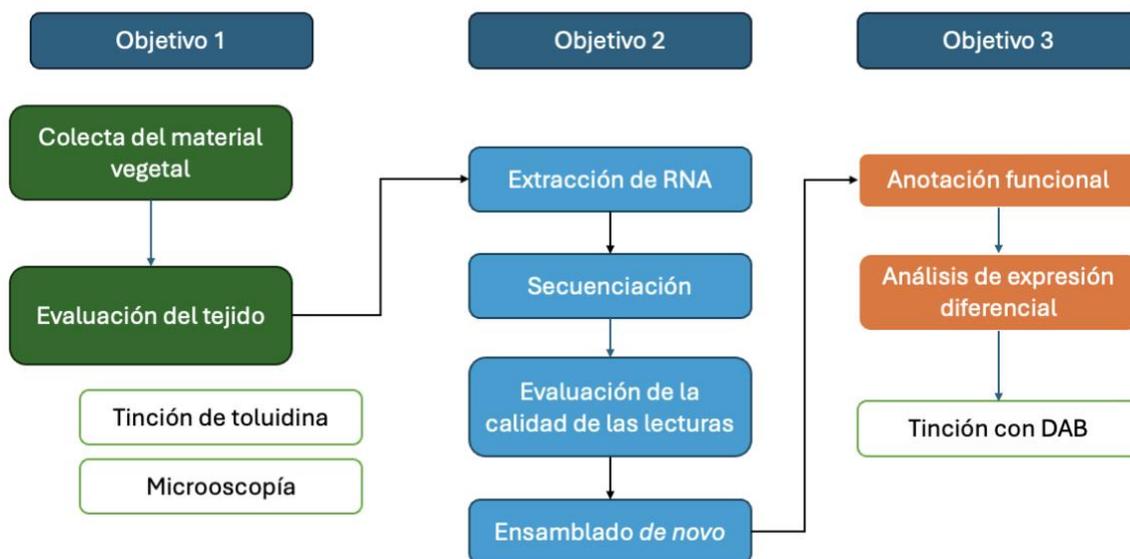
### 1.13.1 Objetivo general

Estudiar la formación de fibras perivasculares en *A. fourcroydes* mediante un enfoque transcriptómico en diferentes etapas de desarrollo.

### 1.13.2. Objetivos específicos

- Determinar mediante análisis histológicos las etapas claves en la formación de las fibras perivasculares.
- Generar un transcriptoma de *novo* a partir de plantas en desarrollo de *A. fourcroydes* que representan el proceso de formación de las fibras perivasculares.
- Definir patrones de expresión transcripcional relacionadas con la biosíntesis de lignina en tejidos adyacentes a las fibras perivasculares en sus diferentes etapas de desarrollo.

## 1.14 Estrategia experimental



**Figura 1.6.** Se muestra la estrategia experimental que se empleó en este trabajo. La caracterización de las fibras y el ensamblado de *de novo* del transcriptoma de *agave fourcroydes*.

## CAPÍTULO II MATERIALES Y MÉTODOS

### 2.1 Determinar mediante análisis histológicos las etapas claves en la formación de las fibras perivasculares

#### 2.1.1 Material Vegetal

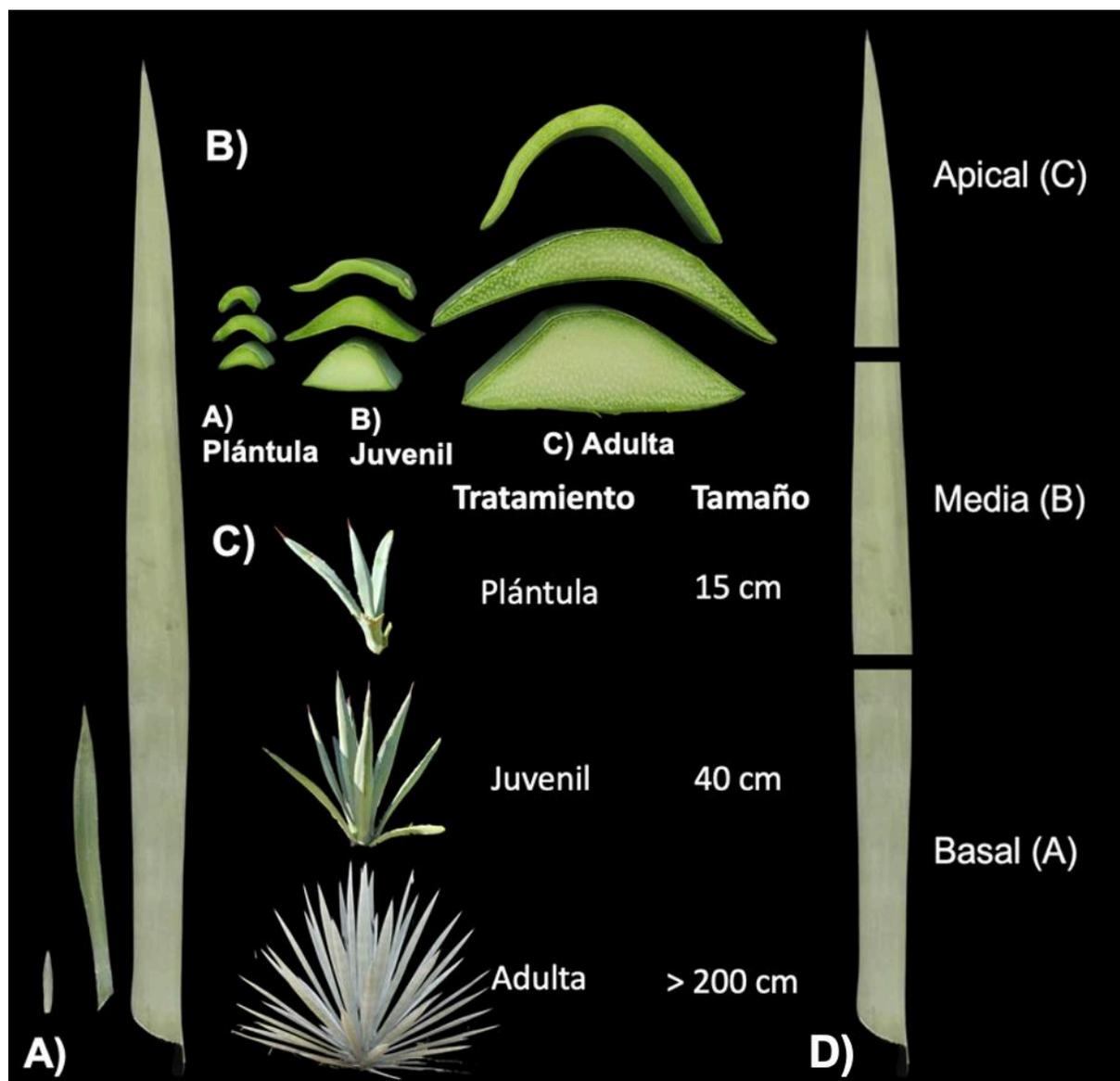
El presente estudio se realizó en la ciudad de Mérida, Yucatán, México (21°01'43"N, 89°33'28"O), donde se colectó el material vegetal empleado para la extracción de ARN. Se seleccionaron individuos de diferentes alturas: 15 cm, 40 cm y 2 m, también llamadas "15 cm", "40 cm" y "2 M", respectivamente. Esta selección de tamaños se hizo basado en resultados previos del grupo de trabajo (Morán *et al.*, 2020) sobre la formación de fibras en *A. fourcroydes*. En la tabla 2.1 podemos se muestran los datos de las plantas muestreadas, en sus tres diferentes edades evaluadas.

Para la extracción del RNA, se recolectó la hoja más joven completamente expandida, como se muestra en la figura 2.1. Las hojas fueron asperjadas con alcohol al 70 % en el lugar de colecta para eliminar contaminantes superficiales. Posteriormente, cada hoja se seccionó en tres zonas: basal (A), media (B) y apical (C) como lo reporta Maceda *et al.*, 2020 para agaves (figura 2.1). Únicamente se recolectó la parte central de la zona A; para el tratamiento de hoja se obtuvieron pequeños cubos de dicha zona y se congelaron inmediatamente en nitrógeno líquido (N<sub>2</sub>) para su posterior almacenamiento a -80°C. Para la obtención de fibras, la extracción se realizó manualmente con pinzas estériles, depositando inmediatamente el material en N<sub>2</sub> y almacenándolo posteriormente a -80°C hasta su procesamiento.

**Tabla 2.1.** Características morfológicas de las plantas muestreadas.

Parámetro morfológico	Muestra		
	15 cm	40 cm	2 m
Largo de la hoja (cm)	11.84 ± 1.06	33.30 ± 2.60	140.72 ± 7.66
Ancho ecuatorial de hoja (cm)	2.23 ± 0.29	4.40 ± 0.36	12.30 ± 0.68

Altura de la planta (cm)	14.43 ± 1.25	39.78 ± 3.66	222.50 ± 13.13
Número total de hojas	5.50 ± 0.85	12.50 ± 2.01	63.50 ± 9.77



**Figura 2.1.** Colecta del material vegetal ) Hojas recolectadas de las tres etapas de desarrollo vegetativo. B) Zona utilizada para la extracción del ARN total. C) Promedio de las plantas que fueron muestreadas. D) Preparación de la muestra para la extracción.

---

### 2.1.2 Análisis de desarrollo y lignificación de fibras

Para analizar el estado de desarrollo de las fibras colectadas para secuenciar sus transcriptomas, se realizó un análisis histológico de las muestras de fibras en sus tres etapas colectadas. Para ello los fragmentos de hoja de *A. fourcroydes* fueron prefijado en FAA (formalina-ácido acético-etanol) durante 24 horas y lavado dos veces con agua destilada durante 20 minutos. Posteriormente, se deshidrató en una serie de gradientes ascendentes de etanol (30 %, 50 %, 70 %, 85 %, 96 % y 100 %), realizando dos lavados cada 2 horas en cada concentración. Después, los tejidos se transfirieron a un medio epoxi que contenía una mezcla [1:1 v/v] de etanol 100 %:resina líquida JB-4 (Embedding kit, Polysciences, Inc.) durante 72 horas, y postertiormente en 100 % resina por una semana. Todos los pasos previos fueron realizados a temperatura ambiente. Después del último paso de infiltración, las muestras se transfirieron a moldes plástico cubriendolas con la solución de incrustación (Solución de infiltración + peróxido de benzoilo). Finalmente, las muestras incluidas en resina polimerizada se seccionaron utilizando un micrótopo rotatorio automático HM340E MICROM, México con un grosor promedio de 5  $\mu\text{m}$  a 12  $\mu\text{m}$ .

Para contrastar los diferentes tejidos en los cortes transversales de hoja, se uso azul de toluidina. Este colorantes es acidofílico y metacromático y se usa para distinguir las paredes celulares primarias de las secundarias; las primeras se tiñen de azul a púrpura, mientras que las secundas se colorean en verde brillante. Para la tinción con azul de toluidina se siguió el protocolo descrito por Bacab (2023) con modificaciones. Los cortes se sumergieron en una solución de azul de toluidina al 0.005% durante 1 minuto y se lavaron con agua destilada para eliminar el excedente de colorante. La toma de micrográficas realizó con un microscopio óptico (Axioplan, Carl Zeiss) acoplado a una cámara digital (AxioCam ICc5, Carl Zeiss). Las imágenes se adquirieron con el software AxioVision SE64 Rel. 4.8 (Carl Zeiss).

Para revisar el estado de lignificación de las fibras en las tres etapas de desarrollo de la hoja secuenciadas, se usó su propiedad de autofluorescencia. Para ello, los cortes histológicos de las 3 edades de hoja fueron observados a una longitud de excitación a 458 nm, usando un microscopio láser confocal (LSM 880, Carl Zeiss) y las imágenes capturadas procesadas usando el software ZEN (Carl Zeiss).

En las estadio “40 cm” donde se evaluó la actividad de lacasas y peroxidasas, el nivel de lignificación también fue revisado en cortes frescos a mano y usando la tinción de fluoroglucinol. Para ello se siguió el protocolo descrito por Jensen (1962) como sigue. Se incubaron cortes

---

---

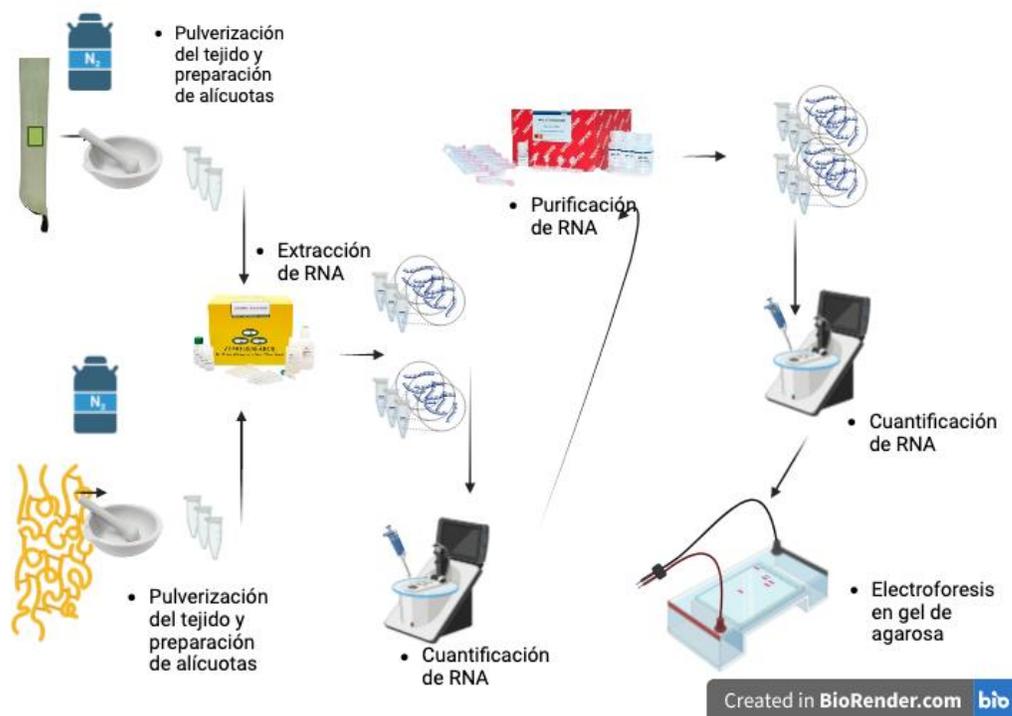
frescos en una solución de floroglucinol al 1 % en etanol 92 % durante 10 minutos, seguido de una inmersión en HCl concentrado al 20%. Posteriormente, los cortes fueron fotografiados en un estereoscopio Carl Zeiss Stemi 2000C.

## **2.2 Generar un transcriptoma de *novo* a partir de plantas en desarrollo de *A. fourcroydes* que representan el proceso de formación de las fibras perivasculares.**

### **2.2.1 Extracción y análisis de calidad de RNA total**

Para la extracción de RNA total, las muestras colectadas (hoja y fibras) se pulverizaron en un mortero con pistilo estériles, manteniendo la temperatura con nitrógeno líquido N<sub>2</sub> para evitar la degradación del RNA. Aproximadamente 100 mg de muestra pulverizada se depositaron en microtubos tipo eppendorf de 2 mL realizando 5 réplicas por muestra. La extracción de RNA total se realizó siguiendo el protocolo recomendado por el fabricante del reactivo QIAzol® (QIAzol Lysis Reagent, Qiagen). El RNA aislado se purificó utilizando el kit RNeasy® Mini Kit (Qiagen), de acuerdo con las instrucciones del fabricante. La pureza del RNA aislado se evaluó mediante la relación de absorbancia A<sub>260</sub>/A<sub>280</sub>, utilizando un espectrofotómetro NanoDrop™ Lite (Thermo Scientific™). En la figura 2.2. se muestra el flujo general de trabajo seguido para la extracción y análisis de los RNAs totales aislados.

La integridad del RNA aislado y purificado se evaluó mediante electroforesis en geles de agarosa al 1.5 % teñidos con bromuro de etidio (0.25 ug/mL de buffer). Los geles se visualizaron utilizando el equipo ChemiDoc™ MP Imaging System (Bio-Rad), con el software Image Lab Versión 5.1 build (Bio-Rad). En los geles se revisó la ausencia de DNA genómico y la integridad del RNA mediante la visualización de dos bandas predominantes de RNA ribosómico 18S y 28S.



**Figura 2.2.** Flujo de trabajo del protocolo de extracción de ARN total.

### 2.2.2 Datos de Secuenciación

Los datos utilizados en este estudio fueron generados como parte de una colaboración entre el Centro de Investigación Científica de Yucatán (CICY) y el Instituto de Ecología A.C. (INECOL). La secuenciación de RNA se realizó en la Unidad de Servicios Genómicos del Laboratorio Nacional de Genómica para la Biodiversidad, ubicado en el Libramiento Norte León, km 9.6, 36821 Irapuato, Guanajuato, México. Para obtener los transcriptomas, se empleó la tecnología de secuenciación de nueva generación (NGS, por sus siglas en inglés) utilizando la plataforma Illumina NexSeq 2x150, la cual genera lecturas pareadas (paired-end reads) de 150 pares de bases (pb). Los identificadores para cada una de las bibliotecas secuenciadas y el número de lecturas obtenidas en cada una de ellas se muestran en la tabla 2.2.

**Tabla 2.2.** Lecturas generadas por biblioteca secuenciada.

Muestra	Réplica	ID	Biblioteca	Lecturas por lado
Hoja 15 cm	1	EI1AF1SS01	Biblioteca 1	19,419,294
Hoja 15 cm	2	EI1AF1SS02	Biblioteca 2	21,707,396
Hoja 15 cm	3	EI1AF1SS03	Biblioteca 3	19,291,473
Fibra 15 cm	1	EI1AF1SS04	Biblioteca 4	17,495,665
Fibra 15 cm	2	EI1AF1SS05	Biblioteca 5	15,813,636
Fibra 15 cm	3	EI1AF1SS06	Biblioteca 6	19,107,631
Hoja 40 cm	1	EI1AF1SS07	Biblioteca 7	20,556,501
Hoja 40 cm	2	EI1AF1SS08	Biblioteca 8	13,695,049
Hoja 40 cm	3	EI1AF1SS09	Biblioteca 9	19,422,282
Fibra 40 cm	1	EI1AF1SS10	Biblioteca 10	14,767,299
Fibra 40 cm	2	EI1AF1SS11	Biblioteca 11	19,360,690
Fibra 40 cm	3	EI1AF1SS12	Biblioteca 12	15,579,003
Hoja 200 cm	1	EI1AF1SS13	Biblioteca 13	30,081,139
Hoja 200 cm	2	EI1AF1SS14	Biblioteca 14	17,917,496
Hoja 200 cm	3	EI1AF1SS15	Biblioteca 15	17,664,788
Fibra 200 cm	1	EI1AF1SS16	Biblioteca 16	19,295,793
Fibra 200 cm	2	EI1AF1SS17	Biblioteca 17	22,994,187
Fibra 200 cm	3	EI1AF1SS18	Biblioteca 18	20,653,773

### 2.2.3 Etapa de procesamiento previo al ensamblado

Para evaluar la calidad de los datos de la secuenciación se utilizó la herramienta FastQC (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>). Con el fin de eliminar bases de baja calidad y

---

secuencias de adaptadores presentes en las lecturas sin procesar, se realizó un recorte utilizando la herramienta Trimmomatic, v.0.39, configurada con los siguientes parámetros: SLIDINGWINDOW: 4:15, MINLEN: 36 y HEADCROP: 13. Todas las lecturas no emparejadas fueron descartadas, tomando únicamente los archivos "paired" para los siguientes pasos. Posterior al recorte con Trimmomatic, se analizó nuevamente la calidad de las lecturas procesadas mediante FastQC, con el objetivo de comprobar que los datos resultantes cumplieran con los estándares de calidad requeridos para los análisis subsecuentes.

