CENTRO DE INVESTIGACIÓN EN QUÍMICA APLICADA

PROGRAMA DE POSGRADO EN AGROPLASTICULTURA

Análisis transcriptómico de plantas silvestres de Guayule (*Parthenium argentatum* Gray) nativas de México

TESIS

Presentada por:

María Brenda Olivia Baylon Palomino

Para obtener el grado de:

MAESTRO EN CIENCIAS EN AGROPLASTICULTURA

Saltillo, Coahuila, México. **Diciembre 2017**

CONACYT
Consejo Nacional de Ciencia y Tecnología

CENTRO DE INVESTIGACIÓN EN QUÍMICA APLICADA

Saltillo, Coahuila a 26 de enero de 2018

Coordinación de Posgrado
Presente

Por este conducto nos permitimos informar a esta coordinación que, el documento de tesis preparado por MARÍA BRENDA OLIVIA BAYLON PALOMINO titulado Análisis transcriptómico de plantas silvestres de Guayule (Parthenium argentatum Gray) nativas de México el cual fue presentado el día 15 de diciembre de 2017, ha sido modificado de acuerdo a las observaciones, comentarios y sugerencias, realizadas por el Comité Evaluador asignado. Por tal motivo, avalamos que el documento adjunto corresponde a la versión final del documento de tesis.

Atentamente,

SINODALES

Dr. José Humberto Valenzuela Soto
Presidente

Dra. Ileana Vera Reyes
Secretario

Antonio S. Ledezma P.

Dr. Antonio S. Ledezma Pérez
Vocal

Vo. Bo. de los Asesores

Dra. Ana Margarita Rodríguez Hernández

Dra. Mona Kassem

CENTRO DE INVESTIGACIÓN EN QUÍMICA APLICADA
Programa de Maestría en Ciencias en Agroplasticultura

TESIS

Análisis transcriptómico de plantas silvestres de Guayule (Parthenium argentatum Gray) nativas de México

Presentada por:

MARÍA BRENDA OLIVIA BAYLON PALOMINO

Para obtener el grado de:

Maestro en Ciencias en Agroplasticultura

Asesorada por:

Dra. Ana Margarita Rodríguez Hernández
Dra. Mona Kassem

SINODALES

Dr. José Humberto Valenzuela Soto
Presidente

Dra. Ileana Vera Reyes
Secretario

Dr. Antonio S. Ledezma Pérez
Vocal

Saltillo, Coahuila **Diciembre, 2017**

TESIS CON CARACTER ABIERTO

PROGRAMA: MAESTRÍA EN CIENCIAS EN AGROPLASTICULTURA

AUTOR: MARÍA BRENDA OLIVIA BAYLON PALOMINO **FIRMA** _____

TITULO: Análisis transcriptómico de plantas silvestres de Guayule (Parthenium argentatum Gray) nativas de México.

ASESORES: Dra. Ana Margarita Rodríguez Hernández **FIRMA** _____

Dra. Mona Kassem **FIRMA** _____

El Centro de Investigación en Química Aplicada clasifica el presente documento de tesis como ABIERTO.

Un documento clasificado como Abierto se expone en los estantes del Centro de Información para su consulta. Dicho documento no puede ser copiado en ninguna modalidad sin autorización por escrito del Titular del Centro de Información o del Director General del CIQA.

Saltillo, Coahuila, a 15 de diciembre de 2017

CENTRO DE INVESTIGACION EN QUIMICA APLICADA

Sello de la Institución

Dr. Oliverio Santiago Rodríguez Fernández
Director General del CIQA

CENTRO DE INVESTIGACIÓN EN QUÍMICA APLICADA

Programa de Maestría en Ciencias en Agroplasticultura

Análisis transcriptómico de plantas silvestres de Guayule (*Parthenium argentatum* Gray) nativas de México

Presentada por:

María Brenda Olivia Baylon Palomino

Para obtener el grado de:

MAESTRO EN CIENCIAS EN AGROPLASTICULTURA

Asesorada por:

Dra. Ana Margarita Rodríguez Hernández

Dra. Mona Kassem

DECLARACIÓN

Declaro que la información contenida en la Parte Experimental así como en la Parte de Resultados y Discusiones de este documento y que forman parte de las actividades de investigación y desarrollo realizadas durante el período que se me asignó para llevar a cabo mi trabajo de tesis, será propiedad del Centro de Investigación en Química Aplicada.

Saltillo, Coahuila a 20 de diciembre de 2017

MARÍA BRENDA OLIVA BAYLON PALOMINO

Nombre y Firma

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por brindarme el apoyo económico durante la maestría; lo que me permitió cumplir con una meta más en el ámbito académico. Además del apoyo financiero recibido a través del proyecto 247278 de la convocatoria “Proyectos de Desarrollo Científico para Atender Problemas Nacionales 2014”

Al Centro de Investigación en Química Aplicada (CIQA) que me abrió las puertas y me brindó las herramientas para finalizar mis estudios.

A mis asesoras la Dra. Ana Margarita Rodríguez Hernández y Dra. Mona Kassem por el apoyo y la confianza brindada a lo largo de este proyecto.

Al Dr. Marco Castillo Campohermoso y al Dr. Jorge Romero por el apoyo brindado para el desarrollo de este proyecto.

Al Dr. Antonio Serguei Ledezma Pérez por ser siempre una persona positiva por lo cual apreció mucho.

A mis sinodales Dra. Ileana Vera Reyes, Dr. José Humberto Valenzuela Soto y Dr. Antonio Serguei Ledezma Pérez por las aportaciones recibidas para la mejora de este documento.

Al Dr. Carlos José Espinoza González que facilitó equipos clave en el desarrollo de este proyecto.

Al Dr. José Román Torres Lubián por su ayuda en este proyecto.

Al Instituto de Ecología (INECOL) por abrirme sus puertas lo que permitió llevar a cabo gran parte de este proyecto.

Al Dr. Enrique Ibarra Laclette por la ayuda recibida ya que fue una pieza clave en el desarrollo de este proyecto en los análisis transcriptómicos.

Al M.C. Emanuel Villafán de la Torre por su valiosa ayuda en cuestiones de bioinformática caracterizado siempre con una excelente actitud.

Al Dr. Eliel Ruiz May por toda la ayuda recibida durante la estancia en el instituto de Ecología.

A Dios y a la Virgen que me han permitido seguir en este camino a pesar de las adversidades.

A toda mi familia que siempre me alentó y jamás permitió que me rindiera.

A los amigos que conocí durante la maestría que permitieron que este proceso fuera más ameno.

A Miriam, Édgar, Yessenia y Evelin ya que los consideró grandes amigos con quienes he pasado momentos inolvidables e increíblemente divertidos. A mi amigo Diego Armando Aguirre a quien estimo y quiero por todos estos años que hemos pasado con tristezas y alegrías pero siempre juntos.

A Luis Alberto Martínez Rodríguez por ser parte importante de mi vida y por toda la comprensión brindada en todos estos años de estar juntos.

DEDICATORIA

A mi madre y hermano a quien amo y admiro sobre todas las cosas

A mi abuela y a mis tíos quienes siempre me apoyaron

ÍNDICE

INDICE DE FIGURAS I
INDICE DE TABLAS III
ABREVIATURAS IV
RESUMEN VI
I. INTRODUCCIÓN 1
1.1. OBJETIVO GENERAL 4
1.1.1. Objetivos específicos 4
II. HIPOTESIS 5
III. ANTECEDENTES 6
3.1. HULE NATURAL: Y su importancia a nivel mundial 6
3.2. Planta de Guayule 11
3.3. MÉTODOS DE EXTRACCIÓN DE HULE Y RESINA 16
3.3.1. Extracción con molino de bolas 17
3.3.2. Extracción Soxhlet 18
3.3.3. Extracción Acelerada con Solventes 20
3.4. RUTA DE SÍNTESIS DEL HULE 22
3.4.1. Vía Mevalónica (MVA) 24
3.4.2. Vía 2C-metil-D-eritrol 4-fosfato (MEP) 24
3.4.3. Genes relacionados con la partícula de hule 26
3.4.3.1. Importancia del transcriptoma 28
IV. METODOLOGÍA 31
4.1. Colecta de material vegetal 31
4.2. Extracción y contenido de hule 32
4.2.1. Identificación de hule 36
4.3. Extracción y cuantificación de ARN total 37
4.3.1. Extracción de ARN total 37
4.4. Secuenciación masiva de transcritos 38
4.5. Integración de datos transcriptómicos 38
4.5.1. Filtrado de calidad 38
4.5.2. Ensamblado 39

4.5.3. Limpieza 39
4.5.4. Identificación de Marcos de Lectura Abierta en los unigenes provenientes de *Parthenium argentatum* 40
4.5.5. Identificación de proteínas homólogas 42
4.5.6. Anotación funcional a los unigenes de *P. argentatum* 43
4.5.7. Identificación de genes ortólogos 43
4.5.7.1. Análisis filogenético 44
4.5.8. Análisis de expresión diferencial 45
V. RESULTADOS Y DISCUSIÓN 46
5.1. Colecta de material vegetal 46
5.2. Estandarización del método de extracción y cantidad de muestra 47
5.2.1. Identificación de hule 49
5.3. Extracciones de hule y resina de plantas provenientes de Durango colectadas en verano e invierno (2016-2017) 51
5.4. Extracciones de hule y resina de plantas provenientes de San Luis Potosí colectadas en verano e invierno (2016-2017) 54
5.5. Extracción y cuantificación de ARN total 60
5.6. Filtrado y ensamblado de las lecturas 61
5.6.1. Identificación de ORF´s 64
5.6.2. Identificación de proteínas homólogas 64
5.6.3. Anotación funcional de los unigenes de *Parthenium argentatum* 65
5.6.4. Identificación de genes ortólogos 67
5.6.5. Análisis de expresión diferencial 73
VI. CONCLUSIONES 80
VII. PERSPECTIVAS 81
VIII. LITERATURA CITADA 82
IX. ANEXOS 91
9.1. Extracción de ARN total (I) 91
9.2. Purificación de ARN total (II) 92

INDICE DE FIGURAS

Figura 1.Panorama del incremento en la producción y hectáreas cosechadas de hule natural a nivel mundial de los años 1994-2014. 7
Figura 2. Producción de hule a nivel mundial. 9
Figura 3. Principales productores de hule a nivel mundial. 10
Figura 4. Planta de *Parthenium argenatum* en condición silvestre 12
Figura 5. Distribución del guayule en la República mexicana..... 13
Figura 6. Micrografía electrónica de barrido de células del parénquima de corteza de tallos *de Parthenium argentatum*, mostrando vacuolas llenas de partículas de látex 14
Figura 7. Planta de guayule seccionada en partes: raíz, tallo y hojas. 15
Figura 8. Imagen ilustrativa de molido de bolas y sus partes 18
Figura 9. Imagen ilustrativa del método de extracción Soxhlet. 19
Figura 10.Ilustración del metabolismo primario y secundario de las plantas 23
Figura 11. Biosíntesis de hule natural 25
Figura 12. Colecta de tallos de *Parthenium argentatum*..... 32
Figura 13. Proceso de molienda de tallos de guayule. a) tallo de guayule. b) tallo de guayule cortado en trozos. c) tallo de guayule molido..... 33
Figura 14. Equipo automatizado Speed Extractor E-916, marca BÜCHI y sus principales componentes (a) Celda de extracción, b) Equipo Speed Extractor E-916, c) Reservorio de solventes, d) Viales de recolección)..... 34
Figura 15. Ubicación de sitios de colecta en el estado de Durango 46
Figura 16. Ubicación de los sitios de colecta en el estado de San Luis Potosí..... 47
Figura 17. Espectro de RMN que comprueba la presencia de hule natural en las muestras analizadas 50
Figura 18. Espectro del C13 en donde se muestran los desplazamientos químicos de los 5 carbonos presentes en el poliisopreno *Cis*..... 50
Figura 19. Porcentajes de resina de la colecta de verano 2016 en el estado de Durango 52
Figura 20. Porcentajes de resina de la colecta de invierno 2017 en el estado de Durango 52
Figura 21. Porcentajes de hule de la colecta de verano 2016 en el estado de Durango .. 53
Figura 22. Porcentajes de hule de la colecta de invierno 2017 en el estado de Durango. 53
Figura 23. Porcentajes de resina de la colecta de verano 2016 en el estado de San Luis Potosí 54
Figura 24. Porcentajes de resina de colecta de invierno 2017 en el estado de San Luis Potosí 55
Figura 25.Porcentajes de hule de colecta verano 2016 en el estado de San Luis Potosí. 58
Figura 26. Porcentajes de hule de colecta invierno 2017 en el estado de San Luis Potosí 59
Figura 27. Gel de agarosa mostrando las bandas de ARN sin presencia de ADN en las muestras de Durango 60
Figura 28. Gel de agarosa mostrando las bandas de ARN extraído de las muestras provenientes del estado de San Luis Potosí sin la presencia de ADN 61

Figura 29. Anotación funcional de los Unigenes de Parthenium argentatum por medio de la base de datos DAVID, asociados a términos GO y relacionados con sus diversas categorías 66
Figura 30. Reconstrucción de rutas de síntesis de IPP: a) MVA y b) MEP y los genes implicados. Las flechas representan los genes, los cuadros presentan las especies que comparten genes ortólogos con *Parthenium argentatum* =PaAr, ArTh=*Arabidopsis thaliana*; HeAn=*Helianthus annus*; HeBr=*Hevea brasiliensis*; LaPa=Ruta del hule; SeIn=*Sesamum indicum*; SoLy=*Solanum lycopersic*um. 69
Figura 31. Reconstrucción de ruta del IPP hasta el FPP. Las flechas representan los genes, los cuadros presentan las especies que comparten genes ortólogos con *Parthenium argentatum*. 70
Figura 32. Árbol filogenético del primer grupo de genes ortólogos a nivel de proteína..... 71
Figura 33. Árbol filogenético del primer grupo de genes ortólogos a nivel de secuencia nucleotídica. 72
Figura 34. Mapa de calor. Representación de los perfiles de expresión genética de las muestras de Guayule en sus diferentes condiciones de colecta. El nombre de las 12 muestras se encuentra en la parte inferior del mapa. Las barras rojas, verde, azul celeste y morado, corresponde a las muestras colectadas en las condiciones A, B, C y D respectivamente. Una relación filogenética agrupa estas muestras según la expresión genética de cada una y muestra en la parte superior del mapa. Cada línea horizontal, corresponde a un gen determinado. Donde el color amarillo refiere a expresión alta y el morado expresión relativamente baja. 75
Figura 35. Unigenes expresados diferencialmente mediante los distintos contrastes analizados con un intervalo del 95% de confianza 76
Figura 36. Ruta del mevalonato (a), síntesis de FPP (b) y componentes del complejo de la partícula de hule (c), las flechas rojas ( ) indican una expresión diferencial negativa de los genes correspondientes en plantas de mayor contenido de hule 79

INDICE DE TABLAS

Tabla 1. Comparación de especies productoras de hule 11
Tabla 2. Condiciones bajo las cuales se realizó la extracción de resina y hule provenientes de muestras de Guayule 35
Tabla 3. Especies elegidas para llevar a cabo la identificación de marcos de lectura abierta en los unigenes de *Parthenium argentatum* 41
Tabla 4. Bases de datos utilizadas en la identificación de proteínas homólogas al realizar la comparación con las proteínas encontradas en *Parthenium argentatum* 43
Tabla 5. Comparación de porcentajes de hule y resina obtenidos en relación a la cantidad de muestra procesada 49
Tabla 6. Desplazamientos químicos obtenidos en el hule proveniente de *P.argentatum* en comparación con *T. kokzaghyz* y lo reportado en la literatura 51
Tabla 7. Concentración de ARN total de las 12 muestras seleccionadas 61
Tabla 8. Muestras secuenciadas (código de identificación, sitio de origen, criterio de selección, total de lecturas crudas, número de lecturas después del filtrado de calidad) 62
Tabla 9. Establecimiento de condiciones y contrastes para el análisis de expresión diferencial. 74

ABREVIATURAS

AACT	Acetil coenzima A tiolasa
ADN	Ácido desoxirribonucleico
aLRT	approximate Likelihhod Ratio Test; Prueba de razón de verosimilitud aproximada
AOS	Aleno óxido sintasa
ARN	Ácido ribonucleico
ASE	Accelerated Solvent Extraction; Extracción Acelerada con Solventes
BLAST	The Basic Local Alignment Search Tool; Herramienta básica de búsqueda de alineación local
cDNA	ADN complementario
CDS	Coding Sequence; secuencias codificantes
CMK	4-difosfocitidil-2-C-metil-D-eritritol quinasa
CPT	Complejo Cis -preniltransferasas
DMAPP	Pirofosfato de dimetilalilo
DXR	1-desoxi-D-xilulosa 5-fosfato reductoisomerasa
EDTA	Ácido etilendiaminotetraacético
EST	Expressed Sequence Tags; Etiquetas de secuencias expresadas
FDPS	Farnesil pirofosfato sintasa
FPP	Difosfato de farnesilo
GGPP	Difosfato de Geranilgeranilo
GHS	Homologa de la Proteína Pequeña Partícula de Hule en Guayule
GO	Gene Ontology; Ontología Génica
GPP	Difosfato de Geranilo
HDR	HMBPP reductasa
HDS	4-hidroxi 3-metil 2-butenil difosfato sintasa
HMGR	3-hidroxi-3-metil-glutaril-CoA reductasa
HMGS	3-hidroxi-3-metil-glutaril-CoA sintasa
IPI	Isopentenil pirofosfato isomerasa.

IPP	Pirofosfato de isopentenilo
KEGG	Kyoto Encyclopedia of Genes and Genomes; Enciclopedia de Genes y Genomas de Kioto
MCS	2-C-metil-D-eritritol 2,4-ciclodifosfato sintasa
MCT	4-MEP citidiltransferasa
MDDC	Difosfomevalonato descarboxilasa
MEP	Vía del 2C-metil-D-eritrol 4-fosfato
mRNA	ARN mensajero
MVA	Vía del Mevalonato
MVK	Mevalonato quinasa
NCBI	Centro Nacional para la Información Biotecnológica
NNI	Nearest Neighbor Interchange; Análisis del Vecino más Cercano
ORF´s	Open Reading Frame; Marcos de lectura abierta
PMVK	Fosfomevalonato quinasa
REF	Rubber Elongation Factor; Factor de Elongación de Hule
RMN	Resonancia Magnética Nuclear
RNA seq	RNA Sequencing; Secuenciación de ARN
RSEM	RNA-seq by Expectation Maximization; RNA-seq por la maximización de expectativas
SALB	South American Leaf Blight; Mal Suramericano de la Hoja
SBS	Secuenciación por síntesis
SPR	Subtree Pruning Regrafting; Poda y re-injerto de subárboles
SRPP	Small Rubber Particle Protein; Proteína Pequeña Partícula de Hule
TPP	Pirofosfato de tiamina
UTR	Untranslated Region; región sin traducir

RESUMEN

Guayule *(Parthenium argentatum)*, es una especie cuya principal característica es producir hule de alto peso molecular e hipoalergénico. El presente estudio se enfocó en encontrar genes en plantas silvestres de guayule que estuvieran relacionados en la síntesis de hule que en un futuro pudieran ser candidatos para la mejora genética y obtener un incremento en el rendimiento de hule. La colecta de plantas silvestres se realizó, en las temporadas de verano (2016) e invierno (2017). Se identificaron 12 sitios de colecta en Durango y 11 en San Luis Potosí. De cada sitio se colectaron 3 muestras de tallos, provenientes de tres plantas diferentes para su análisis. La extracción de hule y resina se realizó por medio de un equipo completamente automatizado que se basa en la técnica de extracción acelerada con solventes (ASE). En las muestras colectadas en Durango se obtuvieron porcentajes de 14 y 16.6 % de resina; 21.86 y 15.3% de hule en verano e invierno, respectivamente. En las muestras colectadas en San Luis Potosí se obtuvo un porcentaje cercano al 14% de resina en ambas temporadas, pero en el caso de hule se obtuvieron porcentajes de 19.4 y 15.3% en verano e invierno, respectivamente. Los 2 sitios de colecta en cada estado que presentaron los resultados más contrastantes en cuanto a porcentaje de hule fueron analizados por secuenciación masiva de transcritos. La secuenciación se realizó por medio de la plataforma de Illumina obteniendo un total de 372,661,284 lecturas crudas de las cuales se identificaron 564,256 unigenes. Dichos unigenes fueron anotados utilizando herramientas de genómica comparativa, 241,176 obtuvieron homología con al menos una de las 7 bases de referencia utilizadas. Se categorizaron funcionalmente 32,844 unigenes asociados a funciones biológicas, funciones moleculares y componentes celulares. El análisis de expresión diferencial de genes de las 12 muestras secuenciadas arrojó un total de 4,684 genes regulados positivamente (up-regulated) y 9,335 genes regulados negativamente (down-regulated) además de 417 genes que no mostraron diferencia en la expresión en ninguno de los contrastes establecidos. Entre los genes regulados negativamente, se encuentran dos enzimas relacionadas en la síntesis del IPP en la ruta del mevalonato: Acetil CoA tiolasa (AACT) e isopentenil pirofosfato isomerasa (IPI) en plantas con mayor contenido de hule en la época de invierno en el estado de Durango.

PALABRAS CLAVE: Transcriptoma, Síntesis de látex, Rutas de síntesis.

I. INTRODUCCIÓN

El caucho o hule natural es un polímero de isopreno con una estructura *Cis* (Cis 1-4 poliisopreno) sus características fisicoquímicas manifiestas de unas moléculas largas y filiformes de un peso molecular alto (>1 MDa), lo hacen incomparable con respecto al hule sintético (Moreno *et al.*, 2005). Sus destacadas características como la firmeza, elasticidad, resistencia a la abrasión, resistencia al impacto y la eficiente disipación del calor (en virtud de la fricción reducida que minimiza la acumulación de calor) (van Beilen y Poirier, 2007) hacen que el hule natural tenga más de 40,000 usos o aplicaciones registradas, en ámbitos tan diferentes que van desde el doméstico hasta el médico; pasando por el automotriz, aeroespacial y muchos otros (van Beilen y Poirier, 2007)

A pesar que existe más de 2,000 especies de plantas productoras de hule natural, en la actualidad *Hevea brasiliensis* es la única especie que se explota comercialmente para la producción de hule natural (George y Jacob, 2000). Otras especies como dandelion (*Taraxacum kok-saghyz*) y guayule (*Parthenium argentatum*) son fuentes alternativas de la producción de hule natural con alto potencial de alcanzar en un plazo medio su comercialización a escala industrial.

El Guayule es un arbusto perenne leñoso, nativo del desierto de Chihuahua en México y el sur de Texas en Estados Unidos, produce hule de alta calidad principalmente en las células del parénquima de la corteza (Bonner, 1943). El descubrimiento que el hule natural producido por *H. brasiliensis* es un material alergénico para la mayoría de las personas, cuando se usó en aplicaciones

médicas o productos de higiene detonó una vez más la investigación y la producción de hule natural a partir de guayule (Wood, 2002). Este biopolímero además de ser hipoalergénico, en la actualidad se ha considerado como una fuente alternativa muy prometedora para complementar o en su caso reemplazar la producción de hule de *Hevea*. El hule natural biosintetizado en el guayule posee cualidades inigualables, en particular para la industria de los neumáticos quien en 2015 tuvo un ingreso de alrededor de los 220.000 millones de USD y cuya demanda de materia prima consume del 70 al 80% de la producción mundial de hule natural.

Por lo anterior, el interés y resurgimiento de la investigación, desarrollo tecnológico e innovación ha cobrado enorme relevancia en los diferentes países y regiones económicas del mundo. Esto se refleja en el incremento sostenido del número de trabajos publicados sobre guayule y la producción de hule. Algunos de estos trabajos se han encaminado a investigar la ruta de biosíntesis del hule de guayule y los genes participantes (Benedict *et al.*, 2008; Ponciano *et al.*, 2012). Para estudios de esta índole existe una gran cantidad de técnicas que pretenden dar respuesta a cada una de ellas; dentro de las que destaca la Transcriptómica o RNA-seq la cual presenta ventajas evidentes ya que es una técnica muy minuciosa y sensible, además que realiza un análisis cuantitativo de la expresión génica y da como resultado el perfil exacto del transcriptoma de las células eucariotas (Ekblom y Galindo, 2011). Ya que no se tienen bien definidos algunos de los elementos participantes en la síntesis de hule de guayule, en el presente trabajo se realizó un estudio exhaustivo en cuanto al transcriptoma de plantas

silvestres de guayule colectadas en diferentes condiciones edafoclimáticas en los estados de Durango y San Luis Potosí.

1.1. OBJETIVO GENERAL

Identificar genes en plantas silvestres de guayule relacionados con los puntos de regulación de la biosíntesis de hule que en un futuro sea la base para el desarrollo de planteamientos de esquemas de mejora genética.

1.1.1. Objetivos específicos

Identificar plantas silvestres de guayule que presenten la mayor y menor cantidad de hule, para el estudio posterior del transcriptoma.

Generar una base de datos con el listado de genes identificados en el transcriptoma de *Parthenium argentatum* que contenga información funcional de los mismos (categoría GO).

II. HIPOTESIS

El análisis del transcriptoma ayudará a la reconstrucción de la ruta de síntesis del hule en *Parthenium argentatum*. Además de mostrar expresión diferencial de genes relacionados directamente en dicha ruta (como es el caso de las proteínas Cis-preniltranferasas (CPT), Aleno óxido sintasa (AOS) y la homóloga de la pequeña partícula de hule (GHS por sus siglas en inglés Guayule homologue Small Rubber Particle Protein/SRPP) entre las plantas de mayor y menor concentración de hule.

III. ANTECEDENTES

3.1. HULE NATURAL: Y su importancia a nivel mundial

El hule natural es un conjunto de monómeros de isopreno $(C_5H_8)_n$ unidos entre sí mediante una configuración 1,4 *Cis*, lo que le permite a este polímero presentar propiedades únicas debido a su estructura intrínseca, su alto peso molecular y contribuciones no definidas de componentes menores como lo son proteínas, minerales, carbohidratos y lípidos, los cuales se encuentran en una suspensión coloidal de partículas de hule comúnmente llamado látex (van Beilen y Poirier, 2007). Estas propiedades incluyen resiliencia, elasticidad, resistencia a la abrasión y al impacto, eficiente dispersión del calor y maleabilidad a bajas temperaturas; las cuales no han podido ser reemplazadas por algún hule sintético producido a partir de petróleo, polímeros o mezclas naturales con ausencia de hule (Cornish, 2001).

