



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



CEIB
CENTRO DE INVESTIGACIÓN
EN BIOTECNOLOGÍA UAEM

Doctorado en Ciencias Naturales

**Comparación de transcriptomas de dos poblaciones de *Galphimia*
spp. para identificar posibles genes candidatos que codifican
enzimas de la ruta de síntesis de las galfiminas**

Tesis para obtener el grado de
Doctora en Ciencias Naturales

Presenta
M.C. Dianella Iglesias Rodríguez

Director de tesis: **Dra. María Luisa Teresa Villarreal Ortega**
Co-director de tesis: **Dr. Ashutosh Sharma**

Cuernavaca, Morelos
2022

Esta tesis fue desarrollada en:

1. Laboratorio de Investigación en Plantas Medicinales, Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos, Cuernavaca, dirigido por la Dra. María Luisa Teresa Villarreal Ortega.



2. Centro de Bioingenierías, Instituto Tecnológico y de Estudios Superiores de Monterrey, Querétaro, dirigido por el Dr. Ashutosh Sharma.



3. Laboratorio de Bioinformática, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Ciudad de México, dirigido por el Dr. Alfonso Méndez Tenorio.



Índice general

Listado de figuras.....	i
Listado de cuadros.....	iii
Listado de abreviaturas.....	iv
Resumen	vii
Abstract.....	ix
1. Introducción.....	1
2. Marco teórico.....	4
2.1. Generalidades de <i>Galphimia</i> spp.	4
2.2. Estudios en <i>Galphimia</i> spp.....	4
2.3. Terpenos	10
2.3.1. Síntesis de terpenos.....	11
2.3.2. Síntesis de triterpenos de estructura pentacíclica.....	13
2.4. Secuenciación de ácidos nucleicos	21
2.5. Justificación	22
2.6. Hipótesis	22
2.7. Objetivo general	22
2.8. Objetivos específicos	22
3. Materiales y métodos	29
3.1. Obtención de plantas de <i>Galphimia</i> spp.....	29
3.2. Determinación de la presencia de galfiminas en los extractos metanólicos de <i>Galphimia</i> spp.	30
3.2.4. Secuenciación de transcriptomas de <i>Galphimia</i> spp. mediante RNAseq con el uso de Illumina Next Seq 550	30
3.2.5. Extracción de RNA	30
3.3. Análisis de los transcriptomas de <i>Galphimia</i> spp. mediante el uso de softwares especializados	37
3.3.1. Análisis de la calidad de las lecturas, depuración y ensamblado <i>de novo</i> de los transcriptomas	37
3.4. Identificación de los posibles genes candidatos relacionados con la biosíntesis de galfiminas.....	38
3.5. Validación de los datos del transcriptoma mediante PCR cuantitativa	38
4. Resultados y discusión.....	41
4.1. Obtención de plantas de <i>Galphimia</i> spp.....	41
4.2. Determinación de la presencia de galfiminas en el extracto metanólico de <i>Galphimia</i> spp.	46

4.2.1. Obtención de los extractos	46
4.2.4. Secuenciación de transcriptomas de <i>Galphimia</i> spp. mediante RNAseq con el uso de Illumina Next Seq 550	49
4.2.5. Extracción de RNA	49
4.3. Análisis de los transcriptomas de <i>Galphimia</i> spp. mediante el uso de softwares especializados	54
4.3.1. Análisis de la calidad de las lecturas, depuración y ensamblado <i>de novo</i> de los transcriptomas	54
4.4. Identificación de los posibles genes candidatos relacionados con la biosíntesis de galfiminas.....	64
4.5. Validación de los datos del transcriptoma mediante una PCR cuantitativa.....	72
5. Conclusiones	76
6. Perspectivas	77
7. Referencias bibliográficas	
8. Anexos	

Listado de figuras

Figura 1. Estructura del ácido tetragaloilquínico.....	5
Figura 2. Estructura de las galfiminas (A-K). Diferencias entre los grupos funcionales y estereoquímica del anillo E.....	7
Figura 3. Estructura de las glaucacetalinas (A-D).....	8
Figura 4. Síntesis de terpenos mediante la ruta del ácido mevalónico (MVA) y del fosfato del metileritritol (MEP).....	12
Figura 5. Modificaciones de triterpenos después la ciclación del 2,3 oxidoescualeno por oxidoescualeno ciclasa (OSC).....	16
Figura 6. Propuesta biogenética para la síntesis de galfiminas a partir de plantas transformadas vía <i>Agrobacterium rhizogenes</i> , con modificaciones.....	18
Figura 7. Síntesis de galfiminas mediante el cation taraxasterol.....	20
Figura 8. Secuenciación en la plataforma Illumina.....	22
Figura 9. Germinación y crecimiento de plantas de dos poblaciones de <i>Galphimia</i> spp. en condiciones semicontroladas de temperatura (25°C) y humedad relativa (60%).....	42
Figura 10. Plantas de dos poblaciones de <i>Galphimia</i> spp. cultivadas durante 17 meses....	45
Figura 11. Separación de compuestos en extractos metanólicos de <i>Galphimia</i> spp. mediante TLC.....	47
Figura 12. Separación de galfiminas por CLAE en extractos metanólicos de dos poblaciones de <i>Galphimia</i> spp.....	48
Figura 13. Electroforesis en gel de agarosa del RNA extraído de hojas y raíces de plantas de dos poblaciones de <i>Galphimia</i> spp.....	50
Figura 14. Electroferogramas de muestras de RNA de hojas y raíces de <i>Galphimia</i> spp.	51
Figura 15. Electroferogramas de librerías de cDNA de <i>Galphimia</i> spp.....	53
Figura 16. Niveles de distribución de ontología génica de plantas de dos poblaciones de <i>Galphimia</i> spp.....	58
Figura 17. Anotación de secuencias de dos poblaciones de <i>Galphimia</i> spp. en los niveles de ontología génica Procesos Biológicos (PB), Función Molecular (FM) y Componente Celular (CC).....	59
Figura 18. Código de distribución de enzimas de secuencias de dos poblaciones de <i>Galphimia</i> spp.....	60
Figura 19. Expresión diferencial de transcritos de plantas de dos poblaciones de <i>Galphimia</i> spp.....	61

Figura 20. Comparación de transcritos con mayores valores de expresión diferencial en dos poblaciones de <i>Galphimia</i> spp.....	62
Figura 21. Comparación de transcritos con mayores valores de expresión diferencial en hojas y raíces de la población productora de galfiminas.....	63
Figura 22. Ruta biosintética de la Friedelina pentacíclica a través de la protonación del oxidoscualeno y la formación de cationes intermediarios por reacciones de ciclación, reordenamientos y desprotonación.....	68
Figura 23. Propuesta de ruta de síntesis de las galfiminas mediada por el citocromo P450 en plantas de <i>Galphimia</i> spp.....	69
Figura 24. Red de co-expresión de transcritos relacionados con la ruta de síntesis de triterpenos y miembros de la familia del citocromo P450 en las poblaciones productoras y no productoras de galfiminas.....	71
Figura 25. Valores de umbral de ciclo (CT) de muestras de hojas de dos poblaciones de <i>Galphimia</i> spp. calculados por una PCR cuantitativa.....	73
Figura 26. Expresión relativa de posibles genes candidatos genes en hojas de plantas productoras de galfiminas (PG) y no productoras (NPG) calculada por el método $2^{-\Delta\Delta Ct}$ en una PCR cuantitativa.....	74

Listado de Cuadros

Cuadro 1. Secuencias de primers sintetizados para qPCR.....	39
Cuadro 2. Desarrollo de plantas de dos poblaciones de <i>Galphimia</i> spp. desde los 14 días hasta los 15 meses posteriores a la germinación.....	43
Cuadro 3. Análisis espectrofotométrico y fluorométrico de muestras de RNA de hojas y raíces de plantas de <i>Galphimia</i> spp.....	49
Cuadro 4. Medidas de calidad del ensamblado <i>de novo</i> de los transcriptomas de plantas de <i>Galphimia</i> spp.....	56
Cuadro 5. Anotación y conteo génico de transcritos relacionados con la síntesis de terpenos en dos poblaciones de <i>Galphimia</i> spp.....	65
Cuadro 6. Transcripción inversa de muestras de hojas de las poblaciones productoras y no productoras de galfiminas de <i>Galphimia</i> spp.....	72

Listado de abreviaturas

DNA	Ácido desoxirribonucleico
RNA	Ácido ribonucleico
RNA-seq	Secuenciación de RNA
NGS	Secuenciación de siguiente generación
SRA	Sequence Read Archive (archivos de lecturas de secuencias)
NCBI	The National Center for Biotechnology Information (Centro Nacional de Información Biotecnológica)
PG	productora de galfiminas
NPG	no productora de galfiminas
m	metro
pH	concentración de iones H ⁺
SNC	Sistema nervioso central
GABA	ácido gamma-aminobutírico
K⁺	potasio
RMN	Resonancia Magnética Nuclear
mat K	maturasa K
rbcl	ribulosa bifosfato carboxilasa
rpoC1	RNA polimerasa de plastidios
psbA-trnH	intergenic non-transcribed spacer regions (región intergénica)
ITS	internal transcribed spacer (región espaciadora intergénica)
MVA	ácido mevalónico
MEP	fosfato del metileritriol
AACT	acetoacetil-CoA tiolasa
HMGS	3-hidroxi-3-metil-glutaril-CoA sintetasa
HMGR	3-hidroxi-3-metilglutaril- CoA reductasa
MK	mevalonato quinasa
PMK	fosfato mevalonato descarboxilasa
MDC	mevalonato descarboxilasa
DXS	5-fosfono-1 desoxi-D-xilulosa sintasa
DXR	1 desoxi-D-xilulosa reductoisomerasa
MCT	2-C-metil-D-eritriol 4-fosfato citidiltransferasa
CMK	4-difosfocitidil-2-C-metil-D-eritritol quinasa
MDS	2-C-metil-D-eritrol-2,4-ciclodifosfato sintasa
HDS	1-hidroxi-2-metil-2-(E)-butenil-4-difosfato sintasa
HDR	1-hidroxi-2-metil-2-(E)-butenil-4-difosfato reductasa
IDI	isopentidil pirofosfato-dimetilalil pirofosfato isomerasa
GPPS	geranil pirofosfato sintasa
FPPS	farnesil pirofosfato sintasa
GGPPS	geranil geranil pirofosfato sintasa
SM	escualeno monoxigenasa
OSC	oxido escualeno ciclasa
CYP	Citocromo P450
A	adenina
G	guanina
T	timina
C	citosina
mRNA	Ácido ribonucleico mensajero
SEMARNAT	Secretaría de Medio Ambiente y Recursos Naturales

L	litros
mg	miligramos
µL	microlitros
min	minutos
rpm	revolución por minuto
v/v	volumen/volumen
g	gramos
ml	mililitros
TLC	Thin-layer chromatography (cromatografía en capa fina)
CLAE	Cromatografía líquida de alta eficiencia
µm	micrómetro
mm	milímetro
TBE	Tris-borato-EDTA
bp	pares de bases
RQN	RNA Quality Number (número de calidad de RNA)
V	voltio
nt	nucleótido
ng	nanogramos
cDNA	DNA obtenido por transcripción inversa
RNAr	Ácido ribonucleico ribosomal
RBB	rRNA Binding Buffer (buffer de unión a RNA ribosomal)
RRMP	rRNA Removal Mix-Plant (Mezcla para remover RNA ribosomal)
g xg	gravidades
RRB	rRNA Removal Beads (perlas de eliminación de RNA)
ELB	Elution Buffer (buffer de elución)
FSA	First Strand Synthesis Act D Mix (Mezcla de síntesis de la primera cadena de DNA)
CTE	End Repair Control (control de reparación de extremos)
RSB	Resuspension Buffer (buffer de resuspensión)
SMM	Second Strand Marking Master Mix (Mezcla para la síntesis de la segunda cadena de DNA)
h	hora
CTA	A-Tailing Control (control de la cola poli A)
ATL	A-Tailing Mix (Mezcla para adicionar la cola poli A)
CTL	Ligation control (control de ligamiento)
LG	Ligation Mix (Mezcla de ligamiento)
STL	Stop Ligation Buffer (buffer para detener el ligamiento)
PMM	PCR Master Mix (Mezcla para PCR)
nM	nanomolar
NaOH	hidróxido de sodio
N	normal
mM	micromolar
pM	picomolar
HT1	Hybridization Buffer (buffer de hibridación)
BLAST	Basic Local Alignment Search Tool (Herramienta de búsqueda de alineamientos locales)
OG	Ontología génica
PB	Proceso biológico
FM	Función molecular

CC	Componente celular
TAIR	The Arabidopsis Information Resource (Recurso de información de Arabidopsis)
CPM	conteos por millón
TPM	transcritos por millón
FDR	False Positive Rates (Tasa de falsos positivos)
logFC	logaritmo de Fold Change
cm	centímetro
RPG	raíces de la población productora de galfiminas
RFU	Unidad de fluorescencia relativa
Q30	quality score (puntuación de calidad de secuenciación)
N50	tamaño promedio de los contigs que cubren la mitad del ensamblado
BAS	beta-amirina sintasa
LS	lanosterol sintasa
SE	escualeno epoxidasa
SS	escualeno sintasa
FT	farnesil transferasa
AMPA	5'-AMP- proteína quinasa activada
Ser/Thr	serina/treonina-proteína quinasa AtPK2/AtPK19
CT	umbral de ciclo