#### 2.2.4 Filtrado y Pre-ensamblado de Secuencias

Los archivos generados por Trimmomatic fueron filtrados para eliminar secuencias de baja calidad utilizando un script de Python disponible en GitHub (<https://github.com/Czh3/NGSTools/blob/master/qualityControl.py>). Los parámetros utilizados son los siguientes: calidad mínima de 20 (formato Phred), cobertura del 90% y valor de calidad promedio de 25 (formato Phred). Todas las lecturas que no cumplieran con los parámetros anteriores fueron descartadas; al final de este paso se mantuvieron 48,6894,486 lecturas limpias, es decir, el 71 % de las lecturas sin procesar.

#### 2.2.5 Ensamblado del Transcriptoma

Para el ensamblaje de *novo* del transcriptoma se utilizó la herramienta Trinity versión 2.6.6 (Chang, 2015; Grabherr *et al.*, 2011), ejecutándolo con los parámetros predeterminados de la herramienta. Los contigs ensamblados fueron procesados con la herramienta SeqClean (<https://www.gnu.org/software/sed/manual/sed.html>) para eliminar regiones de baja complejidad, bases indeterminadas y colas de poli A/T. Posteriormente, se empleó la herramienta DeconSeq (<http://deconseq.sourceforge.net/manual.html>) para identificar y eliminar secuencias contaminantes dentro del ensamblado, usando las bases de datos de secuencias contaminantes de bacteria, virus, insects, repbase, fungi1, fungi2 y elementos transponibles comunes de angiospermas (RepBase: <http://www.girinst.org/repbase/>).

Se evaluó la integridad del ensamblado realizado con Trinity con la ayuda de la herramienta BUSCO versión 5.5.0 a través del portal Galaxy [www.galaxyproject.org](http://www.galaxyproject.org). (Simão *et al.*, 2015), con los siguientes parámetros: modo transcriptoma y linaje Liliopsidas.

---

---

## **2.3 Definir patrones de expresión transcripcional relacionadas con la biosíntesis de lignina en tejidos adyacentes a las fibras perivascularles en sus diferentes etapas de desarrollo.**

### **2.3.1 Identificación de Regiones Codificantes**

Para obtener los unigenes representativos dentro del transcriptoma, se llevó a cabo la identificación y corrección de los marcos de lectura abiertos (ORFs, por sus siglas en inglés) utilizando la herramienta AlignWise (Evans & Loose, 2015). Primeramente, se realizó un alineamiento de las secuencias contra genes homólogos de plantas presentes en la base de datos del clúster de cómputo Huitzillín, esta base de datos curada ha sido usada en el ensamblado del transcriptoma de *Psittacanthus schiedeana* (Ibarra *et al.*, 2022). Posteriormente, esta herramienta emplea BLAST (Altschul *et al.*, 1990) para identificar secuencias homólogas, seguido de un alineamiento múltiple con el paquete MUSCLE (Edgar, 2004). Finalmente, AlignWise utiliza GeneWise (Birney *et al.*, 2004) para delimitar las regiones codificantes e introducir las correcciones necesarias en caso de inserciones o deleciones dentro del marco de lectura. Se conservaron las secuencias traducidas a péptidos/proteínas con un tamaño igual o mayor a 105 nucleótidos para su posterior anotación funcional. Posteriormente, se filtró el transcriptoma original, conservando únicamente los contigs que lograron traducirse con AlingWise.

### **2.3.2 Identificación de homólogos potenciales**

Para la identificación de homólogos potenciales de las secuencias peptídicas de *A. fourcroydes* se utilizó la herramienta DIAMOND, consultando y descargando las proteínas pertenecientes a *Populus trichocarpa*, *Zea mays*, *Oryza sativa japonica* y *Setaria viridis*, *Solanum lycopersicum*, *A. thaliana*, *Glycine max*, *Vitis vinifera* y *Dendrobium catenatum*. Los genomas de todas especies están completamente secuenciado y disponible en la base de datos del NCBI. Las proteínas de *A. thaliana* se descargaron de la base de datos TAIR, actualizada hasta enero de 2024. Durante este paso, únicamente se utilizaron las regiones codificantes obtenidas del AlingWise, excluyendo el resto de las secuencias obtenidas por Trinity transcriptoma final.

### **2.3.3 Anotación funcional**

Para la identificación de genes ortólogos se utilizó la traducción de los uniGenes resultantes de la herramienta alingWise de *A. fourcroydes*, para ello se compararon los proteomas de 10 especies de plantas seleccionadas por su cercanía taxonómica y su interés agrícola, las cuales

---

se describen a continuación: *Zea mays*, *Oryza sativa japonica* y *Setaria viridis* (Poales, Poaceae), *Solanum lycopersicum* (Solanales, Solanaceae), *Arabidopsis thaliana* (Brassicales, Brassicaceae), *Glycine max* (Fabales, Fabaceae), *Vitis vinifera* (Vitales, Vitaceae) y *Dendrobium catenatum* (Asparagales, Orchidaceae). Con una sola excepción (*A. thaliana*; <https://www.arabidopsis.org/>), todos los conjuntos de proteínas se descargaron de la última versión disponible en la base de datos de GenBank (<https://www.ncbi.nlm.nih.gov/>).

### 2.3.4 Análisis del perfil de expresión de uniGenes

Para la cuantificación de la expresión génica y los análisis de expresión diferencial se siguió el flujo de trabajo propuesto por la herramienta Trinity (Haas *et al.*, 2013). Brevemente, las lecturas generadas por Trimmomatic fueron mapeadas de forma independiente a el transcriptoma de referencia de *A. fourcroydes* utilizando la herramienta Bowtie2 (Langmead & Salzberg, 2012). Posteriormente, mediante el paquete RSEM (Li & Dewey, 2011), se generó una matriz de perfiles de expresión que contenía los uniGenes y los valores de los recuentos esperados calculados para cada biblioteca. La normalización de los datos y el análisis de la expresión diferencial se realizaron mediante el paquete R DESeq2 (Love *et al.*, 2014). Los datos se dividieron por comparación en 3 grupos: Fibra-15cm vs Hoja-15cm, Fibra-40cm vs Hoja-40cm, Fibra-200cm-Hoja-200cm. Los mapas calor que incluyen genes regulados se crearon utilizando la herramienta webtool iDEP (integrado Differential Expression and Pathway analysis) (<http://bioinformatics.sdstate.edu/idep/>, consultado el 7 de mayo de 2024), comparando la expresión génica entre ambos tiempos utilizando valores de cambio plegable centrados en la media (Ge *et al.*, 2018). Los genes que exhibieron un valor de ajuste de  $p < 0.001$  y un cambio de al menos el doble en la expresión, representado por un Log<sub>2</sub>FC de  $\pm 1$ , fueron identificados como genes expresados diferencialmente (DEG).

### 2.3.5 Detección de actividad de peroxidasas y lacasas

Para la detección de actividad de lacasas, se siguió el protocolo descrito por Hiraite Hideto *et al.* (2016) con algunas modificaciones. Las muestras de plantas juveniles (cortes frescos y secciones incluidas en metilacrilato) se sumergieron en una solución de catalasa (200  $\mu\text{g/mL}$ ) en buffer Tris salino (TBS; pH 7.6) durante 30 minutos para eliminar  $\text{H}_2\text{O}_2$ . Posteriormente, el tejido se sumergió en una solución de tinción (10 mM 3,3'-diaminobencidina (DAB), 100  $\mu\text{g/mL}$  de catalasa en TBS) a 30°C en oscuridad durante toda la noche. Al finalizar esta incubación, se realizaron dos lavados con agua destilada para eliminar la solución con DAB. Para la detección de actividad de peroxidasas se siguió el protocolo descrito por Daudi (2012) con modificaciones.

El tejido se sumergió en una solución de tinción (10 mM DAB en TBS) a 30°C en oscuridad durante toda la noche, se realizaron dos lavados con agua destilada para eliminar la solución teñidora. Algunos tejidos se incubaron únicamente en TBS o se trataron previamente en un autoclave a 2 psi de presión a 105°C durante 20 minutos, utilizándose como controles negativos.

Tras las pruebas de actividad enzimática, los cortes frescos se sometieron a un proceso de eliminación de clorofila. Se sumergieron en una solución de metanol 20 % y HCl 1.24 N a 57°C durante 45 minutos, se decantó y reemplazó por NaOH 7% y etanol 60% durante 15 minutos a temperatura ambiente, Posteriormente, se deshidrató en una serie de gradientes ascendentes de etanol (60%, 80%, 90% y 100%) hasta eliminar completamente la clorofila. Los cortes frescos se observaron en un estereoscopio Carl Zeiss Stemi 2000C, mientras que las secciones en metilacrilato, el análisis se realizó con un microscopio óptico (Axioplan, Carl Zeiss) (Alemania) acoplado a una cámara digital (AxioCam ICc5, Carl Zeiss). Las imágenes se adquirieron con el software AxioVision SE64 Rel. 4.8 (Carl Zei

## **CAPÍTULO III RESULTADOS**

### **3.1 Resultados y discusión**

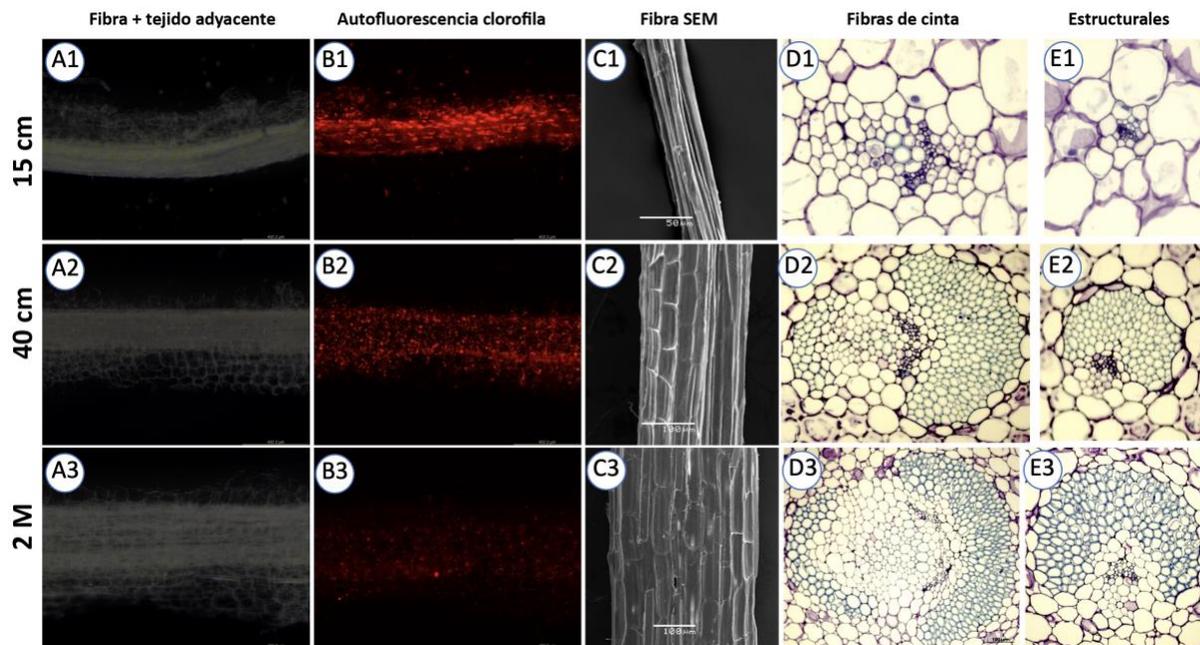
#### **3.1 Determinar mediante análisis histológicos las etapas claves en la formación de las fibras perivasculares**

Se realizó la inspección del tejido adyacente en las fibras, en un estereoscopio Carl Zeiss Stemi 2000c. En la sección A de la figura 3.1 podemos observar una capa celular adherida a la fibra. La capa tiene el grosor necesario para contener el tejido esperado, lo que se cree que está relacionado con la formación de la fibra. Asimismo, se utilizó la lente de autofluorescencia con una longitud de onda de 458 nm. Debido a la longitud de onda y las características del tejido, se propone que los únicos compuestos visualizables son: clorofila y porfirina. A medida que la hoja crece, más fotones son captados por las células parenquimáticas en la epidermis, lo que significa que hay menos fotones capaces de llegar hasta las células parenquimáticas adhesivas a las fibras. Por consiguiente, las células parenquimáticas ubicadas en el centro del mesófilo carecen de actividad fotosintética o se ven reducidas en comparación con las células parenquimáticas de la epidermis.

En la figura 3.1, sección C, se aprecian imágenes de alta resolución generadas por un microscopio electrónico de barrido (JEOL JSM-6360L V Scanning Electron Microscope). Las imágenes mencionadas muestran con detalle la textura de las fibras extraídas en cada estadio de formación, así como el tamaño. Conforme crece, podemos notar diferencias en cada una de ellas. En el estadio que pertenece a plantas adultas, fue posible apreciar los canales de la puntuación de las células esclerenquimáticas en el estadio que corresponde a plantas adultas. Las hojas de agave se encuentran compuestas por fibras de cinta y estructurales; estas presentan diferencias morfológicas notables. Los tejidos de cinta están compuestos por dos capuchones, uno asociado al floema y el otro al xilema, los cuales envuelven los tejidos vasculares, protegiéndolos. Por el contrario, en las fibras estructurales solo es apreciado un capuchón de fibras protegiendo al floema, lo cual se corresponde con el trabajo realizado por Morán Velázquez en 2021. En las secciones D y E, podemos observarlos con detalle; de igual forma, nos ilustran dichas fibras en tres estadios de crecimiento.

El azul de toluidina (ATB) es un colorante metacromático usado para teñir diversos tipos de tejido, incluyendo paredes celulares. Emplea diversos patrones de coloración en los distintos tipos de paredes celulares. Las paredes celulares primarias presentan una tonalidad púrpura debido a su concentración de ácidos pécticos, mientras que las paredes secundarias presentan una tonalidad azul verdosa o azul brillante debido a la presencia de compuestos estructurales. El único inconveniente de este colorante es que no es capaz de teñir ni celulosa ni almidón.

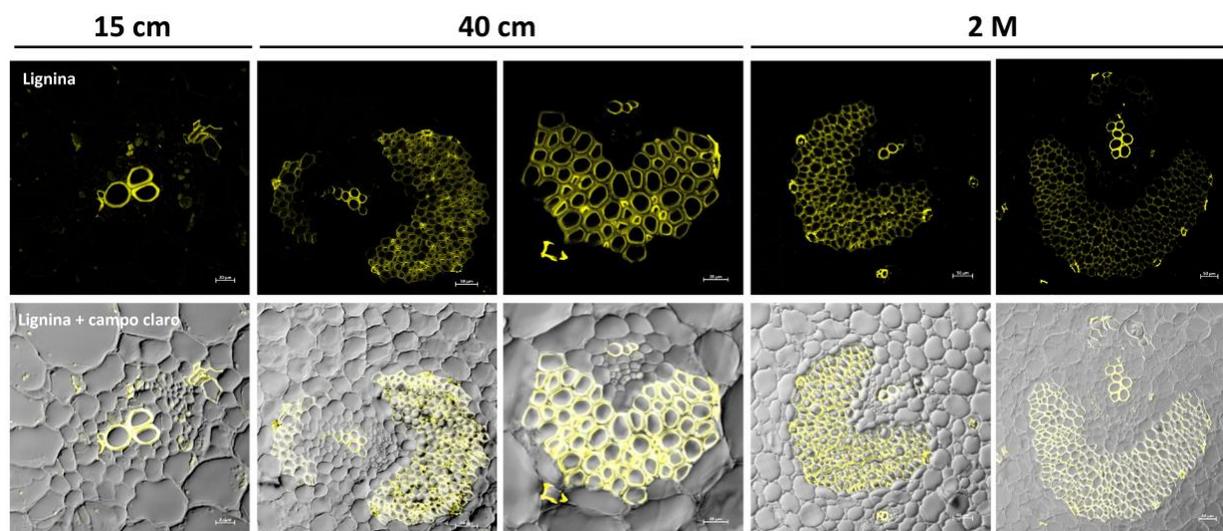
Dentro del estadio “15 cm” se puede apreciar la ausencia de los capuchones. No obstante, son apreciables los tejidos vasculares, solo es apreciable la presencia de la tonalidad azulada que corresponde a la pared secundaria de las células que componen el xilema. En las estructurales se puede apreciar una zona elevada del floema que corresponde al capuchón, quien se encuentra en su fase inicial. En cuanto a las fibras que pertenecen a los estadios “40 cm” y “2 M”, podemos apreciar todas las estructuras que conforman las fibras, previamente mencionadas, sin olvidar el hecho de que las fibras adultas son más grandes y desarrolladas, como se muestra en la figura 3.1



**Figura 3.1.** Cinética temporal del desarrollo de las fibras de *Agave fourcroydes* secuenciadas por RNAseq en este estudio. A) Fibras extraídas. B) Fotografías de autofluorescencia de la clorofila c) Fotografías generadas

por el microscopio electrónico de barrido. D) Teñido de fibras estructurales con azul de toluidina. E) Teñido de fibras estructurales con azul de toluidina.

La superficie del corte fue examinada a través de imágenes obtenidas mediante un microscopio láser confocal (LSM 840, Carl Zeiss). En la figura 3.2 podemos observar la distribución de los depósitos de lignina contenidos en la superficie de las fibras. A pesar de que este método complementa los resultados obtenidos en las pruebas realizadas a través de la tinción de azul de toluidina, este método brinda la distribución de la lignina en la superficie de las fibras y del xilema, de acuerdo con los resultados obtenidos por Morán (2020). A excepción del estadio “15 cm”, se aprecian diversas áreas circundantes destacadas. No obstante, la figura indica que se trata de residuos impregnados en la muestra. Los otros estadios se ajustan con exactitud a los observados previamente.



**Figura 3.2.** Lignificación de fibras de *Agave fourcroydes* en las tres etapas de desarrollo secuenciadas por RNAseq. Arriba autofluorescencia de lignina revelada mediante microscopía láser confocal. Abajo, autofluorescencia de lignina y campo claro..

**3.2 Generar un transcriptoma de *novo* a partir de plantas en desarrollo de *A. fourcroydes* que representan el proceso de formación de las fibras perivasculares.**

Para este trabajo se realizó la extracción de 18 muestras de RNA total de *A. fourcroydes* provenientes de dos tipos diferentes de tejido (células adyacentes a las fibras y fragmentos de hojas de zona basal), en 3 estadios de desarrollo de planta de henequén (15 cm, 40 cm y 2 metros) y usando triplicados biológicos. Las muestras de RNA fueron purificadas y entonces sus valores de concentración y pureza calculados usando espectrofotometría de microvolumenes. Adicionalmente, se les determinó el número de integridad (RIN, por sus siglas en inglés) (tablas 3.1). Este último dato fue proporcionado por el Laboratorio Nacional de Genómica para la Biodiversidad, lugar donde se llevó a cabo la secuenciación de los transcriptomas.