Debido a las excepcionales propiedades del hule natural su utilización es inevitable en más de 40,000 productos de diversos ámbitos incluyendo más de 400 dispositivos de uso médico (Mooibroek y Cornish, 2000). Por lo tanto el aumento en la producción de hule se ha visto reflejada del año de 1994 al 2014 (figura 1), ya que en sólo 20 años se observa el incremento tanto de la producción mundial de hule como la superficie cosechada. (FAO, 2014).

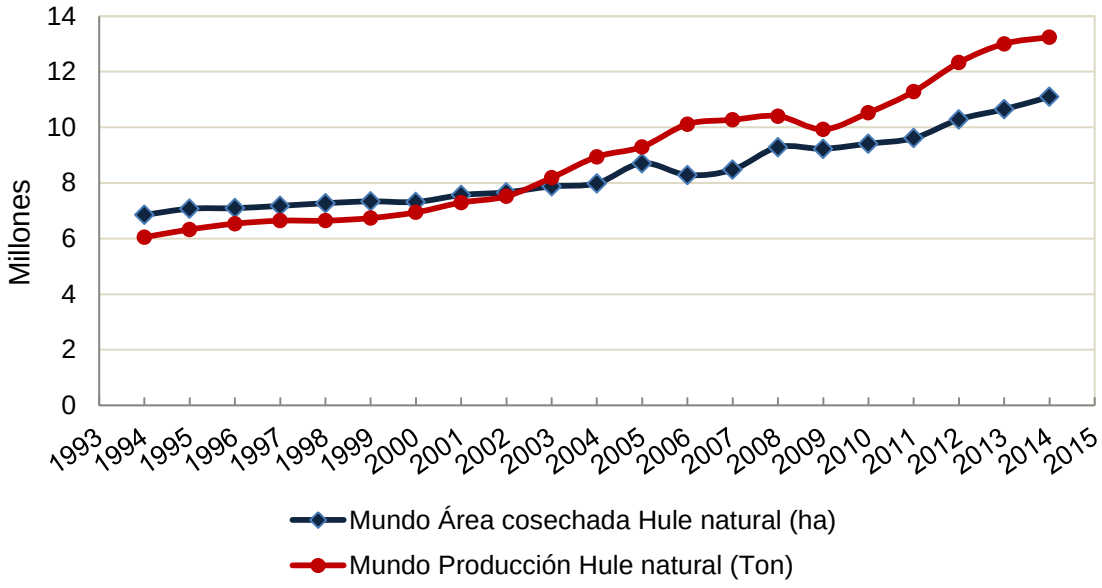

Figura 1.Panorama del incremento en la producción y hectáreas cosechadas de hule natural a nivel mundial de los años 1994-2014. Fuente: FAOSTAT (julio 10, 2017)

Debido al incremento de la producción de hule natural, cuatro continentes se encuentran en la competencia, siendo el principal proveedor Asia (figura 2), ya que en este continente se han encontrado las condiciones óptimas para el desarrollo de la especie productora de hule con mayor importancia: *H. brasiliensis*, ya que alrededor del 80% de la producción mundial es proveniente de esta especie (Tang *et al.*, 2010; van Beilen y Poirier, 2007). Existen ocho familias botánicas, 300 géneros y alrededor de 2,000 especies identificadas como productoras de hule, sólo pocas son conocidas por la producción de hule con alto peso molecular (Bushman *et al.*, 2006; Metcalfe, 1967), por lo tanto, *H. brasiliensis* se ha establecido como una fuente comercial clave de hule natural por su peso molecular, alto rendimiento y excelentes propiedades físicas. *H. brasiliensis* es un miembro de la familia Euphorbiaceae, que mide entre los 10 – 20 metros, con un tallo cilíndrico en plantas injertadas o de forma cónica en la base cuando las

plantas proceden de semilla, comúnmente en las grandes plantaciones el árbol se deja ramificar a partir de los 2.5 m de altura con el objetivo de un tronco liso que permita una buena extracción del látex (Espinal *et al.*, 2005) cabe mencionar que el látex de *Hevea* se encuentra en el citoplasma en células especializadas entrelazadas llamadas laticíferos, donde el hule natural es sintetizado y almacenado en forma de partículas de látex (Hagel *et al.*, 2008). Aproximadamente, el contenido de hule de *H. brasiliensis* representa entre el 30 – 40 % del peso seco del látex, aproximadamente el 2 % del peso seco del árbol (Hayashi, 2009).

A pesar que la especie *H. brasiliensis* es de origen amazónico ha encontrado su máximo esplendor en cuanto a producción de hule en otras partes del mundo (figura 3), ya que la principal causa de baja producción en esta zona es por el azote de mal suramericano de la hoja ó SALB (por sus siglas en inglés South American Leaf Blight) provocado por *Microcyclus ulei* un hongo ascomiceto endémico de dicha región (Lieberei, 2007) que ataca principalmente al género *Hevea* y especialmente a *H. brasiliensis (Junqueira et al., 1989)* causando caída de hojas jóvenes, que en condiciones favorables (períodos de humedad foliar y temperaturas igual o mayor a 20 ^{o}C) pueden llevar a la defoliación; y en caso de ataques constantes el debilitamiento, baja producción de látex y la muerte de árboles infectados (Araújo *et al.*, 2001; Gasparotto *et al.*, 1984)

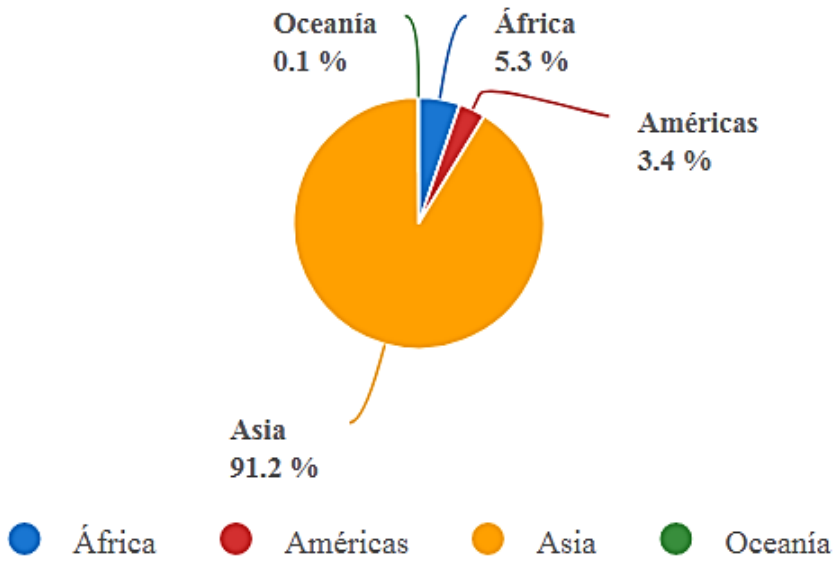

Figura 2. Producción de hule a nivel mundial. Fuente: FAOSTAT (julio 10, 2017)

Además de la problemática que presenta *Hevea* en cuanto al patógeno que lo ataca existe otro gran inconveniente con el látex de *Hevea* ya que se ha demostrado que existe una parte de la población a la cual este látex provoca alergia Tipo I (Siler y Cornish, 1994a) causando hipersensibilidad en pacientes , trabajadores al cuidado de la salud y personas en general que estén en contacto constante con productos derivados del hule como guantes, condones, entre otros (Ruëff, 2014).

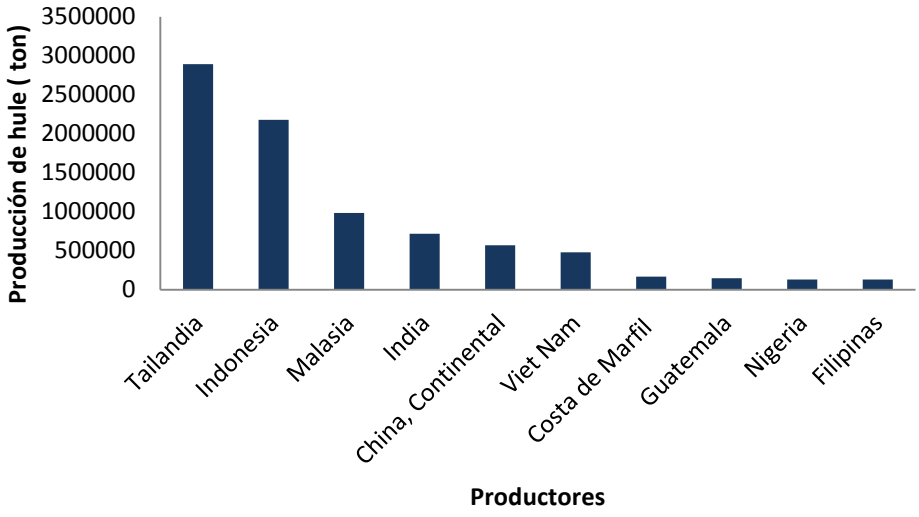

Figura 3. Principales productores de hule a nivel mundial. Fuente: FAOSTAT (julio 10, 2017)

Debido a la problemática en la que se ve envuelta *H. brasiliensis* y que depender de una sola fuente de hule natural no es lo más conveniente, se han buscado cultivos alternativos como fuente de hule, dentro de los que se han destacado *T. kok-saghyz* y *P. argentatum* (Cornish *et al.*, 2015; van Beilen y Poirier, 2007). Demostrando que el hule extraído de estas plantas es similar al que provee *H. brasiliensis* (tabla 1) en cuanto al peso molecular y propiedades.

Tabla 1. Comparación de especies productoras de hule

Fuente	Contenido (%)	Peso Molecular (kD)	Producción Ton/año	Rendimiento kg ha^{-1} a^{-1}	Referencias
Hevea brasiliensis	30-50% de látex 2% de peso seco del árbol	1,310	12,401,000 (2016)	500-3,000	(Blanc *et al.*, 2006; Han *et al.*, 2000; Kim *et al.*, 2004; Mooibroek y Cornish, 2000; Priya *et al.*, 2006)
Parthenium argentatum Gray	3 - 12% del peso seco de la planta	1,280	10,000 (1910)	300-2,000	(Kim *et al.*, 2004; Mooibroek y Cornish, 2000; Polhamus, 1962; Swanson *et al.*, 1979; Van Beilen y Poirier, 2008)
Taraxacum koksaghyz	25 – 40% del peso seco de la planta	2,180	3,000 (1943)	150-500	(Hallahan y Keiper-Hrynko, 2007; Mooibroek y Cornish, 2000; Whaley y Bowen, 1947)

Sin embargo, *P. argentatum* presenta una ventaja sobre *T. kok-saghyz* y *H. brasiliensis* que ha causado la atención de los investigadores*;* esta especie produce hule hipoalergénico debido a que no contiene proteínas que causen algún tipo de reacción alérgica (Cornish, 1996; Cornish *et al.*, 2015)

3.2. Planta de Guayule

P. argentaum Gray (figura 4) es un arbusto perenne comúnmente llamado guayule (Benedict *et al.*, 2008)**.** Es un miembro de la familia Asteraceae, arbusto leñoso que mide entre 0.3 – 0.6 m (Benedict *et al.*, 2008), presenta flores amarillas y hojas verde – grisáceo (Wood, 2002). Este cultivo crece naturalmente en altitudes de 600 – 2,000 msnm el lugares con una precipitación anual de 250 – 380 mm. El rango de temperatura en el que se desarrolla es de 18 – 49.5 ^{o}C, son

consideradas como plantas tolerantes a las altas temperaturas pero se ha demostrado que entran en un proceso de dormancia cuando la temperatura se encuentra por debajo de los 4 ^{o}C (Thompson y Ray, 1989). El pH óptimo para su crecimiento va de 6 – 8 en suelos calcáreos y bien drenados.

Figura 4. Planta de *Parthenium argenatum* en condición silvestre

Presenta una reproducción sexual para las plantas diploides (2n = 2x =36), mientras que las plantas que son triploides (2n = 3x =54), tetraploides (2n = 4x = 72) y octaploides (2n = 8x = 144) se reproducen asexualmente por apomixis (Bekaardt *et al.*, 2010; van Beilen y Poirier, 2007).

La planta de guayule es nativa del suroeste de Texas y México, donde se encuentra distribuida en Chihuahua, Coahuila, Durango, Nuevo León, San Luis Potosí y Zacatecas (figura 5).

Figura 5. Distribución del guayule en la República mexicana (de Rodríguez *et al.*, 2006b)

Se han realizado estudios para identificar variaciones de diferentes características agronómicas entre poblaciones de los estados antes mencionados, además de ubicar áreas potenciales que contengas plantas de guayule con recursos genéticos superiores. Dichos estudios demostraron que las plantas provenientes de Nuevo León, San Luis Potosí y Zacatecas fueron las que presentaron un menor contenido de hule (1.56, 1.56 y 2.13 % respectivamente) con respecto a Durango que presento un promedio de 5.79 % de hule (Kuruvadi *et al.*, 1997a). Datos que fueron confirmados en 2006, ya que dicho estudio demostró que el estado donde se presentaba el mayor contenido de hule (21 %) era Durango específicamente en el área de Mapimí mientras que el menor contenido de hule provenía de las

plantas colectadas en Rocamontes, Zacatecas con 5 % (de Rodríguez *et al.*, 2006b).

Se conoce que la producción de látex en los tallos de *P. argentatum* se origina en las células del parénquima que rodean a los conductos de la resina y después en las células adyacentes del parénquima. Con la maduración, las partículas de látex aparecen en la vacuola de ambas células y es en la vacuola donde son almacenadas (figura 6). Las partículas asumen diferentes formas dependiendo del organelo en el que se encuentren, irregular o globoide si está en el citoplasma o esférico si se encuentra en la vacuola (Backhaus y Walsh, 1983; Goss *et al.*, 1984). Se tienen identificadas las partes en las cuales se encuentra la acumulación de hule en la planta, y se ha demostrado que la parte en la que existe una mayor acumulación de hule es en el tallo principal (de Rodríguez *et al.*, 2001).

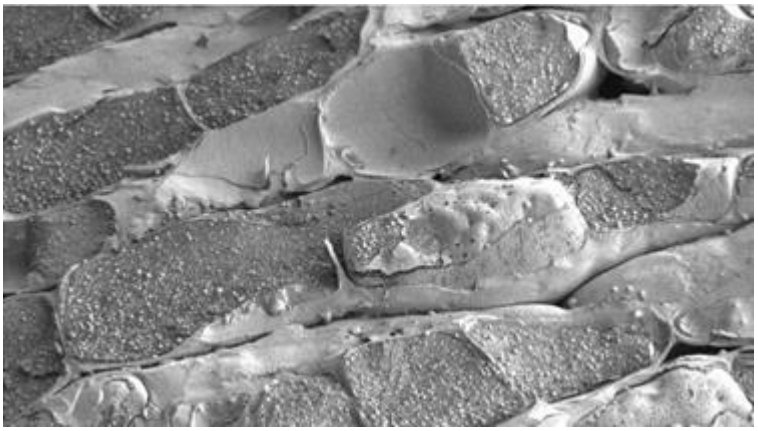

Figura 6. Micrografía electrónica de barrido de células del parénquima de corteza de tallos *de Parthenium argentatum*, mostrando vacuolas llenas de partículas de látex (Fotografía de Delilah Wood, USDA)

En 2001 se reportó el experimento realizado con 35 plantas silvestres de guayule, colectadas en Gómez Farías, Coahuila, las cuales pasaron por un proceso de

secado y división en cuanto a raíz y tallo. Para después hacer la extracción y cuantificación de hule y resina presentes tanto en raíz como en tallo, dando como resultado que la mayor acumulación de hule se encuentra en el tallo con un 10.24%. (de Rodríguez *et al.*, 2001). Dato que sería comprobado años más tarde al realizar un experimento similar, en el cual nuevamente se seccionaron las partes de la planta pero esta vez en raíz, tallo y hojas. Donde el tallo vuelve a presentar un porcentaje mayor de hule (figura 7) (de Rodríguez *et al.*, 2006b; Teetor *et al.*, 2009).

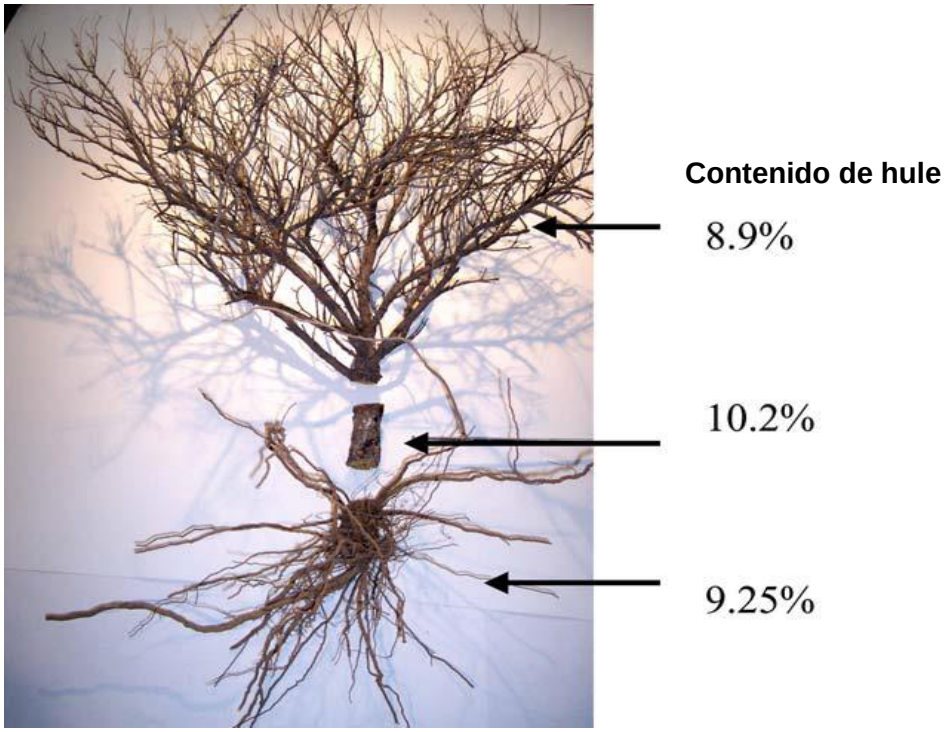

Figura 7. Planta de guayule seccionada en partes: raíz, tallo y hojas (de Rodríguez *et al.*, 2006b).

Desde el punto de vista de las aplicaciones de hule en diversos ámbitos y productos, la resina resulta ser un componente indeseable ya que por sus

características como bajo peso molecular puede actuar como plastificante, de esta forma las películas a base de hule provenientes de guayule presentan el riesgo de ser menos firmes y fácil de distorsionar (Jones, 1948). Sin embargo, este componente puede ser aprovechado en la fabricación de tableros de alta densidad, resistentes a la degradación por termitas; la protección de maderas contra algunos organismos destructivos la cual se realiza por medio de la impregnación de dichas maderas con el extracto de resina; también el bagazo proveniente de guayule que aún contiene resina se ha utilizado en combinación con polímeros epoxi para fabricar recubrimientos desprendibles utilizados para la protección de diversos equipos (Nakayama, 2005). Además se ha encontrado que en los productos activos de la resina se encuentran dos sesquiterpenos y un triterpeno que presentan actividad tóxica para los insectos (Gutiérrez *et al.*, 1999). Ya que la resina ha resultado tener aplicaciones bastante interesantes diversos autores se han enfocado en encontrar nuevas aplicaciones de dicho material además de hacer más rentable el cultivo de guayule. Una de las aplicaciones más recientes es atrapar compuestos volátiles como sabores, perfumes, aceites esenciales y pesticidas, por medio de microcápsulas de alginato o almidón combinados con la resina de guayule para una liberación controlada (Pascual-Villalobos y López, 2013).

3.3. MÉTODOS DE EXTRACCIÓN DE HULE Y RESINA

Existen diversos métodos para la extracción tanto de hule y resina proveniente de *P. argentatum* de los cuales se debe tomar en cuenta los tiempos de extracción,

cantidad de muestra utilizada y lo más importante la cantidad de hule y/o resina obtenida (Salvucci *et al.*, 2009)

Dentro de los más utilizados para la extracción de estos componentes en la especie *argentatum* son: Extracción con molino de bolas (el cual es reconocido por haber sido llevado a escala industrial), Soxhlet, y extracción acelerada con solventes (ASE por sus siglas en inglés accelerated solvent extraction) los cuales se encuentran a una escala de laboratorio. Los métodos como Soxhlet (Benedict *et al.*, 2010) y ASE (Pearson *et al.*, 2013) presentan un factor común, ambos se realizan mediante dos rondas de extracción una para quitar resina y la otra para extraer hule, típicamente se utilizan solventes polares como etanol, metanol, isopropanol y acetona en la primera extracción seguida de la extracción con solventes como tolueno, ciclohexano y hexano para la obtención de hule (Cornish *et al.*, 2007)

3.3.1. Extracción con molino de bolas

Fue uno de los primeros métodos de extracción de hule propuesto por George Miller en 1945, el cual consiste en usar métodos de coagulación del hule en la planta de guayule, seguido de una trituración para llevar al molino con guijarros o comúnmente llamados bolas, la molienda se lleva a cabo colocando los trozos de la trituración en una especie de recipientes o reactores cilíndricos revestidos de porcelana, acero o acero inoxidable (figura 8). Los recipientes giran en sentido horizontal a su eje mayor, al rodar los guijarros, uno sobre otro y contra los costados del reactor en el que se encuentran la pulverización se produce con una

pérdida mínima de material vegetal (Gennaro, 2003). Para después llenar estos reactores con agua que hacía que el bagazo se hundiera y quedará en la superficie el hule y algunas impurezas, para eliminar las impurezas se aplicaba calor y presión para hacer que las impurezas se fueran al fondo del cilindro y el hule se aglomera y flotará, recuperando y lavando los aglomerados de hule, posteriormente estos eran secados y empaquetados para su envió (Byrne, 1975).

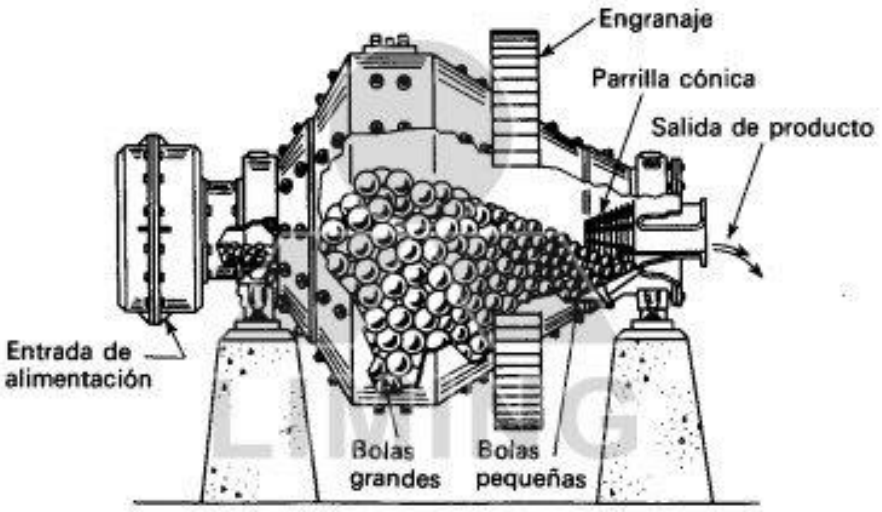

Figura 8. Imagen ilustrativa de molido de bolas y sus partes

3.3.2. Extracción Soxhlet

Es una técnica usada para materiales sólido-líquidos pero que resultó ser eficiente en algún tiempo para la extracción de hule y resina proveniente de guayule, la cual fue denotada como extracción estándar para este tipo de material (Black *et al.*,

1985). El método de extracción consiste en colocar el solvente para la extracción en un matraz bola, el cual se hace hervir por medio de parrillas eléctricas para que los vapores del solvente asciendan por el extractor y se condensen en el refrigerante cayendo gota a gota sobre el cartucho que contiene la muestra molida de partes de la planta de guayule y la parte soluble pasa a través del cartucho (figura 9). Cuando el líquido alcanza el punto de desbordamiento, un sifón es el encargado de tomar el solvente y descárgalo de nuevo en el matraz de destilación (Luque de Castro y García-Ayuso, 1998)

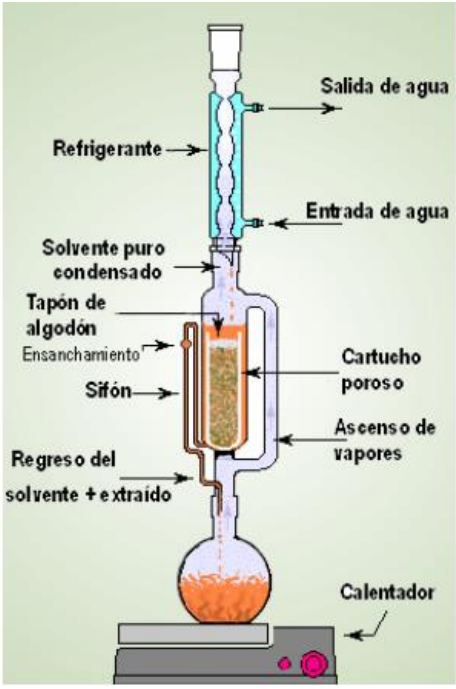

Figura 9. Imagen ilustrativa del método de extracción Soxhlet. Fuente: (Núñez, 2008)

Uno de los primeros experimentos con este tipo de extracción demostró que era relativamente más rápido, ya que la resina que contenía la muestra se había podido eliminar durante 2 horas, para después mantenerse la muestra en hexano

durante una noche completa seguida de una extracción que también duro dos horas (Nurthen *et al.*, 1986). Los productos de la extracción eran evaporados y el contenido de resina como de hule era obtenido por diferencia de pesos. Los autores de este experimento afirmaron que el tiempo de extracción era mucho menor con esta técnica y que la recuperación de hule y resina era más significativa, sin embrago, con este método el hule tuvo una tendencia a la degradación por el hecho de llevar la acetona hasta punto de ebullición para poder extraer la resina (Kopicky, 2014)

Otro de los trabajos que se realizó mediante esta técnica fue con semillas de plantas modificadas otorgadas por la Universidad Autónoma Antonio Narro, las semillas fueron germinadas bajo condiciones de invernadero, a los 15 días las plántulas se trasplantaron a bolsas individuales y mantenidas bajo invernadero hasta su siembra a campo abierto en Ocampo, Coahuila. Cuando las plantas tenía 32 meses de edad se realizó un muestro para determinar la cantidad de hule en diferentes partes de la planta mediante la técnica de Soxhlet, obteniendo porcentajes de hule en el tallo principal que van de 4.80 – 12.05%. La cantidad de muestra utilizada en este trabajo fue de 5 g, cantidad con la que se obtuvieron buenos resultados en cuanto a la extracción (Kuruvadi *et al.*, 1997b)

3.3.3. Extracción Acelerada con Solventes

Esta técnica presenta una importante ventaja sobre los procesos de extracción antes mencionados debido a que reduce el manejo de solventes orgánicos, aumenta la velocidad y eficiencia de la extracción utilizando alta presión, sola o en

combinación de alta temperatura (Thurbide y Hughes, 2000). Además de proteger al operador ya que este presenta un riesgo relativamente bajo de exposición al solvente, es completamente automatizado, tiene una mejor repetibilidad y presenta la opción de métodos personalizados (Lou *et al.*, 1997; Pearson *et al.*, 2010; Pearson *et al.*, 2013)

En este método, la extracción tiene lugar dentro de la celda de extracción, donde se coloca la muestra previamente pulverizada, junto con tierra de diatomeas o arena de cuarzo para eliminar cualquier humedad que pueda interferir en el proceso de extracción. Por las características de la muestra, es necesario colocar uno o más filtros, para evitar el paso de la propia muestra y/o tierra de diatomeas/arena de cuarzo durante la extracción al vial que contiene el extracto final o para evitar que se presenten problemas de taponamiento. El proceso de extracción puede repetirse varias veces (varios ciclos) y a cada ciclo se le asigna un tiempo de contacto entre la muestra y el solvente, de acuerdo a los resultados que se deseen obtener.