Resumen

La calderona amarilla, especie perteneciente al género *Galphimia* (Malpighiaceae), es una planta muy usada en la medicina tradicional mexicana. Sus propiedades ansiolíticas y sedantes la han convertido en el objetivo de estudios farmacológicos, fitoquímicos, metabolómicos y genéticos. Estas propiedades se deben a la presencia de galfiminas que son compuestos triterpenoides modificados del tipo nor-secofriedelanos. La ruta de síntesis de las galfiminas no se encuentra elucidada aun, aunque existen propuestas biogénicas no demostradas experimentalmente que han intentado explicar su origen. Debido a que las galfiminas son metabolitos con una actividad biológica ampliamente demostrada, resulta pertinente identificar los genes que codifican las enzimas que intervienen en su ruta de síntesis. La transcriptómica es una herramienta con la cual se pueden estudiar un conjunto de genes que se expresan bajo una condición o estímulo específico, por lo que un análisis de esta clase permitiría identificar los posibles genes implicados en la ruta de síntesis de las galfiminas. El objetivo de este trabajo fue comparar transcriptomas de dos poblaciones de *Galphimia* spp., una productora y otra no productora de galfiminas, para identificar los posibles genes candidatos relacionados con la ruta de síntesis de estos triterpenoides. Plantas de las dos poblaciones fueron cultivadas bajo condiciones semicontroladas de temperatura y humedad relativa. Se determinó la presencia de galfiminas en los extractos metanólicos de plantas de ambas poblaciones mediante cromatografía en capa fina y cromatografía líquida de alta eficiencia. Se secuenciaron librerías de cDNA en la plataforma Illumina Next Seq 550. Se realizó la anotación funcional de los transcriptomas ensamblados en términos de ontología génica en tres niveles: proceso biológico, función molecular y componente celular. Se realizó un análisis de expresión diferencial con la librería edgeR. Se identificaron posibles genes candidatos relacionados con la síntesis de galfiminas y se construyó una red de co-expresión para establecer relaciones entre los transcritos que codificaron enzimas relacionadas con la síntesis de terpenos. Se validaron los resultados del análisis transcriptómico mediante una PCR cuantitativa. Las plantas de *Galphimia* spp. obtenidas de ambas poblaciones presentaron los caracteres fenotípicos descritos para este género. Se determinó la presencia de

galfiminas en la población productora de estos metabolitos. Las lecturas obtenidas de la secuenciación constaron de elevada calidad, lo que permitió realizar un ensamblado adecuado de los transcriptomas. La mayor parte de los transcritos ensamblados se agruparon en los niveles de ontología génica correspondiente a función molecular y proceso biológico. En las plantas productoras de galfiminas se identificó un mayor número de transcritos con expresión diferencial regulados positivamente. Se identificaron como posibles genes candidatos relacionados con la síntesis de galfiminas a los involucrados en la expresión de los miembros de enzimas de la superfamilia del citocromo P450 que funcionan como monooxigenasas. Durante la validación del transcriptoma de la planta productora se evidenció una mayor expresión relativa de los genes analizados. La comparación de transcriptomas de una población productora de galfiminas y otra no productora permitió la identificación del P450 como posible gen candidato relacionado con la ruta de síntesis de las galfiminas.

Palabras clave: anotación funcional, expresión diferencial, galfiminas, transcriptómica, triterpenoides.

Abstract

Galphimia spp. is a plant widely used in Mexican Traditional Medicine. The anxiolytic and sedative effects of the extract have made this plant a target of pharmacological, phytochemical, metabolomic, and genetic studies. These properties are exerted by modified triterpenoid named as galphimines. The synthetic pathway for galphimines is not yet elucidated, although there are biogenetic proposals not experimentally demonstrated, trying to explain their origin. Since galphimines are metabolites with interesting biological activities, it is important to identify the genes that encode the enzymes involved in their synthesis pathway. Transcriptomic is a technology that studies a set of genes that are expressed under specific conditions or stimulus, so an analysis of this platform will allow for the identification of the possible genes involved in the galphimines synthetic pathway. This work was aimed to compare transcriptomes of two populations of *Galphimia* spp. one galphimine producer and one non-galphimine producer to identify possible candidate genes related to the galphimines synthetic pathway. Plant samples from two populations of *Galphimia* spp (galphimine producer and non-galphimine producer), were obtained under semi-controlled conditions of temperature and relative humidity. The presence of galphimines in the methanolic plant extracts from both populations was-determined by thin layer chromatography and high-efficiency liquid chromatography. cDNA libraries were sequenced on the Illumina Next Seq 550 platform. Functional annotation of the assembled transcriptomes was performed in terms of gene ontology at the biological process, molecular function and cellular component levels. Differential expression analysis was performed with the edgeR library. Putative candidate genes, related to galphimine synthesis were identified and a co-expression network was constructed to establish relationships among transcripts encoding enzymes related to terpene synthesis. The results of the transcriptomic analysis were validated by quantitative PCR. *Galphimia* spp. obtained from both populations presented the phenotypic characters described for this genus. The presence of galphimines was determined in the population that produces these metabolites. The reads obtained from the sequencing were of high quality, which made it possible to assemble the transcriptomes appropriately. Most of the

assembled transcripts were grouped at the gene ontology molecular function and biological process levels. In the galphimine-producer plants, a greater number of transcripts positively regulated was identified. Members of the cytochrome P450 family that function as monooxygenases were regarded as putative candidate genes related to galphimine synthesis. During the validation of the transcriptome of the producer plant, a greater relative expression of the analyzed genes is evidenced. The comparison of transcriptomes of a galphimine producer population with a non-galphimine producer population, made it possible to identify P450 as a putative candidate gene related in the galphimine synthesis pathway.

Keywords: functional annotation, differential expression, galphimines, transcriptomics, triterpenoids.

1. Introducción

Galphimia spp. es una planta muy usada en la medicina tradicional mexicana desde épocas prehispánicas. Esto se debe a su utilidad para bajar la fiebre, aliviar dolores de parto (Tortoriello et al., 1998), contener diarreas (Camacho et al., 2002), su efecto antiinflamatorio (González-Cortazar et al., 2014) y para disminuir alteraciones del sistema nervioso central (SNC) (Campos et al., 2001). La actividad ansiolítica y sedante de los extractos obtenidos a partir de sus partes vegetativas, la han convertido en el objetivo de estudios farmacológicos (Romero-Cerecero et al., 2018), fitoquímicos (Cardoso-Taketa et al., 2004), metabólicos (Cardoso-Taketa et al., 2008; Sharma et al., 2012a) y genéticos (Sharma et al., 2012b; Gesto-Borroto et al., 2019). Estos efectos se deben a la presencia de galfiminas que son compuestos triterpenoides modificados (Toscano et al., 1993).

Las galfiminas constituyen un marcador quimiotaxonómico presente en la especie *Galphimia glauca*. La presencia o ausencia de estos metabolitos ha permitido establecer diferencias entre poblaciones que han sido clasificadas botánicamente como *Galphimia glauca*. Mediante estudios de código de barras de DNA (Sharma et al., 2012b; Gesto-Borroto et al., 2019) se ha demostrado que existen diferencias genéticas entre 11 poblaciones de *Galphimia* analizadas, por lo que se pudo concluir que estas pertenecen a especies diferentes dentro del mismo género. Aunque el genotipo de estas plantas analizadas es diferente, sus caracteres fenotípicos son muy similares.

La ruta de síntesis de las galfiminas no se encuentra elucidada aún. No obstante, se han realizado propuestas biogenéticas para explicar su origen. La primera en plantas transformadas mediante la vía de *Agrobacterium rhizogenes*. En esta se plantea que a partir de la ciclación del 2,3 epoxiescualeno, seguido de varios arreglos estructurales del catión oleanilo, se genera el núcleo de los friedelanos que posteriormente por reacciones de oxidación permitirá la formación de los norfriedelanos y nor-secofriedelanos (Nader, 2004). La segunda relaciona la síntesis de galfiminas al catión taraxasterol. En esta se explica que, a partir de reacciones

de oxidación, metilación, acetilación y O-ciclación, así como de múltiples reordenamientos se obtienen los esqueletos básicos de glaucacetalinas, galfinas y galfiminas (Rios et al., 2020). Las reacciones propuestas por estos autores en la que se produce la transformación del esqueleto básico de los triterpenos para la generación de los nor-secofriedelanos, solo son teóricas, no se encuentran demostradas.

Las galfiminas son metabolitos con un gran número de actividades biológicas, por lo que elucidar su ruta de síntesis resultaría de gran utilidad, con el objetivo de potenciar su producción a partir de herramientas biotecnológicas. Un estudio de transcriptomas de plantas de varias poblaciones del género *Galphimia* permitiría identificar diferencias genéticas relacionadas con la ruta de síntesis de estos triterpenoides. Mediante este análisis se cuantificarían cambios en la expresión de cada transcripto (Raghavachari y Garcia-Reyero, 2018) en las poblaciones analizadas, lo que posibilitaría la identificación de genes relacionados con la síntesis de galfiminas.

La secuenciación masiva de una población de RNA (RNA-seq), realizada en plataformas de secuenciación de próxima generación (NGS), representa una alternativa para aumentar el rendimiento de un estudio de transcriptómica. Si se tiene en cuenta el tamaño del genoma de las plantas, así como su complejidad debido a los diferentes niveles de ploidía (Dirihan et al., 2013), su elevado contenido de secuencias repetitivas y elementos transponibles (Shcherban, 2015), la transcriptómica sería la herramienta menos costosa y más rápida para analizar su contenido genético.

En plantas se han realizado varios estudios de transcriptómica que han permitido identificar genes relacionados con la síntesis de terpenos de estructura acíclica y policíclica (Ma et al., 2016; Aminfar et al., 2019; Lu, et al., 2021). Sin embargo, no se ha realizado un análisis de esta clase en *Galphimia* spp., aunque en la base de datos *Sequence Read Archive* (SRA) del *The National Center for Biotechnology Information* (NCBI) se encuentran depositadas lecturas crudas de la secuenciación del transcriptoma de *Galphimia gracilis*, otra especie del género.

Si se consideran las propiedades biológicas de las galfiminas y el hecho de que no se ha elucidado su ruta de síntesis, la comparación de transcriptomas de dos poblaciones de *Galphimia* spp.; una productora de galfimina (PG) y otra no productora de galfimina (NPG), permitiría identificar los posibles genes candidatos que codifican las enzimas involucradas en la ruta de producción de estos triterpenoides,

2. Marco teórico

2.1. Generalidades de *Galphimia* spp.

Galphimia spp. es una planta de la familia Malpighiaceae conocida comúnmente como “calderona amarilla” o “flor de estrella”. Se caracteriza por ser un arbusto que alcanza una longitud de 2-3 m de altura. Sus partes vegetativas se encuentran cubiertas de pelos pardos rojizos, en su mayoría rectos u ondulados que miden de 0,2-0,6 mm de largo. Las hojas son elípticas u ovadas y presentan venas secundarias prominentes. Las inflorescencias forman un racimo terminal con ejes densamente tomentulosos. Los pétalos son persistentes, desiguales y de color amarillo. Los frutos son pequeños en forma de cápsulas. Las semillas están formadas por un cotiledón externo que presenta una longitud de 8-2 mm de largo y un cotiledón interno que mide de 3-5 mm de largo (Anderson, 2007).

El origen de esta especie se encuentra en México, Guatemala y regiones no muy húmedas de América Central. Se adapta a climas secos y húmedos, así como a temperaturas frías y heladas suaves. Puede crecer en elevaciones de 700 a 2300 m, en suelos bien drenados con un pH preferiblemente de 5.6, aunque también se puede desarrollar en suelos arenosos secos. En México, se puede encontrar en los estados de Aguascalientes, Guanajuato, Hidalgo, Jalisco, Nuevo León, Querétaro, San Luis Potosí, Tamaulipas y Zacatecas (Anderson, 2007). Además, en el estado de Morelos (Herrera-Arellano et al., 2007) y en el de Guerrero (Camacho et al., 2002).

Esta planta se debe sembrar en el mes de agosto. La colecta se realiza de mayo a noviembre que corresponde con el periodo de floración o se puede extender hasta enero para obtener frutos (Anderson, 2007).

2.2. Estudios en *Galphimia* spp.

Los reportes etnomédicos de *Galphimia* spp. aumentaron el interés por el estudio de esta planta y sus metabolitos bioactivos. Su uso en la medicina tradicional mexicana incentivó la investigación del perfil químico de individuos de varias de sus poblaciones, la consecuente identificación de los principales metabolitos con actividad biológica y su mecanismo de acción. Las diferencias metabólicas entre

plantas clasificadas botánicamente como *Galphimia glauca* fueron la causa de la realización de estudios que demostraron las distancias filogenéticas entre individuos que crecen en diferentes localidades mexicanas (Sharma et al., 2012b; Gesto-Borroto et al., 2019).

2.2.1. Estudios farmacológicos

Los estudios farmacológicos de los extractos de *Galphimia* spp. se iniciaron cuando Dorch y Warner (1991) probaron el efecto antiasmático e identificaron como metabolito activo al tetragalioilo del ácido quínico (Figura 1). Posteriormente, se evaluó el efecto de este extracto mediante ensayos neuro-farmacológicos, con lo cual se demostraron sus propiedades sedantes (Tortoriello y Lozoya, 1992) y se identificó como metabolito activo responsable de esta actividad al nor-secofriedelano galfimina B, el cual fue aislado de las partes aéreas de la planta (Toscano et al., 1993). Además, se demostró que la galfimina B actúa sobre neuronas dopaminérgicas del área ventral tegmental del cerebro (Osuna et al., 1999).