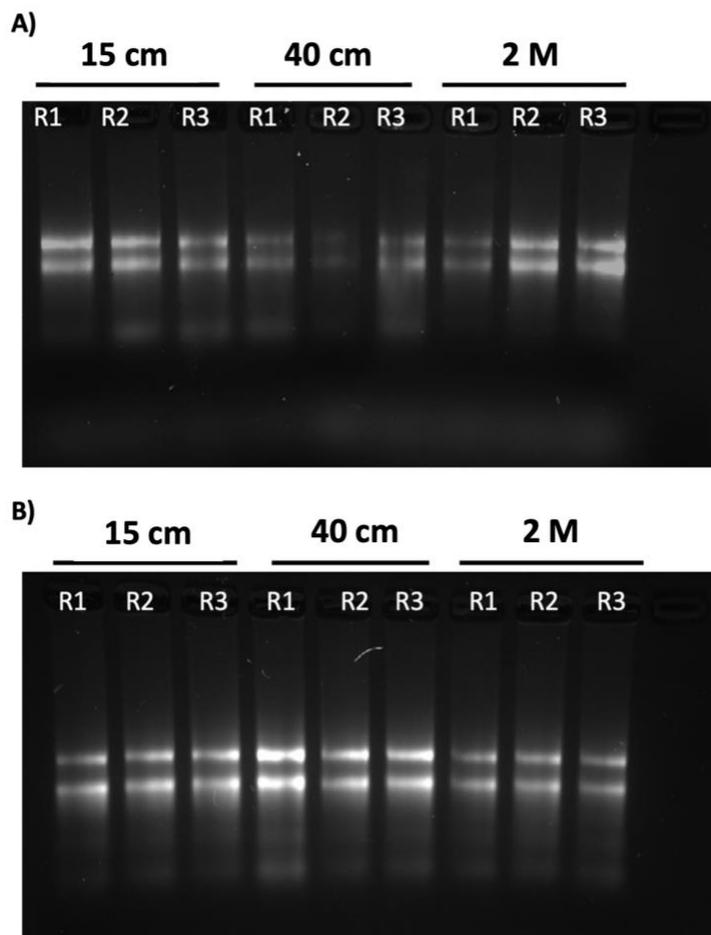
**Tabla 3.1.** Resultados de pureza y concentración de RNA, así como el número de integridad de RNA obtenidos de las muestras de *Agave fourcroydes*.

Tratamiento	Réplica	Biblioteca-Clave	Concentración (ng/ul)	Relación 260/230	Relación 260/280	RI N
Fibra 15 cm	R1	1 - EI1AF1SS04	325	2.05	2.97	7.1
	R2	2 - EI1AF1SS05	310	2.05	2.07	7.8
	R3	3 - EI1AF1SS06	191	1.94	1.97	7.3
Fibra 40 cm	R1	4 - EI1AF1SS10	110	1.82	2.08	7.2
	R2	5 - EI1AF1SS11	125	1.96	2.09	6.6
	R3	6 - EI1AF1SS12	119	1.94	2.08	7.9
Fibra 2 m	R1	7 - EI1AF1SS16	131.1	1.24	1.79	7.9
	R2	8 - EI1AF1SS17	281.6	0.94	1.97	7.7
	R3	9 - EI1AF1SS18	247	0.81	1.92	7.6
Hoja 15 cm	R1	10 - EI1AF1SS01	143.9	0.8	1.84	7.6
	R2	11 - EI1AF1SS02	310	0.75	1.77	7.8
	R3	12 - EI1AF1SS03	312.9	01.05	1.9	7.7
	R1	13 - EI1AF1SS07	388.1	0.79	1.88	8.1

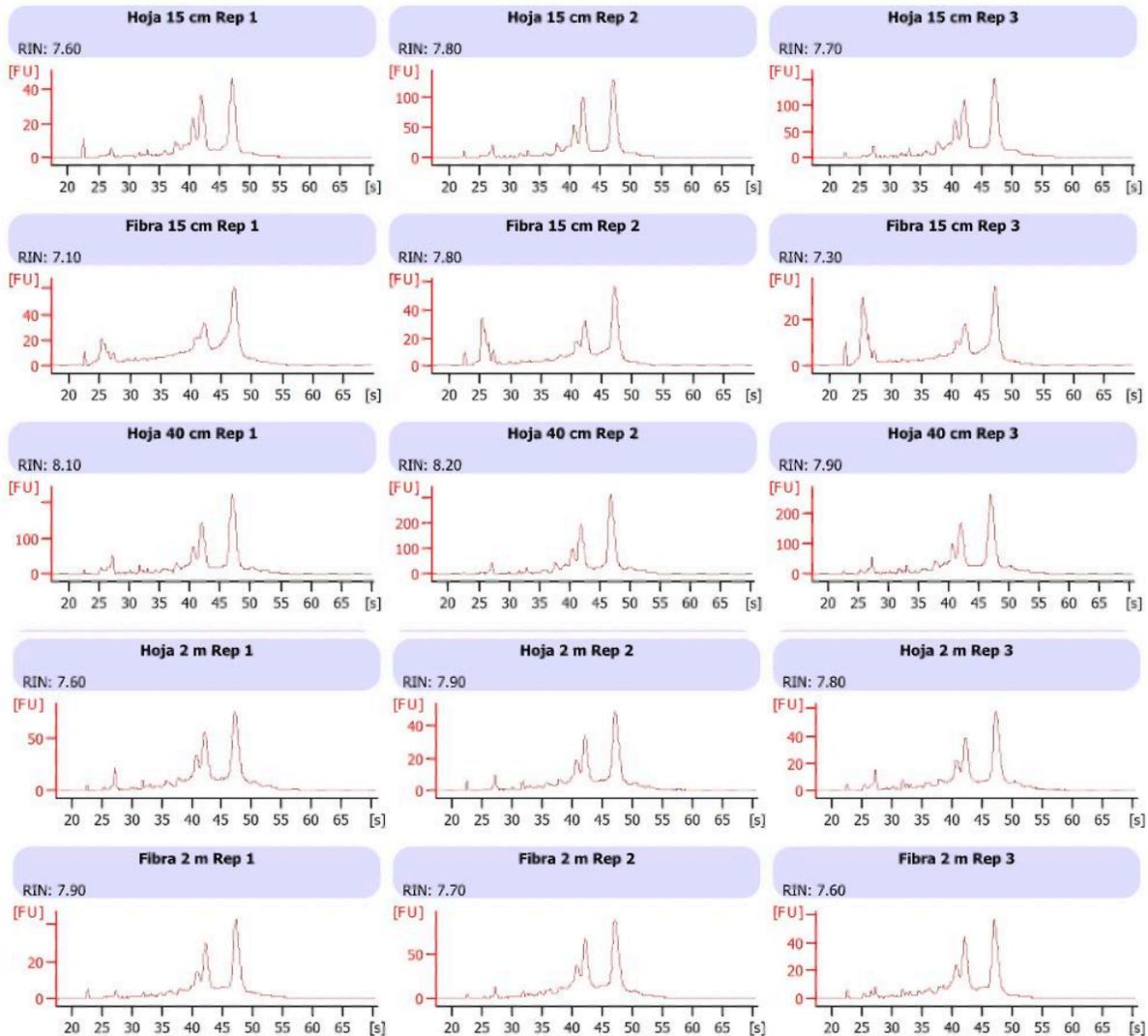
Hoja 40 cm	R2	14 - EI1AF1SS08	665.2	01.06	2	8.2
	R3	15 - EI1AF1SS09	465.5	1.18	1.93	7.9
	R1	16 - EI1AF1SS13	286.2	0.77	1.6	7.6
Hoja 2 m	R2	17 - EI1AF1SS14	140.6	0.77	1.6	7.9
	R3	18 - EI1AF1SS15	265.4	0.52	1.92	7.8

Los valores de pureza mostrados en la tabla 3.1 indican que las muestras de fibra contienen menos impurezas que las de hoja; las de fibras tuvieron un valor promedio de 1.63 para A260/A230 y 2.10 para A260/A280, mientras que los de hoja fueron de 0.85 y 1.82, respectivamente. De acuerdo con estos valores, ambos tipos de muestras que presentan restos de compuestos lignocelulósicos que absorben a 230 nm, lo cual es consistente con el tipo de material vegetal se procesó.

Previo al envío de las muestras al servicio de secuenciación, se determinó la integridad de los RNAs de todas las muestras realizando electroforesis en geles de agarosa al 1.5%. La figura 3.3 muestra los geles realizados, en los cuales se aprecian que las bandas correspondientes a las RNAs ribosomales 28S y 18S están intactas, para ambos tipos de muestras. Tomando en cuenta este resultado, todas las muestras fueron enviadas al servicio de secuenciación. Ahí, el valor de RIN fue calculado, obteniéndose valores muy similares: 7.4 para fibra y 7.8 para hoja, en promedio (Tabla 3.1 y Figura 3.3). Los criterios de la plataforma de microfluidos (Bioanalyzer) indican una concentración mínima de 50 ng/ul y un valor RIN superior a 8, sin embargo los patrones de los electroferogramas del Bioanalyzer, los resultados en geles en agarosa y los valores de RIN muy cercanos a 8, indicaron que era posible iniciar la secuenciación de las muestras.



**Figura 3.3.** Geles de agarosa al 1% teñidos con bromuro de etidio. A) Muestras que corresponden al tejido adyacente a las fibras B) Muestras de tejidos correspondientes las hojas completas de la zona basal.



**Figura 3.4** Análisis de integridad mediante electroforesis de microfluidos (Bioanalyzer). FU: Intensidad de fluorescencia. RIN: Número de integridad. Los picos mas elevados, de derecha a izquierda, corresponden a los RNAs ribosomales 28S y 18S.

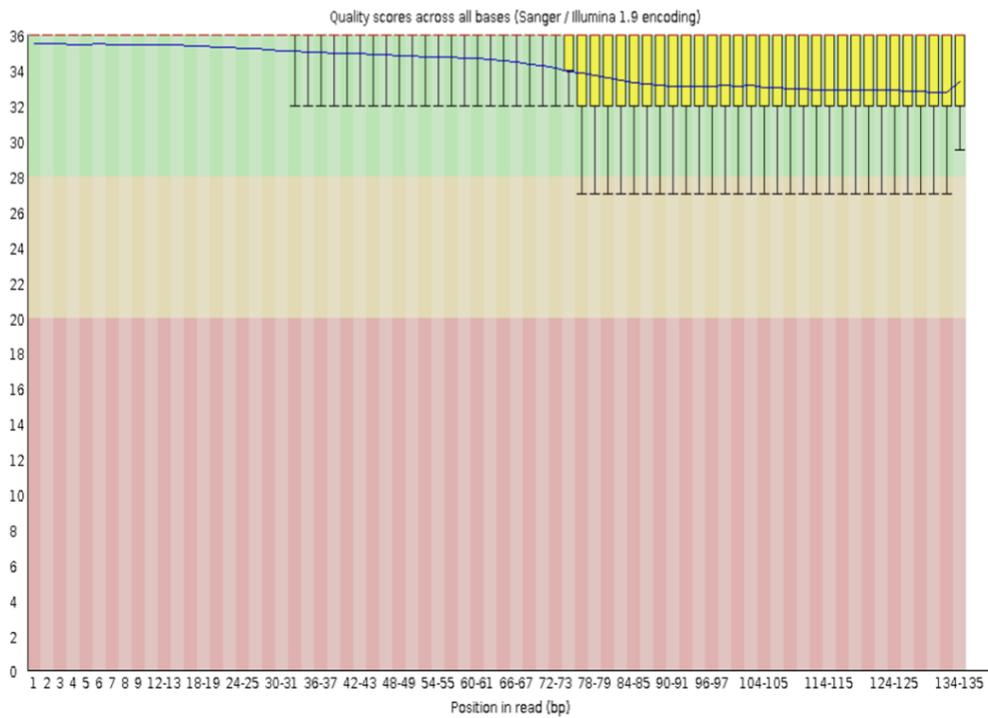
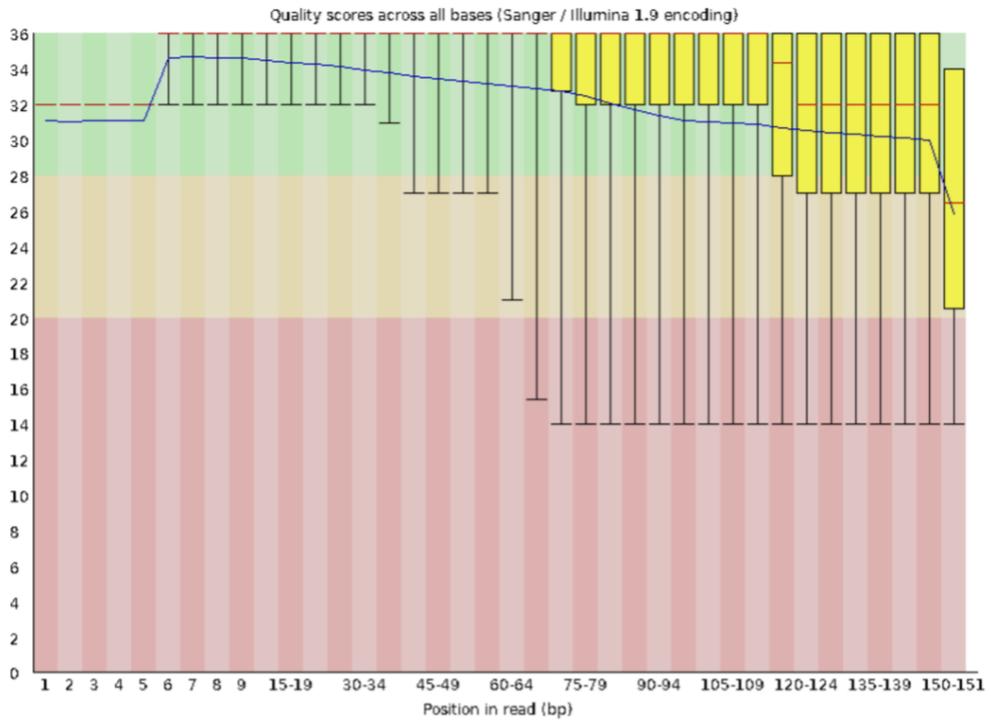
### 3.2.1 Ensamblado *de novo* del transcriptoma de *Agave fourcroydes*.

Para los ensamblados *de novo* se inició evaluando la calidad de las secuencias o lecturas obtenidas en la secuenciación usando FastQC. Este programa es una herramienta muy versátil, con ésta podemos evaluar la calidad de las lecturas, y dentro de sus resultados encontramos los valores de calidad Phred, el cual es un valor de probabilidad de que una base evaluada sea incorrecta. Cuando una base posee una puntuación Q de 30, indica que la probabilidad de error

---

es de una por cada 1000 bases secuenciadas. Por lo general, se considera aceptable un valor igual o superior a Q de 25. En la figura 3.4 se aprecia la fluctuación de las calidades de las lecturas, lo cual genera inconvenientes en el ensamblaje del transcriptoma. Por lo tanto, se decidió utilizar la herramienta Trimmomatic con el fin de procesar las lecturas. Con esta herramienta se recortaron los primeros 10 pares de bases y los últimos pares de las lecturas. Además, se utilizó la función SLIDINGWINDOW:4:20 para eliminar las lecturas con calidad  $Q < 20$ . Finalmente, se aplicó el filtro de *minlen: 75*, conservando únicamente las secuencias con una longitud mayor o igual a 75 nucleótidos.

Trimmomatic genera múltiples archivos de salida. Se procedió a trabajar únicamente con las lecturas que cumplieran con los parámetros previamente pre-establecidos y que mantuvieran su par. En la tabla 3.2 se presentan el número de lecturas originales para cada biblioteca, así como los números finales después de aplicar el proceso de limpieza. Para determinar si el filtrado fue efectivo, se realizó otro análisis con FastQC (figura 5B). Este nuevo reporte del FastQc (figura 5B) indica que se logró eliminar las incongruencias entre las lecturas de baja calidad y los adaptadores que pudieran encontrarse (figura 5A). Desafortunadamente, se eliminó cerca del 30 % de las lecturas originales. Sin embargo, se continuó con el ensamblado *de novo* del transcriptoma, debido a que se determinó que las lecturas eliminadas no significarían un inconveniente, para el ensamblado del transcriptoma se utilizaron dos herramientas, las cuales son MIRA y Trinity.



**Figura 3.5.** Distribución de la calidad de las lecturas para tejido de *A. fourcroydes* generado por la herramienta FastQC. A) Distribución de la calidad de las lecturas generadas por Illumina. B) Distribución de la

calidad de las lecturas filtradas por Trimmomatic. El eje Y indica los puntajes de calidad Phred de las bases secuenciadas, mientras que el eje X representa las pares de base secuenciadas.

**Tabla 3.2.** Lecturas crudas y filtradas con Trimmomatic.

Muestra	Lecturas crudas	Lecturas filtradas	% Recuperación
Biblioteca 1	19419294	11078316	0.57048
Biblioteca 2	21707396	13662220	0.62938
Biblioteca 3	19291473	11278977	0.58466
Biblioteca 4	17495665	20348605	1.16307
Biblioteca 5	15813636	12349494	0.78094
Biblioteca 6	19107631	11865944	0.62101
Biblioteca 7	20556501	12665564	0.61613
Biblioteca 8	13695049	13346082	0.97452
Biblioteca 9	19422282	15402569	0.79304
Biblioteca 10	14767299	12980135	0.87898
Biblioteca 11	19360690	14913162	0.77028
Biblioteca 12	15579003	13214087	0.8482
Biblioteca 13	30081139	12635092	0.42003
Biblioteca 14	17917496	11411710	0.6369
Biblioteca 15	17664788	13659687	0.77327
Biblioteca 16	19295793	15044065	0.77966
Biblioteca 17	22994187	9980146	0.43403

Biblioteca 18	20653773	14611388	0.70744
Total de reads	689646190	486894486	0.70601

Una vez finalizado el ensamblado del transcriptoma, se procedió a obtener las estadísticas (tabla 3.3) con ayuda del comando perl Trinity\_Stats.pl, el cual esta contenido en la plataforma Galaxy. Debido que Trinity genera por defecto una gran cantidad de isoformas, Trinity denomina “gen” a un grupo de transcritos relacionados que comparten parte de su estructura de secuencia.

**Tabla 3.3.** Estadísticas de los transcriptomas ensamblados usando generados por la herramienta Trinity.

Estadísticas generales del transcriptoma ensamblado	
Genes totales	329,067
Transcritos totales	603,560
Porcentaje GC	40.66
Estadísticas basadas en todas las transcripciones	
Contig N10	2,645
Contig N20	1,954
Contig N30	1,536
Contig N40	1,221
Contig N50	962
Longitud media de los contig	438
Promedio de longitud de los contig	678.44
Total de bases ensambladas	409,481,622
Estadísticas basadas en las transcripciones más expresadas	
Contig N10	2,523

Contig N20	1,784
Contig N30	1,336
Contig N40	1,003
Contig N50	757
Longitud media de los contig	363
Promedio de longitud de los contig	574.53
Total de bases ensambladas	189,055,447

Como se mencionó anteriormente, también se realizó el ensamblado mediante la plataforma MIRA, con los mismos datos provenientes del filtrado de Trimmomatic y con los valores preestablecidos. Para comparar ambos tipos ensamblados, se utilizó el script de perl fastx-length.pl para generar las métricas necesarias. En la tabla 3.4 se muestra que el ensamblado por Trinity no solo posee una mayor cantidad de transcritos, sino que también posee, en promedio los transcritos más largos con 15,825 kb, en comparación con el ensamblado de MIRA, que tiene una longitud máxima de 6,615 kb. Incluso, la longitud media de los transcritos en Trinity es superior a aquella hallada con MIRA.