Debido a las prometedoras ventajas que esta técnica presenta sobre los demás tipos de extracción se han realizado varios estudios para determinar la eficiencia de este método. En 2009, se colectaron plantas de guayule de la línea AZ-2 las cuales fueron defoliadas, trituradas y tamizadas. El resultado de este proceso fue congelado a -80 ºC durante toda la noche para después ser secado por medio de liofilización durante 24 horas. El material fue almacenado a temperatura ambiente en un desecador durante 4 semanas, para después ser pulverizado, sellado y mantenido a temperatura ambiente hasta la extracción. El proceso de extracción

fue mediante el equipo Dionex ASE 200 (Bannockburn, IL). Se mezcló 1 g de muestra pulverizada con arena Ottawa. La resina y hule se extrajeron con acetona y ciclohexano respectivamente con temperaturas de 24 y 100 ^{o}C a una presión de 1500 psi; dando como resultado un promedio de 9.1 % para resina y 4.6 % de hule (Salvucci *et al.*, 2009).

Pearson *et al*., (2013), utilizaron la técnica ASE para cuantificar con mayor precisión el hule y resina proveniente de plantas mejoradas de *P. argentaum* de la línea AZ-2, las cuales fueron molidas para su posterior extracción mediante el equipo Dionex 200, a una presión de 10.3 MPa, se utilizaron tres ciclos de 20 min para extraer resina con acetona a 40°C y el hule fue extraído con ciclohexano o hexano a 140 ^{o}C. Dando como resultado una extracción más rápida y con cantidades menores de solvente utilizado (Pearson *et al.*, 2013).

3.4. RUTA DE SÍNTESIS DEL HULE

Los isoprenoides vegetales comprenden un amplio grupo de compuestos; los cuales pueden ser divididos en metabolitos primarios y metabolitos secundarios (figura 10). Los metabolitos primarios son aquellos que resultan ser esenciales para el crecimiento y desarrollo de la planta ya que participan en la integridad de la membrana, la fotoprotección y anclaje de funciones bioquímicas esenciales para los sistemas específicos de la membrana; tal es el caso de esteroles, carotenoides, reguladores de crecimiento, nucleótidos, proteínas, entre otros. Mientras que para algunos metabolitos secundarios aún no se tiene bien definida su función; se ha demostrado que tienen importantes interacciones entre la planta

y el ambiente en el que se desarrolla. Dentro de los metabolitos secundarios se incluyen los monoterpenos, sesquiterpenos y diterpenos (Croteau *et al.*, 2000; Chappell, 1995)

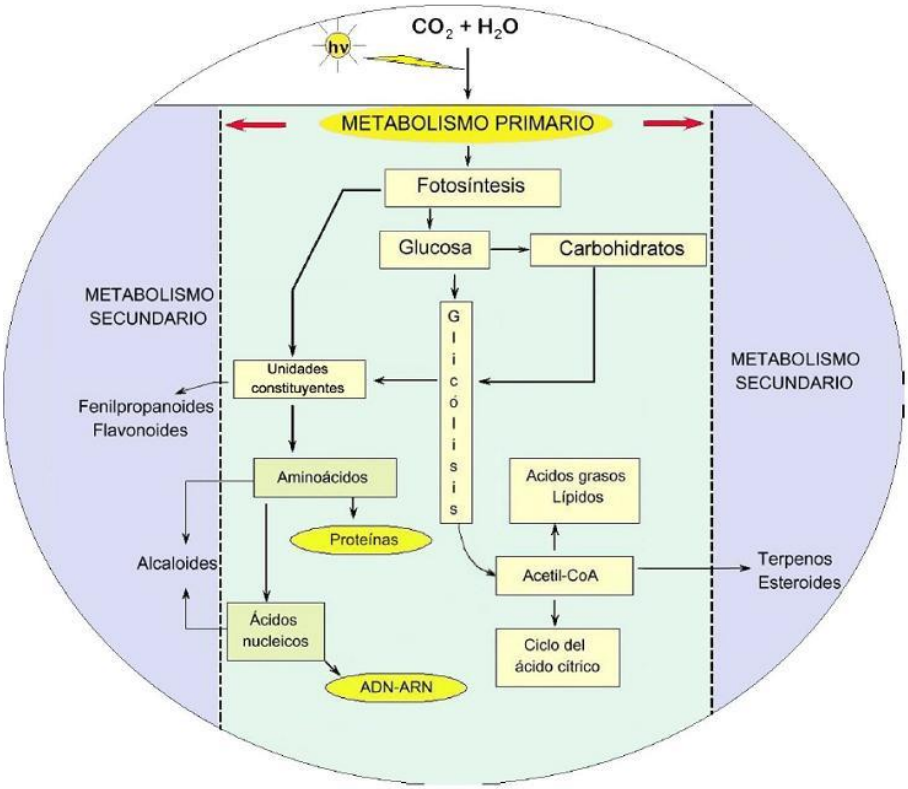

Figura 10.Ilustración del metabolismo primario y secundario de las plantas (García y Carril, 2011)

Todos los isoprenoides se derivan de dos unidades básicas, el isopentenil difosfato (IPP) y su isómero dimetil-alil difosfato (DMAPP). En las plantas, estos precursores se sintetizan por dos vías. La vía mevalónica, presente en el citoplasma y la vía del 2C-metil-D-eritrol 4-fosfato (MEP), que está confinada a los plástidos (Enfissi *et al.*, 2005).

3.4.1. Vía Mevalónica (MVA)

Esta vía consiste en sintetizar IPP a partir de un intermediario de seis carbonos (mevalonato), que surge de la condensación secuencial de tres unidades de acetil-CoA catalizada por tiolasa e hidroximetilglutaril-CoA sintasa para generar 3-hidroxi-3-metilglutaril-CoA (HMG-CoA), posteriormente una reducción por medio de una reacción irreversible catalizada por HMG-CoA reductasa para formar ácido mevalónico. El mevalonato de seis carbonos es secuencialmente fosforilado y descarboxilado para generar IPP (Croteau *et al.*, 2000; Chappell, 1995)

3.4.2. Vía 2C-metil-D-eritrol 4-fosfato (MEP)

Esta vía ha sido demostrada en algas y eubacterias así como en las plantas. Aquí el piruvato reacciona con el pirofosfato de tiamina (TPP) para producir un fragmento de dos carbonos, hidroxieltil-TPP que se condensa por medio del 3-fosfato de gliceraldehído. El TPP es liberado para formar un intermediario de cinco carbonos 1-desoxi-D-xilulosa 5-fosfato, que se reordena y reduce para formar 2-C-metil-D-eritritol-fosfato y posteriormente se transforma para producir IPP (Lois *et al.*, 1998; Rodríguez-Concepción y Boronat, 2002).

Tanto en la vía MVA como la de MEP a partir del IPP se generan otros compuestos por medio de las enzimas preniltransferasas que generan los ésteres de difosfato alílico como es el caso de difosfato de geranilo (GPP), difosfato de farnesilo (FPP) y difosfato de geranilgeranilo (GGPP) (Chappell, 1995).

El hule natural pertenece a la clase de isoprenoides naturales, que utilizan al IPP como precursor principal (figura 11); donde IPP y su isómero difosfato alílico (DMAPP) sufren condensaciones para generar GPP, FPP y GGPP (Tholl, 2006). Siendo FPP el sustrato de elección para la polimerización del hule (da Costa *et al.*, 2005) además de cofactores de iones metálicos como Mg^{2+}. De esta manera varias moléculas de IPP se conjugan en una molécula de FPP en configuración *Cis* -1,4 por medio de reacciones catalizadas por una transferasa de hule (EC 2.5.1.20) (Cornish, 1993; Light y Dennis, 1989).

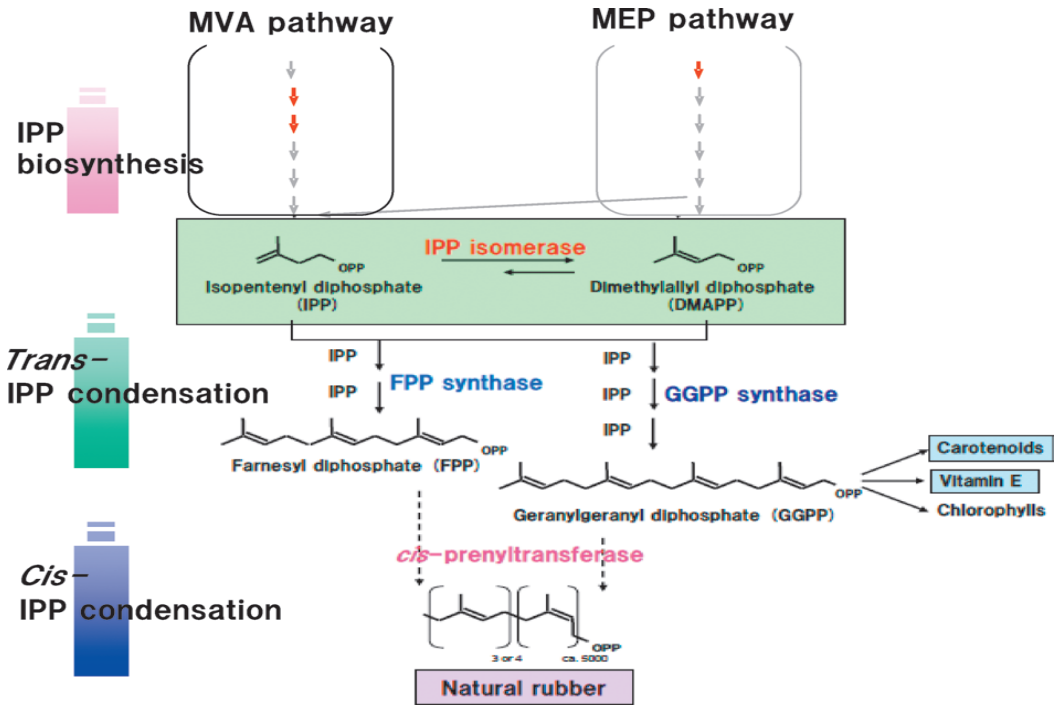

Figura 11. Biosíntesis de hule natural (Hayashi, 2009)

3.4.3. Genes relacionados con la partícula de hule

Como se mencionó anteriormente, el hule natural (*Cis* 1-4 poli-isopreno) se forma dentro de las células vegetales a partir de iniciadores de difosfato alílico e isopentenil difosfato, y se empaqueta en partículas de hule subcelulares. Varias proteínas se encuentran asociadas a estas partículas formando un complejo proteico, la cantidad de estas proteínas varía notablemente entre las especies productoras de dicho material (Siler et al., 1997), sin embargo, todas contienen una enzima cis-preniltransferasa como integrante de un complejo enzimático (transferasa de hule EC 2.5.1.20), la cual ayuda a la síntesis de hule (Siler y Cornish, 1994b)

A pesar de que la actividad y la descripción completa de la tranferasa de hule aún no se ha esclarecido, en *H. brasiliensis* se identificaron dos proteínas, una es el factor de elongación de hule (REF por sus siglas en inglés Rubber Elongation Factor) (Dennis y Light, 1989) y la pequeña partícula de hule (SRPP) (Oh *et al.*, 1999) que intervienen en dicho complejo. Estas proteínas bastante abundantes en las partículas de hule (Asawatreratanakul *et al.*, 2003), parecen ser codificadas por un conjuntos de genes según un estudio genómico recientemente realizado por Lau *et al.*, (2016), que encontraron nueve y ocho genes, correspondiente a REF y SRPP, respectivamente. También identificaron, la familia completa de los genes del complejo de la cis-preniltranferasa (CPT: de 1 a la 11) (Lau *et al.*, 2016; Tang *et al.*, 2016); que como ya se menciono es otro componente importante del complejo enzimático de la transferasa de hule en *H. brasiliensis*.

En *T. kok-saghyz*, se han identificado, las proteínas: CPTs (TbCPT de 1 a 3), SRPP y REF, todos comparten una alta similitud de secuencia con las presentes en *H. brasiliensis*. Las proteínas CPTs en *Taraxacum* también son asociadas con las partículas de hule y expresadas predominantemente en laticíferos (Schmidt *et al.*, 2010a; Schmidt *et al.*, 2010b); el silenciamiento de los genes correspondientes a estas tres proteínas, inhibió la síntesis de hule natural de alto peso molecular y provocó una menor actividad de la hidroximetilglutaril-coenzima A reductasa (HMGR), la enzima clave en la vía del mevalonato (Post *et al.*, 2012). Sin embargo, en un estudio reciente, no encontraron una expresión diferencial de estos genes entre plantas de alto y bajo contenido de hule (Luo *et al.*, 2017); los autores no llegaron a una explicación evidente en cambio, dejaron en duda su papel directo en la biosíntesis del hule. Oh *et al.* (2000) demostraron que CPT de *Arabidopsis* solo, no podía catalizar la formación de hule de alto peso molecular. Por lo cual, Kim *et al.,* (2004) sugirieron que GHS o el SRPP pueden constituir parte del complejo tranferasa de hule y desempeñar un papel en la incorporación de IPP para la obtención de hule de alto peso molecular.

La actividad de la CPT, se describió en guayule por la primera vez en 1984 (Madhavan y Benedict, 1984); la evaluación de la expresión genética de la CPT en plantas de guayule sometidas a condiciones frías no mostró una correlación entre esta y la temperatura. Sin embargo, un aumento brusco en la temperatura nocturna estimulo la expresión de esta proteína (Ponciano et al., 2012). Otra proteína de alta importancia en guayule es la GHS es una proteína pequeña que se localiza alrededor de las partículas de hule en las células epiteliales (Kajiura *et*

al., 2017) comparte 50% de similitud con SRPP y probablemente comparten el mismo papel (Kim et al., 2004). El análisis funcional in vitro de la proteína recombinante sobreexpresada en *E. coli* mostró que el GHS desempeña un papel positivo en la incorporación de isopentenil pirofosfato en hule de alto peso molecular como lo hace SRPP (Kim et al., 2004).

Con el objetivo de incrementar los rendimientos de hule, se han realizado diversos estudio y consiguiente descubrimiento de los genes relacionados a su producción; esto con ayuda de diversas técnicas moleculares utilizadas en la secuenciación tanto del genoma como del transcriptoma de diferentes tejidos de las tres plantas con mayor potencial de explotación *H. brasiliensis* (Chow *et al.*, 2014; Ko *et al.*, 2003; Li *et al.*, 2012; Rahman *et al.*, 2013; Tan *et al.*, 2017; Tang *et al.*, 2016; Triwitayakorn *et al.*, 2011; Xia *et al.*, 2011), *Taraxacum kok-saghyz* (Lin *et al.*, 2017; Luo *et al.*, 2017) y *Parthenium argentatum* (Ponciano *et al.*, 2012). Sin embargo, como se puede evidenciar, el número de trabajos entre las diferentes especies es alto.

3.4.3.1. Importancia del transcriptoma

El análisis del transcriptoma se ha vuelto cada vez más importante a través de los avances en las tecnologías de secuenciación de siguiente generación de compañías como los son Illumina, Roche y Life Technologies. Además de ser una alternativa viable que permite identificar genes presentes en un organismo determinado, ya que la secuenciación y ensamblado de genomas presenta mayor dificultad en la mayoría de las especies vegetales a causa de la alta

heterocigosidad, el alto contenido de elementos transponibles y los eventos de duplicación a los que ha llevado la evolución (Harrison y Kidner, 2011; Wang *et al.*, 2009).

Las mejoras en la secuenciación del transcriptoma en la secuencia química y la longitud de lecturas han permitido una cantidad de secuencias sin precedente, que se ve limitada principalmente por el costo y la disponibilidad de material biológico (Van Verk *et al.*, 2013). En particular, la secuenciación de ARN (RNA-Seq) surge como una valiosa herramienta para perfilar genes expresados en plantas y otros organismos en una etapa determinada (Ozsolak y Milos, 2010; Van Verk *et al.*, 2013; Wang *et al.*, 2009).

A la par del progreso en las tecnologías de secuenciación, numerosos softwares han sido desarrollados para el ensamblado *de novo* del transcriptoma. Entre los softwares más utilizados se encuentran Velvet (Zerbino y Birney, 2008), SOAPdenovo (Li *et al.*, 2010), Trinity (Grabherr *et al.*, 2011), MIRA (Chevreux *et al.*, 2004) y Newbler (perteneciente a la compañía Roche). Adicionalmente al ensamblaje *de novo* del transcriptoma se pueden realizar procedimientos que mejoren la calidad de dicho proceso, como lo es utilizar más de un ensamblador, ya que ayuda en la optimización del número de “contigs” y sus longitudes (Lin *et al.*, 2011; Zhao *et al.*, 2011).

Herramientas como BLAST (por sus siglas en inglés The Basic Local Alignment Search Tool) ayuda en la búsqueda de regiones de alta similitud entre dos secuencias, bien sea a nivel nucleótido o proteico. Lo cual después de la

reconstrucción del transcriptoma es parte vital para la anotación funcional de los contigs generados por los ensambladores, lo que se puede fundamentar por medio de la homología presente entre una secuencia y otra. Las bases más utilizadas en la anotación funcional son Swiss-Prot, TrEMBL, GO y KEGG (Lau *et al.*, 2016; Remm *et al.*, 2001).

IV. METODOLOGÍA

4.1. Colecta de material vegetal

Las colectas del material vegetal silvestre se realizaron en los estados de Durango y San Luis Potosí. Localidades que en los años 1990's se identificaron como unas de las de mayor y menor contenido de hule respectivamente (de Rodríguez *et al.*, 2006b). Se realizaron colectas en la temporada de verano (septiembre, 2016) que es la temporada en la que se tiene reportada la menor acumulación de contenido de hule y en invierno (febrero, 2017) que es cuando se tiene registrado se acumula el mayor contenido de hule (Angulo-Sánchez *et al.*, 2002). La colecta en diferentes localidades y épocas del año, permitió asegurar la identificación de al menos un sitio con material vegetal de alto contenido de hule y otro con bajo, esto para poder contrastar la sobreexpresión de genes involucrados en la biosíntesis de hule en los análisis transcriptómicos.

Los sitios de colecta de material vegetal se eligieron considerando diferentes características edafoclimáticas, las coordenadas geográficas, asociación de plantas endémicas, características fisiológicas y de exposición. De cada sitio identificado se tomaron tres muestras de tallo correspondientes a tres plantas diferentes para su posterior análisis.

Para la extracción de hule y resina se colectaron segmentos en fresco de tallos de 10 a 15 mm de diámetro y 7 cm de longitud (figura 12) para homogeneizar las muestras colectadas (Cornish *et al.*, 2000), las cuales fueron secadas durante 48 h a 45 ^{o}C al llegar al Centro de Investigación en Química Aplicada (CIQA).

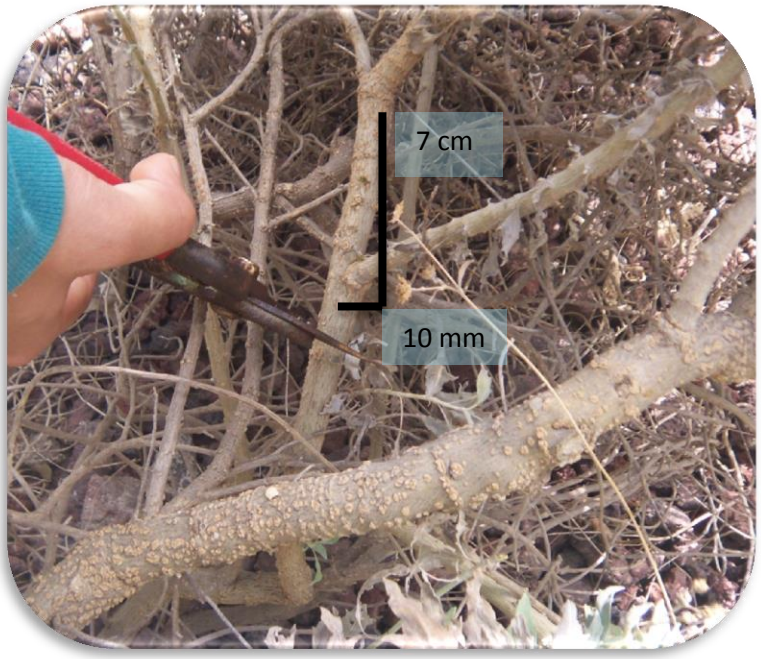

Figura 12. Colecta de tallos de *Parthenium argentatum*

A la par, de las mismas plantas, se realizaron colectas de corteza de tallos, las cuales fueron congeladas de inmediato en nitrógeno líquido y mantenidas a -80 ^{o}C, hasta la extracción de ácido ribonucleico (ARN) para su posterior análisis transcriptómico.

4.2. Extracción y contenido de hule

El hule natural y resina de las plantas de guayule fueron extraídos a partir de muestras de tallo molido en un molino ultra centrifuga (Retsch, ZM 100. USA) a 1400 rpm durante 10 segundos seguido de un proceso de secado a 45 ^{o}C durante 24 h (figura 13).

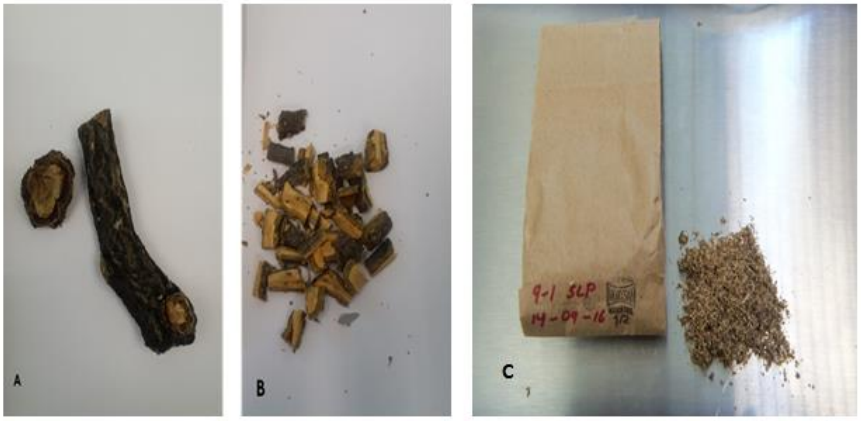

Figura 13. Proceso de molienda de tallos de guayule. a) tallo de guayule. b) tallo de guayule cortado en trozos. c) tallo de guayule molido

La extracción de ambos compuestos se realizó usando un equipo de extracción automatizado Speed extractor (figura 14) (BÜCHI, E-916. Suiza). Este equipo es completamente automatizado y utilizado para la extracción en paralelo de compuestos orgánicos principalmente de una gran variedad de muestras sólidas o semisólidas. Básicamente lo que hace este equipo es aplicar temperatura y presión a un solvente determinado que se encuentra dentro de una celda de extracción aplicando múltiples repeticiones si así se desea, para obtener altos porcentajes de recuperación. Al terminar se realiza un enfriamiento por medio de solvente frío y la eliminación de este se realiza mediante la inyección de nitrógeno (gas).

Figura 14. Equipo automatizado Speed Extractor E-916, marca BÜCHI y sus principales componentes (a) Celda de extracción, b) Equipo Speed Extractor E-916, c) Reservorio de solventes, d) Viales de recolección).

Preparación de la celda de extracción: cada ciclo de extracción puede ser de hasta 5 celdas de 40 ml, cada una se preparó de la siguiente manera:

1. Filtro de celulosa
2. 5 gramos de arena de cuarzo (BÜCHI de 0.3-0.9 mm. Suiza)
3. Filtro de celulosa
4. Mezcla de arena de cuarzo y muestra (2.5 g, 1 g y 0.5 g según corresponda)
5. 5 gramos de arena de cuarzo

6. Filtro de celulosa

Condiciones de extracción programadas en el equipo de extracción: La obtención de resina y hule se realizó con 2 rondas de extracción en base al procedimiento de Pearson *et al.* (2013) con algunas modificaciones. La primera ronda consistió en extraer la resina de la muestra usando acetona como solvente, para después extraer el hule presente utilizando hexano. En la tabla 2 se muestran las condiciones bajo las cuales se realizaron las rondas de extracción.