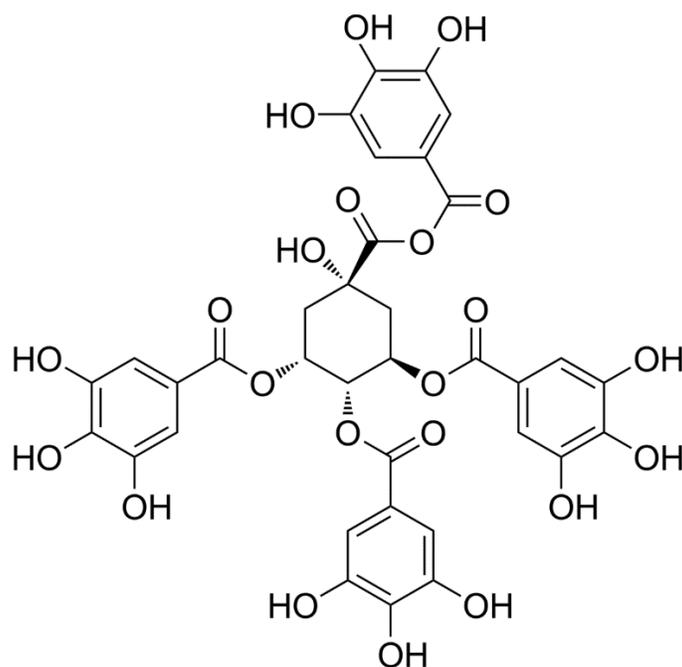


Figura 1. Estructura del tetragalioilo del ácido quínico. Tomado de PubChem

Se propuso un posible mecanismo de acción del extracto de *Galphimia* spp. en el SNC, basado en experimentos que demuestran que triterpenos con estructuras similar a la galfimina B pueden modular un canal iónico de K⁺ localizado en el hipocampo o la corteza cerebral. Los componentes del extracto podrían modificar la excitabilidad neuronal, al intervenir en la regulación de la síntesis de GABA. En las zonas del cerebro donde se encuentran estos canales de K⁺, se produce el procesamiento de la información asociada con la ansiedad, el estrés y la depresión (Herrera-Ruiz et al., 2006).

Además, se ha estudiado que en individuos que padecen de ansiedad, el efecto del extracto de *Galphimia* spp. no difiere de los ansiolíticos sintéticos como el lorazepán (Herrera-Arellano et al., 2007). Posteriormente, se sintetizaron cápsulas que contenían el extracto y se administraron a grupos de individuos, lo que permitió demostrar que el tratamiento con *Galphimia* spp. resultó más efectivo que el de lorazepán (Herrera-Arellano et al., 2012). También, se obtuvieron resultados satisfactorios en jóvenes con trastornos de ansiedad social al administrar oralmente el extracto estandarizado. Esto permitió demostrar la eficacia y tolerabilidad de esta planta medicinal (Romero-Cerecero et al., 2018).

2.2.2. Estudios fitoquímicos

Estudios químicos han permitido la elucidación de la estructura de las galfiminas identificadas en investigaciones previas de *Galphimia* spp. Estos metabolitos son triterpenos modificados del tipo nor-secofriedelanos, constituidos por una lactona α , β insaturada localizada en el anillo A de siete miembros. La pérdida de un carbono (lo que define la nomenclatura nor) provoca la apertura del anillo A de seis miembros (lo que define la nomenclatura seco), el cual mediante oxidaciones del núcleo triperpénico permite la funcionalización para formar la lactona (Cardoso Taketa et al., 2004). Se han identificado 15 galfiminas: el primer miembro caracterizado de esta serie se denominó con la letra B (Toscano et al.1993), seguido por los homólogos A, C-I (Cardoso-Taketa et al., 2004), la J (González-Cortazar et al., 2005), las K y L (González-Cortazar et al., 2014) y, finalmente, las M-O (Ortega et al., 2020). Éstas se diferencian químicamente por la presencia de sustituyentes

hidroxilo y acetatoxilo en C₆ y C₇ así como la estereoquímica en torno al anillo E (Figura 2). La isomería en el anillo E produce tres tipos de estructuras diastereoméricas con dobles enlaces en Δ^{19} , Δ^{20} y $\Delta^{20,29}$ (Cardoso-Taketa et al., 2004).

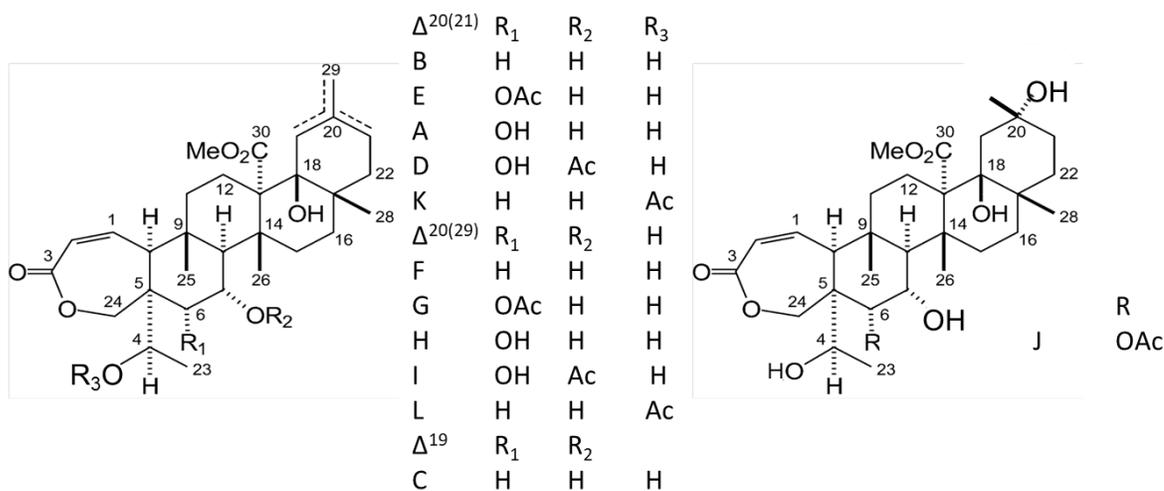


Figura 2. Estructura de las galfiminas (A-L). Diferencias entre los grupos funcionales y estereoquímica del anillo E. Tomado de Cardoso-Taketa et al. (2004).

Por otra parte, Nader y colaboradores (2004) aislaron y caracterizaron tres norfriedelanos nuevos secretados al medio de cultivo, las glaucacetalinas A, B y C, las cuales fueron sintetizadas por raíces transformadas con *A. rhizogenes*. Posteriormente, Ortiz y coautores (2009) aislaron un norfriedelano inédito, la glaucacetalina D del cultivo en suspensión de células de *Galphimia* spp. transformadas y demostraron que estos metabolitos también poseen actividad sedante. Posteriormente, se identificó la glaucacetalina E y se determinó la actividad ansiolítica de la fracción del extracto que contenía a este metabolito (Ríos et al., 2020).

Las glaucacetalinas a diferencia de las galfiminas no presentan anillo lactónico pues el anillo A se encuentra preservado (Figura 3). Sin embargo, poseen un anillo adicional con una función acetal cíclica que se forma entre las posiciones oxidadas en C₂₄ y C₂₅. Estos norfriedelanos presentan un solo doble enlace en Δ^{20} y Δ^{29} y se diferencian químicamente por la sustitución de grupos acetoxilos en C₃ y metoxilos en C₂₅ (Nader, 2004).

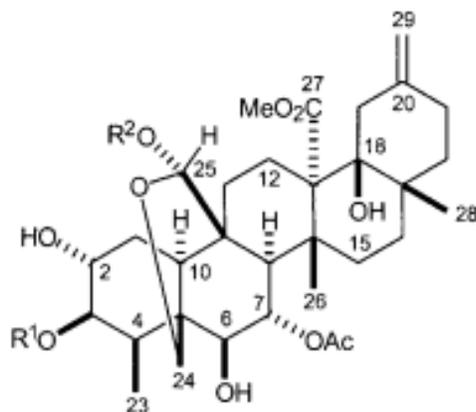


Figura 3. Estructura de las gaulterinas. Tomado de Nader et al. (2004).

2.2.3. Estudios metabólicos

Mediante Resonancia Magnética Nuclear (RMN) asociada a un análisis multivariado, se detectaron diferencias en el contenido de galfiminas y la actividad ansiolítica entre individuos de diferentes poblaciones de *Galphimia* spp. Para ello, se usaron extractos crudos obtenidos de individuos que crecen en las localidades de México; Doctor Mora, Guanajuato; Tepoztlán, Morelos; San Andrés de la Cal, Morelos; Cuernavaca, Morelos; Jalisco, Guadalajara y Jalpan, Querétaro. Las poblaciones estudiadas se clasificaron en activas (productoras de galfimina B y con actividad ansiolítica y sedante) e inactivas (no productoras de galfimina B). Solamente las poblaciones de Querétaro y Guanajuato fueron clasificadas como activas. Sin embargo, la población de Tepoztlán, Morelos, aunque carecía de galfiminas B, tuvo efecto sedante, aunque menor que las de Querétaro y Guanajuato. Lo anterior sugirió que esta población contiene otros metabolitos que poseen esta actividad (Cardoso-Taketa et al., 2008).

Posteriormente, se determinó la consistencia del perfil químico de poblaciones de *Galphimia* spp. colectadas en las mismas localidades que Cardoso-Taketa et al. (2008) e incorporaron a la investigación dos poblaciones nuevas, colectas correspondientes a Tuxtla Gutiérrez, Chiapas y Miacatlán, Morelos (Sharma et al., 2012a). Se realizaron estudios de RMN y se determinó la actividad antiinflamatoria de todos los extractos. En este trabajo se corroboró que las poblaciones de

Querétaro y Guanajuato se diferenciaron del resto atendiendo a la presencia de galfiminas. Además, que la población de Tepoztlán, Morelos, se diferenció de las otras poblaciones no productoras de este metabolito. No obstante, se observaron diferencias entre el contenido de galfiminas y la actividad ansiolítica y sedante de los extractos obtenidos a partir de las poblaciones de Querétaro y Guanajuato en este estudio con respecto a las investigadas por Cardoso-Taketa et al. (2008). Estos resultados pueden deberse a las diferentes condiciones de crecimiento en que se encontraban las plantas colectadas y a los mecanismos de defensa vegetal que podrían estar expresándose en las poblaciones seleccionadas para el estudio, pues las condiciones y épocas de colecta fueron las mismas en ambos casos. Los extractos de todas las poblaciones presentaron actividad antiinflamatoria, lo que sugirió que ésta es independiente de la presencia de galfiminas.

2.2.4. Estudios genéticos

En base a las interrogantes planteadas en el estudio de Sharma et al. (2012a) en cuanto a las diferencias en el contenido de galfiminas entre las siete poblaciones de *Galphimia* spp. analizadas, se realizó un estudio de código de barras de DNA. En este análisis se emplearon como marcadores los genes que codifican para la maturasa K (mat K), la ribulosa bifosfato carboxilasa (rbcL) y la subunidad β de la RNA polimerasa de plástidos (rpoC1), con el objetivo de determinar si los individuos de las diferentes poblaciones pertenecían a la misma especie. Se formaron tres grandes clados y las siete poblaciones se separaron según los resultados de mat K y rpoC1, lo que permitió diferenciar las poblaciones de Querétaro y Guanajuato de las restantes. Estos resultados indicaron que las poblaciones analizadas pueden pertenecer a especies diferentes dentro del género *Galphimia*. Las diferencias en el perfil metabólico de estas poblaciones fueron corroboradas además por cromatografía en capa fina (TLC) (Sharma et al., 2012b). Posteriormente, Gesto-Borroto et al. (2019), realizaron estudios de códigos de barras de ADN donde se incluyeron poblaciones adicionales (Ciudad Valles, San Luis Potosí; Cadereyta, Querétaro, Zimapán, Hidalgo y Santa Catarina, Morelos) para el análisis y utilizaron

además de los marcadores mat K, rbc1 y rpoC1 nuevos marcadores como regiones del espaciador no-transcrito intergénico (psbA-trnH) y del espaciador transcrito interno (ITS1 e ITS2). De esta forma, obtuvieron que, de nueve poblaciones estudiadas, además de las poblaciones de Querétaro y Guanajuato analizadas en el estudio de Sharma et al. (2012b), otra población de Querétaro y una de Hidalgo también poseían galfiminas.

Estos estudios constituyen la respuesta a los diferentes perfiles químicos y actividades biológicas entre las poblaciones de *Galphimia* spp. analizadas. Los análisis filogenéticos sugieren que existen factores genéticos que desempeñan una función importante en la producción de galfiminas y que estos metabolitos constituyen un marcador quimiotaxonómico dentro del género.

2.3. Terpenos

Los terpenos son metabolitos secundarios que se encuentran en flores, tallos, hojas y raíces de numerosas especies de plantas. Se sintetizan a partir de unidades de isopreno con la fórmula general (C₅H₈) (Yadav et al., 2014). Se estima que se han identificado más de 55 000 estructuras de terpenoides, lo cual probablemente es solo una fracción de lo que se puede encontrar en la naturaleza (Bathe y Tissier, 2019).