**Tabla 3.4.** Comparación de las estadísticas de los ensamblados usando MIRA vs Trinity.

Característica	Trinity	Mira
Número total de secuencias	603.560	319,188
Longitud total	409,48 Mb	103.23 Mb
Longitud de la secuencia más larga	15,815 kb	6.615 kb
Longitud de la secuencia más corta	201 bp	31 bp
Longitud media	678 bp	323 bp
Longitud mediana	438 bp	236 bp

N10	11.813 secuencias	6,710 secuencias
L10	2,645 kb	1.191 kb
N50	121.585 secuencias	71,530 secuencias
L50	962 bp	414 bp
N90	435.505 secuencias	229,437 secuencias
L90	295 bp	164 bp

Las lecturas filtradas se alinearon usando como referencia al transcriptoma generado por Trinity, para ellos se utilizó la herramienta RSEM. El porcentaje de alineamiento fue de 31.61-64.95% (tabla 3.5), con un promedio de 50.14% para todo el transcriptoma. En promedio, las bibliotecas que correspondían a los tratamientos de hoja obtuvieron los porcentajes de alineamiento más altos con un 53.48%, contrastando a los tratamientos de fibra con un 46.81%, basado en las lecturas del conteo de unigenes contenidas en la matrix de expresión normalizada, generado por la herramienta RSEM.

**Tabla 3.5.** Porcentajes de alineamiento de las bibliotecas secuenciadas.

Biblioteca	Tratamiento/Réplica biológica	Lecturas crudas	Porcentaje
<b>Biblioteca 1</b>	Hoja 15 cm/R1	11,078,316	61.92
<b>Biblioteca 2</b>	Hoja 15 cm/R2	13,662,220	46.2
<b>Biblioteca 3</b>	Hoja 15 cm/R3	11,278,977	64.95
<b>Biblioteca 4</b>	Fibra 15 cm/R1	20,348,605	42.99
<b>Biblioteca 5</b>	Fibra 15 cm/R2	12,349,494	48.97
<b>Biblioteca 6</b>	Fibra 15 cm/R3	11,865,944	39.68

<b>Biblioteca 7</b>	Hoja 40 cm/R1	12,665,564	31.61
<b>Biblioteca 8</b>	Hoja 40 cm/R2	16,346,082	43.14
<b>Biblioteca 9</b>	Hoja 40 cm/R3	15,402,569	61.89
<b>Biblioteca 10</b>	Fibra 40 cm/R1	12,980,135	43.7
<b>Biblioteca 11</b>	Fibra 40 cm/R2	14,913,162	45.05
<b>Biblioteca 12</b>	Fibra 40 cm/R3	13,214,087	42.69
<b>Biblioteca 13</b>	Hoja 2 M/R1	12,635,092	57.88
<b>Biblioteca 14</b>	Hoja 2 M/R2	11,411,710	57.55
<b>Biblioteca 15</b>	Hoja 2 M/R3	13,659,687	56.19
<b>Biblioteca 16</b>	Fibra 2 M/R1	15,044,065	47.76
<b>Biblioteca 17</b>	Fibra 2 M/R2	9,980,146	55.04
<b>Biblioteca 18</b>	Fibra 2 M/R3	14,611,388	55.38

### 3.2.2 Identificación de regiones codificantes

Para analizar la integridad de los transcriptomas ensamblados, se utilizó la herramienta Benchmarking Universal Single-copy Orthologues (BUSCO). Esta herramienta tiene la capacidad de analizar la integridad del ensamblado identificando el número de genes ortólogos conservados dentro de un linaje, el cual está contenido en la base de datos dentro de la plataforma Galaxy. En nuestro caso, se utilizó el linaje Viridiplantae, el cual contempla 425 genes. La herramienta BUSCO considera como “completo” los contigs que coinciden con el tamaño esperado, mientras que los contigs de menor longitud se consideran fragmentados, y los genes que no se lograron identificar dentro del transcriptoma se considera faltantes. Estos valores para ambos tipos de ensamblados los podemos encontrar en la tabla 3.6

**Tabla 3.6.** Resultado de la herramienta BUSCO.

	Trinity	MIRA
Completos BUSCOs	85.90 %	45.2%
Completos y single-copy BUSCOs	14.60 %	30.6 %
Completos y duplicados BUSCOs	71.30 %	14.6 %
Fragmentados BUSCOs	13.60 %	40 %
Faltantes BUSCOs	0.50 %	14.8 %

Al comparar los dos tipos de ensamblados, con los datos de las estadísticas contenidas en la tabla 3.4 y un análisis de la herramienta BUSCO, se sugiere que el ensamblado de Trinity es superior en múltiples sentidos, al contener los contigs de mayor longitud, promedio de longitud de los transcritos y el porcentaje de genes identificados por la herramienta BUSCO.

Una vez seleccionado el mejor ensamblador, se procedió a la limpieza de los ensamblados. Primero, los contigs ensamblados fueron procesados con la herramienta SeqClean con la finalidad de eliminar las bases indeterminadas y colas de poli A/T; con esta acción se logró eliminar 540 secuencias (tabla 3.7). Después, mediante la herramienta DeconSeq, se procedió a la eliminación de las secuencias contaminantes dentro de los ensamblados, que pudieran provenir de otros organismos diferentes a *A. fourcroydes*. Para ello, se utilizaron las bases de datos de secuencias de bacteria, virus, insectos, fungi, elementos repetitivos (replibase) y transponibles comunes de angiospermas. El uso de DeconSeq eliminó exitosamente 518 secuencias (tabla 3.7).

Finalizada la limpieza de los ensamblados, el siguiente paso fue la identificación de las regiones codificantes. Este paso se realizó para poder eliminar la mayor cantidad de isoformas generadas por Trinity; por ello, se consideraron 3 alternativas, TransDecoder y Augustus.

TransDecoder es una herramienta que se encuentra contenida en la plataforma Galaxy, que determina regiones de los transcritos que podrían codificar a proteínas (open reading frames, ORFs). AlingWise, por su parte, se enfoca en la generación de alineaciones múltiples de secuencias de proteínas, las cuales son optimizadas y refinadas. En contraste, Augustus se concibió como una herramienta de predicción de genes a partir del transcriptoma. En la tabla 3.7

se presenta las secuencias generadas en la salida de las diferentes herramientas utilizadas, identificando exitosamente: 257,491, 264,962 y 78,172 usando TransDecoder, AlingWise y Augustus, respectivamente. Estas cifras son similares a la identificada por los stast de Trinity (tabla 3.3). Sin embargo, fue necesario el análisis de la herramienta BUSCO para poder determinar la mejor alternativa, ya que el número de transcriptos con la capacidad de traducir a proteínas no es suficiente para determinar la viabilidad del transcriptoma.

**Tabla 3.7.** Secuencias conservadas en diferentes herramientas.

Herramienta	Secuencias
Ensamblado <i>de novo</i>	603,560
SeqClean	603,006
Deconseq	602,488
AlingWise	264,962
Transdecoder	257,491
AlingWise Cd-hit 91 %	92,612
AlingWise Cd-hit 95 %	120,491
Augustus	78,172

En la tabla 3.8 se muestra que las opciones de AlingWise y Transdecoder mantienen un porcentaje mayor de genes BUSCO identificados, respecto a la cantidad de genes BUSCO completos, obteniendo 83.7% y 78.2% respectivamente, traduciendo efectivamente un total de 264,962 secuencias y 257,491 secuencias, del transcriptoma filtrado de Deconseq. En contraste, la implementación del predictor de genes Augustus logró únicamente identificar un 26.6% de los genes completos BUSCO, y una traducción exitosa de 78,172 secuencias, descartando esta opción indudablemente. Aunque tanto TransDecoder y AlingWise dieron resultados similares, se seleccionó AlingWise considerando el porcentaje mayor de genes BUSCO identificados.

**Tabla 3.8.** Análisis BUSCO de las diferentes herramientas.

	Ensamblado <i>de novo</i>	Deconseq	AlingWise	Transdecoder	Cd-hit 91%	Cd-hit 95%	Augustus
Completos BUSCOs	85.90 %	85.90 %	83.70 %	78.20 %	83.50 %	83.50 %	26.6 %
Completos single-copy BUSCOs	y 14.60 %	14.60 %	20.90 %	30.40 %	66.65 %	60 %	11.3 %
Completos duplicados BUSCOs	y 71.30 %	71.30 %	62.80 %	47.80 %	16.90 %	23.50 %	15.3 %
Fragmentados BUSCOs	13.60 %	13.60 %	15.80 %	20.20 %	16 %	16 %	22.1 %
Faltantes BUSCOs	0.50 %	0.50 %	0.50 %	1.60 %	0.50 %	0.50 %	51.3 %

### 3.3 Definir patrones de expresión transcripcional relacionadas con la biosíntesis de lignina en tejidos adyacentes a las fibras perivascularles en sus diferentes etapas de desarrollo.

#### 3.3.1 Anotación funcional del transcriptoma de *Agave fourcroydes*

Para poder anotar el proteoma de *A. fourcroydes* se empleó la herramienta Diamond, considerando como proteoma el archivo resultante de la salida que contenía las secuencias de aminoácidos de AlingWise. Posteriormente, las secuencias de aminoácidos se alinearon contra los proteomas de las siguientes especies: *Zea mays*; *Oryza sativa japonica*; *Setaria viridis*; *Solanum lycopersicum* L.; *Arabidopsis thaliana*; *Glycine max*; *Vitis vinifera* L. y *Dendrobium catenatum*, tomando como criterios de selección que fueran especies monocotiledóneas y/o genomas bien anotados. Estos proteomas fueron consultados en la base de datos del GenBank y de la base de datos de *Arabidopsis thaliana* (TAIR). En la tabla 3.9 se muestra el número de secuencias de aminoácidos que lograron un hit, en contraste a la cobertura total de secuencias identificadas. La proporción de secuencias anotadas de *A. fourcroydes* fue similar entre en la mayoría de los proteomas consultados, con la excepción de *A. thaliana* con un 57 % (tabla 3.9). Dentro de las especies consideradas para la anotación, encontramos a tres monocotiledóneas

que obtuvieron los valores más bajos de anotación: *O. sativa japonica*, *S. viritus* y *Z. Mays*, con valores de 70% (185,576 secuencias), 68.3% (180,949 secuencias) y 69.7% (184,590 secuencias), respectivamente. Las anotaciones con la mayor cobertura fueron: *D. catenatum* con un 74.7% (198,002 secuencias), seguido por *Vitis vinifera* con 73.7% (195,372 secuencias).

**Tabla 3.9.** Anotación del proteoma de *Agave fourcroydes*.

Especie	Número de secuencias de <i>A. fourcroydes</i> anotadas	Porcentaje de cobertura
<i>Oryza sativa japonica</i>	185,576	70.0%
<i>Arabidopsis thaliana</i>	111,646	57%
<i>Dendrobium catenatum</i>	198,002	74.7%
<i>Setaria viritus</i>	180,949	68.3%
<i>Zea mays</i>	184,590	69.7%
<i>Glosspium hirsutum</i>	193,501	73.0%
<i>Vitis vinifera</i>	195,372	73.7%
<i>Glycine max</i>	191,830	72.4%
<i>Solanum lycopersicum</i>	186,203	70.3%

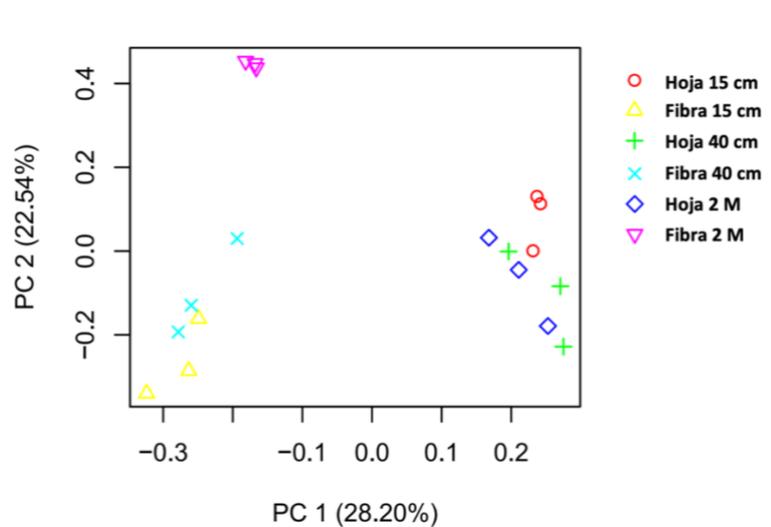
### 3.2.4 Expresión diferencial en tejidos de hoja y tejido adyacente a fibras.

Antes de realizar los comparativos se realizó un análisis de componentes principales (PCA), el cual es una técnica estadística multivariante que se utiliza para reducir la dimensionalidad de un conjunto de datos, identificando las direcciones de máxima varianza de datos analizados.

En la figura 3.6 podemos observar que los datos (transcriptomas secuenciados) conforman tres grandes grupos o clusters. El grupo uno contiene los transcriptomas de Fibra 15 cm y Fibra 40 cm, el grupo 2 a Fibra 2 M, y el grupo 3 abarca todos los transcriptomas de Hoja (Hoja 15 cm, Hoja 40 cm y Hoja 2 M), señalando una alta reproducibilidad de las réplicas biológicas. El primer componente principal (PC 1) captura una ligera variación relacionada con el tipo de tejido, ya que los puntos correspondientes a "Hoja" tienen valores positivos en PC 1, mientras que los puntos

correspondientes a "Fibra" tienen valores negativos. El segundo componente principal (PC 2) podría estar capturando la variación relacionada con la longitud, ya que los puntos correspondientes a 2 M tienden a tener valores más altos en PC 2 que los puntos correspondientes a 15 cm y 40 cm. Una interpretación biológica de este análisis nos indicaría que las muestras de Fibra 15 cm y Fibra 40 cm son transcriptómicamente más similares entre ellas, respecto a Fibra 2 M, lo cual coincide con los tamaños de hoja muestreados. No obstante los transcriptomas de Fibra 2 M conforman un grupo independiente (grupo dos), están más cerca del grupo uno<sup>1</sup>, lo cual es razonable dada la naturaleza de la muestra.

Interesantemente, el cluster tres agrupa a todos los transcriptomas de hoja, mostrando una subagrupación en función del tipo de transcriptoma, acorde con lo esperado. Este resultado sugiere que independientemente de la edad de la hoja, transcriptómicamente hay pocos cambios en las edades muestreadas de hoja completa; este resultado está en línea con el tipo de tejido que conformaba mayoritariamente las muestras de hoja: células de mesófilo o parénquima, con un grado de especialización bajo o nulo.



**Figura 3.6.** Análisis de componentes principales. se muestra la distribución de los datos en el plano definido por la primera componente principal (PC 1) en el eje horizontal, que explica el 28.20% de la varianza total de los datos. El eje vertical representa la segunda componente principal (PC 2), que captura una porción adicional de la varianza.

---

Para los análisis de expresión diferencial entre tipos de muestras y por etapas de desarrollo se consideraron dos tipos de contrastes:

a) Fibra versus hoja completa:

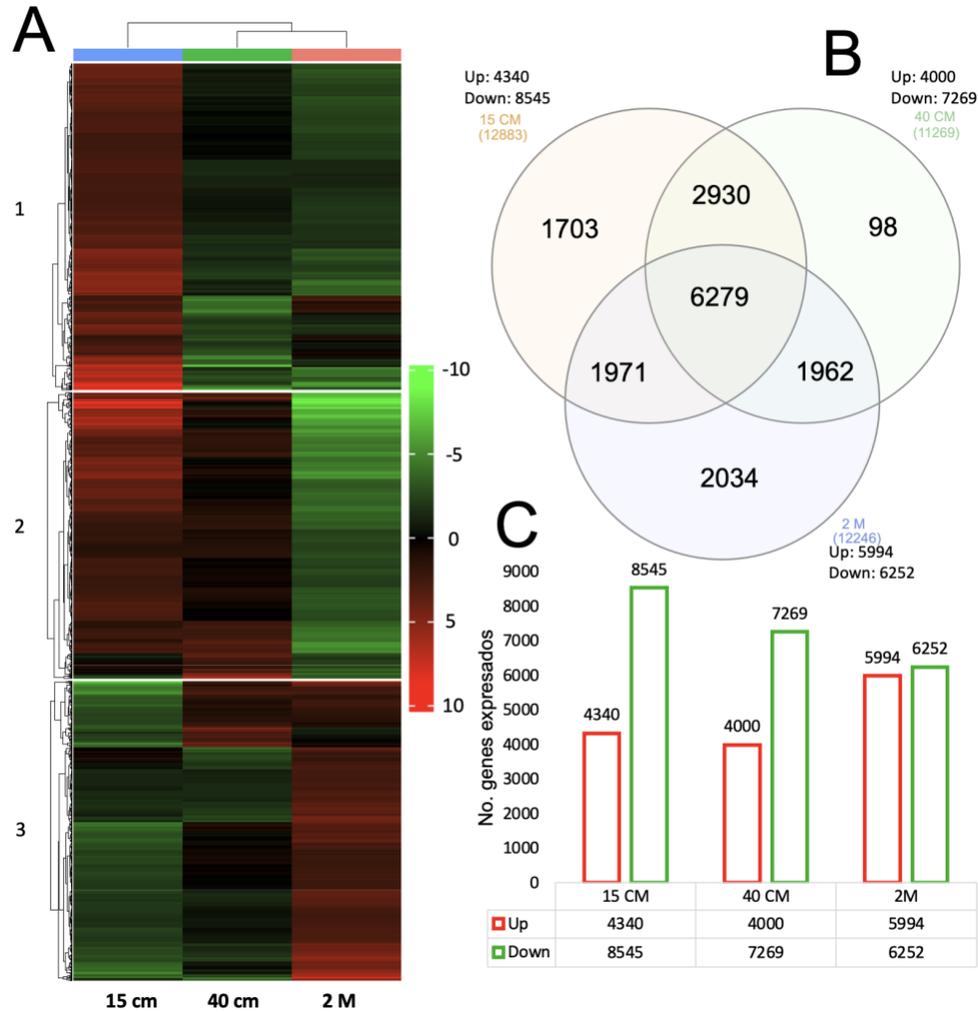
- Fibra 15 cm vs Hoja 15cm (15 cm, de aquí en adelante)
- Fibra 40cm vs Hoja 40cm (40 cm, de aquí en adelante)
- Fibra 2M vs Hoja 2M (2 M, de aquí en adelante)

b) Fibra versus fibra:

- Fibra 2M vs Fibra 40cm (200vs40, de aquí en adelante)
- Fibra 2M vs Fibra 15cm (200vs15, de aquí en adelante)
- Fibra 40 cm vs Fibra 15 cm (40vs15, de aquí en adelante)

### **3.3.2 Análisis de expresión diferencial entre fibra versus hoja completa**

El análisis de expresión diferencial entre fibra versus hoja completa identificó 12,885 DEG (genes expresados diferencialmente, por sus siglas en inglés) en el comparativo 15 cm, con 4,340 y 8,545 genes sobre y sub expresados, respectivamente. Por otro lado, se identificaron 11,269 DEG en la comparación 40 cm, con un total de 4,000 y 7,269 genes regulados a la alta y regulados a la baja. Finalmente, el comparativo 2 M, reveló 6,252 DEG, con un total de 5,994 y 6,252 de genes sobre y sub expresados, respectivamente (figura 3.7). Un diagrama de Venn indica que solo 6,279 genes se compartían entre las tres etapas de fibras analizadas.



**Figura 3.7.** Expresión diferencial para los contrastes de Fibra versus hoja completa. A) Mapa de calor muestra los patrones de niveles de expresión de los DEG encontrados a lo largo de las 3 etapas de desarrollo analizadas. B) Diagrama de Venn de las secuencias putativas identificadas en los análisis de expresión diferencial, mostrando aquellos compartidos y únicos para etapa. C) Número de DEGs identificados en los diferentes contrastes realizados en el análisis de expresión diferencial.