Tabla 2. Condiciones bajo las cuales se realizó la extracción de resina y hule provenientes de muestras de Guayule

	Extracción resina	**Extracción de hule**
Parámetro	**Característica**	**Característica**
Temperatura	40 ^{o}C	60 ^{o}C
Presión	100 bares	100 bares
Solvente	Acetona 100%	Hexano 100%
Celdas	40 ml	40 ml
Vial	240 ml	240 ml
Ciclos	3	3
Calentamiento	3 min	3 min
Retención	5, 10, 15 min	5, 10, 15 min
Descarga	5 min	5 min
Enjuague con solvente	3 min	3 min
Enjuague con gas	3 min	3 min

Cada extracto obtenido, fue colectado en viales de 240 ml previamente etiquetados para su posterior concentración usando un equipo de múltiple evaporación (BÜCHI, Multivapor™ P-6., Suiza) a una temperatura de 40 ^{o}C y 60 ^{o}C para resina y hule respectivamente. Al obtener el extracto lo más concentrado posible (20 ml aproximadamente) este fue recolectado en viales previamente

pesados y etiquetados. Posteriormente los viales fueron puestos en un liofilizador (LABCONCO, Free Dry System/Freezone 4.5. USA) durante 24 h para eliminar el solvente y llevar las muestras hasta la desecación. Los cálculos para determinar el porcentaje de resina o hule presente en las muestras procesadas se realizaron aplicando la siguiente fórmula:

$$Pvm- pv *100/ps$$

Pvm= peso del vial + muestra liofilizada

pv= peso del vial

ps = peso de muestra puesta en las celdas

En esta parte de la investigación se utilizó un diseño experimental completamente al azar, y una comparación de medias con la prueba de Duncan a un intervalo de confianza del 95 %.

4.2.1. Identificación de hule

Para la identificación de hule se utilizó la técnica de Resonancia Magnética Nuclear (RMN) cuyo fundamento es la absorción de energía (radiofrecuencias) por un núcleo magnéticamente activo, que está orientado en el seno de un campo magnético, y que por efecto de esa energía cambia su orientación. La muestra de hule se diluyó en 5 ml de cloroformo a temperatura ambiente, una vez disuelto se colocaron entre 0.5 a 0.8 ml de muestras en tubos de RMN de 8 pulgadas de longitud; estos fueron introducidos en el equipo de RMN (Bruker, 400MHz, Alemania). Los datos obtenidos fueron procesados con el software TopSpin®, 4.0.

4.3. Extracción y cuantificación de ARN total

4.3.1. Extracción de ARN total

Se eligieron 4 sitios que presentaron resultados más contrastantes en cuanto al porcentaje de hule. Las extracciones de ARN total se realizaron para las muestras de la corteza de 12 plantas procedentes de estos 4 sitios y conservadas a -80 °C. Para cumplir este objetivo se completaron dos protocolos uno de extracción y otro de purificación: 1) para la extracción se siguió el protocolo descrito por los fabricantes del reactivo *Tri-Reagent*, a este protocolo se incorporó un paso de centrifugación (12,000 xg, 10 min., 4 °C) tras la homogenización de la muestra con el reactivo *Tri-Reagent,* eso ayudó a liberarse de las grasas y los polisacáridos (ver ANEXO I); 2) la purificación se realizó siguiendo el protocolo de purificación de RNA, usando el kit RNeasy Micro kit (Qiagen) (ver ANEXO II). La comprobación de la calidad del producto obtenido de cada una de las muestras, se realizó por medio de la técnica de electroforesis, preparando un gel de agarosa al 1% en un tampón de corrida de TAE 1X (Tris, acetato y EDTA), corriendo durante una hora y 10 min a 70 volts. Una buena extracción debe generar un producto no degradado y libre de fragmentos de ADN genómico.

Además de la calidad, la técnica elegida para la secuenciación masiva (ver apartado 4.4) exige una cantidad mínima de 3 µg de RNA total y una concentración no menor de 50 ng/µl. Para cuantificar la cantidad de ARN extraído se usó el equipo Quantus^{TM} Fluor (E6150, EUA), siguiendo las instrucciones del

fabricante. Quantus™ Fluor está diseñado para proporcionar una detección fluorescente altamente sensible al cuantificar ácidos nucleicos y proteínas.

4.4. Secuenciación masiva de transcritos

Una vez verificada la concentración y calidad de las extracciones de ARN, las muestras se procesaron usando la plataforma de Secuenciación por Síntesis (SBS), con el sistema NextSeq 500 Medio Rendimiento, de Illumina. El cual tiene la capacidad para generar aproximadamente 35 Gb y un promedio de 110 millones de lecturas pareadas. Esta parte se realizó en el Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO) ubicado en Libramiento Norte, carretera a León km 9.6 código postal 36821, Irapuato, Guanajuato.

4.5. Integración de datos transcriptómicos

Esta parte de la investigación fue realizada en el Instituto de Ecología A.C (INECOL) ubicado en kilómetro 2.5 carretera antigua a Coatepec No. 351, Colonia Congregación el Haya, código postal 91070, Xalapa, Veracruz. En el laboratorio de Estudios Teórico computacionales a cargo del Dr. Enrique Ibarra Laclette y M.C Emmanuel Villafan de la Torre, como parte de una estancia de investigación.

4.5.1. Filtrado de calidad

Lo primero a realizar fue el filtrado de calidad para las secuencias pareadas obtenidas de la plataforma de Illumina. Esto fue realizado por medio de un “script” de Phyton con el cual se pueden automatizar ciertas tareas o acciones en un ordenador (https://github.com/Czh3/NGSTools/blob/master/qualityControl.py). Se

utilizaron parámetros como: -q = 30 que es el valor mínimo de la calidad permitida en formato Phred; -p = 90 siendo el porcentaje mínimo de bases en la secuencia que deben tener una calidad -q; -a = 29 que es el límite de la calidad promedio estimada a lo largo de la secuencia.

4.5.2. Ensamblado

Ya que fueron filtradas las secuencias y se obtuvieron secuencias pareadas con el estatus de calidad asignado en el paso anterior, se prosiguió a evitar la superposición de bases, usando el programa SeqPrep (https://github.com/jstjohn/SeqPrep), que su actividad principal es combinar las secuencias que tengan al menos 25 bases que se solapen en las posiciones inicial o terminal, para obtener una secuencia más larga.

El ensamblado de las secuencias se realizó utilizando el programa TRINITY, un ensamblador de transcriproma de *novo*, el cual emplea un algoritmo basado en la utilización de los gráficos de Brujin (Grabherr *et al.*, 2011). Al resultado del ensamblado de TRINITY se le llama comúnmente “contigs” (también conocidos como unigenes).

4.5.3. Limpieza

Para evitar la presencia de posibles agentes contaminantes en los contigs resultado del ensamblado con TRINITY, se utilizó el programa DeconSeq (http://deconseq.sourceforge.net/manual.html), esto debido a la procedencia de las muestras a las cuales se les extrajo el ARN, ya que como antes se mencionó

provienen de dos distintos estados de la República Mexicana y las cuales fueron colectadas en condición silvestre. Los contigs fueron comparados contra secuencias de referencia (virus, bacterias, hongos, familias de ARN no codificante y elementos de ADN repetitivo). Aquellas secuencias que tuvieron una identidad del 85% a lo largo del 90% de su longitud fueron eliminadas.

La parte final se realizó con el programa SeqClean, el cual fue usado para cortar y/o eliminar regiones terminales de baja complejidad, colas de poli A o poli T de los unigenes (https://sourceforge.net/projects/seqclean/files/). Dejando secuencias igual o mayor a 200 pb de longitud.

Con la finalidad de evitar cualquier redundancia se realizó un reensamblado de las secuencias resultantes de los procesos anteriores, para este fin fue utilizado el ensamblador Newbler el cual redujo considerablemente la cantidad de unigenes.

4.5.4. Identificación de Marcos de Lectura Abierta en los unigenes provenientes de *Parthenium argentatum*

Para realizar la identificación de Marcos de Lectura Abierta (ORF; por sus siglas en ingles Open Reading Frame) en los unigenes de Guayule, lo primero fue crear una base de datos con las secuencias de RNA mensajero (mRNA) de 10 diferentes especies que posteriormente ayudarían en el proceso de la identificación de las secuencias codificantes; las secuencias de los mRNA´s de 9 de las especies fueron descargas del Centro Nacional para la Información Biotecnológica (NCBI) (https://www.ncbi.nlm.nih.gov), se eligieron aquellas que fueran cercanas filogenéticamente y/o que fueran funcionalmente compatibles.

Como un caso especial *Arabidopsis thaliana* fue elegida debido a que se considera como el mejor genoma en existencia para especies vegetales (https://www.arabidopsis.org/). En la tabla 3 se muestran las especies elegidas para realizar la identificación de secuencias codificantes (CDS, por su siglas en inglés, coding sequence).

Tabla 3. Especies elegidas para llevar a cabo la identificación de marcos de lectura abierta en los unigenes de *Parthenium argentatum*

Subclase	**Familia**	**Genero**	**Especie**
Astéridas	Asteraceae	*Helianthus*	*annus*
	Solanaceae	*Solanum*	*lycopersicum*
	Solananceae	*Capsicum*	*annuum*
	Solananceae	*Nicotiana*	*tabacum*
	Pedaliaceae	*Sesamum*	*Indicum*
Rósidas	Euphorbiaceae	*Hevea*	*brasiliensis*
	Vitaceae	*Vitis*	*vinífera*
	Malvaceae	*Gossypium*	*hirsutum*
	Brassicaceae	*Brassica*	*rapa*
	Brassicaceae	*Arabidopsis*	*thaliana*

A cada una de las 10 bases de datos se le realizó la identificación de CDS mediante un "script" de Phyton (https://github.com/peterjc/picogalaxy/blob/master/tools/get_orfs_or_cdss/get_orfsor_cdss.py) (Cock *et al.*, 2009).

La identificación de ORF´s presentes en los unigenes de Guayule se realizó por medio del algoritmo AlingWise el cual está diseñado para identificar regiones de

secuencias codificantes para proteínas biológicamente relevantes mientras realiza una corrección en los cambios de marco (Frame-shift) (Evans y Loose, 2015).

4.5.5. Identificación de proteínas homólogas

Una vez que se tuvieron las CDS encontradas en los unigenes de Guayule y su respectiva traducción proteica. Se procedió a realizar la comparación de las proteínas encontradas en Guayule contra 7 bases de datos (tabla 4) con el objetivo de encontrar proteínas homologas. Esto se realizó por medio del algoritmo de BLASTp utilizando un umbral de e-value 10^{-5}. Las bases de datos fueron reducidas a 6 (en comparación a las 10 utilizadas para ORFs) ya que sólo se utilizaron aquellas que fueran filogenéticamente más cercanas (como es el caso de *H. annus, N. tabacum, S. indicum y S. lycopersicum*), funcionalmente compatibles (*H. brasiliensis*), la especie vegetal mejor anotada hasta el momento (*A. thaliana*) y se adicionó una base de datos de amplio espectro que contiene proteínas de especies vegetales y demás seres vivos (UniProt).

Tabla 4. Bases de datos utilizadas en la identificación de proteínas homólogas al realizar la comparación con las proteínas encontradas en *Parthenium argentatum*

Proteínas identificadas	Bases de datos
Parthenium argetatum	*Arabidopsis thaliana*
	Helianthus annus
	Hevea brasiliensis
	Nicotiana tabacum
	Sesamum indicum
	Solanum lycopersicum
	Base de datos UniProt

4.5.6. Anotación funcional a los unigenes de *P. argentatum*

La anotación funcional de los unigenes de *P.argentaum* se realizó a partir de las proteínas homólogas encontradas en la base de datos de UniProt otorgándole términos de ontología génica (GO) esto mediante la herramienta de anotación funcional en la base de datos David (https://david.ncifcrf.gov/).

4.5.7. Identificación de genes ortólogos

Primero se realizó una concatenación de las proteínas de las diferentes bases de datos: *A. thaliana, H. annus, H. brasiliensis, N. tabacum, S. indicum, S. lycopersicum,* las proteínas *de P. argenatum*, además de las proteínas identificadas en la síntesis de pirofosfato de isopentenilo (IPP), precursor principal en la síntesis de látex en las vías de mevalonato y del metileritrol fosfato, estas proteínas fueron descargadas de PlantCyc (https://www.plantcyc.org/) y MetaCyc

(https://metacyc.org/). Una vez que se tuvo un único archivo se procedió a realizar un BLASTp de todo contra todo con un e-value de 10^{-10}. Ya que se obtuvo el archivo resultante de BLAST, se prosiguió a la identificación de genes ortólogos. Este proceso fue realizado mediante un "script" de OrthoMCL que proporciona un método confiable para la construcción de estos grupos a través de múltiples taxones eucariotas, utilizando un algoritmo de Markov Cluster que ayuda a la agrupación de genes ortólogos y parálogos (putativos) (http://orthomcl.org/orthomcl/) (Li *et al.*, 2003).

Con el fin de contabilizar todos aquellos genes ortólogos que se encontraron en cada una de las especies se utilizó get_species_ortho.pl un "script" de perl.

4.5.7.1. Análisis filogenético

A las proteínas y CDS de los genes ortólogos que se encontraron en el mismo grupo se les realizó el análisis filogenético mediante SeaView 4, un programa que permite realizar dicho análisis por medio de un conjunto de secuencias de ADN o proteína homóloga, hasta la construcción de árboles y la visualización utilizando la alineación actualizada y algoritmos de construcción de árboles de máxima verosimilitud (Gouy *et al.*, 2010). La alineación de secuencias fue tipo MUSCLE (Edgar, 2004) y para la reconstrucción de los arboles filogenéticos de máxima verosimilitud se utilizó PhyML (Guindon y Gascuel, 2003), donde la tipología de todos los árboles se generó por medio del método de distancia BioNJ (por sus siglas en inglés Bio Neighbor Joining) siendo identificado el árbol óptimo después de la aplicación de los métodos de intercambio de ramas adyacentes (NNI por sus

siglas en inglés Nearest Neighbor Interchange) y el de poda y re-injerto de subárboles (SPR por sus siglas en inglés Subtree Pruning Regrafting). Por último, el valor del soporte de las ramas fue estimado por medio de la prueba de razón de verosimilitud aproximada (aLRT; por sus siglas en inglés approximate Likelihhod Ratio Test) (Anisimova y Gascuel, 2006).

4.5.8. Análisis de expresión diferencial

El análisis de expresión diferencial significa tomar los datos de lecturas normalizados y realizar un análisis estadístico para descubrir cambios cuantitativos en los niveles de expresión entre los grupos experimentales. El nivel de expresión se determina a partir de la abundancia de una secuencia en los datos obtenidos.

La expresión diferencial de los unigenes encontrados mediante el ensamblado del transcriptoma de guayule se realizó por medio de RSEM (por sus siglas en inglés RNASeq by Expectation Maximization) un paquete de software de manejo bastante amigable que ayuda a la cuantificación de abundancia de las transcripciones en los datos de RNA-Seq. Ya que no necesita de un genoma de referencia, es ideal para la cuantificación en ensamblados *de novo* del transcriptoma (Li y Dewey, 2011).

Una vez que fueron obtenidos los niveles de abundancia de los transcritos, se llevó a cabo el análisis de expresión diferencial a las 12 muestras; esto realizando contrastes entre los estados, las estaciones de colecta y los porcentajes de hule obtenidos; el análisis se realizó por medio de los paquetes edge R y DESeq.

V. RESULTADOS Y DISCUSIÓN

5.1. Colecta de material vegetal

Como resultado de las colectas de plantas silvestres de guayule, se lograron ubicar 12 sitios diferentes en el estado de Durango clasificados como D1 a D12, mientras que para San Luis Potosí se ubicaron 11 sitios diferentes clasificados como S1 al S11. La figura 15 y 16 muestra las ubicaciones de los sitios de colecta (encerrados en el círculo rojo) para ambos estados. De cada sitio identificado, se procedió a realizar la colecta de material vegetal para la cuantificación de hule y resina y para la extracción de ARN total. De cada sitio se colectaron tallos y cortezas de tres plantas diferentes en dos épocas del año dando un total de 138 muestras a analizar.

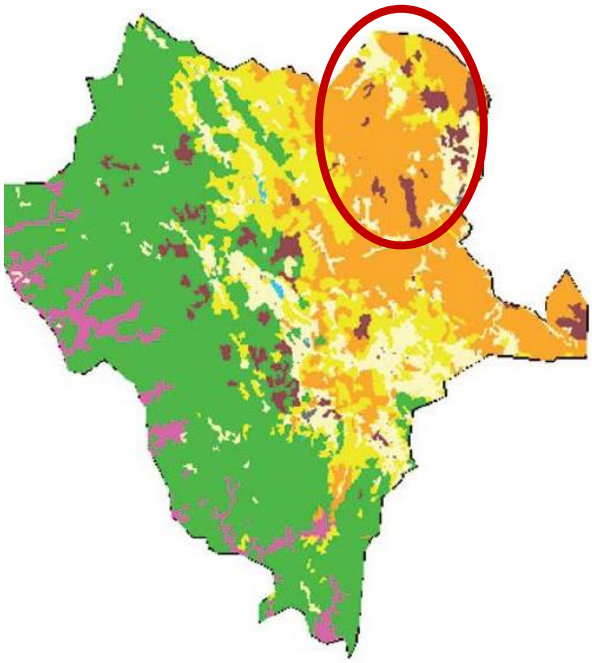

Figura 15. Ubicación de sitios de colecta en el estado de Durango

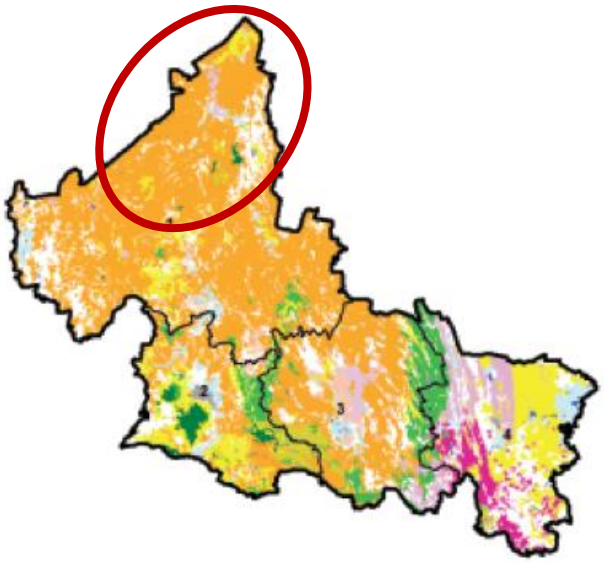

Figura 16. Ubicación de los sitios de colecta en el estado de San Luis Potosí

5.2. Estandarización del método de extracción y cantidad de muestra

El proceso de extracción tanto de resina como de hule fue basado en el procedimiento de Pearson *et al.* (2013). Sin embargo, se realizaron modificaciones en los tiempos de retención y temperatura, adaptándose a los requerimientos del equipo y los puntos de ebullición de los solventes usados, dando como resultado dos rondas de extracción. La primera consistió en la extracción de resina de la muestra utilizando acetona como solvente, con 3 ciclos de 5, 10 y 15 min de retención a una temperatura de 40 ^{o}C a 100 bares de presión.

Esta temperatura fue propuesta por Pearson *et al.* (2013) para evitar que la acetona alcance el punto de ebullición. Una acetona caliente (alrededor de 100 °C) permite una mejor extracción de resina, pero provoca disolución y/o degradación de hule haciendo que el hule sea irrecuperable en la extracción posterior (Salvucci *et al.*, 2009). Lo que se confirmó en el trabajo de Pearson *et al.*

(2013) quienes reportaron una pérdida de 26% de la fracción de hule de alto peso molecular por el contacto directo del hule con la acetona caliente.

Seguido de la remoción de la fracción de resina, las muestras fueron sometidas a una segunda ronda de extracción, específica para extraer el hule utilizando hexano como solvente, 3 ciclos de 5, 10 y 15 min de retención cada uno, a 60 ^{o}C /100 bares de presión. Cabe mencionar que Pearson *et al.* (2013), utilizaron 140 °C para extracción de hule con ciclohexano, sin embargo, bajo nuestras condiciones de trabajo no fue posible usar una temperatura similar sin que se tuvieran problemas de extracción. Revisando las condiciones recomendadas por el fabricante del equipo, mencionan que no es recomendable usar temperaturas de extracción que vayan por encima de la temperatura de ebullición del solvente utilizado, en este caso el hexano tiene una temperatura de ebullición de 68 °C, razón por la cual al bajar la temperatura de extracción a 60 °C no se presentaron problemas de taponamiento o elevada presión en el equipo obteniendo un fluido proceso de extracción.

Otro factor a ajustar fue la cantidad de muestra a usar, principalmente por las limitaciones del material vegetal a analizar por lo que se realizaron varias pruebas con 0.5, 1 y 2.5 g del mismo tejido de corteza molido. Los valores obtenidos de hule y resina se muestran en la tabla 5. Como se puede observar, se presentó una correlación negativa entre la cantidad de muestra analizada y el rendimiento de hule y resina obtenido. Esto es fácil de explicar ya que las condiciones de extracción fueron fijas, principalmente el número de ciclos de extracción y los tiempos de retención, lo cual indica que a mayor concentración de muestra se

tiene mayor cantidad de hule y resina y por tanto se necesitarán más número de ciclos de extracción y/o mayores tiempos de retención. Basado en los resultados se eligió la cantidad de 0.5 g de muestra de corteza molida como punto de partida para todas las muestras a analizar, debido a que esta cantidad dio el mejor rendimiento de hule y resina.

Tabla 5. Comparación de porcentajes de hule y resina obtenidos en relación a la cantidad de muestra procesada

Cantidad de muestra (g)	Hule (%)	Resina (%)
0.5	7 ± 0.98*	17 ± 1.41
1	6 ± 0.89	15.6 ± 1.92
2.5	5 ± 0.61	14 ± 1.58

*Media ± desviación estándar tanto en el porcentaje de hule como de resina

5.2.1. Identificación de hule

A partir de la técnica de RMN y utilizando los procesos estandarizados anteriormente para la extracción de hule y resina bajo las condiciones del presente trabajo, se comprobó la presencia de hule natural en las muestras analizadas (figura 17); y en el espectro del C^{13} se muestran 5 señales dominantes de los desplazamientos químicos que indica la presencia del poliisopreno *Cis* -1,4 pero no para el isómero *Trans* -1,4 (figura 18). Estos datos concuerdan con estudios del hule proveniente de la especie *Taraxacum Koksaghyz* (Schmidt *et al.*, 2010b) y en literatura utilizada como referencia (Brandolini *et al.*, 2001).

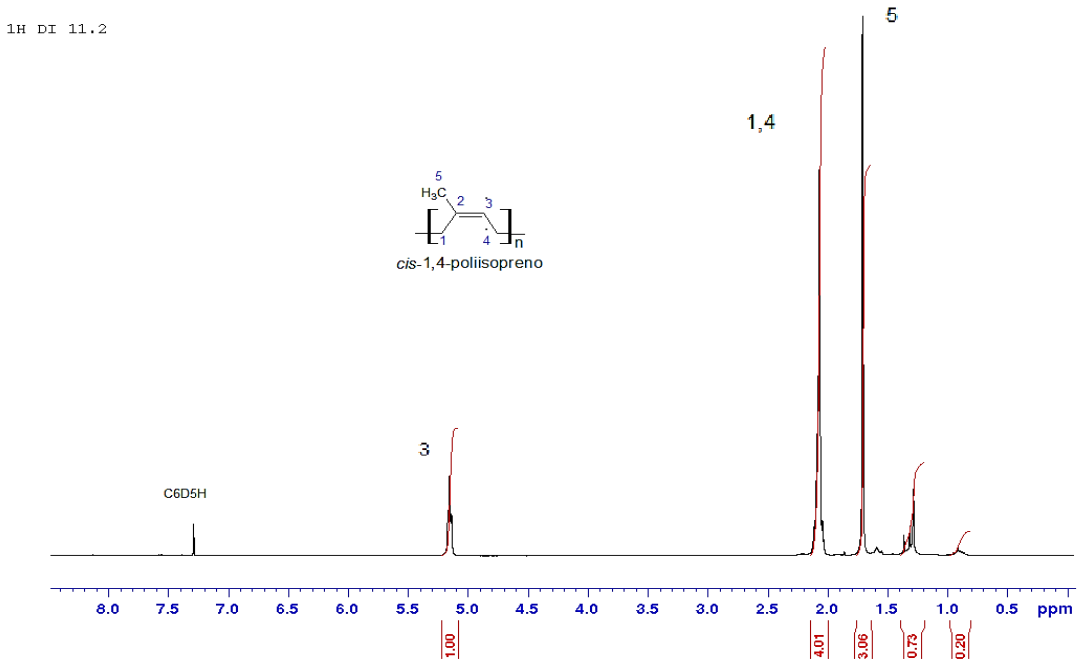

Figura 17. Espectro de RMN que comprueba la presencia de hule natural en las muestras analizadas

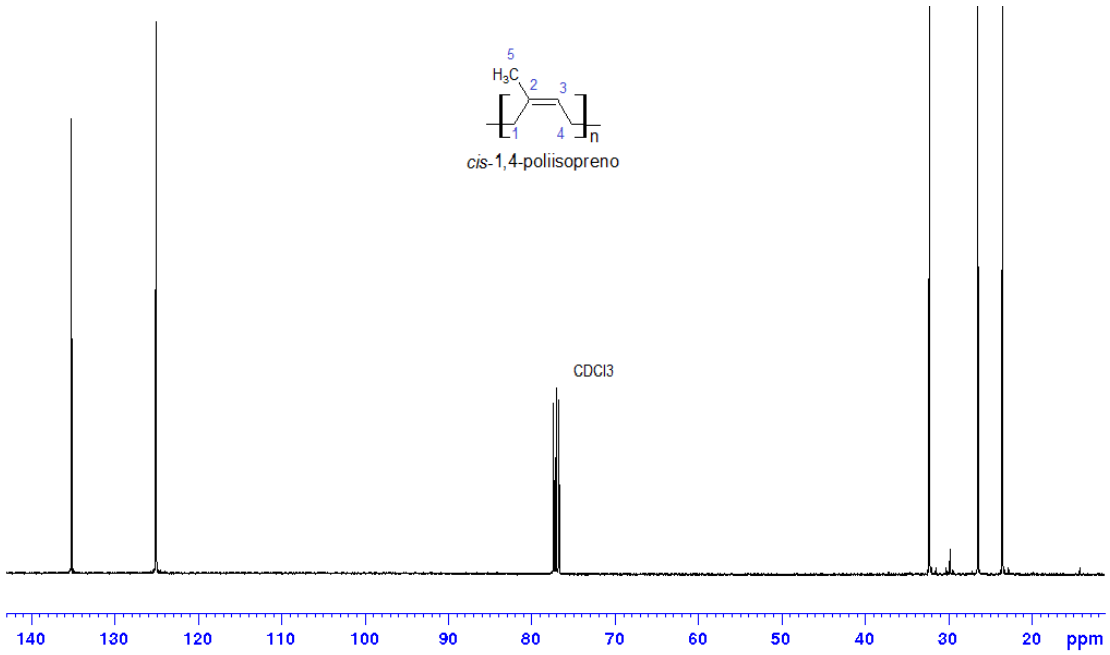

Figura 18. Espectro del C13 en donde se muestran los desplazamientos químicos de los 5 carbonos presentes en el poliisopreno *Cis*

En la tabla 6 se muestran los desplazamientos químicos obtenidos tanto en *P.argentatum* como en *T. kokzaghyz*, además de los desplazamientos reportados en la literatura para el poliisopreno configuración *Cis* y *Trans* para cada una de las posiciones en las cuales se encuentran los carbonos.