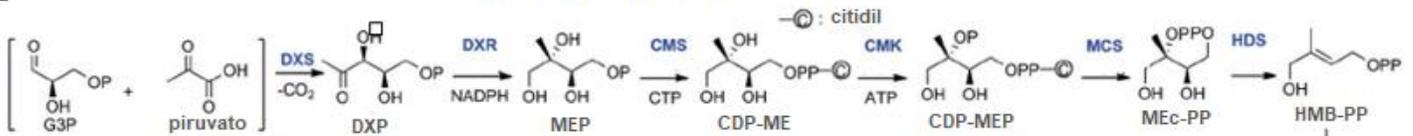
Estos compuestos cumplen funciones importantes en el metabolismo primario pues actúan como pigmentos fotosintéticos (carotenoides), transportadores de electrones (cadenas laterales de ubiquinona y plastoquinona), reguladores del crecimiento y desarrollo (giberelinas, ácido abscísico, estrigolactonas, brasinoesteroides, citoquininas), en la glicosilación de proteínas (dolicoles) y como componentes estructurales y funcionales de la membrana citoplasmática (fitoesteroles). Además, los terpenoides volátiles o semivolátiles de bajo peso molecular, que incluyen isopreno, monoterpenoides, sesquiterpenoides y diterpenoides, se encuentran implicados en la protección de las plantas contra el estrés abiótico, en diversas interacciones bióticas por encima y por debajo del suelo, así como en la polinización al ser constituyentes del aroma floral. Los terpenoides volátiles también cumplen

funciones como compuestos constitutivos o inducidos por patógenos y herbívoros en la defensa de los tejidos fotosintéticos (Tholl, 2015).

2.3.1. Síntesis de terpenos

La síntesis de terpenos puede ocurrir en el citosol mediante la ruta del ácido mevalónico (MVA) o en los cloroplastos por la ruta del fosfato de metileritritol (MEP) (Figura 4). Mediante la ruta del MVA se sintetizan sesquiterpenos y triterpenos. A partir de la condensación de dos moléculas de acetil-coenzima A (CoA) catalizada por la enzima acetoacetil-CoA tiolasa se obtiene la acetoacetil-CoA. La adición de una tercera unidad de acetoacetil-CoA por la enzima 3-hidroxi-3-metil-glutaril-CoA sintasa produce 3-hidroxi-3-metil-glutaril-CoA. El producto de esta reacción se reduce por la enzima 3-hidroxi-3-metilglutaril- CoA reductasa en ácido mevalónico, que constituye el intermediario fundamental de la ruta y la reductasa que lo genera el punto clave de regulación. La enzima mevalonato quinasa fosforila al alcohol primario del ácido mevalónico y se obtiene mevalonato-5-difosfato. De esta forma, en presencia del trifosfato de adenosina (ATP), se facilita la descarboxilación-eliminación mediante la enzima mevalonato difosfato descarboxilasa y se obtiene el pirofosfato de isopentenilo (IPP) que se puede isomerizar estereoespecíficamente en la posición alílica por la isopentenil pirofosfato isomerasa a pirofosfato de dimetilalilo (DMAPP) (Zhang et al., 2019).

Ruta del fosfato de metileritritol



Ruta del ácido mevalónico

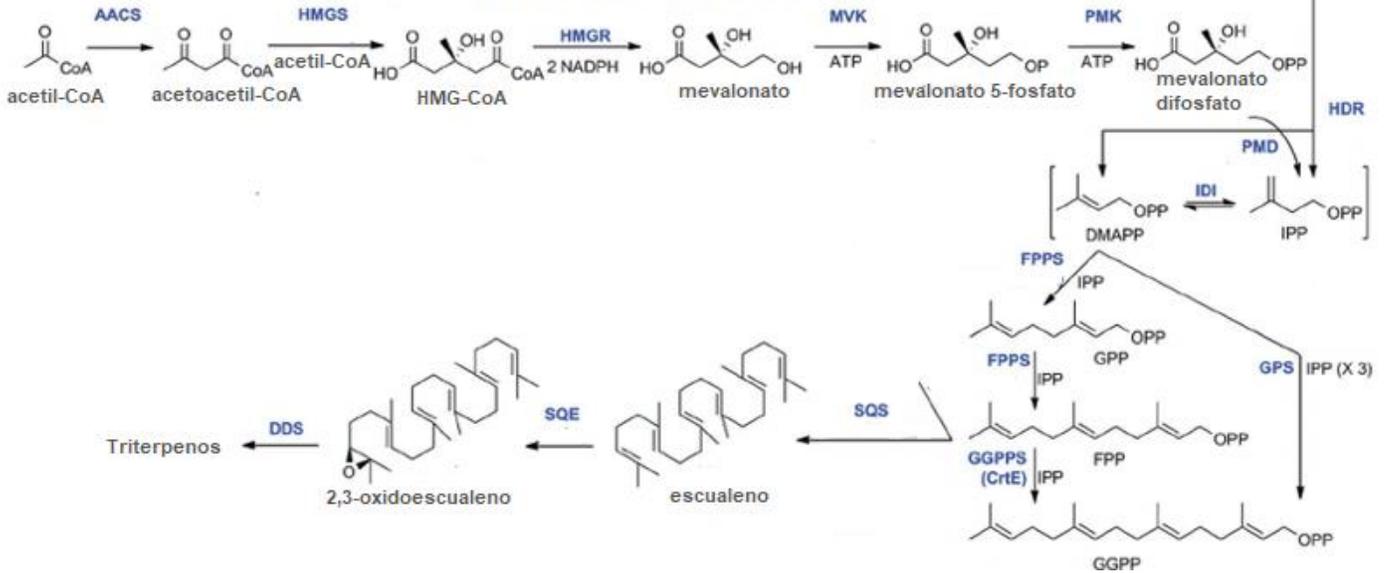


Figura 4. Síntesis de terpenos mediante la ruta del del fosfato del metileritriol (MEP) y del ácido mevalónico. G3P (gliceraldehído 3-fosfato), DXP (1-deoxi-D-xilulosa-5-fosfato), DXS (1-deoxi-D-xilulosa-5-fosfato sintasa), DXR (1 deoxi-D-xilulosa reductoisomerasa), CMS (2-C-metil-D-eritriol 4-fosfato citidiltransferasa), CDP-ME (2-C-metil-D-eritriol 4-fosfato), CMK (4-difosfocitidil-2-C-metil-D-eritritol quinasa), CDP-MEP (4-difosfocitidil-2-C-metil-D-eritritol), MCS (2-C-metil-D-eritrol-2,4-ciclodifosfato sintasa), MEc-PP (4-difosfocitidil-2C metileritritol 2-fosfato), HDS (1-hidroxi-2-metil-2-(E)-butenil-4-difosfato sintasa), HDS (1-hidroxi-2-metil-2-(E)-butenil-4-difosfato sintasa), (1-hidroxi-2-metil-2-(E)-butenil-4-difosfato sintasa) HMB-PP (1-hidroxi-2-metil-2-(E)-butenil-4-difosfato), HDR (1-hidroxi-2-metil-2-(E)-butenil-4-difosfato reductasa), AACS (acetoacetyl-CoA sintasa), HMGS (3-hidroxi-3-metilglutaril-CoA sintetasa), HMGR (3-hidroxi-3-metilglutaril- CoA reductasa), MVK (mevalonato quinasa) PMK (fosfato mevalonato descarboxilasa) PMD (fosfomevalonato descarboxilasa), IDI (isopentenil difosfato delta isomerasa), IPP (isopentenil difosfato), DMAPP (pirofosfato de dimetilalilo), GPPS (geranil pirofosfato sintasa), GPP (geranil pirofosfato), FPPS (farnesil pirofosfato sintasa), FPP (farnesil pirofosfato), GGPPS (geranil geranil pirofosfato sintasa), GGPP (geranil geranil pirofosfato), SQS (escualeno sintasa), SQE (escualeno epoxidasa). Tomado de (Song et al., 2014).

Mediante la ruta del MEP se sintetizan preferentemente hemiterpenos volátiles, monoterpenos, diterpenos, plastoquinonas y carotenos. En ésta se produce la condensación de una molécula de piruvato con el gliceraldehído 3-fosfato (GA3-P) y, a través de la enzima 1-desoxi-D-xilulosa-5-fosfato sintasa, se obtiene el 5 fosfato de 1 desoxi-D-xilulosa. Este compuesto es sometido a una serie de reducciones dependientes de NADPH para formar 4-fosfato de 2-C-metil-D-eritriol, este es el intermediario principal de la ruta. Posteriormente, por la acción de la enzima 2-C-metil-D-eritrol-2,4-ciclodifosfato sintasa, se obtiene 2-C-metil-D-eritrol-2,4-ciclodifosfato, a partir del cual se obtiene 1-hidroxi-2-metil-2-butenil-4-difosfato, que da lugar al IPP por acción de la 4-hidroxi-3-metil-2-enil pirofosfato reductasa (Gutbrod et al., 2019).

La condensación posterior de IPP y DMAPP mediante la enzima isoprenil difosfato sintasa produce intermediarios acíclicos y aquirales del isoprenil difosfato/pirofosfato (ID, C₅), incluidos el pirofosfato de geranilo (GPP, C₁₀), el pirofosfato de farnesilo (FPP, C₁₅) y pirofosfato de geranilgeranilo (GGPP, C₂₀), que se consideran precursores universales de terpenos. Las enzimas terpeno sintasas actúan sobre uno o más de estos precursores, incluido el DMAPP para producir hemiterpenos (C₅), monoterpenos (C₁₀), homoterpenos (C₁₁ y C₁₆), sesquiterpenos (C₁₅), diterpenos (C₂₀), sesterpenos (C₂₅), triterpenos (C₃₀), tetraterpenos (C₄₀) y politerpenos (C_{>40}, terpenos de orden superior) (Boncan et al., 2020).

2.3.2. Síntesis de triterpenos de estructura pentacíclica

En las plantas, los triterpenoides pentacíclicos están representados por varios tipos estructurales; entre los más comunes se destacan el oleanano y el ursano. La cantidad de estos metabolitos varían dependiendo del sistema de enzimas requeridas para su síntesis, así como de la influencia de factores externos. La síntesis de los triterpenoides se produce por la conversión del triterpeno acíclico escualeno en 2,3-oxidoescualeno y su posterior ciclación por complejos enzimáticos específicos de la familia de óxido escualeno ciclasa (OSC) (Luchnikova et al., 2020). Se han caracterizado funcionalmente más de 80 OSC de plantas, principalmente por expresión heteróloga de cDNA en cepas de levadura. Aproximadamente un

tercio de éstas son sintasas de esteroides, un grupo que incluye sintasas del cicloartenol y varias sintasas del lanosterol. La función de las sintasas del lanosterol en las plantas parecen estar involucradas en la síntesis de fitoesteroides y potencialmente también en la de metabolitos derivados de esteroides. Algunas OSC de plantas sintetizan moléculas adicionales al colesterol y el lanosterol a través de los estados de transición de tipo silla-bote-silla para las conformaciones de los anillos A-C para la generación del intermediario catiónico protosterilo durante la biosíntesis del lanosterol que finaliza con la formación de los triterpenos tetracíclicos (Forestier et al. 2019); entre éstas se incluyen a la cucurbitadienol sintasa (CPX) de *Cucurbita pepo* y la parkeol sintasa (OsPS) de *Oryza sativa*. Las OSC restantes sintetizan triterpenos a través de la conformación del sustrato en silla-silla-silla para generar al catión dammarenilo (Forestier et al., 2019). Colectivamente, las OSCs de las plantas son capaces de hacer una amplia diversidad de esqueletos triterpénicos simples como la β -amirina y el lupeol. Además, pueden sintetizar una gran cantidad de productos después de la ciclización del 2,3-oxido escualeno como marneral, talalianol, sinona, glutinol, friedelina, taraxerol, isomultiflorenol, isoarborinol (Thimmappa et al., 2014).

Asimismo, se han realizado estudios con sustratos no naturales que han permitido describir las reacciones enzimáticas en las que intervienen las OSCs. Primeramente, se produce la unión de la enzima a un sustrato poliolefínico que se encuentre en una conformación plegada. Luego, se inicia una reacción por protonación de un doble enlace (escualeno) o un epóxido (2,3-oxido escualeno). Después, se forman los anillos, seguido de un reordenamiento del núcleo estructural por transposiciones de tipo-1,2 de los grupos metilo e hidruros. Finalmente, ocurre una desprotonación o adición de agua (Wendt, 2005).

Los intermediarios generados por las OSCs sufren, en primer lugar, oxidaciones estructurales regio y estereoespecíficas catalizadas por la acción de monooxigenasas pertenecientes a la familia del citocromo P450 (CYP) para producir la diversidad estructural de los triterpenos en las plantas mediante reacciones de hidroxilación iniciales (Yasumoto et al., 2017; Liu et al., 2020) y, de esta forma, generan la funcionalización de los esqueletos triterpénicos básicos como la β -

amirina en las posiciones C₂, C₁₆, C₂₁, C₂₂, C₂₄ y C₂₈ (Figura 5). Enseguida, mediante la intervención de glicosiltransferasas se produce la diversificación de las saponinas triterpénicas mediante la formación de enlaces glicosídicos principalmente con el hidroxilo en C₃ y el carboxilo en C₂₈ (Seki et al., 2015; Ghosh, 2017). Las reacciones de acetilación y metilación de los grupos funcionales oxigenados (OH y CO₂H, respectivamente), que únicamente incrementan la variedad de los sustituyentes de los núcleos triterpénicos oxidados básicos, se producen por la acción de aciltransferasas (Mugford et al., 2013) y metiltransferasas con la S-adenosilmetionina como el donador del grupo CH₃, respectivamente (Chen et al., 2019).