La tabla 3.10 muestra aquellos genes con los niveles de expresión más altos hallados en los contrastes arriba descritos. La tabla incluye genes relacionados con procesos de señalización y transducción de señales, como LAX3 (transporte de auxinas), NOV (histidina quinasa) y APRR2 (regulador de respuesta a dos componentes.) También se hallaron genes involucrados en el plegamiento y degradación de proteínas, como DegP7 (proteasa), un gen codificando una

---

proteína con un dominio RING/U-box (ubiquitinación) y chaperona HSP90. Al menos tres genes (TRINITY\_DN77154\_c0\_g1\_i1, TRINITY\_DN78663\_c0\_g1\_i5 y TRINITY\_DN89404\_c1\_g1\_i1) con altos niveles de expresión a etapas tempranas que descendieron gradualmente pudieran estar involucrados en el metabolismo/remodelación de pared celular, de acuerdo a las funciones reportadas para sus ortólogos en Arabidopsis (De Caroli *et al.*, 2021; Han *et al.*, 2015; Xu *et al.*, 2004).

Hay dos genes relacionados directamente con la lignificación que se hallaron altamente expresados en el tejido adyacente a la fibra: TRINITY\_DN78342\_c1\_g1\_i4, codificando una peroxidasa similar al AtPrx52, y TRINITY\_DN80964\_c0\_g1\_i5 que codifica para una  $\beta$ -glucosidasas, involucrada en lignificación en Arabidopsis. AtPrx52 participa principalmente en la polimerización de subunidades S de lignina que integra a las fibras interfasciculares en Arabidopsis (Ahlawat *et al.*, 2024; Herrero *et al.*, 2014; Fernández *et al.*, 2015). La función de esta peroxidasa en la lignificación también ha sido validada en cañamo (Behr *et al.*, 2018) y en proteomas analizados durante el desarrollo de fibras de algodón (Qin *et al.*, 2017). Es altamente probable que PRX52 de henequén identificada en este trabajo tenga una función conservada en la formación de fibras perivasculares.

La lista de DEG también incluye genes involucrados con regulación genética, como los factores de transcripción (FT) APRR2, SPL1, GRL/NAP1, y un FT de la familia bHLH. El gene APRR2 es un regulador de la pigmentación durante el desarrollo del fruto en pepino, melón, berenjena, y Lagenaria siceraria (Fang *et al.*, 2023 ; Huo *et al.*, 2023; Jiao *et al.*, 2017) y flores como anturios (Jiang *et al.*, 2022); el rol de este FT no ha sido asociado al desarrollo de tejidos de soporte como esclerénquima. En el caso de SPL1, este FT pertenece a una familia muy distribuida en plantas verdes, que controla procesos de desarrollo como el floral en *A. thaliana* (Chao *et al.*, 20179). Se sabe que el FT NAP1 es esencial para la morfogénesis de tricomas de Arabidopsis y kiwi, muy posiblemente a través de un mecanismo que involucra la polimerización de actina (Miao *et al.*, 2023; Li *et al.*, 2004; Deeks *et al.*, 2004). Es interesante el caso del gen (TRINITY\_DN97563\_c4\_g2\_i5) similar a AT1G29950, y que codifica para un FT tipo bHLH. Este gene es altamente expresado en etapas tempranas del desarrollo de la fibra (15 cm); este comportamiento coincide con su ortólogo de Arabidopsis, donde se induce durante ensayos de transdiferenciación in vitro de vasos de xilema (Endo *et al.*, 2015). Es posible que en henequén este FT pudiera tener una función clave en la diferenciación del xilema o bien de células de esclerénquima.

**Tabla 3.10.** Lista de los genes con la expresión diferencialmente más alta de las muestras correspondiente a 15 cm, 40 cm y 2 M.

Código identificador	Identificador <i>Arabidopsis</i> <i>thaliana</i>	Descripción	15 cm	40 cm	2 M
			FC	FC	FC
TRINITY_DN7502 8_c1_g1_i5	AT3G10550.1	MTM1, AtMTM1   Myotubularin-like phosphatases II superfamily	9.320	N/A	N/A
TRINITY_DN7709 6_c4_g1_i3	AT3G14360.1	alpha/beta-Hydrolases superfamily protein	12.472	10.584	N/A
TRINITY_DN7715 4_c0_g1_i1	AT5G66150.1	Glycosyl hydrolase family 38 protein	10.053	5.364	N/A
TRINITY_DN7770 4_c6_g2_i9	AT3G22320.1	ATRPABC24.3, RPB5A, NRPB5, NRPD5   Eukaryotic rpb5 RNA polymerase subunit family protein	9.932	8.308	N/A
TRINITY_DN7834 2_c1_g1_i4	AT5G05340.1	Peroxidase superfamily protein	10.703	N/A	N/A
TRINITY_DN7837 5_c1_g1_i7	AT2G01275.1	RING/FYVE/PHD zinc finger superfamily protein	9.176	N/A	N/A
TRINITY_DN7866 3_c0_g1_i5	AT1G18070.1	Translation elongation factor EF1A/initiation factor IF2gamma family protein	10.568	9.747	N/A
TRINITY_DN7909 4_c4_g1_i6	AT5G40280.1	ERA1, ATFTB, WIG   Prenyltransferase family protein	11.124	10.591	N/A
TRINITY_DN8079 5_c1_g3_i6	AT1G60030.1	ATNAT7, NAT7	10.155	N/A	N/A
TRINITY_DN8081 9_c2_g1_i2	AT1G09560.1	GLP5   germin-like protein 5	8.991	N/A	N/A
TRINITY_DN8096 4_c0_g1_i5	AT1G26560.1	BGLU40   beta glucosidase 40	9.395	6.651	N/A
TRINITY_DN8127 5_c2_g2_i1	AT2G34660.1	ATMRP2, EST4, MRP2, ABCC2, AtABCC2   multidrug resistance-associated protein 2	10.002	5.347	N/A
TRINITY_DN8184 6_c1_g2_i2	AT4G03200.2	catalytics	9.223	8.684	N/A

TRINITY_DN8418 2_c0_g1_i2	AT4G00990.1	Transcription factor jumonji (jmjC) domain- containing protein	9.772	8.848	N/A
TRINITY_DN8510 4_c0_g2_i2	AT1G06550.1	ATP-dependent caseinolytic (Clp) protease/crotonase family protein	9.049	N/A	N/A
TRINITY_DN8778 8_c1_g2_i2	AT5G41580.1	RING/U-box superfamily protein	9.438	N/A	N/A
TRINITY_DN8845 5_c2_g1_i8	AT3G52640.1	Zn-dependent exopeptidases superfamily protein	9.264	5.354	N/A
TRINITY_DN8876 8_c3_g1_i7	AT2G37250.1	ADK, ATPADK1   adenosine kinase	10.172	N/A	N/A
TRINITY_DN8940 4_c1_g1_i1	AT1G63000.1	NRS/ER, UER1   nucleotide-rhamnose synthase/epimerase-reductase	9.329	N/A	N/A
TRINITY_DN8985 9_c0_g1_i4	AT1G17450.2	B-block binding subunit of TFIIC	9.845	4.287	N/A
TRINITY_DN9001 7_c3_g4_i9	AT4G39280.1	phenylalanyl-tRNA synthetase, putative / phenylalanine--tRNA ligase, putative	9.242	N/A	N/A
TRINITY_DN9090 1_c1_g1_i1	AT4G04340.2	ERD (early-responsive to dehydration stress) family protein	10.251	9.618	N/A
TRINITY_DN9101 4_c0_g1_i5	AT1G27680.1	APL2   ADPGLC-PPase large subunit	9.011	N/A	N/A
TRINITY_DN9116 6_c1_g1_i12	AT1G19715.1	Mannose-binding lectin superfamily protein	10.179	N/A	N/A
TRINITY_DN9153 2_c0_g1_i10	AT5G01360.1	TBL3   Plant protein of unknown function (DUF828)	9.251	N/A	N/A
TRINITY_DN9279 9_c0_g1_i5	AT1G06840.1	Leucine-rich repeat protein kinase family protein 3	9.224	N/A	N/A
TRINITY_DN9310 8_c1_g2_i4	AT5G35910.1	Polynucleotidyl transferase, ribonuclease H fold protein with HRDC domain	9.027	N/A	N/A
TRINITY_DN9323 4_c1_g1_i2	AT5G26742.3	emb1138   DEAD box RNA helicase (RH3)	11.039	10.488	N/A
TRINITY_DN9347 6_c0_g1_i7	AT1G77690.1	LAX3   like AUX1 3	9.696	N/A	N/A

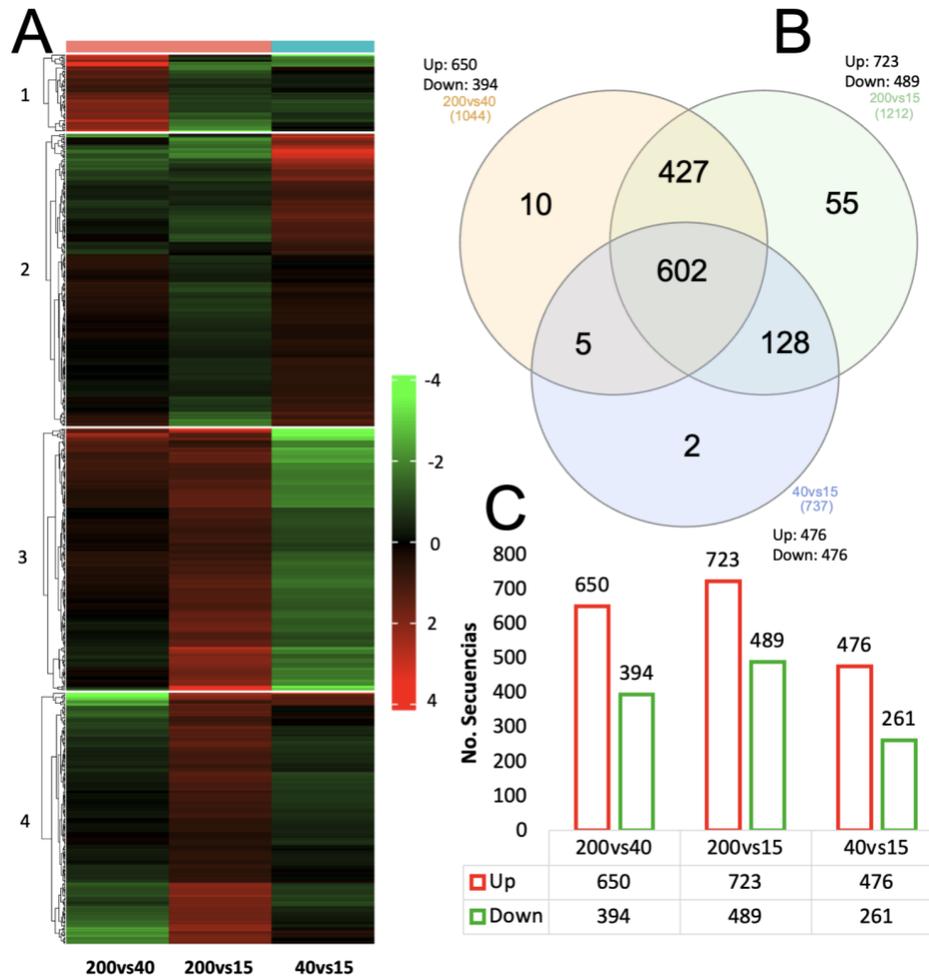
TRINITY_DN9374 4_c1_g1_i2	AT1G18490.1	Protein of unknown function (DUF1637)	9.104	N/A	N/A
TRINITY_DN9378 2_c1_g1_i4	AT5G03940.1	FFC, 54CP, CPSRP54, SRP54CP	10.218	6.902	N/A
TRINITY_DN9398 8_c4_g1_i11	AT3G52990.2	Pyruvate kinase family protein	10.441	N/A	N/A
TRINITY_DN9398 8_c4_g1_i6	AT3G52990.2	Pyruvate kinase family protein	9.502	N/A	N/A
TRINITY_DN9425 4_c2_g1_i6	AT4G18020.1	APRR2, PRR2   CheY-like two-component responsive regulator family protein	9.125	N/A	N/A
TRINITY_DN9431 7_c3_g1_i2	AT4G36960.1	RNA-binding (RRM/RBD/RNP motifs) family protein	9.619	N/A	N/A
TRINITY_DN9487 7_c1_g1_i6	AT5G52920.1	PKP1, PKP-BETA1, PKP2   plastidic pyruvate kinase beta subunit 1	9.172	5.657	N/A
TRINITY_DN9490 1_c2_g1_i1	AT1G30680.1	toprim domain-containing protein	9.212	N/A	N/A
TRINITY_DN9522 9_c3_g3_i2	AT1G68650.1	Uncharacterized protein family (UPF0016)	9.992	6.647	N/A
TRINITY_DN9535 4_c1_g1_i6	AT2G46560.1	transducin family protein / WD-40 repeat family protein	9.292	N/A	N/A
TRINITY_DN9544 2_c2_g2_i3	AT3G03380.1	DegP7   DegP protease 7	10.858	9.981	N/A
TRINITY_DN9564 5_c1_g1_i2	AT2G47070.1	SPL1   squamosa promoter binding protein-like 1	9.514	N/A	N/A
TRINITY_DN9672 6_c2_g1_i7	AT4G26850.1	VTC2   mannose-1-phosphate guanylyltransferase (GDP)s;	9.655	8.770	N/A
TRINITY_DN9672 8_c3_g2_i3	AT1G64110.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein	10.276	6.061	N/A
TRINITY_DN9698 5_c4_g1_i1	AT4G13750.1	NOV   Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase family protein	10.031	4.697	N/A
TRINITY_DN9705 5_c4_g1_i8	AT2G35110.1	GRL, NAP1, NAPP   transcription activators	9.021	5.863	N/A

TRINITY_DN9740 1_c4_g1_i1	AT1G48090.1	calcium-dependent lipid-binding family protein	9.784	N/A	N/A
TRINITY_DN9756 3_c4_g2_i5	AT1G29950.1	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	9.536	N/A	N/A

### 3.2.6 Análisis de expresión diferencial entre fibra versus fibra

El análisis de expresión diferencial contrastando fibra versus fibra identificó un total de 1,044 DEG en el comparativo 200vs40. De éstos, 650 y 394 genes estaban sobre y sub expresados en Fibra 2 M comparado con Fibra 40 cm, respectivamente. Por otro lado, se identificaron 1,212 DEG en la comparación 200vs15; de este total de genes regulados, un total de 723 y 489 genes estaban sobre y sub expresados en Fibra 2 M en comparación con Fibra de 15 cm. Finalmente, en el comparativo 40vs15 se identificaron 737 DEG; de ellos, 476 y 261 estaban regulados a la alta y regulados a la baja en Fibra de 40cm en comparación con Fibra de 15 cm (figura 3.8C).

La tabla 3.11 muestra aquellos genes con los niveles de expresión más altos hallados en los contrastes anteriores de Fibra versus Fibra. Dicha tabla incluye genes involucrados en los diversos procesos biológicos. Por ejemplo, respuesta a estrés (heat shock proteins, proteínas ricas en glicina, factores de transcripción WRKY), metabolismo de la pared celular (celulosa sintasas, xiloglucano endotransglucosilasas), señalización (proteínas quinasas, proteínas fosfatasa), y otros procesos aún no determinados. Algunos de los genes descritos en la tabla 3.11 están relacionados con una respuesta ante el estrés, por ejemplo: las heat shock HSP81-1, HSP90.1 y los factores de transcripción, como es el caso de HSFA6B, CBF1 (respuesta a frío). Mientras que WRKY40 y WRKY30 son conocidos reguladores implicados en respuestas de defensa en las plantas y modificaciones epigenéticas involucradas en las respuestas de estrés ambiental (Javed *et al.*, 2023). Otros genes están relacionados con el metabolismo de la pared celular, como CSLD3 (síntesis de celulosa) (Qiet *et al.*, 2013). En Arabidopsis, AtCSLD3 afecta al proceso de morfogénesis y al crecimiento de la raíz (Hu *et al.*, 2021). En cambio, AtCSLD5 está implicado en la regulación de la síntesis de la pared celular xilano y homogalacturonano (Bernal *et al.*, 2007).



**Figura 3.8.** Expresión diferencial para las contras 200vs40, 200vs15, 40vs15. A) Mapa de calor muestra la correlación entre los perfiles de expresión de los tejidos de *A. fourcroydes*. B) Diagrama de Venn de las secuencias putativas identificadas en la expresión diferencial. C) Número de DEGs identificados en los diferentes contras comparados.

En *Oryza sativa*, OsCSLD4 está implicado en una serie de funciones, que van desde la formación de la pared celular hasta la regulación ante estrés salino y como regulador de la osmosis dentro de la planta, aunque las investigaciones sugieren que aún se desconocen por completo todas sus funciones (Zhao, 2022). Con respecto a XTR6 (remodelación de xiloglucanos), está relacionada con el desarrollo de la pared celular secundaria y la formación de tejidos vasculares, de acuerdo con estudios realizados en *Arabidopsis thaliana* y el híbrido aspen (*Populus tremula*

× *P. tremuloides*) (Bourquin *et al.*, 2002; Perrin, 2001). Finalmente, algunos genes tienen funciones poco conocidas o no anotadas, como los identificados como "unknown protein".

Los DEGs hallados sugieren que se están activando múltiples vías de respuesta a estrés, remodelación de la pared celular y regulación transcripcional, dentro de los que presentaron mayor actividad. Sin embargo, se identificaron 1,445 que cumplen con un fold change mayor de superior a 1 e inferior a -1, con un p adjust, inferior a 0.001. Esta cantidad de DEGs hallados contrastando Fibra versus Fibra, son considerablemente menor a los 20,105 DEGs de los comparativos "fibra y hoja completa", tanto sobre expresados como subexpresados. Una explicación a ello pudiera ser una mayor cantidad de genes compartidos entre las diferentes etapas de la formación de la fibra analizadas, cuyo niveles de expresión pudieran ser muy similares para ser considerados desregulados.

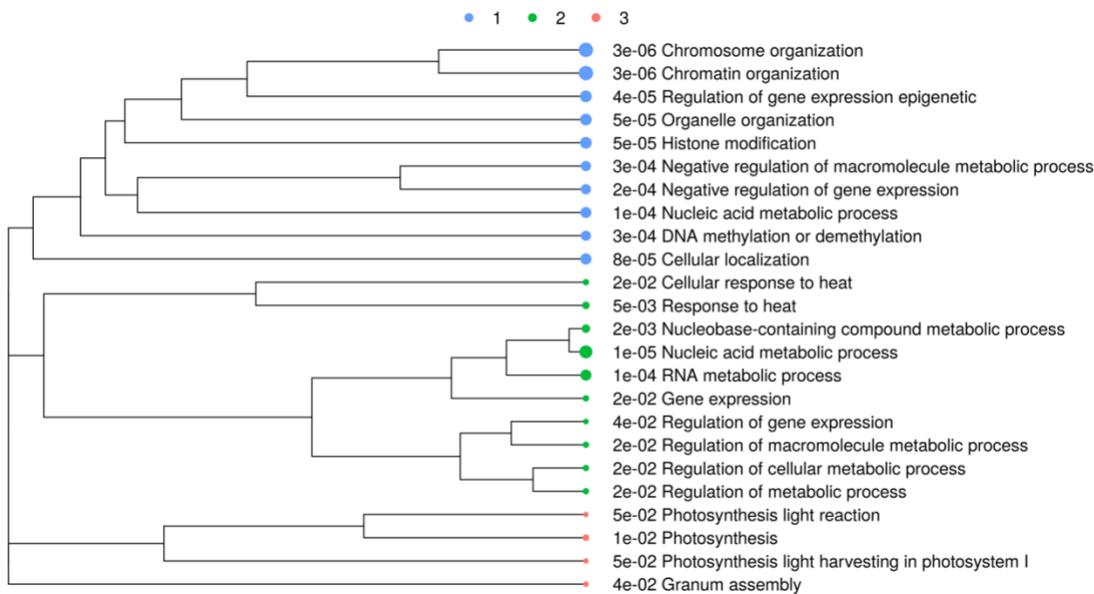
**Tabla 3.11.** Lista de los genes con la expresión diferencialmente más alta de las muestras correspondiente a 200vs40, 200vs15 y 40vs15.