Tabla 6. Desplazamientos químicos obtenidos en el hule proveniente de *P.argentatum* en comparación con *T. kokzaghyz* y lo reportado en la literatura

Posición	*P. argentatum* (ppm)	*T. koksaghyz.* (ppm)	Poliisopreno *Cis* (ppm)	Poliisopreno *Trans* (ppm)
1	26.41	27.0	26.4	26.9
2	125.04	124.0	125.6	124.5
3	135.22	134.2	135.00	134.9
4	32.22	31.4	32.2	39.9
5	23.45	22.3	22.9	16.0

5.3. Extracciones de hule y resina de plantas provenientes de Durango colectadas en verano e invierno (2016-2017)

A partir de las muestras de tallo colectadas en los 12 sitios diferentes identificados en dos épocas del año en el estado de Durango, se realizaron las extracciones de hule y resina y se calcularon sus porcentajes.

En las figuras 19 y 20 se presentan los contenidos de resina (en porcentaje) extraídos de las muestras colectadas en verano e invierno respectivamente. En verano, los valores de porcentaje de resina oscilaron entre 9-14% y en invierno entre 11 y 17%. Estos resultados muestran un incremento de resina en todas las muestras en la época de invierno frente al verano. Es importante resaltar el sitio D7 que se mantiene en ambas épocas con el mayor porcentaje de resina con 14.40% para verano y 16.6% en invierno.

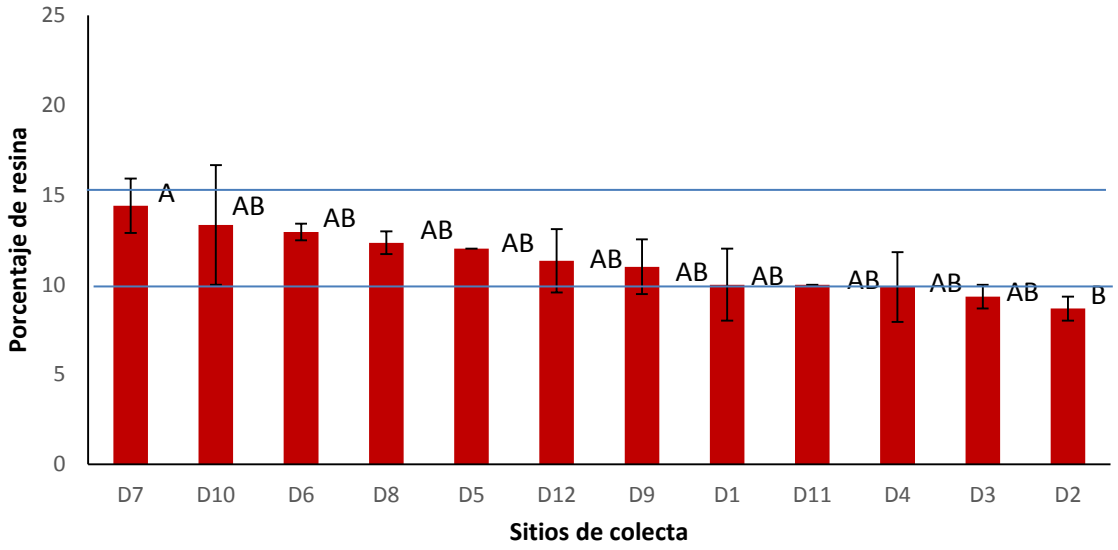

Figura 19. Porcentajes de resina de la colecta de verano 2016 en el estado de Durango.
Medias ± error estándar (n=3), con la misma letra no son significativamente diferentes (Duncan *p*>0.05)

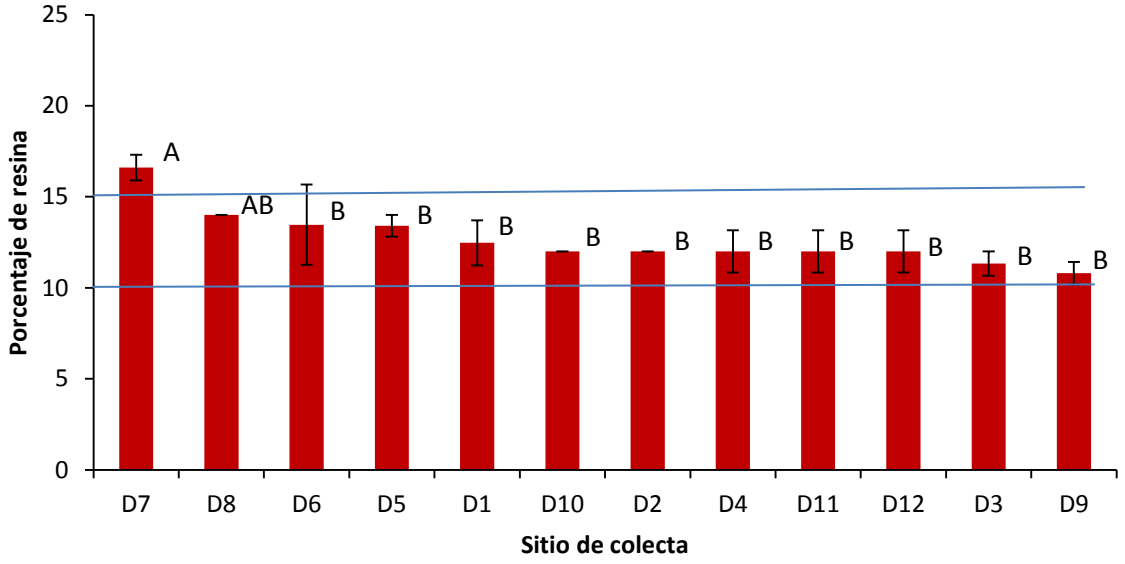

Figura 20. Porcentajes de resina de la colecta de invierno 2017 en el estado de Durango
Medias ± error estándar (n=3), con la misma letra no son significativamente diferentes (Duncan *p*>0.05)

Al contrario de los resultados obtenidos con resina, el hule muestra mayores porcentajes en verano con respecto a la colecta realizada en invierno. El rango de

porcentajes de hule en verano oscila entre 12 y 21.86% y en invierno entre 9 y 15.3%. En algunos sitios de colecta como D4, D6 y D7 este cambio fue muy contrastante (figura 21 y 22).

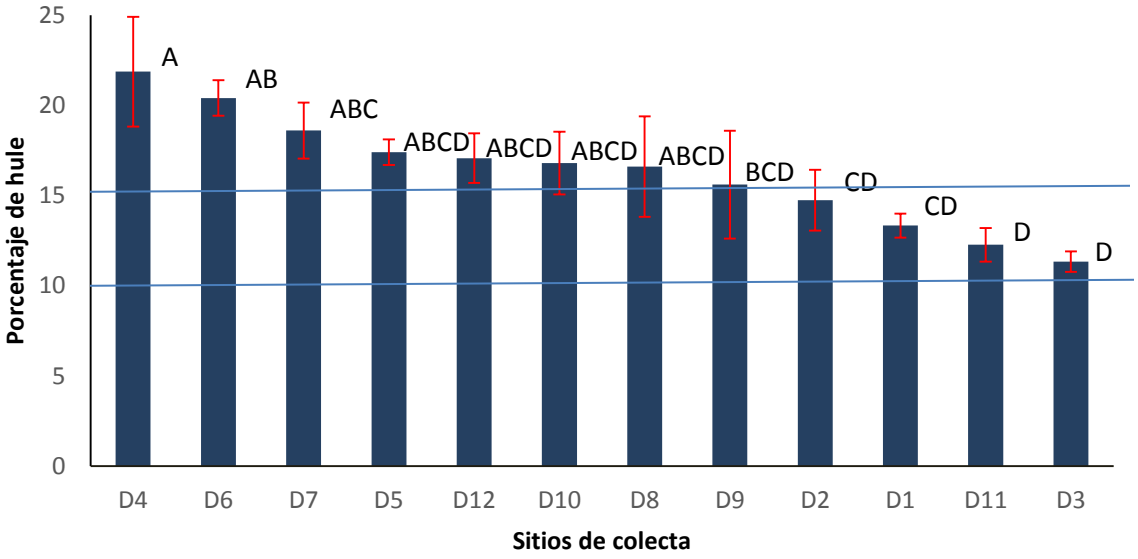

Figura 21. Porcentajes de hule de la colecta de verano 2016 en el estado de Durango
Medias ± error estándar (n=3), con la misma letra no son significativamente diferentes (Duncan *p*>0.05)

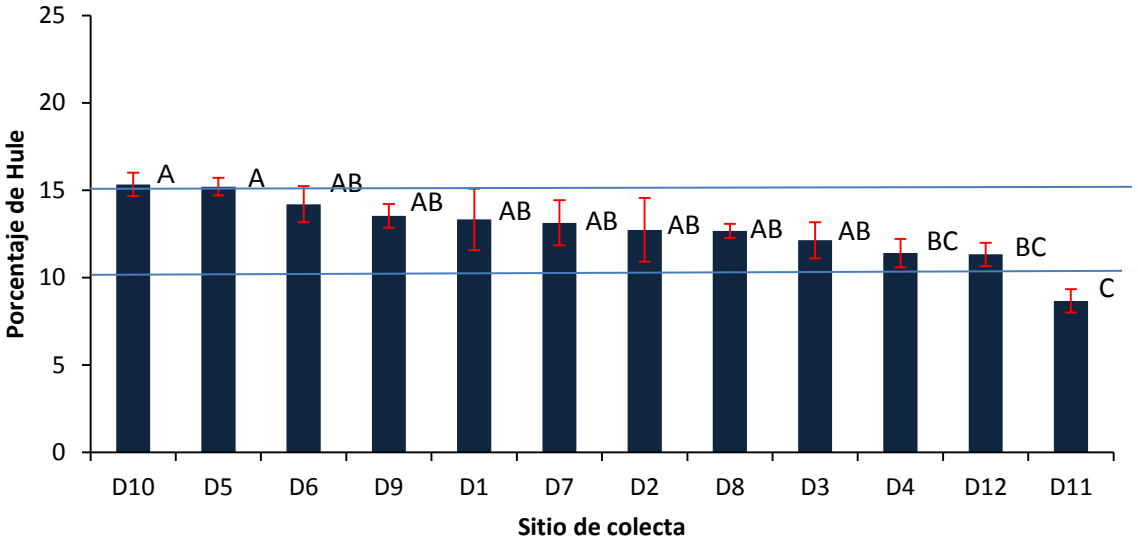

Figura 22. Porcentajes de hule de la colecta de invierno 2017 en el estado de Durango
Medias ± error estándar (n=3), con la misma letra no son significativamente diferentes (Duncan *p*>0.05).

5.4. Extracciones de hule y resina de plantas provenientes de San Luis Potosí colectadas en verano e invierno (2016-2017)

En el caso de las colectas realizadas en el estado de San Luis Potosí los porcentajes de resina en las muestras analizadas, en invierno no tuvieron el mismo comportamiento que en Durango en el que todas las muestras fueron a la alza, sin embargo, de 11 sitios identificados en San Luis Potosí, 6 si incrementaron los porcentajes de resina en invierno comparado con verano, 4 de 11 se mantuvieron igual y solo uno el S3 si bajo sus porcentajes de resina en invierno respecto a verano (figura 23 y 24).

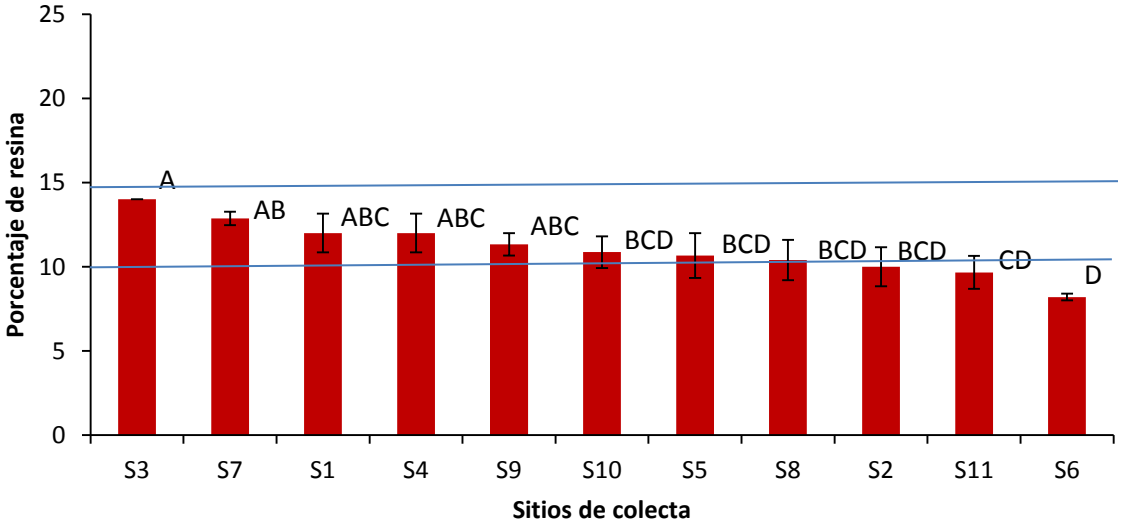

Figura 23. Porcentajes de resina de la colecta de verano 2016 en el estado de San Luis Potosí. Medias ± error estándar (n=3), con la misma letra no son significativamente diferentes (Duncan *p*>0.05)

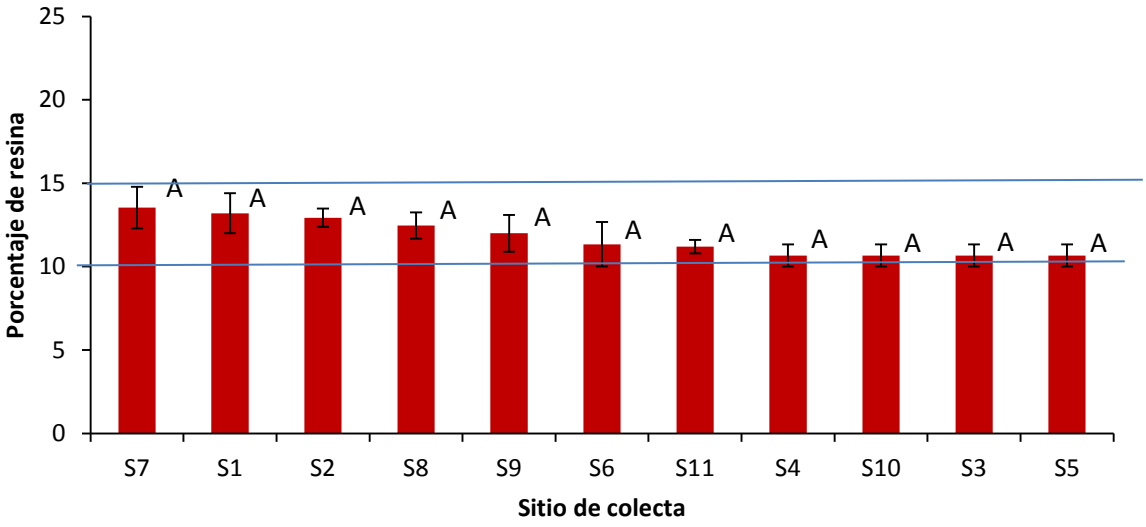

Figura 24. Porcentajes de resina de colecta de invierno 2017 en el estado de San Luis Potosí. Medias ± error estándar (n=3), con la misma letra no son significativamente diferentes (Duncan *p*>0.05)

Por lo que respecta a los porcentajes de hule obtenidos de las muestras de verano, estos oscilan entre el 14 y 19.4% y los de invierno entre 7 y 15.3%. Por lo cual, el hule demostró un comportamiento similar a lo observado en Durango, debido a que su acumulación se ve más alta en la época del verano que la del invierno (figuras 25 y 26).

De forma general, en este trabajo se obtuvieron valores altos de concentración de hule en guayule alcanzando casi el 22% (muestra D4 de verano en Durango). Este porcentaje de hule se reportó también en plantas silvestres en el año 2004 (de Rodríguez *et al.* 2004, citado (de Rodríguez *et al.*, 2006a). Sin embargo, el promedio general del contenido de hule (10 -15%) en las plantas analizadas en este trabajo supera lo que esta reportado (8-10%), en 1997, para plantas silvestres (Cantú *et al.*, 1997). Cabe mencionar que estos datos no pueden ser comparables

considerando el método de extracción usado en su momento. El método ASE usado en este trabajo, tiene una alta repetibilidad, alta recuperación de la extracción, operación totalmente automática, bajo consumo y exposición a los solventes y menor tiempo de extracción en comparación a las técnicas previamente usadas como Soxhlet, y gravimétrica (Suchat *et al.*, 2013).

En 2013 dos grupos usaron la extracción acelerada de solventes para extraer el hule de guayule. El grupo de Suchat *et al.* (2013) analizaron 240 plantas procedentes de 6 variedades y cultivadas en condiciones del mediterráneo, el valor más alto obtenido de hule fue 13.42%. Los valores obtenidos por Pearson *et al*., en el mismo año era de 0.85% pero el trabajo se realizó con plantas de una misma línea mejorada. Es importante resaltar que el trabajo aquí presentado es la primera vez que se usa la técnica de ASE para la extracción de hule y resina de plantas silvestres maduras.

Comparando la producción de hule frente a la resina en invierno y verano, en nuestros datos vemos que en verano la producción global de hule es mayor que la de resina, mientras que en invierno la acumulación de hule y resina es prácticamente igual. Estos resultados en principio no coinciden con lo que demostró Macrae *et al.* (1986), quienes encontraron que durante el verano las plantas de guayule producen más resina que hule y en invierno ocurre lo contrario. Sin embargo, ellos concluyen que dado que la resina y el hule se producen por la misma ruta biosintética del poliisopreno, cuando uno es más alto el otro es más bajo, tomando esto en cuenta, si podemos observar este comportamiento en los datos en verano donde tenemos un mayor porcentaje de hule y por lo tanto una

menor concentración de resina. Macrae *et al.* (1986), también mencionan que la producción de resina se ve favorecida a los 40 °C, este dato ayudo a la relación hule resina encontrada en nuestros datos en verano, podría indicar que las temperaturas en verano del año de colecta de nuestras muestras no alcanzaban los 40 °C (Macrae *et al.*, 1986)

Así mismo, muchos autores coinciden con la hipótesis que las bajas temperaturas favorecen la producción de hule. Vaetch-Blohm *et al. (*2006), han demostrado que la síntesis de hule parece ser altamente estimulada por temperaturas inferiores a 10 °C, específicamente durante la noche. Macrae *et al.* (1986) proponen que las temperaturas bajas estimulan la producción de hule sin eliminar el posible efecto de otros factores ambiental tal como la intensidad de luz, los autores llegaron a determinar que la producción óptima de hule se produce a una temperatura de 27 °C durante el día seguidas por 7 °C durante la noche. Angulo *et al.* (2002) concluyeron que temperaturas entre -4 ^{o}C y 2 ^{o}C aumenta la síntesis de hule. La tolerancia a la temperatura fría puede ser un factor en la producción de hule (Agrawal y Konno, 2009). Sin embargo, varios autores adicionan otros factores, que pueden tener igual importancia, solos o en combinación con la temperatura como las precipitaciones, así que, Angulo-Sánchez (2002) ve en sus experimentos que los bajos niveles de lluvia mejoraron la acumulación de hule, mientras que los altos niveles de lluvia no lo hicieron a pesar que la alta precipitación también mejoró la producción de biomasa. De igual forma Veatch-Blohm *et al.,* (2006) demostraron que plantas de guayule sometidas a estrés hídrico tuvieron mayores concentraciones de hule que las plantas bien regadas y de igual forma que

nuestros datos ellos obtuvieron mayores concentraciones de hule en plantas sometidas a estrés hídrico en la época de verano comparada con invierno (Veatch-Blohm *et al.*, 2006)

Todo lo anterior, pone de manifiesto que el hecho de que nuestros datos indiquen que en 23 sitios diferentes de dos estados diferentes en un mismo año de colecta la concentración de hule sea mayor en verano que invierno puede ser posible bien sea por que las condiciones de estrés hídrico en ese año fueron suficientes para promover la producción de hule y/o porque las temperaturas de día y noche en esa época de ese año fueron favorables para la producción de hule entre muchos otros factores.

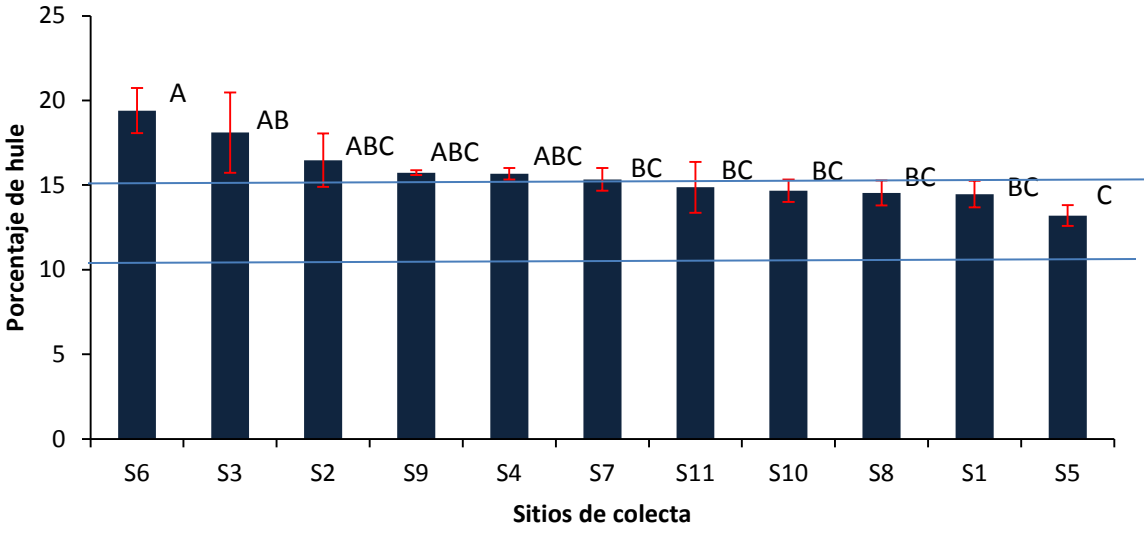

Figura 25.Porcentajes de hule de colecta verano 2016 en el estado de San Luis Potosí. Medias ± error estándar (n=3), con la misma letra no son significativamente diferentes (Duncan *p*>0.05).

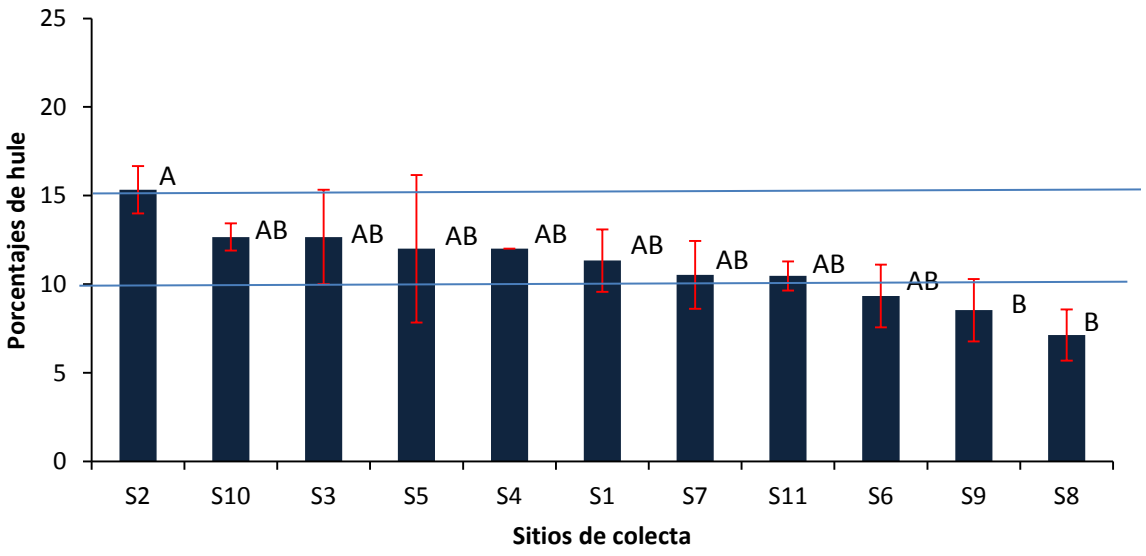

Figura 26. Porcentajes de hule de colecta invierno 2017 en el estado de San Luis Potosí. Medias ± error estándar (n=3), con la misma letra no son significativamente diferentes (Duncan *p*>0.05)

En base a estos resultados y tomando en consideración el criterio inicial de elegir plantas con la mayor y menor sobreexpresión de proteínas involucradas en la biosíntesis de hule. Para los análisis del transcriptoma, se eligieron 12 muestras de corteza almacenadas a -80°C correspondientes a 3 sitios y 2 épocas de colecta diferentes (con sus 3 repeticiones). De San Luis Potosí, se eligió el sitio S6 por ser el que presentó los más altos porcentaje de contenido de hule en verano y de los más bajos porcentaje de contenido de hule en invierno. En Durango, por problemas técnicos que se tuvieron con las extracción de ARN, se perdieron las muestras colectadas en verano y por lo tanto se decidió elegir los sitios D10 y D11 de la colecta de invierno pues fueron los que tuvieron el mayor y el menor contenido de hule.