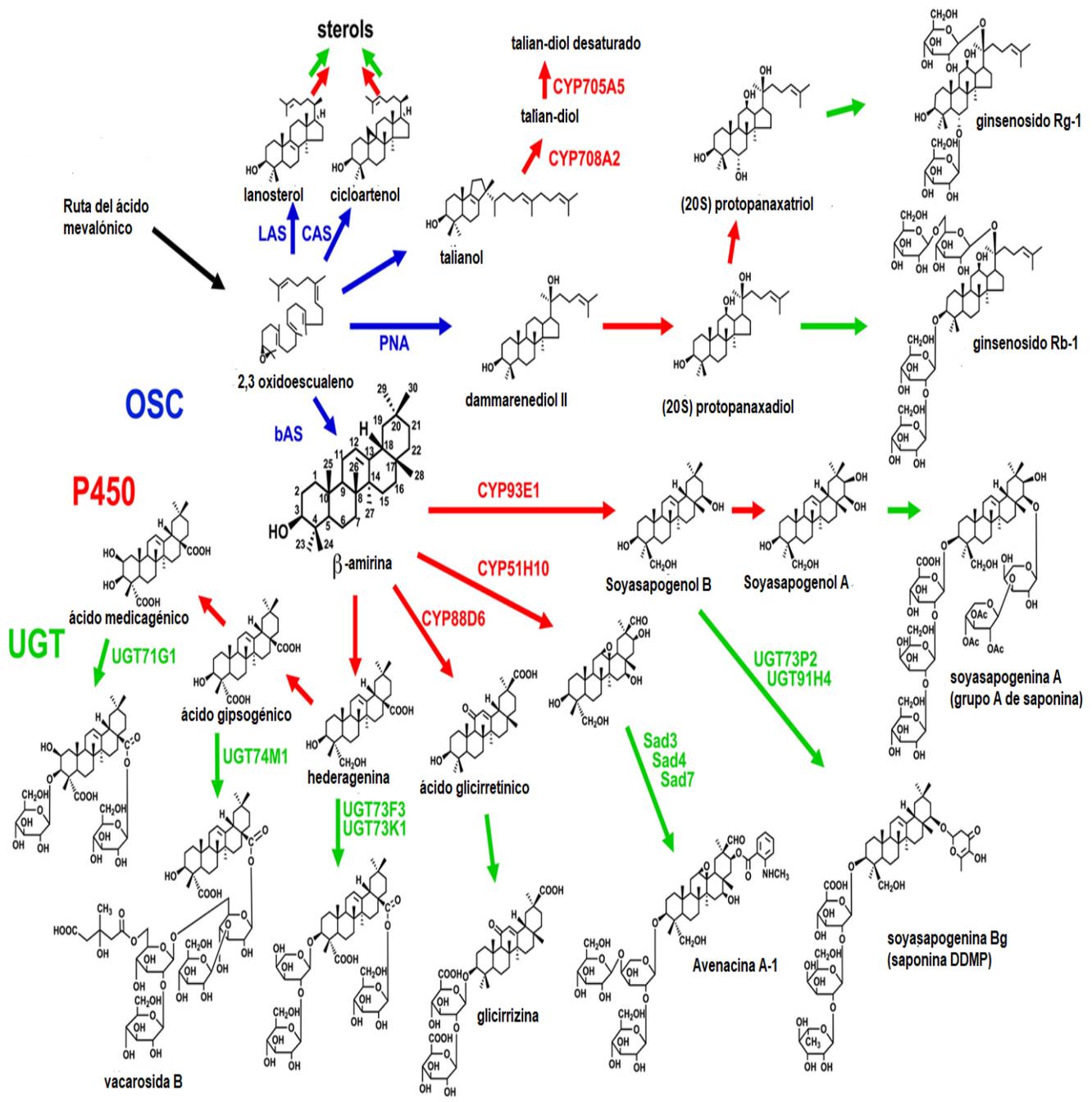


Figura 5. Modificaciones de triterpenos después la ciclación del 2,3 oxidoesqualeno por oxidoesqualeno ciclasas (OSC). Flechas azules indican reacciones catalizadas por OSC, flechas rojas indican reacciones de oxidación catalizadas por P450 y flechas verdes reacciones de glicosilación. Tomado de Sawai y Saito. (2011).

El citocromo P450 (CYP) constituye la familia más grande de monooxigenasas involucradas en el metabolismo de las plantas y representa alrededor del 1% de los genes codificantes de proteínas en sus genomas. Estas enzimas generalmente se clasifican en diferentes familias y subfamilias según la homología de secuencia y los criterios filogenéticos. Se han identificado 127 familias de P450 en plantas las cuales se han clasificado en 11 clanes (Ghosh, 2017). Aproximadamente 55 CYPs actúan sobre intermediarios de triterpenos pentacíclicos de plantas. La mayoría de estas pertenecen a los clanes CYP51 (miembro: CYP51H), CYP71 (miembros: CYP71A, CYP71D, CYP81Q, CYP93E y CYP705A), CYP72 (miembro: CYP72A) y CYP85 (miembros: CYP87D, CYP88D, CYP88L, CYP708A, CYP716A, CYP716C, CYP716E, CYP716S, CYP716U y CYP716Y) (Sandeep et al., 2019).

Las reacciones de oxidación catalizadas por CYP son las más comunes en la oxidación de los núcleos triterpenoides de estructura pentacíclica. Estas pueden producirse secuencialmente en una posición específica para dar lugar a ácidos carboxílicos o conducir a reordenamientos de las estructuras básicas de los triterpenos mediante la creación de nuevos enlaces C-C. Los CYPs de las plantas transfieren dos electrones de los cofactores NADPH y el oxígeno atmosférico a sus sustratos. Generalmente, en las plantas se requiere una reductasa de CYP para contrarrestar la transferencia de electrones (Bathe y Tissier, 2019).

El número limitado de clanes enzimáticos del citocromo P450, hasta el momento caracterizados, dificulta la generación de conclusiones acerca de los orígenes de las enzimas de esta familia que participan en la modificación oxidativa de las estructuras triterpénicas básicas. Sin embargo, como resultado de los análisis filogenéticos realizados con esta enzima, se ha propuesto que la selección de las enzimas del citocromo P450 para la biosíntesis de triterpenos haya ocurrido varias veces durante la evolución (Thimmappa et al., 2014).

2.3.2.1. Síntesis de galfiminas

Las galfiminas son triterpenoides modificados del tipo nor-secofriedelanos. Se han realizado propuestas biogenéticas para explicar su mecanismo de síntesis, pero estas son teóricas y aún no se encuentran demostradas. La primera propuesta fue

en plantas transformadas por la vía *A. rhizogenes* (Nader, 2004) (Figura 6). En estas propuestas biosintéticas, a partir de la ciclación del 2,3 epoxiescualeno, se genera el bacaranilo, el catión tetracíclico intermediario, que genera el anillo terminal de la estructura pentacíclica, mediante la ciclización de la cadena lateral genera el intermediario común de la β -amirina y de los friedelanos. Múltiples arreglos del catión bacaranilo generan el núcleo del friedelano. Éste sufre la pérdida de CO_2 y se forma la nor-friedelina, que constituye el primer precursor con el núcleo norfriedelánico de las galfiminas. El esqueleto estructural del nor-friedelano al sufrir oxidaciones del anillo A, provocaría la ruptura del enlace $\text{C}_3\text{-C}_4$, lo que da lugar a los dos grupos funcionales necesarios para la formación de la lactona de siete miembros característica de la galfiminas.

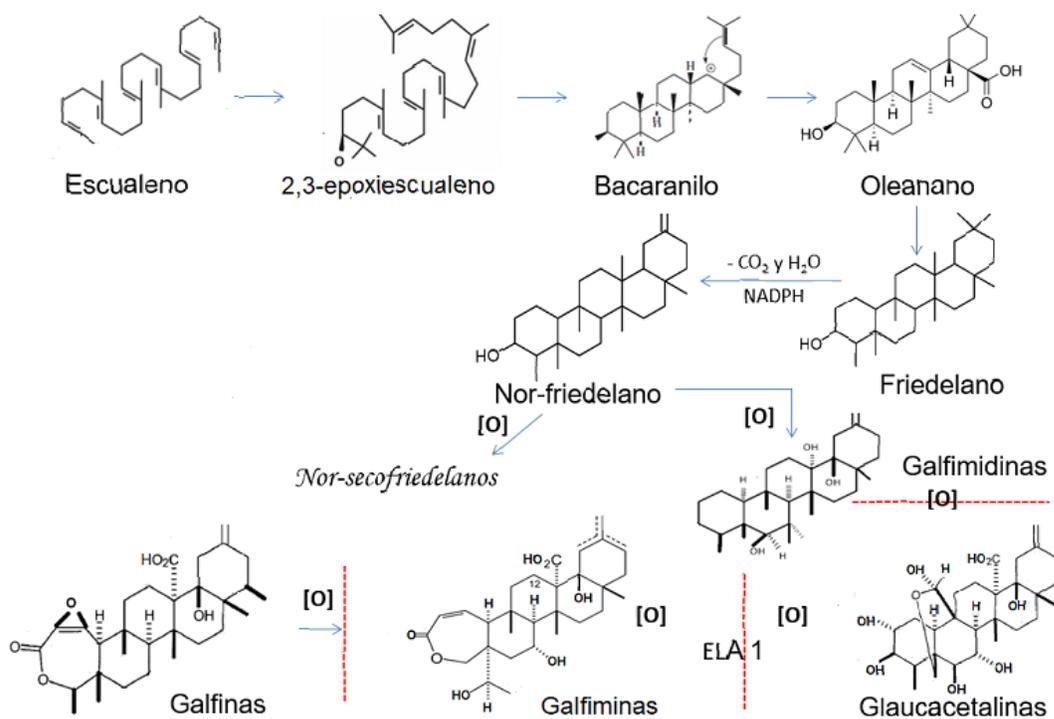


Figura 6. Propuesta biogenética para la síntesis de galfiminas a partir de plantas transformadas vía *Agrobacterium rhizogenes*, con modificaciones. Tomado de Nader (2004).

La segunda propuesta fue realizada por Rios et al. (2020) se basó en que la síntesis de galfiminas estaba asociada al catión taraxasterol (Figura 7). Se planteó que el

cación taraxasterol pierde un protón en C₂₁ o C₂₉, seguido por la oxidación del grupo metilo C₃₀ a un ácido carboxílico. Esta oxidación y la pérdida del OH-19 establecen un carbocatión en C₁₉ (intermediario A). Posteriormente, se produce un reordenamiento 1,2 de hidruro de tipo Wagner-Meerwein, seguido de una eliminación, lo que conduce a la formación de un doble enlace C₁₃-C₁₈ (intermediario B). La epoxidación de este intermediario y la apertura posterior de este epóxido forman el intermediario C. A partir de esta reacción tienen lugar varios reordenamientos de Wagner-Meerwein: CH₃ (C₂₇) de C₁₄ a C₁₃; CH₃ (C₂₆) de C₈ a C₁₄; H de C₉ a C₈; CH₃ (C₂₅) de C₁₀ a C₉; H de C₅ a C₁₀; CH₃ (C₂₄) de C₄ a C₅. Una adición de hidruro en C₄ conduce a la formación del intermediario D. Una serie de oxidaciones, metilaciones, acetilaciones y una O-ciclización del intermediario D producen las glaucacetalinas. La oxidación de las glaucacetalinas en C₃ da lugar a una cetona que por medio de una oxidación de Baeyer-Villiger produce galfinas a través del intermedio E, así como galfiminas mediante una adición SN₂ de un alcohol en C-24.