Código identificador	ID Arabidopsis	Descripción	200vs40	200vs15	40vs15
			FC	FC	FC
TRINITY_DN88999_c1_g1_i10	AT4G22740.1	glycine-rich protein	6.523	11.675	N/A
TRINITY_DN87479_c1_g2_i7			1.876	9.025	7.083
TRINITY_DN70226_c0_g2_i3	AT2G30020.1	Protein phosphatase 2C family protein	3.557	8.571	4.955
TRINITY_DN71863_c1_g1_i1	AT4G08950.1	EXO   Phosphate-responsive 1 family protein	2.967	8.403	5.379
TRINITY_DN78286_c0_g2_i1	AT1G80840.1	WRKY40, ATWRKY40   WRKY DNA-binding protein 40	1.115	8.055	6.882
TRINITY_DN96313_c2_g4_i7	AT5G52640.1	HSP81-1, ATHS83, HSP81.1, HSP83, ATHSP90.1, AtHsp90-1, HSP90.1   heat shock protein 90.1	3.547	8.015	4.408
TRINITY_DN70226_c0_g1_i1	AT1G07160.1	Protein phosphatase 2C family protein	4.733	7.950	3.156
TRINITY_DN71940_c1_g2_i1	AT4G25490.1	CBF1, DREB1B, ATCBF1   C-repeat/DRE binding factor 1	1.076	7.841	6.707
TRINITY_DN86476_c3_g3_i1			4.886	7.738	2.794
TRINITY_DN86746_c4_g2_i4	AT5G52640.1	HSP81-1, ATHS83, HSP81.1, HSP83, ATHSP90.1, AtHsp90-1, HSP90.1   heat shock protein 90.1	3.898	7.685	3.713
TRINITY_DN95561_c1_g1_i9	AT2G30020.1	Protein phosphatase 2C family protein	3.724	7.677	3.896
TRINITY_DN97438_c1_g2_i1	AT3G03050.1	CSLD3, KJK, ATCSLD3   cellulose synthase-like D3	1.817	7.584	5.707
TRINITY_DN75869_c1_g2_i1	AT4G29780.1	unknown protein;	2.227	7.561	5.276
TRINITY_DN95185_c2_g2_i1			3.889	7.484	3.536
TRINITY_DN95561_c1_g1_i10	AT2G40180.1	ATHPP2C5, PP2C5   phosphatase 2C5	4.456	7.468	2.956
TRINITY_DN85580_c0_g5_i9	AT5G52640.1	HSP81-1, ATHS83, HSP81.1, HSP83, ATHSP90.1, AtHsp90-1, HSP90.1   heat shock protein 90.1	4.588	7.442	2.806
TRINITY_DN80514_c1_g2_i1	AT4G08950.1	EXO   Phosphate-responsive 1 family protein	1.761	7.394	5.574

TRINITY_DN72004_c0_g5_i2				2.767	7.361	4.535
TRINITY_DN86476_c3_g1_i2				2.759	7.352	4.537
TRINITY_DN97438_c4_g2_i1	AT3G03050.1	CSLD3, KJK, ATCSLD3   cellulose synthase-like D3		3.338	7.342	3.946
TRINITY_DN75869_c1_g1_i3	AT5G12010.1	unknown protein;		2.385	7.311	4.866
TRINITY_DN72004_c0_g5_i1				2.382	7.300	4.861
TRINITY_DN86477_c3_g2_i2				4.255	7.261	2.944
TRINITY_DN85035_c2_g1_i4	AT3G22830.1	AT-HSFA6B, HSFA6B   heat shock transcription factor A6B		2.921	7.259	4.277
TRINITY_DN96274_c3_g3_i1	AT1G17420.1	LOX3   lipoxygenase 3		2.779	7.252	4.409
TRINITY_DN78288_c1_g2_i1	AT3G03050.1	CSLD3, KJK, ATCSLD3   cellulose synthase-like D3		3.431	7.251	3.759
TRINITY_DN73454_c0_g1_i3	AT5G12010.1	unknown protein;		2.419	7.232	4.754
TRINITY_DN87271_c1_g1_i7	AT2G43790.1	ATMPK6, MPK6, MAPK6, ATMAPK6   MAP kinase 6		4.193	7.204	2.956
TRINITY_DN72952_c4_g2_i1				4.569	7.162	2.530
TRINITY_DN87524_c2_g1_i8	AT5G47710.1	Calcium-dependent lipid-binding (CaLB domain) family protein		6.045	7.158	N/A
TRINITY_DN97435_c1_1_g1_i1	AT3G11130.1	Clathrin, heavy chain		4.550	7.135	N/A
TRINITY_DN79095_c4_g2_i2	AT5G24110.1	WRKY30, ATWRKY30   WRKY DNA-binding protein 30		3.264	7.118	3.796
TRINITY_DN69574_c2_g1_i1	AT5G09590.1	MTHSC70-2, HSC70-5   mitochondrial HSO70 2		2.651	7.092	4.380
TRINITY_DN97584_c6_g1_i3	AT5G47430.2	DWNN domain, a CCHC-type zinc finger		5.911	7.024	N/A
TRINITY_DN81748_c0_g1_i1				6.868	7.014	N/A
TRINITY_DN81780_c5_g1_i1	AT3G23640.1	HGL1   heteroglycan glucosidase 1		2.971	6.972	3.949
TRINITY_DN80113_c2_g5_i2	AT4G25810.1	XTR6, XTH23   xyloglucan endotransglycosylase 6		1.005	6.961	5.895
TRINITY_DN73454_c0_g1_i1	AT5G12010.1	unknown protein		2.361	6.957	4.537
TRINITY_DN83998_c2_g1_i3	AT3G51250.1	Senescence/dehydration-associated protein-related		5.839	6.951	N/A
TRINITY_DN73454_c0_g1_i4	AT5G12010.1	unknown protein;		2.090	6.947	4.799
TRINITY_DN76123_c0_g1_i2				3.613	6.931	N/A
TRINITY_DN62143_c0_g1_i1	AT5G06600.3	UBP12   ubiquitin-specific protease 12		4.534	6.903	2.317

### 3.3.4 Análisis de enriquecimiento funcional.

Se realizó un análisis de enriquecimiento de Gen Ontology (GO) en la categoría de Procesos biológicos (figura 3.9) con ayuda de la herramienta webtool iDEP (integrado Differential Expression and Pathway analysis). El análisis de enriquecimiento de genes fue ejecutado en los comparativos 15cm, 40cm y 2M (Fibra versus hoja completa), donde se obtuvo la mayor cantidad de DEGs.



**Figura 3.9.** Términos GO en la categoría de Proceso Biológicos para el transcriptoma a partir de la notación funcional del proteoma de *Arabidopsis thaliana*.

En la figura 3.9 podemos observar como los principales términos GO se encuentran distribuidos en tres grupos o clústers representativos. En el primer grupo encontramos los genes encargados del mantenimiento de la integridad genómica. Esto incluye los genes relacionados con la reparación del ADN, la recombinación homóloga y la desmetilación del ADN. Interesantemente, se hallan enriquecidos los procesos de control del desarrollo, diferenciación celular, reguladores del desarrollo y diferenciación celular, incluso se enriquecieron términos GO enfocados a expresiones epigenéticas y modificaciones de histonas

En el segundo grupo se encuentran enriquecidos los procesos relacionados con la regulación de la expresión genica; procesos metabólicos de compuestos que contienen nucleobases; los términos GO encargados de diversos procesos metabólicos de macromoléculas, y diversos tipos de respuesta celular ante el estrés. Finalmente, en el grupo tres podemos encontrar procesos relacionados con fotosíntesis, específicamente los ligados a la reacción luminosa y la captura de luz.

En los comparativos entre 200vs15, 200vs40 y 40vs15 (tabla 3.12) se presenta una distribución de los términos GO en cuatro grupos, los cuatro grupos se centran en funciones claves dentro

del funcionamiento de la planta. El grupo dos se centra en la respuesta ante diferentes tipos de estrés biótico y abiótico, en particular a estrés y otros tipos de estímulos. Por ejemplo, en el término GO:0006950 respuesta a estrés se identificaron 118 genes involucrados, aunque solo obtuvo un valor de fold enriched de 1.86. Dentro de los términos GO con un porcentaje de enriquecimiento significativo podemos encontrar el GO:0006457 y el GO:0042026 con 6.35 y 15.19 respectivamente. Estos términos se encargan del proceso de plegamiento de proteínas.

Al igual que en el comparativo anterior en el grupo tres, se encuentran los términos GO relacionados en la regulación y control del proceso de fotosíntesis. En el grupo cuatro observamos una gran variedad de términos GO involucrados, desde respuesta a estrés hasta proceso de regulación catabólicos. Por último, encontramos al grupo uno, en el se distribuyen los términos Go involucrados en la formación de la pared celular, con valores de fold enriched superior a 5, en este grupo sobresalen los términos GO involucrados en el metabolismo de pectina, biogénesis y formación de la pared celular. Con respecto al término GO:0009698 phenylpropanoid metabolic process, entran los siguientes genes: FAH1 (AT4G36220), 4CL (AT1G51680), 4CL 3 (AT1G65060), IRX12 (AT2G38080), PAL3 (AT3G53260), LAC 11 (AT5G03260), citocromo p450 (AT2G40890), feruloyl-CoA (AT5G41040).

**Tabla 3.12.** Análisis de enriquecimiento de genes, con referencia a la expresión diferencial realizado Fibra entre Fibra

Grupo	FDR	nGenes	Fold enriched	Ruta metabólica
1	8.76E-09	14	10.850	GO:0005976 polysaccharide metabolic process
1	3.57E-08	15	8.220	GO:0071554 cell wall organization or biogenesis
1	5.61E-08	18	5.837	GO:0005975 carbohydrate metabolic process
1	1.04E-06	8	18.053	GO:0009698 phenylpropanoid metabolic process
1	1.04E-06	8	18.268	GO:0010393 galacturonan metabolic process
1	1.04E-06	8	18.377	GO:0045488 pectin metabolic process
1	2.37E-06	10	10.312	GO:0016051 carbohydrate biosynthetic process
1	3.69E-06	11	8.274	GO:0044262 cellular carbohydrate metabolic process
1	7.42E-06	9	10.656	GO:0044264 cellular polysaccharide metabolic process
1	7.68E-06	8	12.895	GO:0000271 polysaccharide biosynthetic process
2	3.12E-10	118	1.868	GO:0006950 response to stress
2	7.92E-10	25	5.928	GO:0009408 response to heat
2	8.73E-10	78	2.250	GO:0009628 response to abiotic stimulus
2	8.73E-10	98	1.984	GO:0042221 response to chemical
2	5.99E-09	21	6.351	GO:0006457 protein folding
2	1.98E-08	36	3.478	GO:0009266 response to temperature stimulus

2	2.10E-08	63	2.349	GO:1901700 response to oxygen-containing compound
2	8.66E-08	70	2.136	GO:0010033 response to organic substance
2	1.57E-07	51	2.510	GO:0033554 cellular response to stress
2	6.22E-07	9	15.651	GO:0042026 protein refolding
3	8.30E-37	47	13.646	GO:0015979 photosynthesis
3	1.06E-21	43	7.101	GO:0006091 generation of precursor metabolites and energy
3	6.23E-19	25	13.781	GO:0019684 photosynthesis light reaction
3	3.21E-15	12	42.567	GO:0009768 photosynthesis light harvesting in photosystem I
3	6.82E-14	13	28.666	GO:0009765 photosynthesis light harvesting
3	2.39E-10	12	18.828	GO:0006006 glucose metabolic process
3	2.62E-10	12	18.472	GO:0018298 protein-chromophore linkage
3	3.65E-09	16	9.128	GO:0009642 response to light intensity
3	3.77E-09	63	2.368	GO:0009628 response to abiotic stimulus
3	9.08E-09	8	31.081	GO:0019253 reductive pentose-phosphate cycle
4	3.56E-05	15	5.717	GO:0009266 response to temperature stimulus
4	4.87E-05	10	9.356	GO:0009408 response to heat
4	3.88E-03	7	8.352	GO:0006457 protein folding
4	3.88E-03	7	8.312	GO:0009894 regulation of catabolic process
4	3.88E-03	2	247	GO:1904350 regulation of protein catabolic process in the vacuole
4	7.32E-03	2	164.666	GO:0007039 protein catabolic process in the vacuole
4	7.32E-03	5	12.227	GO:0009615 response to virus
4	9.52E-03	10	4.410	GO:0051246 regulation of protein metabolic process
4	1.13E-02	6	7.925	GO:0031329 regulation of cellular catabolic process
4	2.88E-02	9	4.131	GO:0032268 regulation of cellular protein metabolic process

**Tabla 3.13** Genes identificados del término GO:0009498, compartidos en todos los comparativos.

Código identificador	ID <i>A. thaliana</i>	Descripción	40	200	40 vs	15	40	2M
			vs 15	vs 15	200	CM	CM	
			FC	FC	FC	FC	FC	FC
TRINITY_DN67956_c2_g2_i2	AT5G05170.1	CESA3	-2.68	-3.64	-1.02	4.11	-0.34	0.6
TRINITY_DN67956_c2_g2_i3	AT5G05170.1	CESA4	-2.58	-2.74	-0.22	3.12	-0.79	1.22
TRINITY_DN70358_c1_g1_i2	AT4G34100.2		1.49	4.26	2.72	-3.45	-2.36	0.26
TRINITY_DN76681_c1_g2_i1	AT2G38080.1	LAC4	-6.22	-6.11	N/A	2.65	-4.28	-4.93
TRINITY_DN78342_c1_g1_i4	AT5G05340.1	POX	-3.15	-6.9	-3.8	10.7	1.47	N/A
TRINITY_DN80609_c1_g1_i7	AT4G34050.1	CCoAOMT1	-2.93	-4.56	-1.68	4.78	1.82	0.46
TRINITY_DN80689_c0_g1_i3	AT4G36220.1	FAH1	-3.89	-6.17	-2.33	5.29	4.79	1.8
TRINITY_DN81440_c1_g1_i3	AT3G19820.1	DWF1,	0.24	-2.41	-2.71	-0.03	0.31	-2.96
TRINITY_DN82415_c1_g1_i1	AT5G41040.2		-4.22	-7.78	-3.62	5.09	6.02	N/A
TRINITY_DN82415_c1_g1_i2	AT5G41040.2		-3.77	-7.36	-3.64	5.62	5.34	N/A
TRINITY_DN82415_c1_g2_i2	AT5G41040.2		-4.02	-6.36	-2.38	5.33	4.04	N/A
TRINITY_DN82911_c1_g2_i1	AT1G05850.1	POM1,	-1.94	-3.94	-2.06	3.09	1.23	-0.29
TRINITY_DN82911_c1_g2_i3	AT1G05850.1	POM1	-2.19	-2.26	-0.12	1.49	0.25	-0.4

TRINITY_DN82911_c1_g2_i5	AT1G05850.1	POM1,	-2.24	-3.08	-0.89	2.23	-0.05	-0.52
TRINITY_DN83028_c1_g3_i5	AT1G15950.1	CCR1	-0.74	-2.03	-1.34	3.98	1.57	0
TRINITY_DN83028_c1_g4_i1	AT1G15950.1	CCR2	-2.5	-5.83	-3.38	4.43	5.88	N/A
TRINITY_DN85815_c0_g1_i4	AT5G54160.1	ATOMT1, OMT1	-1.9	-2.79	-0.94	5.51	2.88	1.28
TRINITY_DN85815_c0_g2_i3	AT5G54160.1	ATOMT1, OMT2	-1.35	-2	-0.71	3.26	1.94	1.22
TRINITY_DN85815_c0_g4_i1	AT5G54160.1	ATOMT1, OMT3	-2.81	-4.35	-1.59	4.57	2.04	0.64
TRINITY_DN86862_c4_g1_i1	AT3G53260.1	PAL2	-3.33	-3.79	-0.52	2.37	-0.56	-0.52
TRINITY_DN86862_c4_g3_i2	AT3G53260.1	PAL3	-4.35	-5.46	-1.16	4.34	1.11	0.19
TRINITY_DN87072_c2_g3_i2	AT4G39230.1		-2.61	-3.89	-1.35	3.79	2.61	1.35
TRINITY_DN87838_c1_g1_i11	AT5G03260.1	LAC11	-6.02	-5.91	N/A	6.19	-6.03	-5.68
TRINITY_DN91573_c1_g2_i1	AT2G40890.1	CYP98A3	-3.59	-3.48	0.05	3.86	0.9	0.63
TRINITY_DN93823_c1_g1_i6	AT4G36220.1	FAH1	-0.95	-3.88	-2.99	-2.12	-2.43	-4.62
TRINITY_DN94692_c0_g1_i1	AT4G34050.2	CCoAOMT1	-0.58	-2.1	-1.57	4.48	2.3	1.13
TRINITY_DN95248_c1_g1_i1	AT5G05170.1	CESA3	0.16	2.45	2.24	-3.44	-2.84	-0.2
TRINITY_DN95248_c1_g1_i3	AT5G05170.1	CESA4	0.28	2.36	2.03	-3.4	-2.56	-0.21
TRINITY_DN97083_c2_g2_i5	AT5G05170.1	CESA5	0.05	1.91	1.81	-2.93	-2.2	-0.03

En la tabla 3.12, se logró la correcta identificación de ocho genes relacionados en dicho término GO, el cual presenta un valor de fold enriched de 18.05. En este caso, el valor FDR fue de 0.00000104, indicando un nivel de significancia estadística muy bajo. Considerando que existe una probabilidad extremadamente baja de que el enriquecimiento se deba al azar.

Se consultó la base de datos de Términos GO, para poder realizar un filtrado de los genes involucrados en dicho término y así evaluar el comportamiento de la expresión diferencial del mismo.

En la tabla 3.13, se muestra los 29 contings identificados que juegan un rol dentro del término mencionado con anterioridad. Sin embargo, es importante recalcar que un existe un porcentaje significativo de los contings que no se pudieron anotar dentro del transcriptoma. Retornando a la tabla anterior, se identificaron exitosamente múltiples conting anotados con genes involucrados activamente durante el proceso de lignificación de la pared celular secundaria, en esta aproximación podemos encontrar genes como PAL, lacasas, peroxidasas, CCR, OMT y FAH. Cabe mencionar que también se identificaron el resto de los genes involucrados en esta ruta, pero únicamente en los comparativos entre 15cm, 40cm y 2m.

### 3.3.5 Análisis histoquímico de la actividad de lacasa y peroxidasa en fibras

Algunos de los principales contings que se lograron identificar son los correspondientes a genes que codifican potencialmente para lacasas y peroxidasas, las cuales juegan un rol primordial en la polimerización de los monolignoles durante la formación de lignina.