5.5. Extracción y cuantificación de ARN total

Las figuras 27 y 28 muestran los geles de agarosa realizados para verificar la calidad de los extractos de ARN total obtenido, a partir de las muestras provenientes de los estados de Durango y San Luis Potosí, respectivamente. Se nota que, en ambas figuras, el ARN no se encuentra degradado ni contaminado de ADN, por lo cual estas extracciones fueron buenas candidatas para una secuenciación masiva posterior.

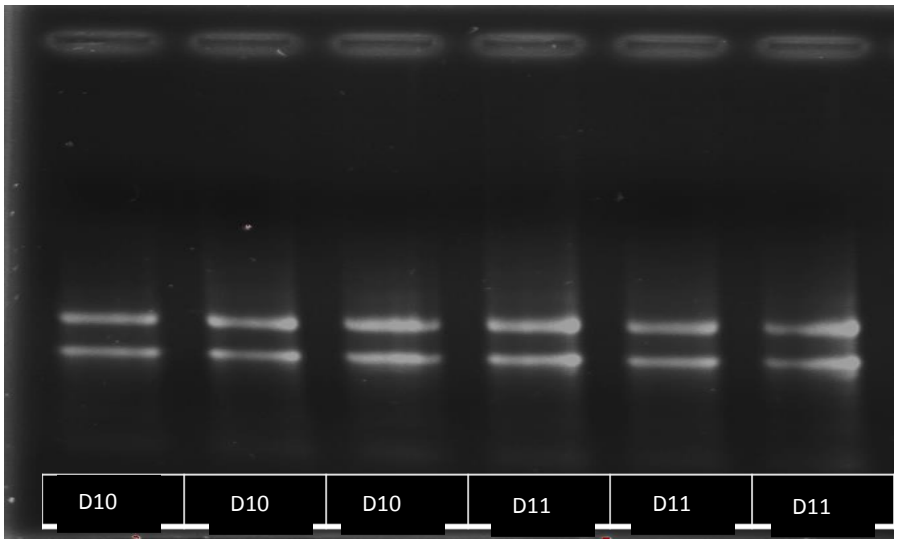

Figura 27. Gel de agarosa mostrando las bandas de ARN sin presencia de ADN en las muestras de Durango

Con respecto a la cuantificación, dadas las especificaciones del sistema de secuenciación (LANGEBIO), se trató de tener como mínimo 3 µg de ARN, por cada muestra, en la tabla 7 se muestran las cantidades de ARN por µl de muestra leída en el equipo Quantus^{TM}Fluor

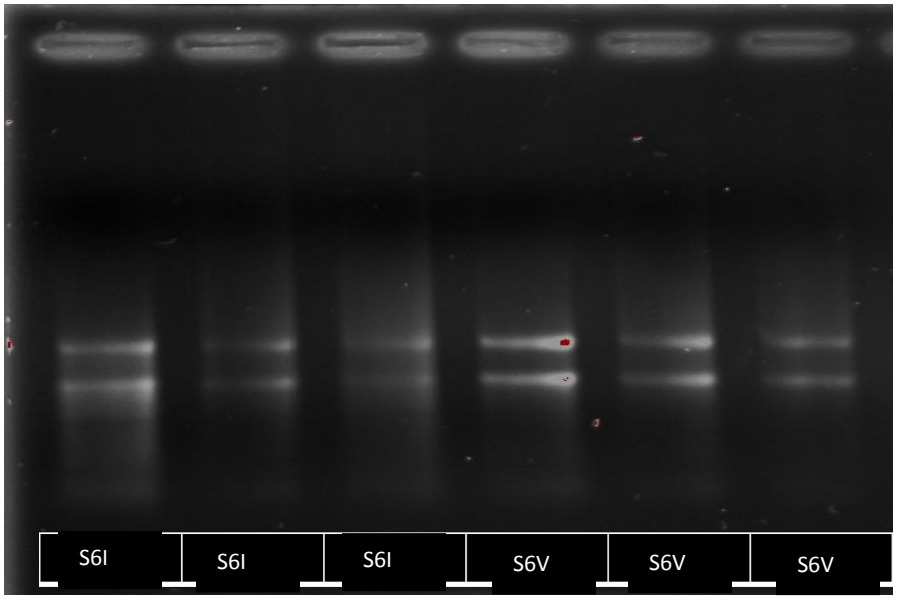

Figura 28. Gel de agarosa mostrando las bandas de ARN extraído de las muestras provenientes del estado de San Luis Potosí sin la presencia de ADN

Tabla 7. Concentración de ARN total de las 12 muestras seleccionadas

Nombre Muestra	D10	D10	D10	D11	D11	D11	S6I	S6I	S6I	S6V	S6V	S6V
Concentración (ng/µl)	171	166	418	266	196	402	244	60	88	91	76	60

Las muestras se encontraban resuspendidas en 50 µl de agua libre de RNasa. Por el tanto todas las muestras cumplieron con la mínima concentración requerida.

5.6. Filtrado y ensamblado de las lecturas

Se generó un total de 372,661,284 lecturas en bruto. Posteriormente, se realizaron pasos de filtración de calidad siguiendo los criterios mencionados en materiales y métodos. La tabla 8 presenta las 12 muestras y el número de secuencias pareadas de lectura correspondiente antes y después de la filtración.

El filtrado consistió en eliminar adaptadores, bases defectuosas al principio y al final de las lecturas y lecturas de baja calidad. Este proceso aumentó el umbral de calidad por base para todas las muestras por encima de Q30, lo que resulto en la eliminación del 25% del número de secuencias obtenidas para cada muestra dando un total de 282,486,060 secuencias. El siguiente filtrado eliminó, unas 36, 561,287 secuencias por no alcanzar la longitud deseada (75 bases), y por ultimo unos 36,295,743 pares de secuencias se fusionaron en una de longitud mayor, para tener finalmente un total de 245,659,229 lecturas.

Tabla 8. Muestras secuenciadas (código de identificación, sitio de origen, criterio de selección, total de lecturas crudas, número de lecturas después del filtrado de calidad)

Código de identificación de la muestra	**Sitio de origen**	**Criterio de selección**	**Época**	**N° de lecturas (R1 - R2)**	**N° de lecturas filtradas**
AR2HG1SS01	Durango	Mayor contenido de hule.	Invierno	18,437,463	14,342,000
AR2HG1SS02	Durango	Mayor contenido de hule.	Invierno	14,640,509	10,966,283
AR2HG1SS03	Durango	Mayor contenido de hule.	Invierno	16,334,818	12,319,123
AR2HG1SS04	Durango	Menor contenido de hule.	Invierno	18,732,649	14,474,993
AR2HG1SS05	Durango	Menor contenido de hule.	Invierno	14,118,581	10,599,075
AR2HG1SS06	Durango	Menor contenido de hule.	Invierno	18,031,605	13,784,482
AR2HG1SS07	San Luis Potosí	Menor contenido de hule.	Invierno	13,981,550	10,977,059
AR2HG1SS08	San Luis Potosí	Menor contenido de hule.	Invierno	14,541,818	10,971,626
AR2HG1SS09	San Luis Potosí	Menor contenido de hule.	Invierno	15,049,504	11,479,433
AR2HG1SS10	San Luis Potosí	Mayor contenido de hule.	Verano	14,650,513	11,114,304
AR2HG1SS11	San Luis Potosí	Mayor contenido de hule.	Verano	13,625,397	9,627,402
AR2HG1SS12	San Luis Potosí	Mayor contenido de hule.	Verano	14,186,235	10,587,250
		Total		**372,661,284**	**282,486,060**

Con estas lecturas obtenidas de las 12 muestras analizadas, se realizó un ensamblado de *novo* con TRINITY, el algoritmo primero realizó un montaje ambicioso, donde una lectura se agregaba iterativamente a una lectura o contig hasta que ya no era posible. Luego, se generaron secuencias únicas a partir de conjuntos de secuencias que se superponen, lo que resultó en la obtención de 1,083,355 contigs de longitudes variables. A estos contigs se realizó otro filtrado de limpieza lo que redujo el número a 1,069,960 contigs. Dado que la especie de *P. argentatum* es una especie heterocigota se pretendió evitar cualquier redundancia por lo que se realizó un reensamblado con Newbler de lo cual resultaron 564,256 contigs.

Este número de contigs que corresponde al número de transcrito únicos supera altamente lo que se obtuvo en el primer análisis de transcriptoma realizado con plantas de guayule (de 2 años de edad), donde se generaron 12,682 lecturas brutas de los cuales consiguieron sólo 9,021 unigenes (Ponciano *et al.*, 2012). Estos resultados demuestran la alta sensibilidad de la técnica de Illumina en comparación a la técnica de la secuenciación de “etiquetas de secuencias expresadas” (EST, expressed sequence tags) que se aplicó con Ponciano *et al.* (2012). Así mismo, los valores de esta investigación superan los obtenidos en los trabajos realizados con muestras de *H.brasiliensis* y *T. kok-saghyz*, usando la misma técnica, los valores obtenidos son de 48,780 y 57,760 para *Hevea* y 55,532 unigenes para *T. kok-saghyz* (Luo *et al.*, 2017; Tan *et al.*, 2017; Tang *et al.*, 2016). Una explicación a eso puede ser que las plantas con que se trabajó fueron plantas silvestres que presentaron una variabilidad entre ellas.

5.6.1. Identificación de ORF´s

De los 564,256 contigs obtenidos, se identificaron 301,427 secuencias codificantes (CDS). Probablemente la pérdida de 261,806 contigs es debido a que tienen regiones no codificantes UTR (por sus siglas en inglés untraslated regions).

5.6.2. Identificación de proteínas homólogas

Se compararon las 301,427 CDS generadas de *P. argentatum,* de manera individual, con las seis bases de datos citadas en la tabla 3 (ver materiales y métodos) para identificar proteínas homologas.

Se identificaron aquellas secuencias (hit, término usado por el programa) que encontraron una homología mayor de 90% en al menos una secuencia de las 6 bases de datos comparadas.

Al final se identificaron 241,176 hits que expresan proteínas homologas (lo que corresponde a más del 80% de los CDS generados). Entre ellos, 79.53% tienen homólogos en la base de datos de *H. annus,* 75.66% en UniProt, 67.48%, en *S. lycopersicum*, 67.45% en *N. tabacum*, 66.4% en *S. indicum*, 67% en *H. brasiliensis,* 64.5% e*n A. thaliana.*

H. annus igual que guayule forma parte de la familia Asteraceae, lo que explica esta mayor homología en proteínas. *H. brasiliensis* (67%) conocida por su alta producción de hule, probablemente comparte con *P. argentatum* algunos mecanismos de biosíntesis de hule.

5.6.3. Anotación funcional de los unigenes de *Parthenium argentatum*

De las proteínas homologas de la base de datos UniProt se obtuvieron los términos GO (por sus siglas en ingles Gene Ontology) para los unigenes de *P. argentatum,* 38,834 unigenes fueron subidos a la base de datos DAVID de los cuales 32,844 unigenes heredaron al menos un término GO, en una de las categorías funcionales: 17,818 unigenes (54%) asociados a funciones biológicos, 12,209 unigenes (37%) asociados a funciones moleculares y 20,547 unigenes (62.5%) asociados a la categoría de componente celular. La figura 29 muestra los primeros 10 procesos con el mayor número de proteínas homólogas en cada una de las categorías funcionales. Es necesario un estudio más profundo para identificar la abundancia de estas proteínas en la planta y su posible relación o efecto en el rendimiento de hule. Ya que se sabe que guayule es una planta desértica, sus proteínas relacionadas con respuestas a estrés abiótico y biótico son relativamente abundantes (Ponciano *et al.*, 2012) razón por la cual, proteínas relacionadas con el estrés salino se presentan en estas gráficas, al igual que proteínas relacionada con la actividad de defensa contra el estrés hídrico (respuesta al ácido abscísico), (Jiang y Zhang, 2004) y proteínas con actividad de quinasa, que controlan, en eucariotas, la señalización de defensa (Lee *et al.*, 2015). En cambio proteínas relacionadas con metabolismos de plantas, no se detectan, tal vez por falta de información en las bases de datos usadas como referencia para ser comparadas en este estudio. Lo que exige un trabajo manual revisando secuencia por secuencia, los CDS identificados en este trabajo y compararlos con genes o proteínas de guayule descritas anteriormente.

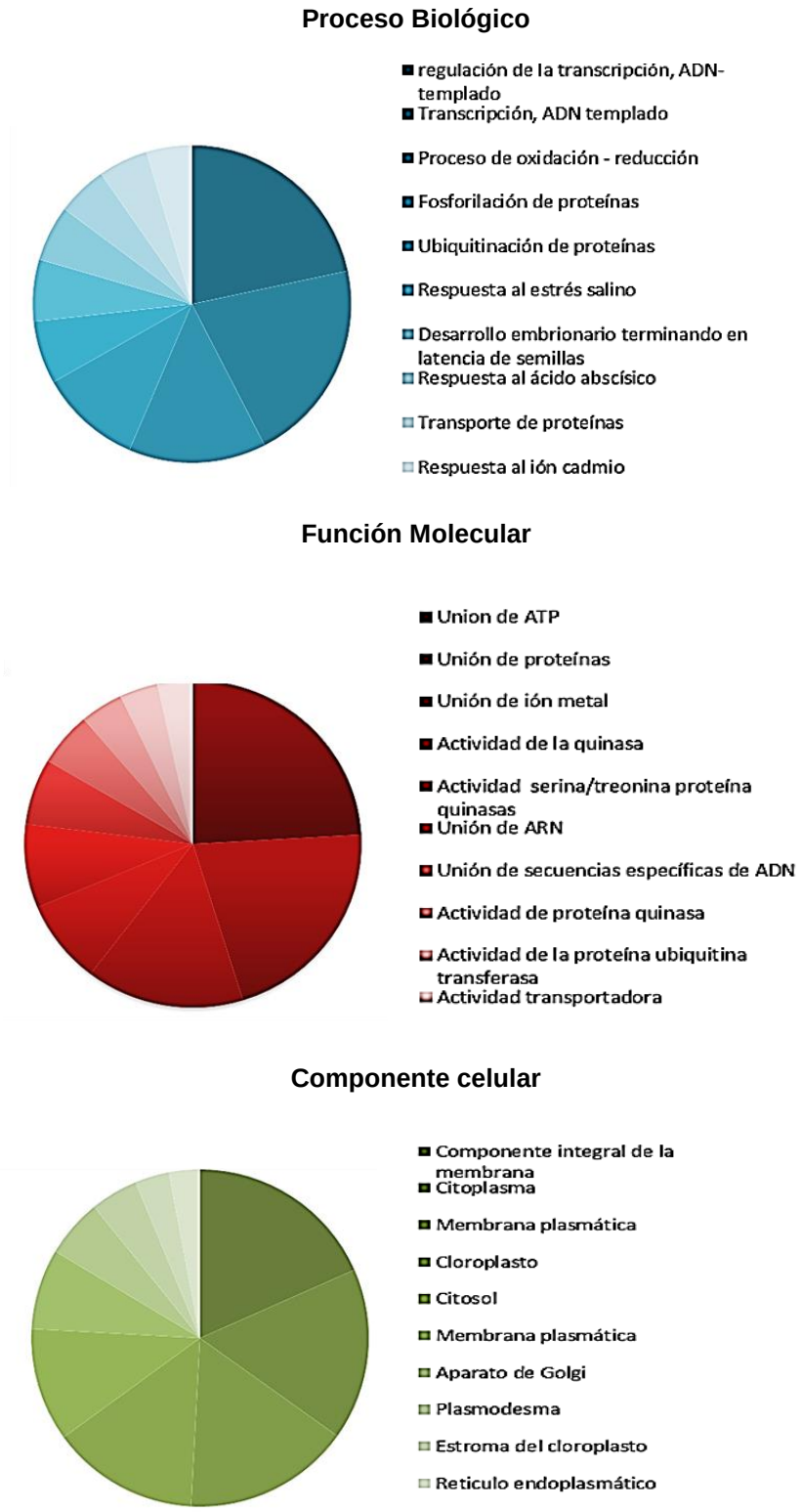

Figura 29. Anotación funcional de los Unigenes de *Parthenium argentatum* por medio de la base de datos DAVID, asociados a términos GO y relacionados con sus diversas categorías

5.6.4. Identificación de genes ortólogos

A nivel genético se identificaron 340,718 genes ortólogos entre *P. argentatum,* las especies *A. thaliana, H. annus, H. brasiliensis, S. indium, S. lycopersicum,* y proteínas relacionadas en las dos rutas de síntesis de IPP conocidas: MEP y MVA, las cuales fueron descargadas de las bases de datos PlantCyc y MetaCyc. Estos genes se reparten en 58,925 grupos.

En base a las proteínas descargadas de PlantCyc y MetaCyc se realizó la reconstrucción de estas rutas (figura 30). Todos los genes codificantes de las enzimas de la ruta MVA se han identificado dentro del transcriptoma generado. En el caso de la ruta MEP se identificaron 6 enzimas de las 7 conocidas.

Se sabe que la biosíntesis del hule se produce a través de las ruta MVA y probablemente MEP (Sando *et al.*, 2008). Estas rutas son catalizadas por muchas enzimas y cofactores así se comienza con la síntesis de IPP. Las secuencias que no encontraron homólogas en la base de datos, podrían resultar de la presencia de ARNm no traducido, genes nuevos, genes únicos de guayule o errores de ensamblaje.

Las bases de datos relacionados con la segunda fase de la biosíntesis, que incluye la ruta de la condensación de IPP en partículas de hule no estuvieron disponibles en el momento de la presente investigación, ya que las secuencias que fueron usadas para este análisis resultaron no contar con aquellas secuencias de las *cis*-preniltransferasas, a las que se les atribuye la síntesis de látex, las cuales ya se encuentran reportadas para *Hevea brasiliensis* y que los autores del

artículo en el que fueron reportadas (Lau *et al.*, 2016) no subieron los datos de esas secuencias, aunque en su trabajo hacen mención de las bases públicas en las que se podían encontrar y descargar. Razón por la cual no se cumplió esta parte del estudio. Aunque no se pudo realizar la reconstrucción completa de la ruta del hule; en concreto de los elementos participes en el complejo de la partícula de hule, si se pudo reconstruir pasos posteriores al IPP hasta el FPP (figura 31) que se conoce como el sustrato de elección para la polimerización del hule (da Costa et al., 2005)

Se tiene conocimiento que todos los isoprenoides se derivan de dos unidades básicas, siendo una de ellas el IPP y que en diversas especies vegetales este precursor es sintetizado por dos vías: MVA y MEP. Por lo que varios autores se han enfocado en aquellos compuestos que se encuentren involucrados en la síntesis del hule (Lau et al., 2016; Tan et al., 2017), lo que ha permitido la reconstrucción de dichas rutas y que estas reconstrucciones son similares a las aquí presentadas.

Una vez que se tuvieron los grupos con los genes ortólogos se procedió a realizar un análisis filogenético a nivel de proteína y a nivel de secuencia nucleotídica esto con el fin de demostrar que realmente a nivel proteína los genes eran ortólogos y a nivel de secuencia nucleotídica para ver las distancias filogenéticas entre las especies utilizadas, en total se obtuvieron 30 árboles filogenéticos (15 a nivel proteína y 15 a nivel de secuencia nucleotídica, en el presente escrito sólo se presenta uno de los 15 árboles, el resto es presentado como material suplementario).

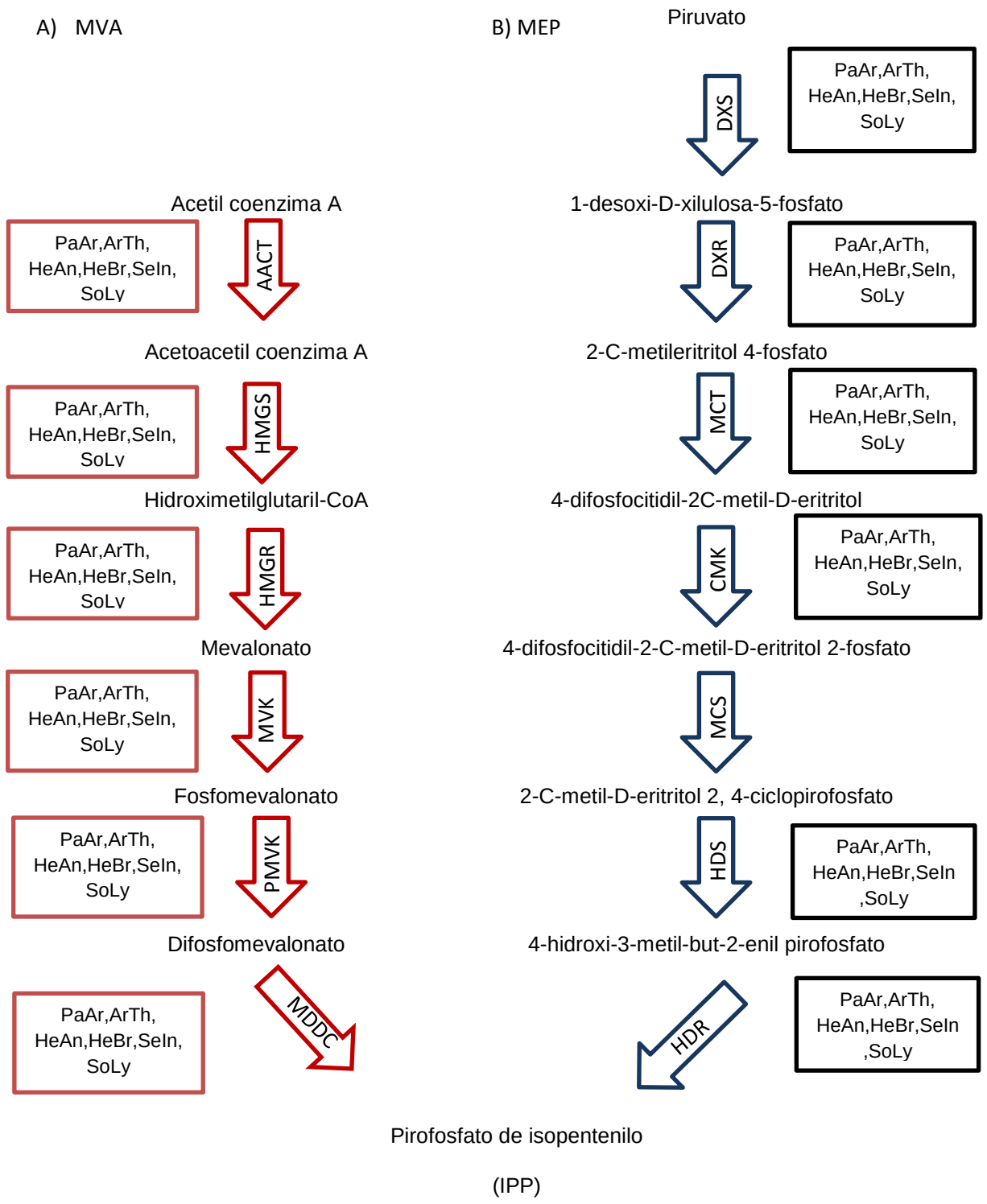

Figura 30. Reconstrucción de rutas de síntesis de IPP: a) MVA y b) MEP y los genes implicados. Las flechas representan los genes, los cuadros presentan las especies que comparten genes ortólogos con *Parthenium argentatum* =PaAr, ArTh=*Arabidopsis thaliana*; HeAn=*Helianthus annus*; HeBr=*Hevea brasiliensis*; LaPa=Ruta del hule; SeIn=*Sesamum indicum*; SoLy=*Solanum lycopersic*um.

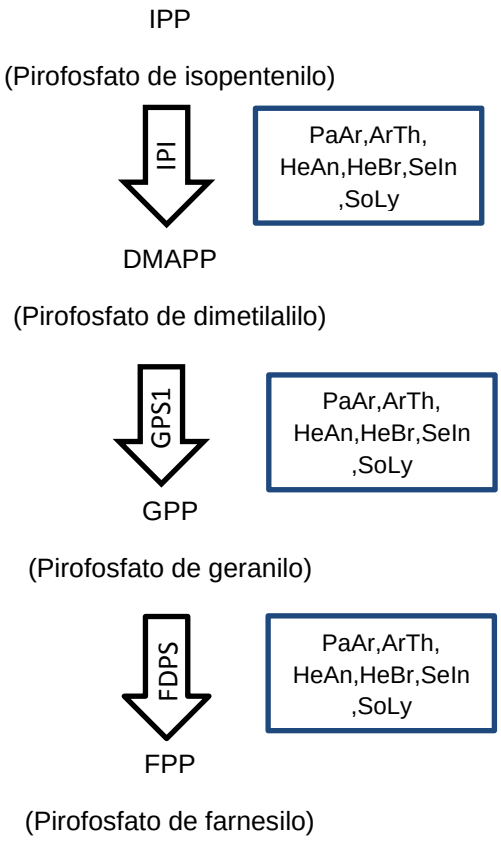

Figura 31. Reconstrucción de ruta del IPP hasta el FPP. Las flechas representan los genes, los cuadros presentan las especies que comparten genes ortólogos con *Parthenium argentatum*.

La figura 32 muestra los porcentajes de identidad entre una proteína y otra, independientemente de la especie que fue utilizada para la construcción del árbol filogenético. Se observa la formación de tres grupos en los cuales las proteínas involucradas en la síntesis del precursor primario del látex (IPP) se agrupan con la proteína de *P. argentatum* y *H. brasiliensis*. A nivel nucleotídico se vuelven a formar los mismos tres grupos donde se puede observar que a este nivel las especies se agrupan por su afinidad filogenética (figura 33), ya que los porcentajes de identidad que se muestran en la figura 33 presentan una visión global de que

tan cercanas o lejanas son las especies utilizadas para la construcción de este árbol filogenético.