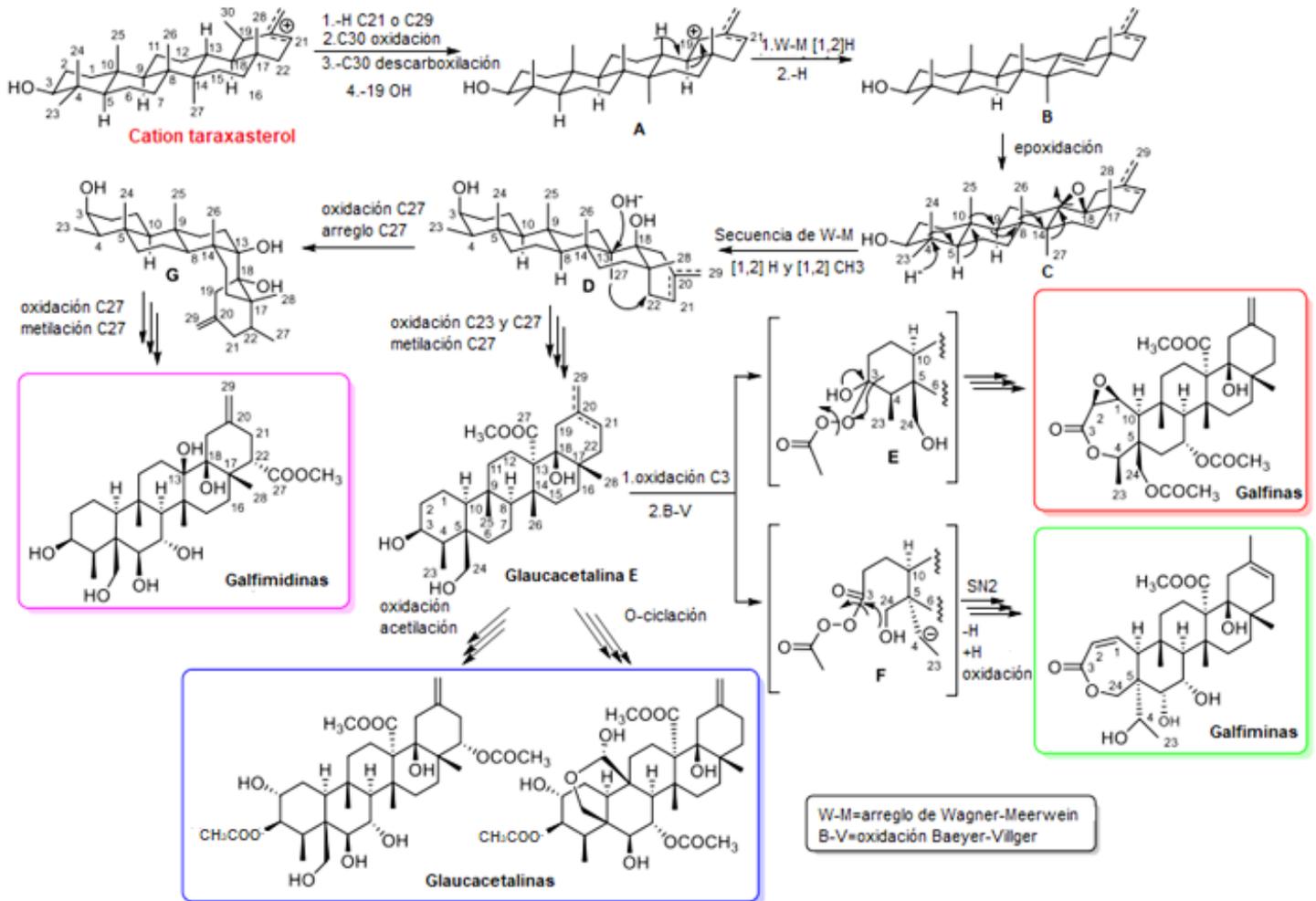


Figura 7. Síntesis de galfiminas mediante el cation taraxasterol. Tomado de Rios et al. (2020).

Las rutas propuestas para explicar la biosíntesis de las galfiminas son muy diferentes. Esto puede estar relacionado con las condiciones y objetivos experimentales que constituyen las bases de estos planteamientos. En ambos casos se aborda la producción de compuestos relacionados con las galfiminas, como las galfimidinas y glaucacetalinas, aunque se explica su origen a partir de metabolitos intermedios diferentes. Estas hipótesis no han sido suficientes para elucidar y demostrar una ruta de biosíntesis de estos triterpenoides. Sería conveniente el uso de herramientas como genómica o transcriptómica que permitan

un estudio a nivel genético que posibilite la identificación de genes que codifiquen las enzimas involucradas.

2.4. Secuenciación de ácidos nucleicos

La secuenciación de ácidos nucleicos en sus inicios se realizaba mediante el método de terminación de cadena propuesto por Sanger. Mediante este procedimiento fue posible secuenciar fragmentos de DNA de aproximadamente 1 kilobase de longitud de una sola vez. El método de Sanger fue durante años uno de los mejores para la secuenciación de genes clonados en un sistema heterólogo debido a su precisión, solidez y facilidad de uso. Sin embargo, su aplicación era costosa, laboriosa y requería de un largo tiempo para su ejecución (Türktaş et al., 2015).

Posteriormente, surgió la tecnología de secuenciación de siguiente generación (NGS), que forma parte de la secuenciación de segunda generación. Esta incluye plataformas como Roche 454, Illumina Solexa y ABI-SOLiD, las cuales han transformado el campo de la ómica. NGS fue capaz de producir una enorme cantidad de datos a un costo muy económico y con rapidez. Además, es más rápida que el método tradicional, pues se puede secuenciar el genoma completo de organismos pequeños en un solo día (Gupta y Verma, 2019). Mediante esta tecnología que utiliza amplificación y secuenciación en paralelo se producen longitudes de lectura más cortas, con tasas de error bruto promedio de 1-1,5% (Hari y Parthasarathy, 2018). Los flujos de trabajo de las plataformas de NGS incluyen pasos similares: extracción de ácido nucleico; obtención de librerías, preparación de plantillas y secuenciación automatizada. Sin embargo, estas plataformas de secuenciación difieren en: ingeniería, química de secuenciación, salida (longitud de lecturas, número de secuencias), precisión y costo (Besser et al., 2018).

En los últimos años, la plataforma Illumina ha contribuido notablemente a los secuenciadores de segunda generación (Gupta y Verma, 2019). Esta se basa en una técnica conocida como "amplificación en puente" (Figura 8). Las moléculas de DNA (alrededor de 500 pb) con adaptadores apropiados ligados en cada extremo, se utilizan como sustratos para reacciones de síntesis de amplificación repetida en

un soporte sólido (celda de flujo) que contiene secuencias de oligonucleótidos complementarias a un adaptador ligado. Los oligonucleótidos en la celda de flujo están espaciados de tal manera que el ADN se somete a rondas repetidas de amplificación y se forman "clusters" clonales que consisten en aproximadamente 1000 copias de cada fragmento del oligonucleótido. Cada celda de flujo puede admitir millones de reacciones de agrupamiento paralelas. Durante las reacciones de síntesis, se incorporan y luego se detectan nucleótidos correspondientes a cada una de las cuatro bases del DNA, cada una con una etiqueta fluorescente diferente. Los nucleótidos también actúan como terminadores de la síntesis de cada reacción que se desbloquean después de la detección para la siguiente ronda de síntesis. Las reacciones se repiten durante 300 o más rondas (Slatko et al., 2018).

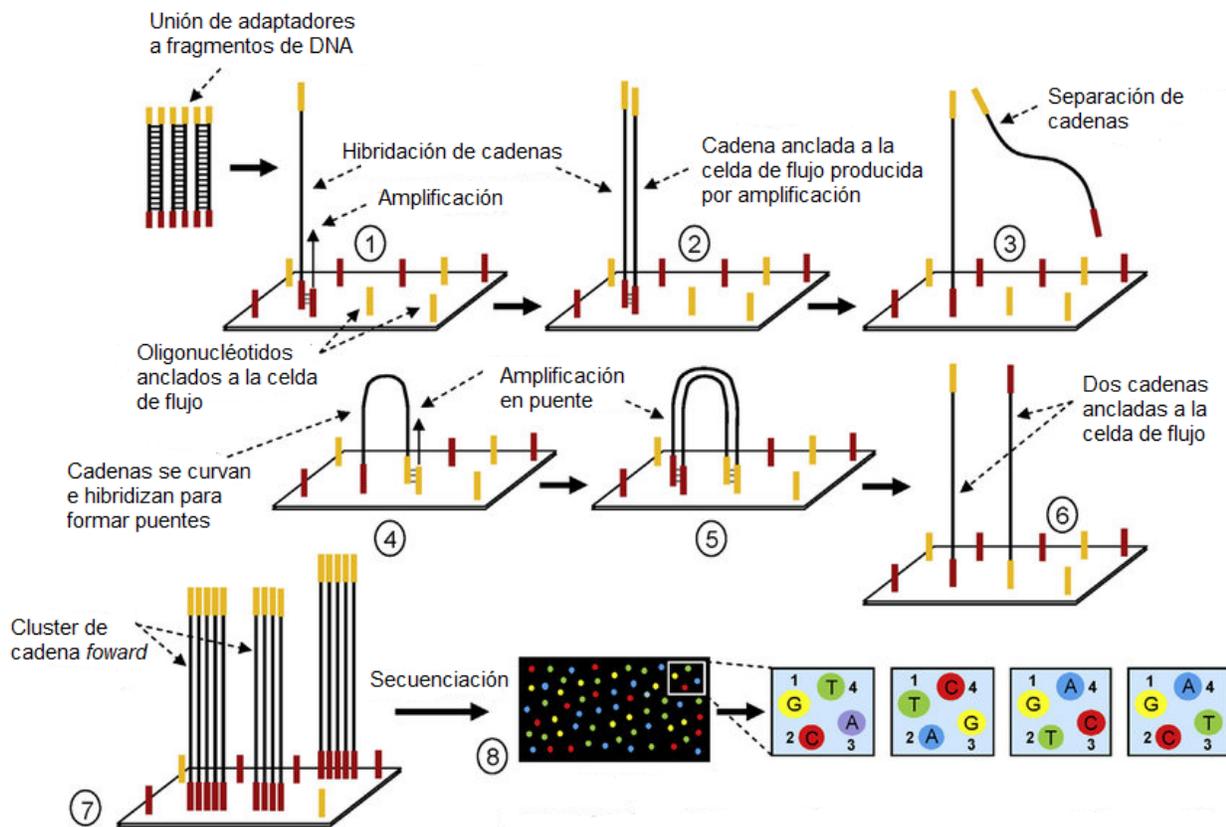


Figura 8. Secuenciación en la plataforma Illumina. Tomado de Ateto. (2014).

Esta plataforma de secuenciación en comparación con otras de nueva generación permite un incremento en el tamaño de las lecturas. Esto resulta esencial para aumentar la profundidad de la secuenciación, aspecto importante cuando se va a realizar un ensamblado *de novo* de las especies en las que no se conoce un genoma de referencia (López de Heredia, 2016).

Las NGS constituyen un avance tecnológico en las áreas de genómica, transcriptómica y metagenómica. La posibilidad de obtener lecturas de tamaño adecuado en dependencia del proyecto de secuenciación en muy corto tiempo, donde el llamado de bases se realice con una tasa de error muy baja, se traduce en la obtención rápida de resultados satisfactorios. Las principales desventajas de las NGS se encuentran en la infraestructura que se requiere para su implementación, como la capacidad y almacenamiento de las computadoras, la preparación de los usuarios para analizar, interpretar y manejar los datos obtenidos con destreza y, posteriormente, extraer la información importante y mostrarla de manera adecuada.

2.4.1. Genómica

La secuenciación de DNA es el método experimental para determinar la disposición secuencial de las bases de los ácidos nucleicos (A, T, G y C) en un polinucleótido que codifica diferentes proteínas que son funcionales en una célula viva. El conjunto completo de secuencias codificantes y no codificantes en el DNA se conoce como genoma. Las secuencias biológicas muestran patrones complejos de similitud entre sí los cuales se pueden identificar investigando similitudes entre las secuencias. Para ello, se emplean herramientas como la genómica (Shetty et al., 2019).

La genómica posibilita determinar el orden de las secuencias de DNA con el objetivo de estudiar la estructura de los genes, su evolución y realizar inferencias sobre filogenia molecular (Unamba et al., 2015). En plantas luego del ensamblado del genoma se puede obtener una representación del espacio génico, que a su vez revelaría el contexto genómico que incluye grupos, duplicaciones en tándem, subfuncionalización y neofuncionalización de la expresión (Kellner et al., 2015). Después, es posible realizar la anotación de genes y la identificación de ortólogos

candidatos para genes potencialmente relacionados con la síntesis de metabolitos de interés (Zhao et al., 2017).

2.4.2. Transcriptómica

La transcriptómica tiene como objeto de estudio la secuenciación y análisis de transcriptomas (conjunto completo de ARNm, o transcripciones primarias que tienen lugar bajo una condición específica en la célula) (Sanfilippo et al., 2017). El análisis de transcriptomas posibilita interpretar el contenido de genes de un organismo cuando no se encuentra disponible la secuencia de su genoma. El objetivo principal de estos estudios es identificar todos los transcritos, incluidos mRNAs, RNAs no codificantes y RNAs pequeños para así determinar la estructura de cada gen; sus sitios de inicio 5' y final 3', patrones de splicing y otras modificaciones postranscripcionales. Además, cuantificar cambios en los niveles de expresión de cada transcrito bajo diferentes condiciones (Raghavachari y Garcia-Reyero, 2018).

En la secuenciación de transcriptomas la tecnología de microarreglos ha sido muy usada para medir el nivel de expresión promedio de cada gen en una población de células. Sin embargo, los microarreglos solo detectan RNA conocidos, por lo que no se pueden utilizar para el descubrimiento. Además, la hibridación de fondo interfiere con la detección de expresión de bajo nivel y la saturación de la sonda provoca variaciones en los genes altamente expresados que no son detectables. Sin embargo, los avances en las tecnologías de secuenciación han hecho posible cuantificar el transcriptoma de una población de células mediante plataformas de NGS que cuentan con técnicas como RNA-seq (Srivastava et al., 2018).