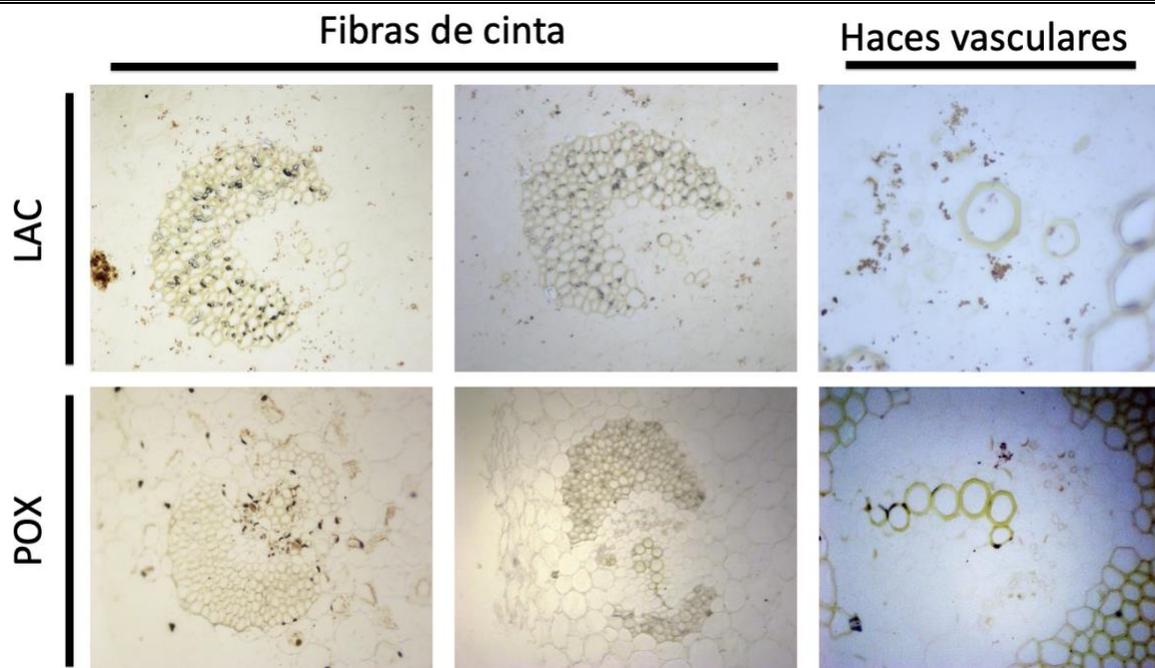
---

Dentro de los análisis de expresión diferencial se logró identificar a *LACASA 4 (LAC4)* y *LACASA 11 (LAC11)* de los contigs TRINITY\_DN76681\_c1\_g2\_i1 y TRINITY\_DN87838\_c1\_g1\_i11, respectivamente. Los dos contigs se comportan de manera similar con niveles de FC comparables en los comparativos 200vs 15 y 40vs15, con la excepción del comparativo 200vs40, en donde no se logró cuantificar. Sin embargo el comparativo 15cm se halló una sobreexpresión génica de dichos genes, mientras el comparativo 40 cm una sub-expresión. Es importante comentar que los niveles de expresión génica fueron aproximadamente del doble en el gen *LAC4* a diferencia de *LAC 11*.

A fin de determinar la concordancia entre los datos de expresión de estos genes relacionados con la formación de lignina y la actividad de sus respectivas enzimas codificadas, se realizaron ensayos histoquímicos en cortes de fibras de 40 cm, donde ya es posible identificar fibras perivasculares.

La peroxidasa (POX) es una enzima que cataliza la oxidación de diversos sustratos en presencia de peróxido de hidrógeno ( $H_2O_2$ ). La tinción con 3,3'-diaminobenzidina (DAB) se basa en la oxidación del DAB por  $H_2O_2$  mediada por la acción de POX; el DAB oxidado adquiere una coloración marrón, es insoluble en agua o alcohol, y que puede ser registrado por microscopía óptica. Durante el proceso de tinción con DAB, este compuesto se transforma en un precipitado marrón insoluble. Dicho residuo marrón se distribuye en los sitios donde se encuentra presente la enzima. Si bien la tinción con DAB puede detectar tanto POX como lacasas (*LAC*), para una visualización precisa de las lacasas es necesario eliminar las peroxidases del tejido. Ésto se logra mediante una incubación previa con catalasa, una enzima que degrada el  $H_2O_2$ .

En la figura 3.10 se ilustra la tinción con DAB. Los ensayos histoquímicos permitieron detectar las actividades de POX y LAC principalmente en los capuchones de fibra y células de xilema. Estos patrones de actividad coinciden con la deposición de lignina, previamente observado por microscopía confocal en los mismos tejidos. Esto resultados también están en línea con los niveles de expresión de genes tipo POXs y LACs hallados por RNAseq. No se observó actividad de POX y LAC en tejido parenquimatoso asociado a las fibras, sugiriendo que la lignificación está ocurriendo de manera autónoma o semi-cooperativa.



**Figura 3.10** Actividad de peroxidasas (POX) y lacasas (LAC) en fibras de cinta y haces vasculares de *Agave fourcroydes*. La detección de la actividad de ambas enzimas se realizó usando el reactivo DAB en cortes transversales de hoja.

---

## CAPITULO IV: DISCUSIÓN GENERAL

### 5.1 Discusión

Los estudios enfocados en la formación de fibras perivasculares en agaves son muy pocos. La mayoría de las investigaciones se enfocan en cultivos comerciales, como son el caso del lino y el algodón, debido a su rápido crecimiento a diferencia de los agaves. Sin embargo, las fibras de estos últimos también se suelen utilizar como refuerzos de estructuras o para la fabricación de utensilios o artesanías por la dureza de sus fibras; también tienen potencial en biopolímeros, y dado su bajo porcentaje de lignina y alto contenido de azúcares en forma de celulosa, las fibras pueden ser materia prima para la producción de bioetanol.

A diferencia de otros componentes de las fibras, la lignina es un compuesto sumamente regulado por la incapacidad de la planta para degradarlo. Este biopolímero se sintetiza a partir de la fenilalanina. Este proceso complejo aún no se comprende completamente en monocotiledóneas, y los principales avances se han logrado en maíz, arroz y algunos pastos. Las rutas biosíntesis de lignina y las redes genéticas que controlan su producción se han estudiado de forma intensiva en plantas como *Arabidopsis thaliana*. (Boerjan *et al.*, 2003). Trabajos como Whetten & Sederoff, (1995) y Vanholme *et al.* (2010) describen con precisión la ruta de biosíntesis de la lignina a partir de la fenilalanina. Sin embargo, a diferencia de la celulosa y la hemicelulosa, la lignina puede cambiar las concentraciones de sus monolignoles canónicos, dependiendo del tipo de fibra, la ubicación dentro de la hoja y la función que sustentara, lo cual dificulta su estudio.

Este trabajo tiene como objetivo el estudio de los procesos involucrados en la formación de las fibras. Centrándose en el proceso de lignificación, dicho proceso es fundamental para la formación de pared secundaria. La lignificación puede presentarse de forma autónoma, semi-autónoma y cooperativa. En diversos estudios realizados por el grupo del Dr. Alatorre en *A. fourcroydes* y *A. tequilana*, identificaron un aumento en la expresión de genes involucrados en la biosíntesis de la pared secundaria. De hecho, los estudios realizados por Morán (2021) y Maceda (2022) señalan que los mecanismos involucrados en la lignificación de las fibras en agaves pueden ser cooperativos o semi-cooperativos. En dichos trabajos se cuantificaron niveles de expresión en genes como *CESA7* y *CAD5*, hasta cinco veces mayores en comparación con la hoja. Sin embargo, para el gen *CCR1*, solo se observó un ligero aumento en fibras en comparación a hoja completas. Estos estudios también reportan la sobreexpresión de *CESA7* y *CAD5* en tejidos como la raíz.

---

Otros grupos de estudios han realizado el ensamblaje de novo de transcriptoma para especies comerciales de agaves. Investigaciones realizadas por Raya *et al.*, (2021) fueron los primeros el ensamblado de transcriptomas de *novo* de especies como son el caso de *Agave sisalana* e híbrido H11648. El estudio anterior se centró en la identificación de genes relacionados con los componentes estructurales de la fibra como son el caso de la celulosa, hemicelulosa y lignina, a diferencia ensamblado de la misma especie, cuyo enfoque se centró en el estudio de fructanos y del proceso de floración de la planta (Simpson *et al.*, 2011).

Los análisis histológicos identificaron las etapas de desarrollo relevantes para la elaboración de la tesis. La utilización de toluidina posibilitó la identificación de las paredes celulares, así como las estructuras fundamentales que conforman las fibras perivasculares. Este análisis se fundamentó en el trabajo realizado por Morán (2021), y fue la base para determinar dichas etapas. Se consideró el primer estadio correspondiente a "15 cm", ya que las fibras todavía no presentan los capuchones asociados a las haces vasculares. No obstante, si es posible identificar las células asociadas al xilema, las cuales son las primeras estructuras que desarrollan una pared secundaria, en contraste con el capuchón de microfibrillas, dicho comportamiento también es descrito por Raya *et al.* (2021). El segundo estadio llamado "40 cm" muestra cómo se desarrollan todas las estructuras, pero al compararlo con el segundo estadio "2 M", podemos ver que todavía están en proceso de desarrollo. Es interesante apreciar que en *A. tequilana* también es apreciable dichas estructuras, como lo muestra el trabajo de Corbin, *et al.*, (2015), sin embargo, no es posible apreciar las fibras de cinta. También se logró identificar las zonas enriquecidas con depósitos de lignina (figura 3.2) mediante microscopía confocal, los cuales concuerdan con las tinciones histoquímicas de pared primaria y secundaria hechas con azul de toluidina.

El presente proyecto ha logrado el ensamblaje de 264,962 transcritos, un número significativamente superior al compararlo con trabajos llevados a cabo con especies similares, como los tres transcriptoma reportados por Raya *et al.* (2021) consta en un rango de 136,692 a 170,472 transcritos para cada uno. Mientras que Grosset *et al.*, (2013) han reportado que su transcriptoma consta de 16,331 y 16,214 transcritos para las especies *Agave tequilana* y *Agave deserti*, respectivamente. Asimismo, Andrade ha reportado en el año 2023 el ensamblaje del transcriptoma de *A. angustifolia*, con un total de 270,163 transcritos. Enfocándonos en el mapeo de las lecturas al transcriptoma, se logró el alineamiento de aproximadamente el 50 % de las lecturas. En estudios comparables, el mapeo de lecturas suele oscilar entre el 65 % y el 69 % (Andrade *et al.*, 2023). Sin embargo, Maceda (2021) ha reportado valores superiores al 90 % en términos de cobertura. Estos resultados sugieren que hay margen de mejora en nuestro

---

proceso de alineamiento, posiblemente mediante la optimización de los parámetros de mapeo o la mejora de la calidad del transcriptoma de referencia.

El trabajo también contó con tinciones para detección de POX y LAC. Los datos transcriptómicos revelaron por primera vez altos niveles de expresión de genes como *LAC4* y *PRX52*, que están relacionados con la lignificación de la pared celular secundaria, de acuerdo a lo hallado en otros modelos experimentales. Nuestros análisis de actividad *in situ* de *POX* y *LAC* soportan los datos hallados por RNAseq; se requiere validar si esta actividad corresponde a enzimas codificadas por *PRX52* y *LAC4*, respectivamente. Ambos tipos de enzimas son vitales para la lignificación. Jeong *et al.* (2022) señalan que al reducirse la actividad de estas proteínas decrece y altera la composición de la lignina en *A. thaliana*. Por ejemplo, *PRX35* y *PRX73* afectan de manera negativa el proceso de elongación, crecimiento y formación de las estructuras vasculares. De igual forma, Somssich, (2020) señala la importancia de *LAC4* y *PRX52* dentro del proceso de lignificación en *Arabidopsis*.

Uno de los aportes de este trabajo fue la identificación de genes regulados a la alta asociados a lignificación, muchos de ellos reportados por primera vez para agaves; por ejemplo, genes como *CESA3*, *CESA4*, *CESA5*, *LAC4*, *LAC11*, *PRX52*, *CCoaOMT*, *FAH1*, *DWF1*, *POM*, *CCR1* y *CCR2*. Entre ellos destaca *LAC4*, esta lacasa se expresa en haces y fibras vasculares; contribuye a la biosíntesis de lignina y, por tanto, a la biosíntesis de la pared celular. En *Arabidopsis*, las mutantes *lac4/irx12* tienen un fenotipo de xilema levemente irregular evidenciando el papel central de este gen en la polimerización de subunidades de lignina (Liu, 2015.). También se halló por primera vez en agaves a *PRX52*, esta proteína con secuencia similar a las peroxidasa, participa en la biosíntesis de lignina. Este gen se puede considerar un marcador para polimerización de subunidades de lignina, casi exclusivo de fibras, como se ha reportado en *Arabidopsis* donde su mutación genera un desarrollo anormal de las fibras del xilema y niveles reducidos de lignina (Jang *et al.*, 2022).

El enfoque experimental usado en este trabajo fue novedoso; por primera vez se secuenciaron transcriptomas tejido-específico asociados a fibras en agave. Los resultados mostrados aquí indican que su aplicación fue exitosa. Sin embargo, resta la validación de perfiles de expresión por otras técnicas como qRT-PCR y la caracterización funcional de genes de agave henequén, potencialmente controlando la formación de fibras duras.

## CAPITULO V: CONCLUSIONES Y PERSPECTIVAS

### 5.1 Conclusiones

- Los análisis histológicos y de microscopía permitieron identificar tres etapas de desarrollo de fibras perivasculares en *A. fourcroydes*. Estas etapas de desarrollo estuvieron en línea con los niveles de lignificación determinados por autofluorescencia.
- Aplicando un enfoque de RNAseq de fibras y sus células parenquimatosas asociadas en *A. fourcroydes* fue posible una identificación de múltiples genes involucrados en la remodelación de pared (*XTR6*) y en la síntesis de monolignoles como *CCR*, *PAL*, *FAH*, *CESA*, *CCoaOMT*, *POMT* y *CYP98A3*.
- Fue posible identificar por primera vez en agave genes como *LAC4*, *LAC11* y *PRX52*, los cuales han sido previamente reportados tener un papel central en la polimerización de subunidades de lignina.
- Los análisis de actividad *in situ* de *LAX* y *POX* muestran que dichas enzimas están siendo activas principalmente en fibras y células de xilema. Se requiere determinar *in situ* los patrones de expresión de enzimas involucradas en la biosíntesis de subunidades de lignina para determinar si los mecanismos de lignificación de fibras perivasculares en agave son cooperativos o autónomos.

---

**5.2 Perspectivas**

- Determinar la composición química de las fibras en las diferentes etapas de desarrollo analizadas, específicamente celulosa, hemicelulosa, lignina y ceras.
- Efectuar un análisis de ortología para los DEGs hallados en este estudio para determinar sus relaciones evolutivas respecto a otras monocotiledóneas.
- Seleccionar un grupo de genes de interés (remodelación de pared, biosíntesis de lignina, POXs y LACs, FTs) para validar su expresión mediante ensayos de RT-PCR e hibridaciones de RNA *in situ*.
- Validar la función *in planta* de algunos genes involucrados en la formación de fibras perivasculares mediante ensayos de expresión heteróloga en plantas modelo como *A. thaliana*.
- Determinar la actividad *in situ* de POXs y LACs en cortes de las etapas de desarrollo analizadas en este estudio.

**BIBLIOGRAFÍA**

- Andrade Marcial. (2023). Estudio transcriptómico y proteómico de la biogénesis del cloroplasto en el modelo albino de *Agave angustifolia* Haw. Tesis de Doctorado. Centro de Investigación Científica de Yucatán, A.C.,
- Ahlawat, Y. K., Biswal, A. K., Harun, S., Harman-Ware, A. E., Doeppke, C., Sharma, N., ... & Hankoua, B. B. (2024). Heterologous expression of *Arabidopsis* laccase2, laccase4 and peroxidase52 driven under developing xylem specific promoter DX15 improves saccharification in populus. *Biotechnology for biofuels and bioproducts*, 17(1), 5.
- Alejandro, S., Lee, Y., Tohge, T., Sudre, D., Osorio, S., Park, J., ... & Martinoia, E. (2012). AtABCG29 is a monolignol transporter involved in lignin biosynthesis. *Current Biology*, 22(13), 1207-1212.
- Altschul, Stephen F., et al. "Basic local alignment search tool." *Journal of molecular biology* 215.3 (1990): 403-410.
- Adler, E. (1977). Lignin chemistry—past, present and future. *Wood science and technology*, 11(3), 169-218.
- Bacab-Caamal, D. V., & HERNANDEZ, G. D. C. G. (2023). Caracterización de semillas de palma jipi (*Carludovica palmata* Ruiz & Pavón)
- Behr, M., Lutts, S., Hausman, J. F., & Guerriero, G. (2018). Expression analysis of cell wall-related genes in *Cannabis sativa*: the “ins and outs” of hemp stem tissue development. *Fibers*, 6(2), 27.
- Barros, J., Serk, H., Granlund, I., & Pesquet, E. (2015). The cell biology of lignification in higher plants. *Annals of botany*, 115(7), 1053-1074.
- Bernal, A. J., Jensen, J. K., Harholt, J., Sørensen, S., Moller, I., Blaukopf, C., ... & Willats, W. G. (2007). Disruption of ATCSLD5 results in reduced growth, reduced xylan and homogalacturonan synthase activity and altered xylan occurrence in *Arabidopsis*. *The Plant Journal*, 52(5), 791-802.

## BIBLIOGRAFÍA

- Balam-Cocom, R. J., Duarte-Aranda, S., & Canché-Escamilla, G. (2006). Obtención y caracterización de materiales compuestos de fibras de la "piña" de henequén y polipropileno. *Revista Mexicana de ingeniería química*, 5(1), 39-44.
- Birney, E., Clamp, M., & Durbin, R. (2004). GeneWise and genomewise. *Genome research*, 14(5), 988-995.
- Boerjan, W., Ralph, J., & Baucher, M. (2003). Lignin biosynthesis. *Annual review of plant biology*, 54(1), 519-546.
- Bourquin, V., Nishikubo, N., Abe, H., Brumer, H., Denman, S., Eklund, M., ... & Mellerowicz, E. J. (2002). Xyloglucan endotransglycosylases have a function during the formation of secondary cell walls of vascular tissues. *The Plant Cell*, 14(12), 3073-3088
- Chokshi, S., Parmar, V., Gohil, P., & Chaudhary, V. (2022). Chemical composition and mechanical properties of natural fibers. *Journal of Natural Fibers*, 19(10), 3942-3953.
- Chao, L. M., Liu, Y. Q., Chen, D. Y., Xue, X. Y., Mao, Y. B., & Chen, X. Y. (2017). Arabidopsis transcription factors SPL1 and SPL12 confer plant thermotolerance at reproductive stage. *Molecular plant*, 10(5), 735-748.
- Chang, Z., Li, G., Liu, J., Zhang, Y., Ashby, C., Liu, D., ... & Huang, X. (2015). Bridger: a new framework for de novo transcriptome assembly using RNA-seq data. *Genome biology*, 16, 1-10.
- Corbin, K. R., Byrt, C. S., Bauer, S., DeBolt, S., Chambers, D., Holtum, J. A., ... & Burton, R. A. (2015). Prospecting for energy-rich renewable raw materials: Agave leaf case study. *PLoS One*, 10(8), e0135382.
- Chávez-Sifontes, M., & Domine, M. E. (2013). Lignina, estructura y aplicaciones: métodos de despolimerización para la obtención de derivados aromáticos de interés industrial. *Avances en ciencias e Ingeniería*, 4(4), 15-46.
- Castillo, A., Robert, M., & Larqueé, A. (2010). CICY, Treinta años de labor científica y educativa. Mérida: Centro de Investigación Científica de Yucatán, AC.
- Chernova, T. E., & Gorshkova, T. A. (2007). Biogenesis of plant fibers. *Russian Journal of Developmental Biology*, 38, 221-232.