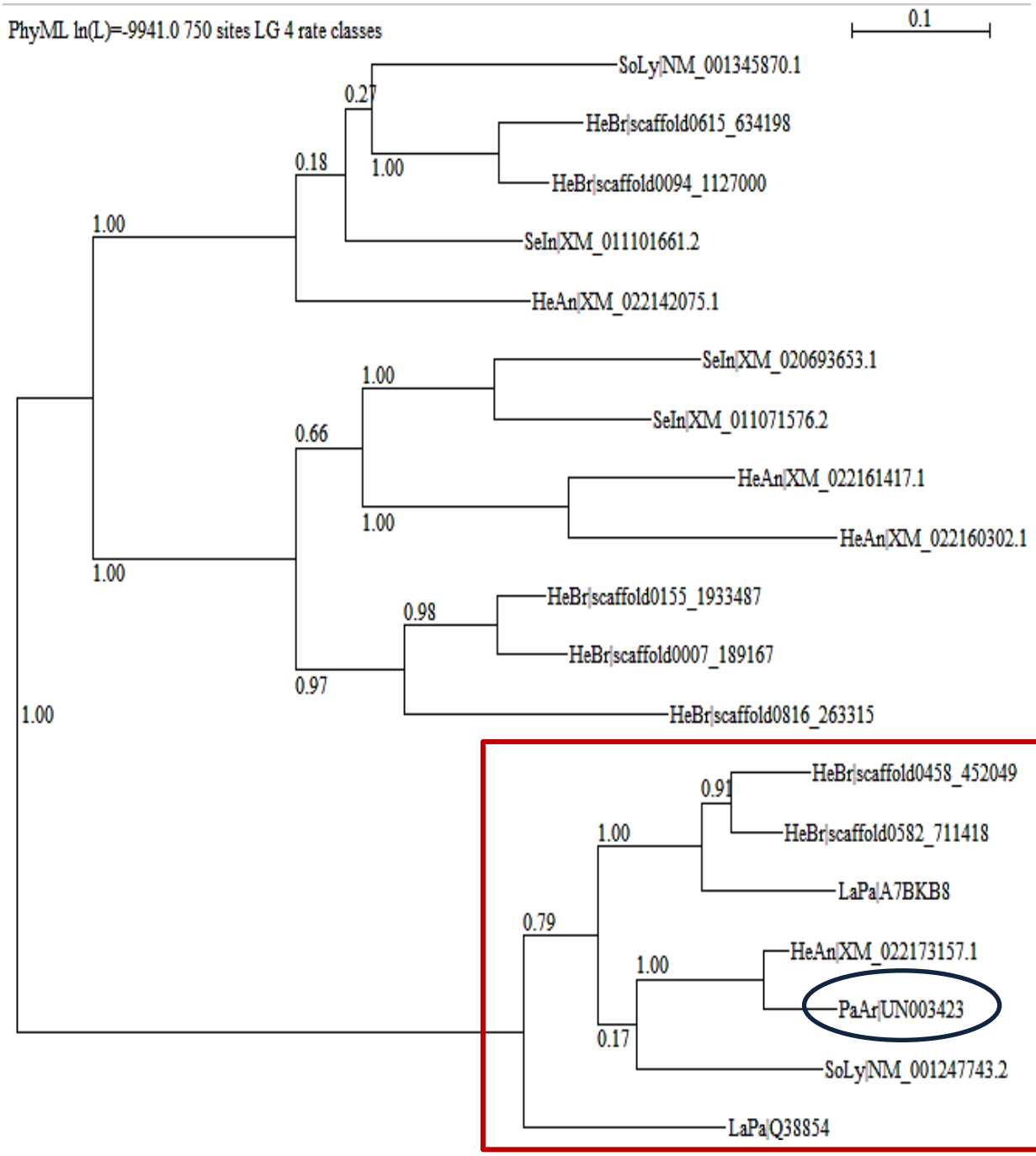

Figura 32. Árbol filogenético del primer grupo de genes ortólogos a nivel de proteína

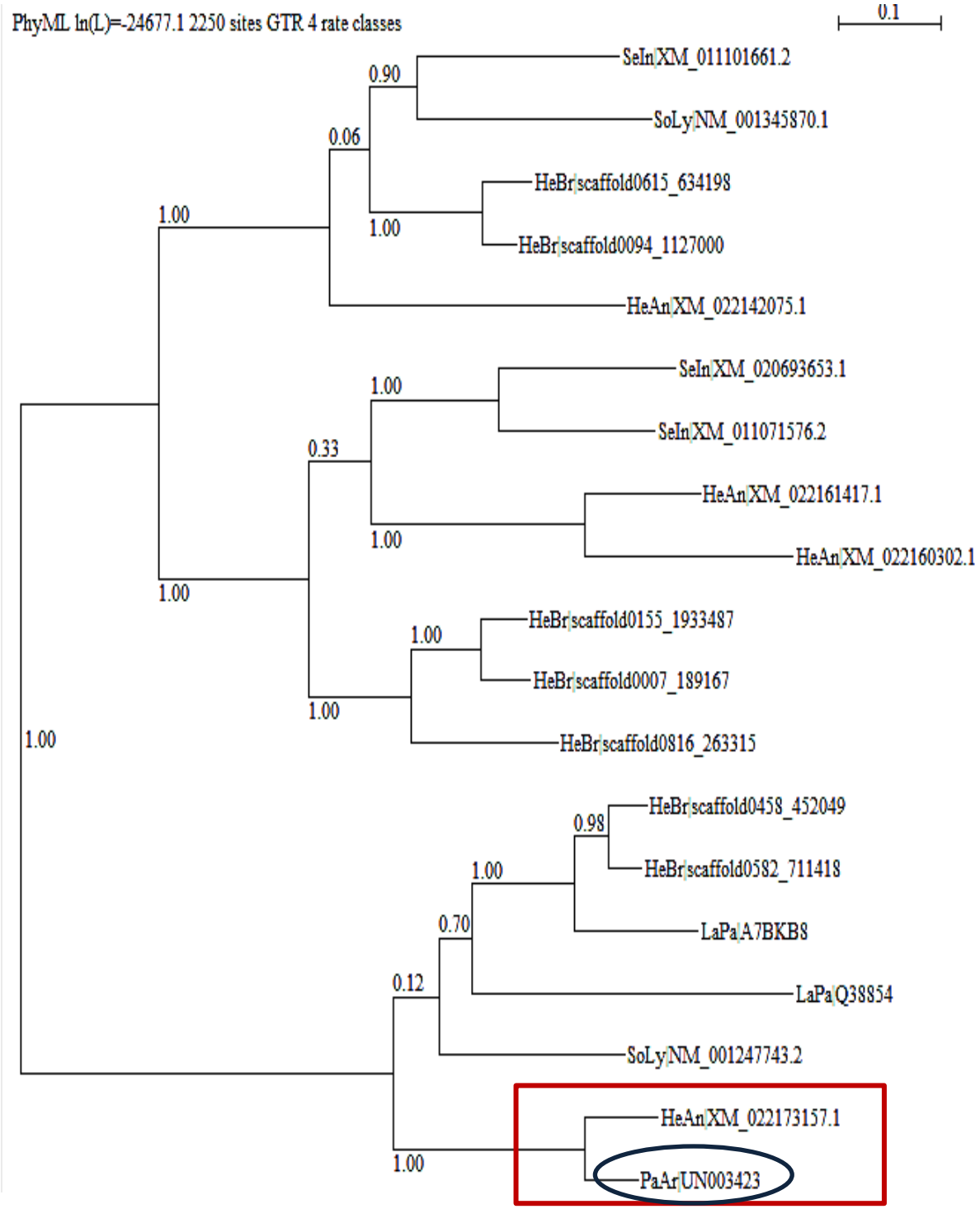

Figura 33. Árbol filogenético del primer grupo de genes ortólogos a nivel de secuencia nucleotídica.

Los datos anteriores sirven para validar la información ya que *P. argentatum* se agrupa por función con *H. brasiliensis* pues son dos especies productoras de hule natural y filogenéticamente se agrupa con *H. annus* ya que ambas pertenecen a la familia Asteraceae.

5.6.5. Análisis de expresión diferencial

Los valores de la expresión diferencial se calcularon con el programa DESeq versión 1.28.0; el cual mostró resultados más concluyentes que edgeR.

Las comparaciones se realizaron para encontrar genes particulares que pueden diferenciar condiciones contrastantes. Dadas las diferentes condiciones con las que se contó en este estudio (Tabla 9) se optó por establecer 3 diferentes combinaciones: 1) El primer contraste fue mayor/menor contenido en hule, la comparación se realizó con datos procedentes de la misma región (Durango) y la misma época de colecta (invierno), de esta manera se intentó reducir los parámetros variables; 2) El segundo contraste fue Durango/San Luis potosí, en este caso se compararon datos de dos regiones diferentes pero en la misma época de colecta (Invierno) y con plantas de contenido de hule relativamente bajo; 3) El tercer contraste fue invierno-verano con plantas con el menor y mayor contenido de hule respectivamente, con datos procedentes de San Luis Potosí.

Tabla 9. Establecimiento de condiciones y contrastes para el análisis de expresión diferencial.

Condición A	**Condición B**	**Condición C**	**Condición D**	**Contrastes**
Invierno	Invierno	Invierno	Verano	Condición A_vs_B
Menor porcentaje de hule	Mayor contenido de hule	Menor contenido de hule	Mayor contenido de hule	Condición A_vs_C
Durango	Durango	San Luis Potosí	San Luis Potosí	Condición C_vs_D

La figura 34, presenta los datos de expresión de los genes identificados, agrupados según las 4 condiciones antes definidas.

Se puede observar que las plantas procedentes de Durango durante el invierno del año 2017 (condiciones A y B) se comportan de una manera homogénea entre ellas. Además, se ve claramente el grupo de genes que se expresan altamente (color amarillo) frente a los de menos expresión (color morado) y el cambio inverso de la expresión entre plantas de mayor y menor contenido de hule. Esta homogeneidad puede ser relacionada a la probable naturaleza apomíctica de las plantas muestreadas. Las plantas de naturaleza apomíctica y las semillas que producen son clones de la planta de la que se obtuvieron y tienen el mismo genotipo (Bicknell y Koltunow, 2004; de Rodríguez *et al.*, 2006b). Específicamente en Durango, se detectó además de la presencia de plantas diploides una dominancia de plantas tri- y tetra-ploides. Plantas triploides y niveles de ploidía

más altos generalmente son de naturaleza apomíctica (Bicknell y Koltunow, 2004; de Rodríguez *et al.*, 2006b). .

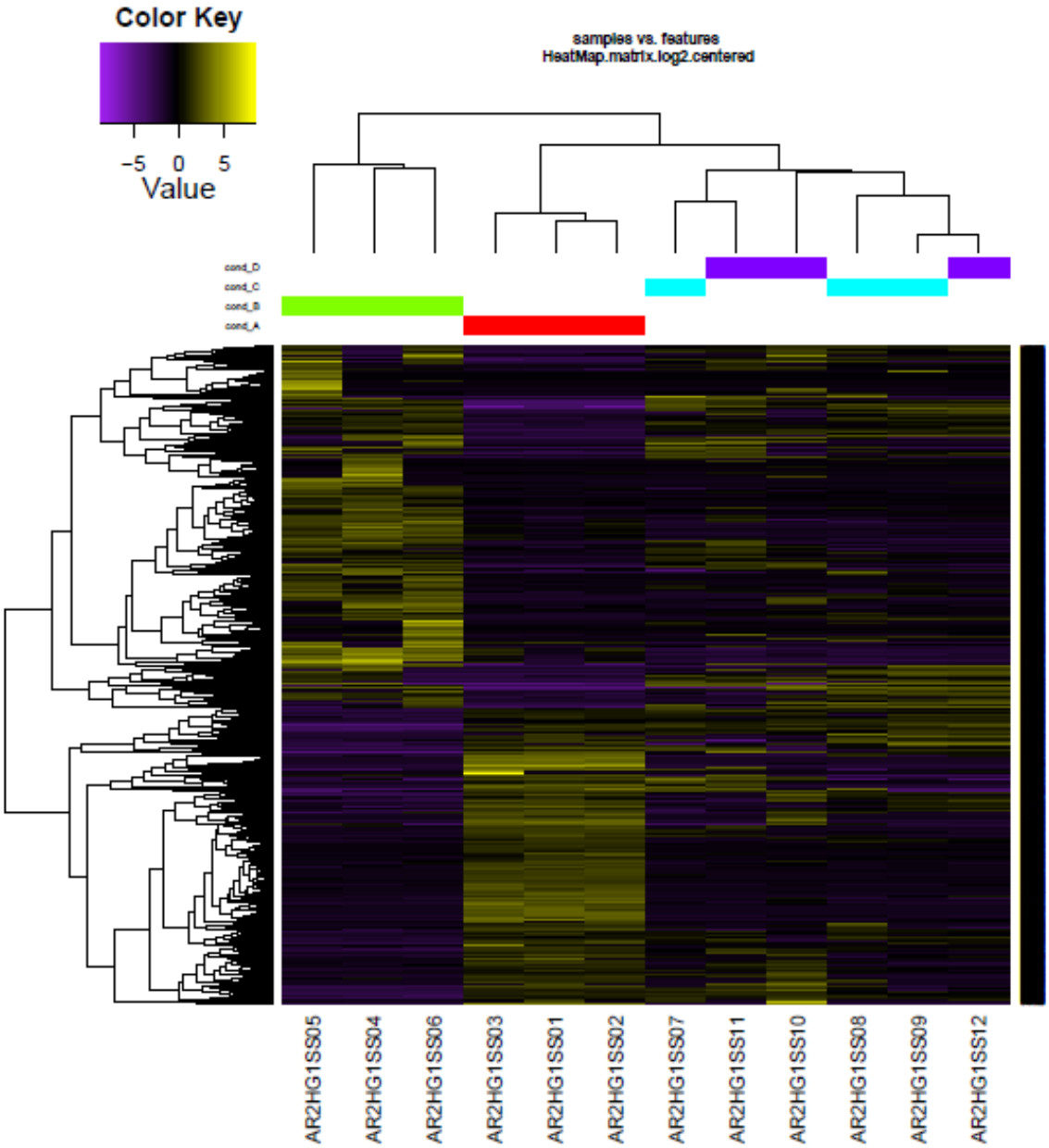

Figura 34. Mapa de calor. Representación de los perfiles de expresión genética de las muestras de Guayule en sus diferentes condiciones de colecta. El nombre de las 12 muestras se encuentra en la parte inferior del mapa. Las barras rojas, verde, azul celeste y morado, corresponde a las muestras colectadas en las condiciones A, B, C y D respectivamente. Una relación filogenética agrupa estas muestras según la expresión genética de cada una y muestra en la parte superior del mapa. Cada línea horizontal, corresponde a un gen determinado. Donde el color amarillo refiere a expresión alta y el morado expresión relativamente baja.

En cambio las muestras procedentes de San Luis Potosí, muestran un patrón de expresión heterogéneo, reflejo de la naturaleza silvestre y probablemente la alta variabilidad genética en la región.

Debido a lo anterior, se realizó un punto de corte en los valores de expresión génica de e-value al 0.05, esto con el objetivo de quedarse sólo con aquellos unigenes que se expresaran diferencialmente con un intervalo de confianza de 95%, lo cual redujo considerablemente el número de unigenes expresados diferencialmente. Al seguir trabajando con los contrastes establecidos en la figura 34, se encontró un total de 4684 genes regulados positivamente (up-regulated) y 9335 genes regulados negativamente (down-regulated) además de 417 genes que no mostraron diferencia en la expresión en ninguno de los contrastes establecidos (figura 35).

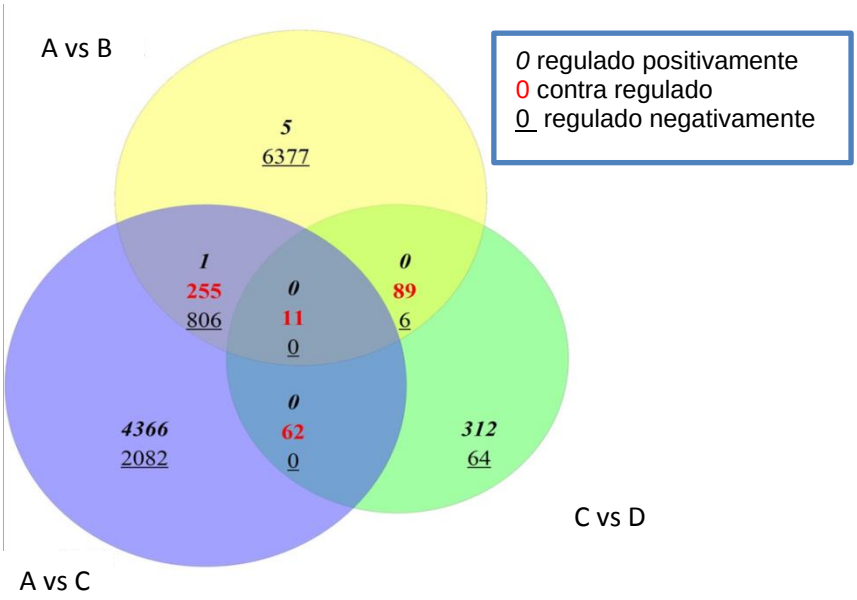

Figura 35. Unigenes expresados diferencialmente mediante los distintos contrastes analizados con un intervalo del 95% de confianza

La mayoría de las enzimas que se identificaron en las rutas MVA y MEV en este trabajo (figura 30 y 31), no mostraron una diferencia en su expresión entre plantas de alto y bajo contenido en hule. En el trabajo de Sando *et al.* realizado con *H. brasiliensis,* las proteínas AACT, HMGS, HMGR, MVK, PMK y MVD se encontraron altamente expresadas en el látex en comparación con otros tejidos de la planta, lo que llevo a la conclusión del papel importante que juegan en la biosíntesis de hule dichas enzimas. En plantas de *T. brevicorniculatum*, la sobre expresión del gen HMGR de *A. thaliana* y otros dos genes de la ruta MVA, aumentó su actividad enzimática y afectó a la acumulación de productos isoprenoides finales (Pütter *et al.*, 2017). Sin embargo, la sobreexpresión de HMGR en plantas de guayule no produjo un aumento en los niveles de resinas o hule natural (Dong *et al.*, 2013). Benedict *et al*., sugieren que la biosíntesis de esterol en las plantas de guayule en el momento de la aparición de nuevos brotes vegetativos, reduce la expresión de HMGR y la polimerización de hule (Benedict *et al.*, 2014).

En el presente estudio y otros estudios recientes en guayule (Ponciano *et al.*, 2012) y en *Taraxacum kok-saghyz*, (Luo *et al.*, 2017), no se encontraron genes expresados diferencialmente que se relacionaran directamente con las proteínas de partículas de hule. Los autores de Luo *et al.,* (2017) dejan en duda que estas proteínas putativas en realidad tengan un efecto directo en el rendimiento de hule y proponen la posibilidad de la existencia de nuevos genes no descritos.

Una expresión diferencial se observó solo en el caso de las enzimas Acetil CoA tiolasa (AACT) en la ruta MVA e isopentenil pirofosfato isomerasa (IPI) en la ruta

posterior a la síntesis de IPP, estas dos proteínas mostraron una regulación negativa (figura 36) en plantas de mayor contenido de hule. Existen varios estudios, de los cuales se destacan el de Pütter *et al*. (2017), los autores describen 3 enzimas claves en la síntesis de hule presentes en la ruta del MVA una de ellas la AACT, que al estar regulada positivamente influye en un aumento en la cantidad de hule. Sin embargo, los resultados obtenidos en este trabajo no son muy contradictorios, ya que las proteínas correspondientes a estos dos genes, junto a otros más, se encuentran más acumulados en las plantas del invierno (resultados de un estudio proteómico obtenidos con las mismas muestras/ comunicación personal A. Rodríguez). Cabe mencionar que el análisis de los genes expresados en un tejido determinado puede proporcionar información de mecanismos moleculares que ocurren en el tejido en el momento de la investigación. Una manera de explicar esto es que la expresión de estos genes se realizó en un tiempo previo a la toma de muestras y probablemente se realizó un almacenamiento post-transcripcional de las proteínas correspondientes, en cantidades suficientes para cumplir sus funciones en la polimerización de hule.

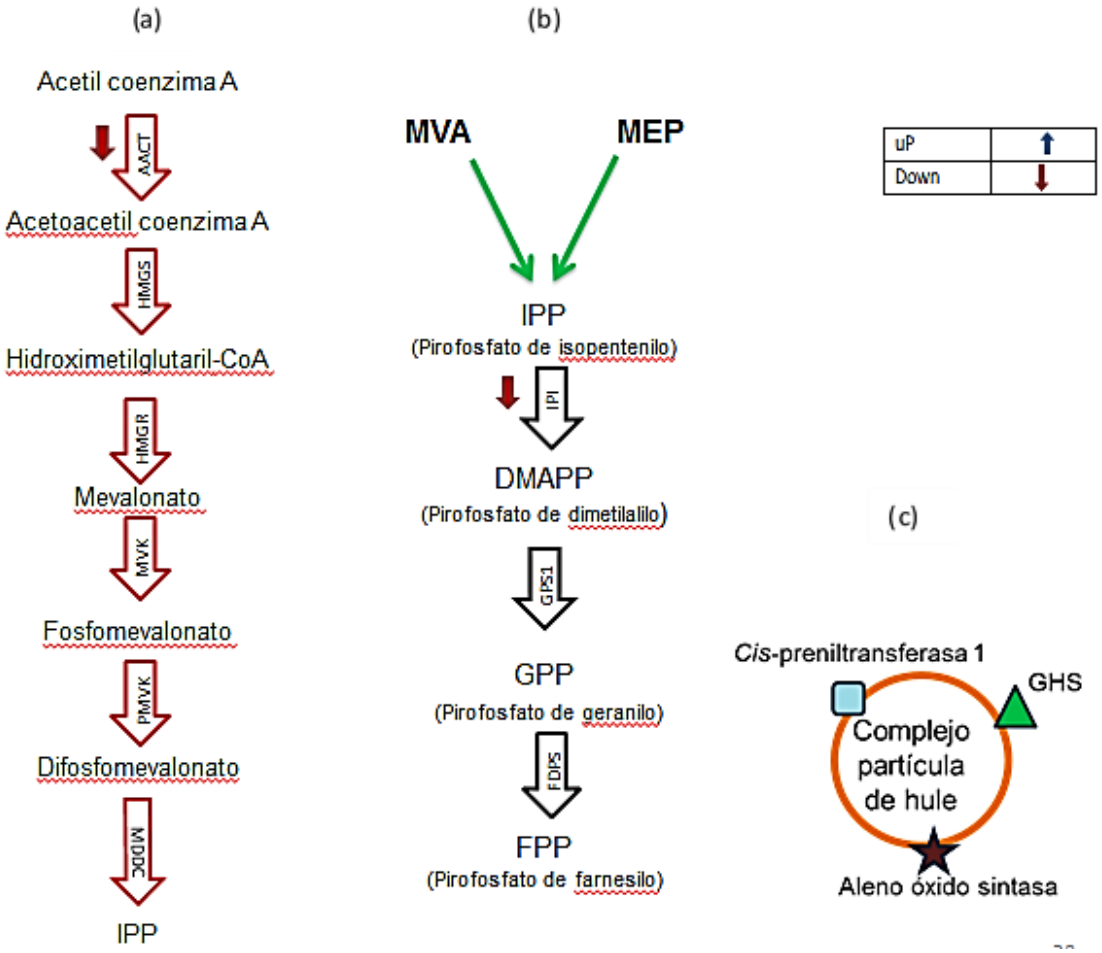

Figura 36. Ruta del mevalonato (a), síntesis de FPP (b) y componentes del complejo de la partícula de hule (c), las flechas rojas (⬇) indican una expresión diferencial negativa de los genes correspondientes en plantas de mayor contenido de hule

VI. CONCLUSIONES

1- Plantas de guayule con contenidos de hule cercano al 20% colectadas en Durango y San Luis Potosí podrían facilitar programas futuros de mejora genética.

2- Las plantas de guayule colectadas en verano presentaron un mayor contenido de hule lo cual puede ser atribuido al estrés hídrico que estas presentaron tanto en el estado de Durango como en San Luis Potosí.

3- Las enzimas Acetil CoA tiolasa (AACT) e isopentenil pirofosfato isomerasa. (IPP) muestran una regulación negativa en plantas de mayor contenido de hule en la época de invierno provenientes del estado de Durango.

VII. PERSPECTIVAS

- Realizar la reconstrucción de la ruta de síntesis de látex en guayule basados en los datos generados en este proyecto comparándolos con la base de datos de *H. brasiliensis.* Lo cual podría dar una mayor comprensión de lo que se sucede después de la síntesis del IPP en *Parthenium argentatum*.
- Una vez que el genoma de *P. argentatum* sea liberado realizar una comparación con la base de datos obtenida en este proyecto con el fin de generar mayor información respecto a la ruta de biosíntesis de hule y de los procesos que la regulan.
- En ensayos futuros someter plantas de guayule a estrés hídrico con el fin de obtener mayor acumulación de hule y de este modo buscar los genes involucrados en su biosíntesis a diferentes tiempos.