RNA-seq se ha convertido en una alternativa a las técnicas de secuenciación convencionales. Mediante esta es posible la creación de perfiles de todo el transcriptoma, lo que permite revelar alteraciones en las redes de señalización completas y conducir potencialmente a la identificación de nuevos genes de interés (Ayturk, 2019). Además, el desarrollo de la secuenciación de una población de mRNA resulta efectivo para construir y caracterizar el transcriptoma de un organismo, así como obtener información sobre los niveles de expresión génica en múltiples condiciones y una resolución elevada (McDermaid et al., 2019). Esta

técnica ofrece una cobertura completa de los transcritos, genera información no solo de la secuencia, sino también de la estructura de exones y posibles eventos de splicing alternativo (Soto Sedano y López Carrascal, 2014).

Un gran número de proyectos de investigación se han centrado en el estudio de transcriptomas. Esto puede deberse a que la secuenciación del genoma es un proceso costoso, laborioso y complejo. Aunque existe una tercera alternativa que pudiera ser la proteómica, solamente del 1-2% de los genes están codificados y del 80-90% de los genes transcritos no se traducen a proteína (Anamika et al., 2016). Por estas razones, es más conveniente realizar un análisis transcriptómico, ya que en la mayoría de los estudios es más probable encontrar genes que se estén expresando bajo un estímulo o condición determinada.

2.4.3. Estudios de transcriptómica relacionados con la síntesis de terpenos

En plantas se han realizado estudios de transcriptómica que han permitido la identificación de enzimas que intervienen en las diferentes etapas de la síntesis de terpenos. Relacionados con la ruta del ácido mevalónico se han identificado en *Artemisia annua* genes que codifican para la 3-hidroxi-3-metil-glutaril CoA reductasa y la acetoacetil CoA tiolasa (Soetaert et al., 2013), en *Asparagus racemosos* genes que codifican para la hidroximetil-glutaril CoA sintasa y la mevalonato pirofosfato decarboxilasa (Upadhyay et al., 2014) y en *Platycodon grandiflorum* genes que codifican para la mevalonato quinasa, fosfomevalonato quinasa, geranil pirofosfato sintasa e isopentenil difosfato isomerasa (Ma et al., 2016).

En varias plantas de interés medicinal se han identificado genes que codifican enzimas que intervienen en la ruta MEP. Entre estos se pueden destacar el estudio de la síntesis de ginsenoside en *Panax ginseng* para lo cual se identificaron todas las enzimas de la ruta de síntesis de terpenos que tiene lugar en los cloroplastos (Xue et al., 2019).

Además, se han identificado enzimas que intervienen en la producción de triterpenos de estructuras más complejas. Entre estas se encuentran una enzima clave en la biosíntesis de witanólidos, un grupo de lactonas esteroidales derivadas del 24-metilencolesterol en *Withania somnifera* (Knoch et al., 2018). En tres

especies de la familia Lamiaceae (*Rosmarinus officinalis*, *Salvia officinalis* y *Thymus persicus*) mediante un análisis de red de coexpresión y expresión de metabolitos específicos de tejidos, se identificaron posibles genes involucrados en las vías biosintéticas de ácidos triterpénicos (Aminfar et al., 2019). En cuatro plantas de la familia Celastraceae productoras de celastrol se usaron datos transcriptómicos para identificar 21 supuestas OSC. Mediante la caracterización funcional de estas ciclasas se elucidaron reacciones enzimáticas relacionadas con la síntesis de friedelina (Hansen et al., 2020).

La generalidad de los estudios de transcriptómica, relacionados con la síntesis de terpenos que se encuentran publicados, abordan rutas donde se sintetiza un terpenoide en particular. Esto ha permitido determinar que enzimas como las OSCs catalizan un gran número de reacciones hasta obtener el esqueleto básico triterpénico de estructura pentacíclica. Sin embargo, no nos muestran la diversidad enzimática y especificidad metabólica en la funcionalización de distintos esqueletos triterpénicos, por lo cual elucidar una ruta de síntesis de estos metabolitos puede resultar una labor compleja.

2.4.4. Estudios de transcriptómica en el género *Galphimia*

En la base de datos SRA del NCBI se encontraron datos crudos de lecturas de secuenciación del transcriptoma de *Galphimia gracilis*. Estas fueron enviadas por el Department of Biological Sciences luego de formar parte de un proyecto de secuenciación de 100 transcriptomas de plantas, que realizó un consorcio multidisciplinario internacional. Este consorcio adquiere secuencias de genes a gran escala para Viridiplantae (plantas verdes), incorporando en algún nivel filogenético/taxonómico casi todas las especies conocidas, desde angiospermas hasta algas. La secuenciación de este transcriptoma se realizó en la plataforma Illumina mediante RNA-seq en Julio de 2017, programada para la obtención de lecturas de extremos pareados. En esta base de datos se encuentran disponibles 1.1GB de información. Las lecturas de *G. gracilis* serán de utilidad para el análisis transcriptómico de *Galphimia spp.*

Los estudios referidos con anterioridad acreditan las actividades biológicas que poseen especies del género *Galphimia*. En concordancia con lo sucedido con otras plantas de importancia medicinal, las rutas por las cuales se producen los metabolitos bioactivos no se encuentran explicadas en su totalidad. Aunque se han realizado trabajos que involucran procedimientos de ciencia básica, así como técnicas espectroscópicas y de biología molecular que han intentado explicar sus especificidades genéticas, resulta necesario abordar este tema desde la perspectiva de expresión de genes. De esta forma se podrán analizar patrones de expresión coordinada de genes relacionados con una determinada ruta metabólica o la influencia de otras rutas sobre la que es objeto de análisis.

Justificación

En varias plantas se han estudiado diferentes etapas de la síntesis de los triterpenos. Sin embargo, no hay estudios publicados sobre el desarrollo de esta ruta en *Galphimia spp.*, ni tampoco de la vía de síntesis y de las etapas intermedias para la producción de los triterpenos modificados galfiminas del tipo nor-secofriedelano. Considerando las propiedades biológicas demostradas de las galfiminas, sería importante realizar un análisis transcriptómico para identificar posibles transcritos y genes funcionales útiles en la ruta de síntesis de estos triterpenoides. Esta información permitirá comprender reacciones bioquímicas de la producción de estos metabolitos de importancia medicinal, lo cual posteriormente permitirá regular la producción de galfiminas utilizando herramientas biotecnológicas.

Hipótesis

El estudio y comparación de transcriptomas de dos poblaciones de *Galphimia spp.* una productora y otra no productora de galfiminas, permitirán identificar posibles genes candidatos que codifican enzimas involucradas en la ruta de producción de estos triterpenoides.

Objetivo general

Comparar transcriptomas de dos poblaciones de *Galphimia* spp., una productora y otra no productora de galfiminas para identificar posibles genes candidatos que codifican enzimas de la ruta de síntesis de las galfiminas.

Objetivos específicos

- 1) Obtener plántulas de *Galphimia* spp. en condiciones semicontroladas de temperatura y humedad.
- 2) Determinar la presencia de galfiminas en extractos metanólicos de *Galphimia* spp.
- 3) Secuenciar transcriptomas de *Galphimia* spp. mediante RNAseq con el uso de Illumina Next Seq 550.
- 4) Analizar los transcriptomas de *Galphimia* spp. mediante el uso de software especializados.
- 5) Identificar los posibles genes candidatos relacionados con la biosíntesis de galfiminas.
- 6) Validar el transcriptoma mediante una Reacción en Cadena de la Polimerasa cuantitativa en tiempo real (qRT-PCR).

3. Materiales y métodos

3.1. Obtención de plantas de *Galphimia* spp.

La colecta de plantas y semillas se realizó de forma manual bajo permiso de la Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT) en la población productora de galfiminas (PG) Doctor Mora (Guanajuato) (W 100° 19.22; N 21° 08.7; 2120 m de altura) y la no productora de galfiminas (NPG) Tepoztlán (Morelos) (W 99° 06.97 N 18° 59.35, 1700 m de altura). Las muestras se depositaron en el Herbario HUMO del Centro de Educación Ambiental e Investigación Sierra de Huatla, de la Universidad Autónoma del Estado de Morelos (UAEM) donde fueron debidamente identificadas, con número de voucher 15189 y 15485, respectivamente.

A partir de 10 semillas de cada población se obtuvieron y cultivaron plántulas de *Galphimia* spp. en condiciones controladas de temperatura (25°C) y humedad relativa (60%), durante 75 días. Posteriormente las plántulas se transfirieron a macetas de 2 L y se cultivaron en condiciones naturales. El experimento se realizó de septiembre de 2019 a enero de 2021 en el laboratorio de cultivo de tejidos vegetales del Centro de Bioingenierías del Instituto Tecnológico y de Estudios Superiores de Monterrey, campus Querétaro. Durante el periodo de cultivo se evaluaron los siguientes parámetros: altura de las plántulas, número de ramas laterales, área foliar (ecuación de la elipse) y aparición de inflorescencias.

Se realizó un análisis estadístico con el paquete PASW Statistics (v.23) para Windows. Se verificaron los supuestos de normalidad Sapiro Wilk. Los datos que no seguían distribución normal se procesaron mediante análisis de varianza para dos variables no relacionadas H deKruskal Wallis y varias muestras independientes, U de Mann Whitney. Las diferencias se consideraron significativas para un valor de $p < 0,05$ en la prueba H de Kruskal Wallis y $p < 0,01$ en la prueba U de Mann Whitney.

3.2. Determinación de la presencia de galfiminas en los extractos metanólicos de *Galphimia* spp.

3.2.1. Preparación de los extractos

De plántulas de las poblaciones PG y NPG se obtuvieron 100 mg de tejido seco pulverizado, los cuales se mezclaron independientemente con 500 μ L de metanol. La mezcla se sonicó durante 10 min a 25 °C y se centrifugó a 1008 \times g por 5 min. El sobrenadante se colectó y el residuo vegetal fue reprocesado en una serie de 6 extracciones consecutivas.

3.2.2. Cromatografía en capa fina (TLC)

Se aplicaron 2 μ L de los extractos correspondientes de cada población directamente a la placa cromatográfica para lo cual se usó un capilar de vidrio. Posteriormente, la placa se colocó en una fase móvil de cloroformo: acetato de etilo (2:1 v/v) y se dejó correr durante 20 min. Después, se asperjó con una solución de vainillina (1 g de vainillina en 100 mL de ácido sulfúrico) y se sometió a una temperatura de 120 °C hasta la obtención de bandas que indicaron la presencia o ausencia de galfiminas, mediante la comparación con los R_f de muestras auténticas de control.

3.2.3. Cromatografía líquida de alta eficiencia (CLAE)

Los extractos de cada población se ajustaron a una concentración de 5 mg/mL. Se usaron 10 μ L de cada muestra para realizar la inyección en un equipo Agilent. Para la corrida se empleó una columna Symmetry C-18 (5 μ m, 4.6 \times 150 mm), una elución isocrática de acetonitrilo-agua (1:1) y un flujo de 1 mL/min. Se usó como control una fracción de galfiminas pura. Las determinaciones se realizaron por triplicado.

3.2.4. Secuenciación de transcriptomas de *Galphimia* spp. mediante RNAseq con el uso de Illumina Next Seq 550

3.2.5. Extracción de RNA

De hojas de plantas de ambas poblaciones, así como de raíces de la población PG se extrajo RNA. La extracción se realizó en una campana de flujo laminar horizontal para garantizar condiciones de esterilidad y asepsia. Las muestras se maceraron

con nitrógeno líquido y para la extracción se usaron 100 mg de tejido. Se utilizó el RNA easy Mini Kit, (Qiagen).

3.2.5.1. Análisis de la concentración, pureza e integridad del RNA

La concentración del RNA se determinó mediante un análisis fluorométrico en el Qubit 4^{Td} (Invitrogen, ThermoFisher). Se añadió al Buffer RNA HS el fluoróforo RNA HS reagent 200x a razón de 1 μ L por determinación. Se usaron 10 μ L de los estándares 1 y 2 disueltos en 190 μ L de Buffer RNA HS, como control. En las determinaciones de la concentración de RNA se usaron 2 μ L de muestra y 198 μ L de Buffer RNA HS.

Para determinar la pureza del RNA se realizó un análisis espectrofotométrico en el Nanodrop One (Thermo Scientific). Para ello se empleó como blanco 2 μ L de agua libre de nucleasas. En las determinaciones se analizaron 2 μ L de cada una de las muestras de RNA extraído. Se determinaron los índices de puridades a partir de la medición de valores de absorbancia 260/280 y 260/230. Además, se detectaron otras impurezas como la presencia de isotiocianato de guanidina.

La integridad del RNA extraído se analizó mediante una electroforesis en gel de agarosa. En la preparación del gel se usaron 0.5 g de agarosa y 50 mL del buffer TBE 1x. Se añadió 1 μ L del colorante GelRed previo a la solidificación del gel para garantizar la posterior visualización bajo luz ultravioleta. En cada pocillo del gel se aplicaron 6 μ L de cada una de las ocho muestras que contenían 1 μ L del colorante 5x DNA Loading Dye para marcar el avance de la electroforesis. Además, se aplicó en uno de los pocillos el marcador de peso molecular HyperLadder 100 bp (Bioline). La corrida se realizó durante 60 min con una corriente de 60 V. Una vez finalizada la corrida, el gel se observó bajo luz ultravioleta en un transiluminador ibright con el objetivo de determinar la presencia del RNA 18s y 28s.