- Cazaurang-Martinez, M. N., Herrera-Franco, P. J., Gonzalez-Chi, P. I., & Aguilar-Vega, M. (1991). Physical and mechanical properties of henequen fibers. *Journal of applied polymer science*, 43(4), 749-756.
- Data México (2023), Otras Fibras Textiles Vegetales <https://www.economia.gob.mx/datamexico/es/profile/product/other-vegetable-textile-fibers> [Acceso 10 marzo 2024].
- De Caroli, M., Manno, E., Piro, G., & Lenucci, M. S. (2021). Ride to cell wall: Arabidopsis XTH11, XTH29 and XTH33 exhibit different secretion pathways and responses to heat and drought stress. *The Plant Journal*, 107(2), 448-466.
- Daudi, A., & O'brien, J. A. (2012). Detection of hydrogen peroxide by DAB staining in Arabidopsis leaves. *Bio-protocol*, 2(18), e263-e263.
- Deeks, M. J., Kaloriti, D., Davies, B., Malhó, R., & Hussey, P. J. (2004). Arabidopsis NAP1 is essential for Arp2/3-dependent trichome morphogenesis. *Current Biology*, 14(15), 1410-1414.
- Espinosa-Andrews, H., Urías-Silvas, J. E., & Morales-Hernández, N. (2021). The role of agave fructans in health and food applications: A review. *Trends in Food Science & Technology*, 114, 585-598.
- Endo, H., Yamaguchi, M., Tamura, T., Nakano, Y., Nishikubo, N., Yoneda, A., ... & Demura, T. (2015). Multiple classes of transcription factors regulate the expression of VASCULAR-RELATED NAC-DOMAIN7, a master switch of xylem vessel differentiation. *Plant and Cell Physiology*, 56(2), 242-254.
- Evans, Teri, and Matthew Loose. "AlignWise: a tool for identifying protein-coding sequence and correcting frame-shifts." *BMC bioinformatics* 16 (2015): 1-11.
- Evert, R. F. (2006). *Esau's plant anatomy: meristems, cells, and tissues of the plant body: their structure, function, and development*. John Wiley & Sons.
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5), 1792-1797.
- Esau, K. (1953). *Plant anatomy* (Vol. 75, No. 5, p. 407). LWW.

- Fang, H., Wang, P., Wang, W., Peng, J., Zheng, J., Zhu, G., ... & Yu, W. (2023). Fine mapping and identification of SmAPRR2 regulating rind color in eggplant (*Solanum melongena* L.). *International Journal of Molecular Sciences*, 24(4), 3059.
- Fernández-Pérez, F., Pomar, F., Pedreño, M. A., & Novo-Uzal, E. (2015). The suppression of AtPrx52 affects fibers but not xylem lignification in *Arabidopsis* by altering the proportion of syringyl units. *Physiologia Plantarum*, 154(3), 395-406.
- Fraga, J. (2004). Los habitantes de la zona costera de Yucatán: entre la tradición y la modernidad. El manejo costero en México. Universidad Autónoma de Campeche, SEMARNAT, CETYS-Universidad, Universidad de Quintana Roo, 497-506.
- Fujita, M., & Harada, H. (1979). Autoradiographic investigations of cell wall development. II. Tritiated phenylalanine and ferulic acid assimilation in relation to lignification.
- Gross, S. M., Martin, J. A., Simpson, J., Abraham-Juarez, M. J., Wang, Z., & Visel, A. (2013). De *novo* transcriptome assembly of drought tolerant CAM plants, *Agave deserti* and *Agave tequilana*. *BMC genomics*, 14, 1-14.
- García-Castillo, M. J., Rodríguez-Zapata, L. C., & Teyer, L. F. S. (2022). Differential expression of *CesA* genes and the relationship with fiber content in henequen. *Tropical and Subtropical Agroecosystems*, 25(3).
- Ge, S. X., Son, E. W., & Yao, R. (2018). iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC bioinformatics*, 19, 1-24.
- Gorshkova, T., Brutch, N., Chabbert, B., Deyholos, M., Hayashi, T., Lev-Yadun, S., ... & Pilate, G. (2012). Plant fiber formation: state of the art, recent and expected progress, and open questions. *Critical Reviews in Plant Sciences*, 31(3), 201-228.
- Huo, Y., Zhang, G., Yu, W., Liu, Z., Shen, M., Zhao, R., ... & Yang, Y. (2023). Forward genetic studies reveal *LsAPRR2* as a key gene in regulating the green color of pericarp in bottle gourd (*Lagenaria siceraria*). *Frontiers in Plant Science*, 14, 1130669.
- Huang, X., Xu, B., Tan, S., Huang, Y., Xi, J., Qin, X., ... & Yi, K. (2022). Transcriptome sequencing of *Agave angustifolia* reveals conservation and diversification in the expression of cinnamyl alcohol dehydrogenase genes in *Agave* species. *Agriculture*, 12(7), 1003.

## BIBLIOGRAFÍA

- Hu, H., Zhang, R., Dong, S., Li, Y., Fan, C., Wang, Y., ... & Peng, L. (2018). AtCSLD3 and GhCSLD3 mediate root growth and cell elongation downstream of the ethylene response pathway in Arabidopsis. *Journal of Experimental Botany*, 69(5), 1065-1080.
- Hiraide, H., Yoshida, M., Sato, S., & Yamamoto, H. (2016). In situ detection of laccase activity and immunolocalisation of a compression-wood-specific laccase (CoLac1) in differentiating xylem of *Chamaecyparis obtusa*. *Functional Plant Biology*, 43(6), 542-552.
- Han, W. Q., Chang, F. J., Wang, Q. R., & Pan, J. Q. (2015). Microparticles from patients with the acute coronary syndrome impair vasodilatation by inhibiting the Akt/eNOS-Hsp90 signaling pathway. *Cardiology*, 132(4), 252-260.
- Herrero, J., Carrasco, A. E., & Zapata, J. M. (2014). *Arabidopsis thaliana* peroxidases involved in lignin biosynthesis: in silico promoter analysis and hormonal regulation. *Plant physiology and biochemistry*, 80, 192-202.
- Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., ... Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, 8(8), 1494–1512
- Higuchi, T. (2006). Look back over the studies of lignin biochemistry. *Journal of Wood Science*, 52, 2-8.
- Ibarra-Laclette, E., Venancio-Rodríguez, C. A., Vásquez-Aguilar, A. A., Alonso-Sánchez, A. G., Pérez-Torres, C. A., Villafán, E., ... & Ornelas, J. F. (2022). Transcriptional basis for haustorium formation and host establishment in hemiparasitic *Psittacanthus schiedeanus* Mistletoes. *Frontiers in Genetics*, 13, 929490.
- Javed, T., & Gao, S. J. (2023). WRKY transcription factors in plant defense. *Trends in Genetics*.
- Jang, J. H., Seo, H. S., & Lee, O. R. (2022). Overexpression of pPLAIIIγ in Arabidopsis reduced xylem lignification of stem by regulating peroxidases. *Plants*, 11(2), 200.

## BIBLIOGRAFÍA

- Jeong, Y. J., Kim, Y. C., Lee, J. S., Kim, D. G., & Lee, J. H. (2022). Reduced expression of PRX2/ATPRX1, PRX8, PRX35, and PRX73 affects cell elongation, vegetative growth, and vasculature structures in *Arabidopsis thaliana*. *Plants*, 11(23), 3353.
- Jiang, L., Fu, Y., Tian, X., Ma, Y., Chen, F., & Wang, G. (2022). The Anthurium APRR2-like gene promotes photosynthetic pigment accumulation in response to salt stress. *Tropical Plant Biology*, 1-10.
- Jarvis, M. C. (2012). Sclerenchyma. eLS.
- Simpson, J., Martinez Hernandez, A., Jazmín Abraham Juárez, M., Delgado Sandoval, S., Sanchez Villarreal, A., & Cortes Romero, C. (2011). Genomic resources and transcriptome mining in *Agave tequilana*. *Gcb Bioenergy*, 3(1), 25-36.
- Jensen WA (1962) Botanical Histochemistry – principles and practice. University of California, Berkeley. W.H Freeman and Company
- Kurien, R. A., Anil, M. M., Mohan, S. S., & Thomas, J. A. (2023). Natural fiber composites as sustainable resources for emerging applications-a review. *Materials Today: Proceedings*.
- Katz, F. (1962). El sistema de plantación y la esclavitud. *Ciencias políticas y sociales*, 8(27), 102-35.
- Liu, Q., Luo, L., & Zheng, L. (2018). Lignins: biosynthesis and biological functions in plants. *International journal of molecular sciences*, 19(2), 335.
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, 15, 1-21.
- Langmead, B. y S. L. Salzberg (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9, 357-359.
- Liu, C. J. (2012). Deciphering the enigma of lignification: precursor transport, oxidation, and the topochemistry of lignin assembly. *Molecular Plant*, 5(2), 304-317.
- Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics*, 12, 1-16.

- Liu, C. J., Miao, Y. C., & Zhang, K. W. (2011). Sequestration and transport of lignin monomeric precursors. *Molecules*, 16(1), 710-727
- Li, Y., Sorefan, K., Hemmann, G., & Bevan, M. W. (2004). Arabidopsis NAP and PIR regulate actin-based cell morphogenesis and multiple developmental processes. *Plant Physiology*, 136(3), 3616-3627.
- Lev-Yadun, S. (2001). Intrusive Growth. The Plant Analog of Dendrite and Axon Growth in Animals. *New Phytologist*, 508-512.
- Li, Y., Mai, Y. W., & Ye, L. (2000). Sisal fibre and its composites: a review of recent developments. *Composites science and technology*, 60(11), 2037-2055.
- Liu, Y. (2015). *Functional analysis of homeodomain transcription factors in secondary cell wall formation in Arabidopsis thaliana* (Doctoral dissertation, University of British Columbia).
- Miao, T., Bao, H., Ling, H., Li, P., Zhang, Y., He, Y., ... & Wang, S. (2023). Comparative transcriptomic analysis revealed the suppression and alternative splicing of kiwifruit (*Actinidia latifolia*) NAP1 gene mediating Trichome development. *International Journal of Molecular Sciences*, 24(5), 4481.
- Morán Velázquez, D.C., Luis F. Maceda-López, Lorena Rodríguez-López, Francisco Vilaplana, Fulgencio Alatorre-Cobos. (2023), ANATOMÍA, COMPOSICIÓN QUÍMICA Y GENÓMICA DE LAS FIBRAS DE *Agave* spp, *Agave* y derivados: Tendencias científicas, uso sostenible y patrimonio, Jalisco. pp. 93-109
- Maceda-López, L. F., Góngora-Castillo, E. B., Ibarra-Laclette, E., Morán-Velázquez, D. C., Girón Ramírez, A., Bourdon, M., ... & Alatorre-Cobos, F. (2022). Transcriptome mining provides insights into cell wall metabolism and fiber lignification in *Agave tequilana* Weber. *Plants*, 11(11), 1496.
- Maceda, A., Soto-Hernández, M., Peña-Valdivia, C. B., Trejo, C., & Terrazas, T. (2021). Lignina: composición, síntesis y evolución. *Madera y bosques*, 27(2).
- Morán Velázquez, D. C. (2021). Análisis del proceso de acumulación de compuestos lignocelulósicos en fibras de hojas de *Agave fourcroydes* Lem.

## BIBLIOGRAFÍA

- Maceda López, L. F. (2020). Identificación y análisis de los genes involucrados en las rutas biosintéticas de celulosa y lignina en *Agave tequilana* Weber.
- Mónica Moraes R., Benjamin øllgaard, Lars Peter Kvist, Finn Borchsenius & Henrik balslev (2007). *Botánica económica de los andes centrales.*, Universidad Mayor De San Andrés.
- Micic, M., Radotic, K., Jeremic, M., & Leblanc, R. M. (2003). Study of self-assembly of the lignin model compound on cellulose model substrate. *Macromolecular bioscience*, 3(2), 100-106.
- Martín, A. M. V. (1973). Factores que afectan la formación del tejido de esclerenquima en castanea sativa mill: relaciones con su rizogenesis (Doctoral dissertation, Universidade de Santiago de Compostela).
- Nava-Cruz, N. Y., Medina-Morales, M. A., Martínez, J. L., Rodríguez, R., & Aguilar, C. N. (2015). Agave biotechnology: an overview. *Critical reviews in biotechnology*, 35(4), 546-559.
- Núñez, H. M., Rodríguez, L. F., & Khanna, M. (2011). Agave for tequila and biofuels: an economic assessment and potential opportunities. *Gcb Bioenergy*, 3(1), 43-57.
- Peña-Ramírez, Y. J., Labrín-Sotomayor, N. Y., Becerra-Lucio, P. A., Ruiz-González, H., Fortuny-Fernández, N. M., Alatorre-Cobos, F., & Pirro, S. (2024). The Complete Genome Sequences of 12 Species of Plants from the Yucatan Peninsula, Mexico. *Biodiversity genomes*.
- Perrin, R. M. (2001). Xyloglucan fucosyltransferase, a plant cell wall biosynthetic enzyme. Michigan State University.
- Qin, Y., Wei, H., Sun, H., Hao, P., Wang, H., Su, J., & Yu, S. (2017). Proteomic analysis of differences in fiber development between wild and cultivated *Gossypium hirsutum* L. *Journal of Proteome Research*, 16(8), 2811-2824.
- Qi, G., Hu, R., Yu, L., Chai, G., Cao, Y., Zuo, R., ... & Zhou, G. (2013). Two poplar cellulose synthase-like D genes, PdCSLD5 and PdCSLD6, are functionally conserved with Arabidopsis CSLD3. *Journal of plant physiology*, 170(14), 1267-1276
- Pickett-Heaps, J. D. (1968). Xylem wall deposition: Radioautographic investigations using lignin precursors. *Protoplasma*, 65, 181-205.

## BIBLIOGRAFÍA

- Raya, F. T., de Carvalho, L. M., José, J., da Cruz, L. P., Almeida, R. L., Delevatti, H. A. D. A., ... & Pereira, G. A. G. (2023). Rescuing the Brazilian Agave breeding program: morphophysiological and molecular characterization of a new germplasm. *Frontiers in Chemical Engineering*, 5, 1218668.
- Raya, F. T., Marone, M. P., Carvalho, L. M., Rabelo, S. C., De Paula, M. S., Campanari, M. F. Z., ... & Pereira, G. A. G. (2021). Extreme physiology: Biomass and transcriptional profiling of three abandoned Agave cultivars. *Industrial Crops and Products*, 172, 114043.
- Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P. F., ... & Boerjan, W. (2004). Lignins: natural polymers from oxidative coupling of 4-hydroxyphenyl-propanoids. *Phytochemistry reviews*, 3, 29-60.
- Somssich, M. (2020). Lignification and oxidative enzymes: localization, localization, localization!.
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19), 3210-3212.
- Smith, R. A., Schuetz, M., Roach, M., Mansfield, S. D., Ellis, B., & Samuels, L. (2013). Neighboring parenchyma cells contribute to Arabidopsis xylem lignification, while lignification of interfascicular fibers is cell autonomous. *The Plant Cell*, 25(10), 3988-3999.
- Sibout, R., & Höfte, H. (2012). Plant cell biology: the ABC of monolignol transport. *Current biology*, 22(13), R533-R535.
- Snegireva, A. V., Ageeva, M. V., Amenitskii, S. I., Chernova, T. E., Ebskamp, M., & Gorshkova, T. A. (2010). Intrusive growth of sclerenchyma fibers. *Russian Journal of Plant Physiology*, 57, 342-355.
- Silva, F. D., & Chawla, N. (2008). Tensile behavior of high performance natural (sisal) fibers. *Composites Science and Technology*, 68(15), 3438-3443.
- The Brainy Insights (2024), Natural fibers Market Size by Type (Plant-based and Animal-based), By Application (Automotive, Textile, Medical and Others), Regions, Global Industry Analysis, Share, Growth, Trends, and Forecast 2023 to 2032

<https://www.thebrainyinsights.com/report/natural-fibers-market-13957> [Acceso 1 junio 2024].

- Trujillo-Ramírez, D., Bustos-Vázquez, M. G., Martínez-Velasco, A., & Torres-de los Santos, R. (2023). INTEGRAL USE OF HENEQUEN (*Agave fourcroydes*): APPLICATIONS AND TRENDS—A REVIEW. *Tropical and Subtropical Agroecosystems*, 26(2).
- Tavares, T. D., Antunes, J. C., Ferreira, F., & Felgueiras, H. P. (2020). Biofunctionalization of natural fiber-reinforced biocomposites for biomedical applications. *Biomolecules*, 10(1), 148.
- Tobimatsu, Y., & Schuetz, M. (2019). Lignin polymerization: how do plants manage the chemistry so well?. *Current opinion in biotechnology*, 56, 75-81.
- Valadez-Moctezuma, E., Samah, S., Mascorro-Gallardo, J. O., Reynoso, J. J. L., Marbán-Mendoza, N., Aranda-Osorio, G., & Brito-Nájera, G. (2023). The *de novo* transcriptome assembly of *Agave angustifolia* Haw, mining for carbohydrates and cellulose synthesis genes profiling, and molecular markers development. *Plant Biotechnology Reports*, 17(6), 889-904.
- Vermaas, J. V., Dixon, R. A., Chen, F., Mansfield, S. D., Boerjan, W., Ralph, J., ... & Beckham, G. T. (2019). Passive membrane transport of lignin-related compounds. *Proceedings of the National Academy of Sciences*, 116(46), 23117-23123.
- Voxeur, A., Wang, Y., & Sibout, R. (2015). Lignification: different mechanisms for a versatile polymer. *Current opinion in plant biology*, 23, 83-90.garcia
- Vanholme, R., Demedts, B., Morreel, K., Ralph, J., & Boerjan, W. (2010). Lignin biosynthesis and structure. *Plant physiology*, 153(3), 895-905.
- Vázquez Pasos, L.A. (1999). *Identidad, henequén y trabajo: los desfibradores de Yucatán*. El Colegio de México.
- Vidal, D. C. (1988). El sector henequenero en Yucatán. *Revista CIDOB D'Afers Internacionals*, 77-93.
- Weng, J. K., & Chapple, C. (2010). The origin and evolution of lignin biosynthesis. *New Phytologist*, 187(2), 273-285.

- Whetten, R., & Sederoff, R. (1995). Lignin biosynthesis. *The plant cell*, 7(7), 1001.
- Wenham, M. W., & Cusick, F. (1975). The growth of secondary wood fibres. *New Phytologist*, 74(2), 247-261.
- Xu, Z., Escamilla-Treviño, L., Zeng, L., Lalgondar, M., Bevan, D., Winkel, B., ... & Esen, A. (2004). Functional genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 1. *Plant molecular biology*, 55, 343-367
- Yano, H., & Fu, W. (2023). Hemp: A sustainable plant with high industrial value in food processing. *Foods*, 12(3), 651.
- Yokoyama, R., Shinohara, N., Asaoka, R., Narukawa, H., & Nishitani, K. (2014). The biosynthesis and function of polysaccharide components of the plant cell wall. *Plant cell wall patterning and cell shape*, 1-34.
- Zhao, H., Li, Z., Wang, Y., Wang, J., Xiao, M., Liu, H., ... & Zhang, Z. (2022). Cellulose synthase-like protein OsCSLD4 plays an important role in the response of rice to salt stress by mediating abscisic acid biosynthesis to regulate osmotic stress tolerance. *Plant Biotechnology Journal*, 20(3), 468-484.
- Zhao, Q. (2016). Lignification: flexibility, biosynthesis and regulation. *Trends in plant science*, 21(8), 713-721.
- Zhong, R., Lee, C., McCarthy, R. L., Reeves, C. K., Jones, E. G., & Ye, Z. H. (2011). Transcriptional activation of secondary wall biosynthesis by rice and maize NAC and MYB transcription factors. *Plant and Cell Physiology*, 52(10), 1856-1871