VIII. LITERATURA CITADA

Agrawal, A. A. y Konno, K. (2009). Latex: a model for understanding mechanisms, ecology, and evolution of plant defense against herbivory. Annual Review of Ecology, Evolution, and Systematics. 40: 311-331

Angulo-Sánchez, J. L., De Rodríguez, D. J. y Rodriguez-García, R. (2002). Relationship between guayule biomass production, rubber synthesis, and climatic conditions. Trends in New Crops and New Uses, American Society for Horticultural Science., Alexandria, VA. 234-237

Anisimova, M. y Gascuel, O. (2006). Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Systematic biology. 55(4): 539-552

Araújo, A. E. d., Kalil Filho, A. N., Nóbrega, M. B., Sousa, N. R. y Santos, J. W. d. (2001). Reaction of ten clones of rubber (*Hevea benthamiana*) to three Microcyclus ulei isolates. Acta Amazónica. 31(3): 349-349

Asawatreratanakul, K., Zhang, Y. W., Wititsuwannakul, D., Wititsuwannakul, R., Takahashi, S., Rattanapittayaporn, A. y Koyama, T. (2003). Molecular cloning, expression and characterization of cDNA encoding cis-prenyltransferases from *Hevea brasiliensis*. The FEBS Journal. 270(23): 4671-4680

Backhaus, R. A. y Walsh, S. (1983). The ontogeny of rubber formation in guayule, Parthenium argentatum Gray. Botanical Gazette. 144(3): 391-400

Bekaardt, C. R., Coffelt, T. A., Fenwick, J. R. y Wiesner, L. E. (2010). Environmental, irrigation and fertilization impacts on the seed quality of guayule (*Parthenium argentatum* Gray). Industrial Crops and Products. 31(3): 427-436

Benedict, C., Goss, R., Greer, P. J. y Foster, M. (2010). The formation of rubber-producing cortical parenchyma cells in guayule (Parthenium argentatum Gray) by low temperature. Industrial crops and products. 31(3): 516-520

Benedict, C., Greer, P. J. y Foster, M. (2008). The physiological and biochemical responses of guayule to the low temperature of the Chihuahuan Desert in the biosynthesis of rubber. industrial crops and products. 27(3): 225-235

Benedict, C., Rosenfield, C., Ji, W. y Foster, M. (2014). The low temperature induced rate of 3-hydroxy-3-methylglutaryl-CoA reductase in *Parthenium argentatum* Gray limits the theoretical rate of rubber formation. Industrial Crops and Products. 61: 176-179

Bicknell, R. A. y Koltunow, A. M. (2004). Understanding apomixis: recent advances and remaining conundrums. The Plant Cell Online. 16(suppl 1): S228-S245

Black, L., Hamerstrand, G. y Kwolek, W. (1985). Analysis of rubber, resin, and moisture content of guayule by near infrared reflectance spectroscopy. Rubber chemistry and technology. 58(2): 304-313

Blanc, G., Baptiste, C., Oliver, G., Martin, F. y Montoro, P. (2006). Efficient Agrobacterium tumefaciens-mediated transformation of embryogenic calli and regeneration of *Hevea brasiliensis* Müll Arg. plants. Plant cell reports. 24(12): 724-733

Bonner, J. (1943). Effects of temperature on rubber accumulation by the guayule plant. Botanical Gazette. 105(2): 233-243

Brandolini, A. J., Hills, D. D. y Vickers, T. J. (2001). NMR Spectra of Polymers and Polymer Additives. Applied Spectroscopy. 55: 108A

Bushman, B. S., Scholte, A. A., Cornish, K., Scott, D. J., Brichta, J. L., Vederas, J. C., Ochoa, O., Michelmore, R. W., Shintani, D. K., *et al.* (2006). Identification and comparison of natural rubber from two *Lactuca* species. Phytochemistry. 67(23): 2590-2596

Byrne, J. Review of World War II extracting and processing guayule rubber. En International Conference on the Utilization of Gyayule; Proceedings of Meeting, 1975.

Cantú, D. J., Angulo-Sánchez, J., Rodríguez-García, R. y Kuruvadi, S. (1997). Seasonal growth, rubber and resin yield characteristics of guayule under natural environmental conditions. Industrial Crops and Products. 6(2): 131-137

Cock, P. J., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., *et al.* (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics. 25(11): 1422-1423

Cornish, K. (1993). The separate roles of plant cis and trans prenyl transferases in cis-1, 4-polyisoprene biosynthesis. The FEBS Journal. 218(1): 267-271

Cornish, K. 1996. Hypoallergenic natural rubber products from parthenum argentatum (gray) and other non-*hevea brasiliensis* species. Google Patents.

Cornish, K. (2001). Similarities and differences in rubber biochemistry among plant species. Phytochemistry. 57(7): 1123-1134

Cornish, K., Chapman, M. H., Brichta, J. L., Vinyard, S. H. y Nakayama, F. S. (2000). Post-harvest stability of latex in different sizes of guayule branches. Industrial crops and products. 12(1): 25-32

Cornish, K., Martin, J. A., Marentis, R. T. y Plamthottam, S. 2007. Extraction and fractionation of biopolymers and resins from plant materials. Google Patents.

Cornish, K., Xie, W., Kostyal, D., Shintani, D. y Hamilton, R. G. (2015). Immunological analysis of the alternate rubber crop *Taraxacum kok-saghyz* indicates multiple proteins cross-reactive with *Hevea brasiliensis* latex allergens. Journal of Biotechnology & Biomaterials. 5(4): 1

Croteau, R., Kutchan, T. M. y Lewis, N. G. (2000). Natural products (secondary metabolites). Biochemistry and molecular biology of plants. 24: 1250-1319

Chappell, J. (1995). The biochemistry and molecular biology of isoprenoid metabolism. Plant Physiology. 107(1): 1

Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A. J., Müller, W. E., Wetter, T. y Suhai, S. (2004). Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome research. 14(6): 1147-1159

Chow, K.-S., Ghazali, A.-K., Hoh, C.-C. y Mohd-Zainuddin, Z. (2014). RNA sequencing read depth requirement for optimal transcriptome coverage in *Hevea brasiliensis*. BMC research notes. 7(1): 69

da Costa, B. M., Keasling, J. D. y Cornish, K. (2005). Regulation of Rubber Biosynthetic Rate and Molecular Weight in *Hevea brasiliensis* by Metal Cofactor. Biomacromolecules. 6(1): 279-289

de Rodríguez, D. J., Angulo-Sánchez, J. y Rodríguez-García, R. (2006a). An overview of guayule research and development in Mexico. Industrial crops and products. 24(3): 269-273

de Rodríguez, D. J., Ângulo-Sanchez, J. y Rodriguez-Garcia, R. (2006b). Mexican high rubber producing guayule shrubs: A potential source for commercial development. Journal of Polymers and the Environment. 14(1): 37-47

de Rodríguez, D. J., Angulo-Sánchez, J. L. y Rodríguez-García, R. (2001). Correlation and path coefficient analyses of the agronomic trait of a native population of guayule plants. Industrial Crops and Products. 14(2): 93-103

Dennis, M. S. y Light, D. (1989). Rubber elongation factor from Hevea brasiliensis. Identification, characterization, and role in rubber biosynthesis. Journal of Biological Chemistry. 264(31): 18608-18617

Dong, N., Ponciano, G., McMahan, C., Coffelt, T., Johnson, L., Creelman, R., Whalen, M. y Cornish, K. (2013). Overexpression of 3-hydroxy-3-methylglutaryl coenzyme A reductase in *Parthenium argentatum* (guayule). Industrial crops and products. 46: 15-24

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research. 32(5): 1792-1797

Ekblom, R. y Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity. 107(1): 1

Enfissi, E. M. A., Fraser, P. D., Lois, L.-M., Boronat, A., Schuch, W. y Bramley, P. M. (2005). Metabolic engineering of the mevalonate and non-mevalonate isopentenyl diphosphate-forming pathways for the production of health-promoting isoprenoids in tomato. Plant Biotechnology Journal. 3(1): 17-27

Espinal, C. F., Martínez Covaleda, H. J., Salazar Soler, M. y Barrios Urrutia, C. A. 2005. La cadena del caucho en Colombia: una mirada global de su estructura y dinamica 1991-2005. Ministerio de Agricultura y Desarrollo Rural-MINAGRICULTURA.

Evans, T. y Loose, M. (2015). AlignWise: a tool for identifying protein-coding sequence and correcting frame-shifts. BMC bioinformatics. 16(1): 376

García, A. Á. y Carril, E. P.-U. (2011). Metabolismo secundario de plantas. Reduca (biología). 2(3):

Gasparotto, L., Lieberei, R. y Trindade, D. (1984). Germination of *Microcyclus ulei* conidia and their sensitivity to fungicides, in vitro [rubber tree, *Hevea* spp, leaf blight]. Fitopatologia Brasileira (Brazil).

Gennaro, A. R. (2003). Remington farmacia. Vol. 2. ed. Ed. Médica Panamericana.

George, P. y Jacob, C. K. (2000). Natural rubber agromanagement and crop processing.

Goss, R. A., Benedict, C. R., Keithly, J. H., Nessler, C. L. y Stipanovic, R. D. (1984). cis-Polyisoprene synthesis in guayule plants (*Parthenium argentatum* Gray) exposed to low, nonfreezing temperatures. Plant physiology. 74(3): 534-537

Gouy, M., Guindon, S. y Gascuel, O. (2010). SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. Molecular Biology and Evolution. 27(2): 221-224

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., *et al.* (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology. 29(7): 644-652

Guindon, S. y Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic biology. 52(5): 696-704

Gutiérrez, C., Gonzalez-Coloma, A. y Hoffmann, J. J. (1999). Antifeedant properties of natural products from *Parthenium argentatum*, P. argentatum×P. tomentosum (Asteraceae) and Castela emoryi (Simaroubeaceae) against Reticulitermes flavipes. Industrial Crops and Products. 10(1): 35-40

Hagel, J. M., Yeung, E. C. y Facchini, P. J. (2008). Got milk? The secret life of laticifers. Trends in plant science. 13(12): 631-639

Hallahan, D. L. y Keiper-Hrynko, N. M. 2007. Cis-prenyltransferases from the rubber-producing plants russian dandelion (*Taraxacum kok-saghyz*) and sunflower (*Helianthus annus*). Google Patents.

Han, K.-H., Shin, D. H., Yang, J., Kim, I. J., Oh, S. K. y Chow, K. (2000). Genes expressed in the latex of *Hevea brasiliensis*. Tree physiology. 20(8): 503-510

Harrison, N. y Kidner, C. A. (2011). Next-generation sequencing and systematics: what can a billion base pairs of DNA sequence data do for you? Taxon. 60(6): 1552-1566

Hayashi, Y. (2009). Production of natural rubber from Para rubber tree. Plant Biotechnology. 26(1): 67-70

Jiang, M.-Y. y Zhang, J.-H. (2004). Abscisic acid and antioxidant defense in plant cells. ACTA BOTANICA SINICA-ENGLISH EDITION-. 46(1): 1-9

Jones, E. P. (1948). Recovery of rubber latex from guayule shrub. Industrial & Engineering Chemistry. 40(5): 864-874

Junqueira, N., Gasparotto, L., Lieberei, R., Normando, M. y Lima, M. (1989). Especialição fisiològica de *Microcyclus ulei* em diferentes especies de seringueira: Identifição de grupos de patotipos. Fitopatologia Brasileira. 14(2): 147

Kajiura, H., Suzuki, N., Mouri, H., Watanabe, N. y Nakazawa, Y. (2017). Elucidation of rubber biosynthesis and accumulation in the rubber producing shrub, guayule (Parthenium argentatum Gray). Planta. 1-14

Kim, I. J., Ryu, S. B., Kwak, Y. S. y Kang, H. (2004). A novel cDNA from *Parthenium argentatum* Gray enhances the rubber biosynthetic activity in vitro. Journal of experimental botany. 55(396): 377-385

Ko, J.-H., Chow, K.-S. y Han, K.-H. (2003). Transcriptome analysis reveals novel features of the molecular events occurring in the laticifers of *Hevea brasiliensis* (para rubber tree). Plant molecular biology. 53(4): 479-492

Kopicky, S. E. 2014. The Use of Near Infrared Spectroscopy in Rubber Quantification. Tesis de The Ohio State University.

Kuruvadi, S., Benitez, A. L. y de Rodríguez, D. J. (1997a). Rubber and resin contents of native guayule germplasm of Mexico. Industrial Crops and Products. 6(2): 139-145

Kuruvadi, S., Jasso Cantu, D. y Angulo-Sanchez, J. L. (1997b). Rubber content in different plant parts and tissues of Mexican guayule shrubs. Industrial Crops and Products. 7(1): 19-25

Lau, N.-S., Makita, Y., Kawashima, M., Taylor, T. D., Kondo, S., Othman, A. S., Shu-Chien, A. C. y Matsui, M. (2016). The rubber tree genome shows expansion of gene family associated with rubber biosynthesis. Scientific Reports. 6:

Lee, J., Eschen-Lippold, L., Lassowskat, I., Böttcher, C. y Scheel, D. (2015). Cellular reprogramming through mitogen-activated protein kinases. Frontiers in plant science. 6:

Li, B. y Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics. 12(1): 323

Li, D., Deng, Z., Qin, B., Liu, X. y Men, Z. (2012). De novo assembly and characterization of bark transcriptome using Illumina sequencing and development of EST-SSR markers in rubber tree (*Hevea brasiliensis* Muell. Arg.). BMC genomics. 13(1): 192

Li, L., Stoeckert, C. J. y Roos, D. S. (2003). OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. Genome Research. 13(9): 2178-2189

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., Li, Y., Li, S., Shan, G.*, et al.* (2010). De novo assembly of human genomes with massively parallel short read sequencing. Genome research. 20(2): 265-272

Lieberei, R. (2007). South American leaf blight of the rubber tree (*Hevea* spp.): new steps in plant domestication using physiological features and molecular markers. Annals of botany. 100(6): 1125-1142

Light, D. y Dennis, M. (1989). Purification of a prenyltransferase that elongates cis-polyisoprene rubber from the latex of *Hevea brasiliensis*. Journal of Biological Chemistry. 264(31): 18589-18597

Lin, T., Xu, X., Ruan, J., Liu, S., Wu, S., Shao, X., Wang, X., Gan, L., Qin, B.*, et al.* (2017). Genome analysis of *Taraxacum kok-saghyz* Rodin provides new insights into rubber biosynthesis. National Science Review.

Lin, Y., Li, J., Shen, H., Zhang, L., Papasian, C. J. y Deng, H. W. (2011). Comparative studies of de novo assembly tools for next-generation sequencing technologies. Bioinformatics. 27(15): 2031-2037

Lois, L. M., Campos, N., Putra, S. R., Danielsen, K., Rohmer, M. y Boronat, A. (1998). Cloning and characterization of a gene from *Escherichia coli* encoding a transketolase-like enzyme that catalyzes the synthesis of D-1-deoxyxylulose 5-phosphate, a common precursor for isoprenoid, thiamin, and pyridoxol biosynthesis. Proceedings of the National Academy of Sciences. 95(5): 2105-2110

Lou, X. X., Janssen, J. H.-G. y Cramers, C. (1997). Parameters affecting the accelerated solvent extraction of polymeric samples. Analytical Chemistry. 69(8): 1598

Luo, Z., Iaffaldano, B., Zhuang, X., Fresnedo-Ramírez, J. y Cornish, K. (2017). Analysis of the first *Taraxacum kok-saghyz* transcriptome reveals potential rubber yield related SNPs. Scientific reports. 7(1): 9939

Luque de Castro, M. D. y García-Ayuso, L. E. (1998). Soxhlet extraction of solid materials: an outdated technique with a promising innovative future. Analytica Chimica Acta. 369(1): 1-10

Macrae, S., Gilliland, M. G. y Van Staden, J. (1986). Rubber production in guayule: determination of rubber producing potential. Plant physiology. 81(4): 1027-1032

Madhavan, S. y Benedict, C. R. (1984). Isopentenyl pyrophosphate cis-1, 4-polyisoprenyl transferase from guayule (*Parthenium argentatum* Gray). Plant physiology. 75(4): 908-913

Metcalfe, C. (1967). Distribution of latex in the plant kingdom. Economic Botany. 21(2): 115-127

Mooibroek, H. y Cornish, K. (2000). Alternative sources of natural rubber. Applied microbiology and biotechnology. 53(4): 355-365

Moreno, R. M. B., Ferreira, M., Gonçalves, P. d. S. y Mattoso, L. H. C. (2005). Technological properties of latex and natural rubber of *Hevea brasiliensis* clones. Scientia Agricola. 62(2): 122-126

Nakayama, F. (2005). Guayule future development. Industrial Crops and Products. 22(1): 3-13

Núñez, C. E. (2008). Extracciones con equipo Soxhlet. Obtenido el. 15:

Nurthen, E. J., McCleary, B. V., Milthorpe, P. L. y Whitworth, J. W. (1986). Modified soxhlet procedure for the quantification of resin and rubber content of guayule. Analytical Chemistry. 58(2): 448-453

Oh, S. K., Kang, H., Shin, D. H., Yang, J., Chow, K.-S., Yeang, H. Y., Wagner, B., Breiteneder, H. y Han, K.-H. (1999). Isolation, characterization, and functional analysis of a novel cDNA clone encoding a small rubber particle protein from *Hevea brasiliensis*. Journal of Biological Chemistry. 274(24): 17132-17138

Ozsolak, F. y Milos, P. M. (2010). RNA sequencing: advances, challenges and opportunities. Nature reviews genetics. 12(2): nrg2934

Pascual-Villalobos, M. J. y López, M. D. (2013). New application of guayule resin in controlled release formulations. Industrial Crops and Products. 43(Supplement C): 44-49

Pearson, C., Cornish, K., McMahan, C., Rath, D. y Whalen, M. (2010). Natural rubber quantification in sunflower using an automated solvent extractor. Industrial crops and products. 31(3): 469-475

Pearson, C. H., Cornish, K. y Rath, D. J. (2013). Extraction of natural rubber and resin from guayule using an accelerated solvent extractor. Industrial Crops and Products. 43: 506-510

Polhamus, L. G. (1962). Rubber: Botany, Production and Utilization. Rubber: Botany, Production and Utilization.:

Ponciano, G., McMahan, C. M., Xie, W., Lazo, G. R., Coffelt, T. A., Collins-Silva, J., Nural-Taban, A., Gollery, M., Shintani, D. K.*, et al.* (2012). Transcriptome and gene expression analysis in cold-acclimated guayule (*Parthenium argentatum*) rubber-producing tissue. Phytochemistry. 79: 57-66

Post, J., van Deenen, N., Fricke, J., Kowalski, N., Wurbs, D., Schaller, H., Eisenreich, W., Huber, C., Twyman, R. M.*, et al.* (2012). Laticifer-specific cis-prenyltransferase silencing affects the rubber, triterpene, and inulin content of T*araxacum brevicorniculatum*. Plant physiology. 158(3): 1406-1417

Priya, P., Venkatachalam, P. y Thulaseedharan, A. (2006). Molecular cloning and characterization of the rubber elongation factor gene and its promoter sequence from rubber tree (*Hevea brasiliensis*): A gene involved in rubber biosynthesis. Plant science. 171(4): 470-480

Pütter, K. M., van Deenen, N., Unland, K., Prüfer, D. y Schulze Gronover, C. (2017). Isoprenoid biosynthesis in dandelion latex is enhanced by the overexpression of three key enzymes involved in the mevalonate pathway. BMC Plant Biology. 17(1): 88

Rahman, A. Y. A., Usharraj, A. O., Misra, B. B., Thottathil, G. P., Jayasekaran, K., Feng, Y., Hou, S., Ong, S. Y., Ng, F. L.*, et al.* (2013). Draft genome sequence of the rubber tree Hevea brasiliensis. BMC genomics. 14(1): 75

Remm, M., Storm, C. E. y Sonnhammer, E. L. (2001). Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. Journal of molecular biology. 314(5): 1041-1052

Rodríguez-Concepción, M. y Boronat, A. (2002). Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. Plant Physiology. 130(3): 1079-1089

Ruëff, F. (2014). Sensitization and Allergy to Natural Rubber Latex. En Encyclopedia of Medical Immunology. (Eds.). ed. Vol. Springer. pp. 615-619

Salvucci, M. E., Coffelt, T. A. y Cornish, K. (2009). Improved methods for extraction and quantification of resin and rubber from guayule. Industrial Crops and Products. 30(1): 9-16

Sando, T., Takeno, S., Watanabe, N., Okumoto, H., Kuzuyama, T., Yamashita, A., Hattori, M., Ogasawara, N., Fukusaki, E.*, et al.* (2008). Cloning and characterization of the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway genes of a natural-rubber producing plant, *Hevea brasiliensis*. Bioscience, biotechnology, and biochemistry. 72(11): 2903-2917

Schmidt, T., Hillebrand, A., Wurbs, D., Wahler, D., Lenders, M., Gronover, C. S. y Prüfer, D. (2010a). Molecular cloning and characterization of rubber biosynthetic genes from *Taraxacum koksaghyz.* Plant molecular biology reporter. 28(2): 277-284

Schmidt, T., Lenders, M., Hillebrand, A., van Deenen, N., Munt, O., Reichelt, R., Eisenreich, W., Fischer, R., Prüfer, D.*, et al.* (2010b). Characterization of rubber particles and rubber chain elongation in *Taraxacum koksaghy*z. BMC biochemistry. 11(1): 11

Siler, D. J. y Cornish, K. (1994a). Hypoallergenicity of guayule rubber particle proteins compared to *Hevea* latex proteins. Industrial Crops and Products. 2(4): 307-313

Siler, D. J. y Cornish, K. (1994b). Identification of *Parthenium argentatum* rubber particle proteins immunoprecipitated by an antibody that specifically inhibits rubber transferase activity. Phytochemistry. 36(3): 623-627

Suchat, S., Pioch, D., Palu, S., Tardan, E., van Loo, E. N. y Davrieux, F. (2013). Fast determination of the resin and rubber content in *Parthenium argentatum* biomass using near infrared spectroscopy. Industrial Crops and Products. 45: 44-51

Swanson, C., Buchanan, R. y Otey, F. (1979). Molecular weights of natural rubbers from selected temperate zone plants. Journal of applied polymer science. 23(3): 743-748

Tan, D., Hu, X., Fu, L., Kumpeangkeaw, A., Ding, Z., Sun, X. y Zhang, J. (2017). Comparative morphology and transcriptome analysis reveals distinct functions of the primary and secondary laticifer cells in the rubber tree. Scientific reports. 7(1): 3126

Tang, C., Huang, D., Yang, J., Liu, S., Sakr, S., Li, H., Zhou, Y. y Qin, Y. (2010). The sucrose transporter HbSUT3 plays an active role in sucrose loading to laticifer and rubber productivity in exploited trees of *Hevea brasiliensis* (para rubber tree). Plant, cell & environment. 33(10): 1708-1720

Tang, C., Yang, M., Fang, Y., Luo, Y., Gao, S., Xiao, X., An, Z., Zhou, B., Zhang, B.*, et al.* (2016). The rubber tree genome reveals new insights into rubber production and species adaptation. Nature plants. 2(6): 16073-16073

Teetor, V., Ray, D. y Schloman, W. (2009). Evaluating chemical indices of guayule rubber content: Guayulins A and B. industrial crops and products. 29(2): 590-598

Tholl, D. (2006). Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. Current opinion in plant biology. 9(3): 297-304

Thompson, A. E. y Ray, D. T. (1989). Breeding guayule. Plant Breeding Reviews, Volume 6. 93-165

Thurbide, K. B. y Hughes, D. M. (2000). A rapid method for determining the extractives content of wood pulp. Industrial & engineering chemistry research. 39(8): 3112-3115

Triwitayakorn, K., Chatkulkawin, P., Kanjanawattanawong, S., Sraphet, S., Yoocha, T., Sangsrakru, D., Chanprasert, J., Ngamphiw, C., Jomchai, N.*, et al.* (2011). Transcriptome sequencing of *Hevea brasiliensis* for development of microsatellite markers and construction of a genetic linkage map. DNA research. 18(6): 471-482

van Beilen, J. B. y Poirier, Y. (2007). Guayule and Russian dandelion as alternative sources of natural rubber. Critical reviews in biotechnology. 27(4): 217-231

Van Beilen, J. B. y Poirier, Y. (2008). Production of renewable polymers from crop plants. The Plant Journal. 54(4): 684-701

Van Verk, M. C., Hickman, R., Pieterse, C. M. J. y Van Wees, S. C. M. (2013). RNA-Seq: revelation of the messengers. Trends in Plant Science. 18(4): 175-179

Veatch-Blohm, M. E., Ray, D. T. y McCloskey, W. B. (2006). Water-stress-induced changes in resin and rubber concentration and distribution in greenhouse-grown guayule. Agronomy journal. 98(3): 766-773

Wang, Z., Gerstein, M. y Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nature reviews genetics. 10(1): 57-63

Whaley, W. y Bowen, J. (1947). Russian dandelion (*kok-saghyz*): an emergency source of natural rubber (US Department of Agriculture).

Wood, M. (2002). Desert shrub may help preserve wood. Agricultural Research. 50(4): 10

Xia, Z., Xu, H., Zhai, J., Li, D., Luo, H., He, C. y Huang, X. (2011). RNA-Seq analysis and de novo transcriptome assembly of *Hevea brasiliensis*. Plant molecular biology. 77(3): 299

Zerbino, D. R. y Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome research. 18(5): 821-829

Zhao, Q.-Y., Wang, Y., Kong, Y.-M., Luo, D., Li, X. y Hao, P. (2011). Optimizing de novo transcriptome assembly from short-read RNA-Seq data: a comparative study. BMC bioinformatics. 12(14): S2

IX. ANEXOS

9.1. Extracción de ARN total (I)

1. Por cada 100 g de muestra macerada agregar 1 ml de TRI Reagent y homogeneizar.
2. Centrifugar a 12,000 g durante 10 min a 4^0C.
3. Recuperar sobrenadante, pasarlo a un tubo eppendorf limpio dejar reposar durante 5 min a temperatura ambiente.
4. Agregar 200 µl de cloroformo por cada ml de TRI Reagent. Mezclar vigorosamente 15 s. Y dejar a temperatura ambiente durante 10 min.
5. Centrifugar 12,000 g durante 15 min a 4^0C.
6. Recuperar sobrenadante, pasarlo a un tubo eppendorf limpio y agregar 500 µl de isopropanol mezclar suavemente y agregar 200 µl de cloruro de litio volver a mezclar. Dejar reposar 10 min a temperatura ambiente.
7. Centrifugar 12,000 g durante 15 min a 4^0C.
8. Desechar sobrenadante y agregar 1 ml de etanol al 75%.
9. Centrifugar a 7,500 g 5 min a 4^oC.
10. Tirar sobrenadante y dejar secar de 5 – 10 min.
11. Resuspender el pelet en 50 µl de agua libre de RNasas.

9.2. Purificación de ARN total (II)

1. Ajustar el volumen de la muestra a 100 µl con agua libre de RNasas.
2. Adicionar 350 µl de buffer RLT y mezclar.
3. Adicionar 250 µl etanol (96 – 100%) y mezclar bien.
4. Transferir la muestra (700µl) a la columna color rosa, cerrar la tapa y centrifugar por 15 s a 8,000 rcf a temperatura ambiente. Desechar sobrenadante.
5. Adicionar 500 µl de buffer RPE a la columna. Centrifugar por 15 s a 8,000 rcf a temperatura ambiente.
6. Adicionar 500 µl de buffer RPE. Centrifugar a 8,000 rfc durante 2 min.
7. Opcional: Secar la columna. Centrifugar 1 min a toda velocidad.
8. Cambiar la columna a un tubo de recolección de 1.5 ml y directamente sobre la columna adicionar 50 µl de agua libre de RNasas. Centrifugar por 1 min a 8,000 rcf.