Para el análisis de integridad del RNA se realizó además una electroforesis capilar en el Qsep 100 Advance Bioptric. Para ello se usó un cartucho de RNA (R1-O-220217-2 & 84) y 5 mL de Buffer de separación. Las muestras de RNA y del marcador de alineamiento (RNA-LM, 20 nt) se prepararon con Buffer de dilución a una concentración de 50 ng/ μ L. La corrida se realizó mediante el método R-2-10-

04-480. Se calculó la integridad del RNA como la proporción entre las subunidades de ARN ribosómico grandes (28S) y pequeño (18S).

3.2.6. Obtención de las librerías

Para la obtención de las librerías se siguió el protocolo TruSeq Stranded Total RNA de Illumina. Se realizó la reducción y fragmentación del RNA, la síntesis de la primera cadena de cDNA, la síntesis de la segunda cadena, la adenilación del extremo 3', la unión de índices y adaptadores y finalmente el enriquecimiento de los fragmentos de cDNA.

3.2.6.1. Reducción y fragmentación del RNA

Se eliminó el rRNA del RNA total. Para ello, en tubos de 200 μ l se diluyó el RNA total en agua ultrapura libre de nucleasas para obtener un volumen final de 10 μ l. A cada muestra se le añadieron 5 μ l del rRNA Binding Buffer (RBB), posteriormente se le añadió igual cantidad del rRNA Removal Mix-Plant (RRMP). Esta mezcla se centrifugó a 280 g xg durante 1 min y se transfirió al termociclador para su desnaturalización donde fue sometida a 68 °C durante 5 min y después permaneció a 4 °C. Luego, se incubó a temperatura ambiente durante 1 min.

Para la separación del rRNA, se aplicó un 1 min de Vortex a una suspensión de rRNA Removal Beads (RRB) de la cual se añadieron 35 μ l a cada muestra. Se transfirieron 20 μ l de cada muestra a tubos nuevos y se mezclaron a 112 xg durante 1 min. Posteriormente, se incubaron a temperatura ambiente durante 1 min. Transcurrido este tiempo, las muestras se transfirieron a un soporte magnético donde permanecieron durante 1 min, periodo en el cual las perlas se adhirieron a la pared del tubo. Luego, se transfirió todo el volumen de cada muestra a tubos nuevos. Estos permanecieron en el soporte magnético durante 1 min.

Para la purificación del RNA, se aplicó 1 min de Vortex a una suspensión de RNAClean XP Beads de la cual se añadieron 99 μ l a cada tubo. Esta mezcla se agitó a 363 xg por 2 min y se incubó a temperatura ambiente durante 15 min. Los tubos se transfirieron a un soporte magnético en donde permanecieron durante 5 min, tiempo en el que el líquido se tornó transparente. Posteriormente, se eliminó el sobrenadante y se realizaron los lavados. Para ello, a cada tubo se añadieron 200

µl de etanol 70% fresco, estos se incubaron en el soporte magnético durante 30 segundos y después se eliminó el sobrenadante. El contenido de los tubos se dejó secar en el soporte magnético durante 15 min a temperatura ambiente. Transcurrido este tiempo, los tubos de cada muestra se separaron del soporte magnético. Se añadieron 11 µl del Elution Buffer (ELB) a cada tubo, estos se agitaron a 363 xg por 2 min y se incubaron a temperatura ambiente durante 2 min. Luego, se centrifugaron a 280 xg por 1 min y se transfirieron al soporte magnético, donde se incubaron por 5 min hasta que el líquido se tornó transparente. Se transfirieron 8.5 µl del sobrenadante de cada muestra a tubos nuevos, el contenido de cada tubo se mezcló a 287 xg por 20 segundos. Posteriormente, se colocaron en el termociclador donde se mantuvieron a 94 °C por 8 min y luego permanecieron a 4 °C.

3.2.6.2. Síntesis de la primera cadena de cDNA

Se centrifugó el First Strand Synthesis Act D Mix (FSA) a 600 x g por 5 segundos y se le añadieron 50 µl de SuperScript II Reverse Transcriptase. A cada una de las muestras se añadieron 8 µl de FSA con SuperScript II, luego se agitaron a 287 xg por 20 segundos. Se centrifugaron a 280 xg durante 1 min. Posteriormente, se colocaron en el termociclador a 25°C por 10 min, luego a 42°C por 15 min, 70°C por 15 min y finalmente permanecieron a 4°C.

3.2.6.3. Síntesis de la segunda cadena de cDNA

Para la síntesis de la segunda cadena de cDNA a cada muestra se le adicionaron 5 µl de CTE diluido en una proporción 1:50 en Resuspension Buffer (RSB). Después, se añadieron 5 µl de RSB y 20 µl de Second Strand Marking Master Mix (SMM). Se agitaron a 287 xg por 20 segundos y se centrifugaron a 280 xg por 1 min. Posteriormente, se transfirieron al termociclador donde se incubaron a 16°C durante 1 h.

Se purificó el cDNA, para lo cual se añadieron 90 µl de AMPure XP beads a cada muestra. Se mezclaron a 363 xg por 2 min, se incubaron a temperatura ambiente durante 15 min y luego se centrifugaron a 280 xg por 1 min. Después se colocaron en el soporte magnético por 5 min hasta que el líquido se tornara claro. Se extrajeron 135 µl de sobrenadante, los cuales fueron desechados. Se realizaron dos lavados

para lo cual se agregaron 200 µl de etanol 80% fresco a cada muestra. Se incubaron en el soporte magnético durante 30 segundos, luego se desechó el sobrenadante y se dejaron secar a temperatura ambiente por 15 min, evitando el sobre secado de las perlas. Las muestras se retiraron del soporte magnético y se le añadieron 17.5 µl de RSB. Posteriormente, se mezclaron a 363 xg por 2 min, se incubaron a temperatura ambiente por 2 min y se centrifugaron a 280 xg por 1 min. Se colocaron nuevamente en el soporte magnético durante 5 min hasta que el líquido se tornó claro. Finalmente, se transfirieron 15 µl del sobrenadante a tubos nuevos.

3.2.6.4. Adenilación del extremo 3'

Se añadieron a cada muestra 2.5 µl de A-Tailing Control (CTA) diluido 1:100 en RSB y 12.5 µl de A-Tailing Mix (ATL). Se mezclaron a 363 xg por 2 min y se centrifugaron a 280 x g por 1 min. Posteriormente, se incubaron en el termociclador a 37°C por 30 min, luego a 70°C por 5 min y después permanecieron a 4°C. Finalmente, se centrifugaron a 280 x g por 1 min.

3.2.6.5. Unión de índices adaptadores

A cada muestra se agregaron 2.5 µl de CTL diluido, 2.5 µl de Ligation Mix (LIG) y 2.5 µl de RNA adapters. Se mezclaron a 363 xg por 2 min y se centrifugaron a 280 x g por 1 min. Posteriormente, se incubaron en el termociclador a 30°C durante 10 min y después permanecieron a 4°C. Transcurrido este tiempo, se agregaron 5 µl de Stop Ligation Buffer (STL) a cada muestra, se mezclaron a 363 xg por 2 min y se centrifugaron a 280 x g por 1 min.

Se purificaron los fragmentos ligados. Para ello se adicionaron 42 µl de AMPure XP beads, se mezclaron a 363 xg por 2 min, se incubaron a temperatura ambiente por 15 min y se centrifugaron a 280 x g por 1 min. Los tubos con las muestras se transfirieron a un soporte magnético durante 5 min hasta que el líquido se tornó claro. Luego, se desechó el sobrenadante y se realizaron dos lavados con 200 µl de etanol 80% fresco cada uno. Durante los lavados, las muestras permanecieron en el soporte magnético por 30 min. Después, se desechó el sobrenadante y se dejaron secar a temperatura ambiente en el soporte magnético durante 15 min. Transcurrido este tiempo, las muestras se retiraron del soporte magnético y se les

adicionaron 52.5 µl de RSB. Estas se mezclaron a 363 xg por 2 min, se incubaron a temperatura ambiente por 2 min y se centrifugaron a 280 x g durante 1 min. Se colocaron sobre el soporte magnético por 5 min hasta que el líquido se tornó claro. Se transfirieron 50 µl del sobrenadante a tubos nuevos y se realizó el mismo procedimiento con la diferencia de que el volumen de AMPure XP beads añadido fue de 50 µl y el de RSB de 22.5 µl. Finalmente, se transfirieron 20 µl del sobrenadante a tubos nuevos.

3.2.6.6. Enriquecimiento de los fragmentos de cDNA

Las muestras se colocaron en hielo y se le añadieron 25 µl de PCR Master Mix (PMM). Posteriormente, se mezclaron a 287 xg por 20 segundos y se centrifugaron a 280 x g por 1 min. Se transfirieron al termociclador y se corrió el siguiente programa de PCR: 98°C por 30 segundos, 15 ciclos de: 98°C por 10 segundos, 60°C por 30 segundos, 72°C por 30 segundos, 72°C por 5 min y finalmente permanecieron a 4°C.

Se purificaron los fragmentos de DNA para lo cual se agregaron a cada muestra 47.5 µl de AMPure XP beads, se mezclaron a 363 xg por 2 min, se incubaron a temperatura ambiente durante 15 min y se centrifugaron a 280 x g por 1 minuto. Después, se transfirieron al soporte magnético donde permanecieron durante 5 min hasta que el líquido se tornó claro. Se desechó el sobrenadante. Se realizaron dos lavados para lo cual se añadieron 200 µl de etanol fresco al 80% a cada muestra. Estas se incubaron en el soporte magnético durante 30 segundos. Posteriormente, se desechó el sobrenadante y se dejaron secar a temperatura ambiente por 15 min. Las muestras se separaron del soporte magnético y se les agregaron 32.5 µl de RSB, se mezclaron a 363 xg por 2 min, se incubaron a temperatura ambiente durante 2 min y se centrifugaron a 280 x g por 1 min. Después, se transfirieron al soporte magnético y se incubaron por 5 min hasta que el líquido se tornó claro. Finalmente, se transfirieron 30 µl del sobrenadante de cada muestra a tubos nuevos.

3.2.6.7. Determinación de la calidad de las librerías

Se cuantificaron las librerías mediante un análisis fluorométrico en el Qubit 4^{td} (Invitrogen, Thermofisher). Se añadió al Buffer ds DNA HS el fluoróforo ds DNA HS

reagent 200x a razón de 1 μ L por determinación. Se usaron 10 μ L de los estándares 1 y 2 disueltos en 190 μ L de Buffer ds DNA HS, como control. En las determinaciones de la concentración de dsDNA se usaron 2 μ L de muestra y 198 μ L de Buffer ds DNA HS.

Se determinó el tamaño de los fragmentos de cDNA obtenidos mediante una electroforesis capilar en el Bioanalyzer Agilent 2100 según el método DNF-474-HS-NGS. A dos placas de 96 pocillos se añadieron 1 mL y 200 μ L de los buffer Inlet y Rinser, respectivamente. Posteriormente, 2 μ L de DNA de cada librería diluido en RSB 1:1 se mezclaron por separado con 22 μ L de HS NGS Diluent Marker (1-6000 bp) en la primera fila de otra placa de 96 pocillos. En la posición H8 se añadieron 2 μ L del HS NGS DNA Ladder diluido en 22 μ L de HS NGS Diluent Marker (1-6000 bp). Al resto de los pocillos de la placa se le añadieron 24 μ L de una solución blanco. El tamaño de las librerías se comprobó mediante la aparición de bandas de aproximadamente 260 pb.

3.2.6.8. Desnaturalización y Dilución de las librerías

Para la desnaturalización, se mezclaron 10 μ L de las librerías a una concentración de 1.6 nM con 10 μ L de NaOH 0.2 N. Se centrifugó a 280 x g por 1 min y se incubó a temperatura del cuarto durante 5 min. Posteriormente, se le añadieron 10 μ L de Tris-HCl 200 mM, se mezcló brevemente con vortex y se centrifugó a 280 x g por 1 min.

Las librerías previamente desnaturalizadas se diluyeron hasta una concentración de 20 pM. Para ello, se agregaron 970 μ L de (Hybridization Buffer) HT1 a las librerías desnaturalizadas. Se mezclaron brevemente con el vortex y se centrifugaron a 280 x g durante 1 min. Se mantuvieron en hielo hasta realizar la dilución final a 1.5 pM. Se transfirieron a un tubo nuevo 97.5 μ L de las librerías, se añadieron 1202.5 μ L de HT1. Se mezclaron con un breve pulso de centrifuga. Se obtuvo un volumen final de 1.3 mL que se depositó en el secuenciador.

3.2.7. Secuenciación de las librerías

Se secuenciaron 6 librerías (una de hojas de las plantas de la población PG, una de hojas de las plantas de la población NPG y una de raíces de la planta PG, con sus

correspondientes réplicas) en la plataforma Illumina Next Seq 550. Se programaron 2 x 76 ciclos de secuenciación y la obtención de lecturas de extremos pareados. Se uso como control el Phix.

3.3. Análisis de los transcriptomas de *Galphimia* spp. mediante el uso de softwares especializados

3.3.1. Análisis de la calidad de las lecturas, depuración y ensamblado *de novo* de los transcriptomas

La calidad de las lecturas obtenidas se analizó mediante el software FastQC (v.0.11.8). Posteriormente, se realizó un recorte al inicio y final de cada lectura con el objetivo de mejorar la calidad total de estas. Para ello se empleó el Trimmomatic, (v.0.38.0), con los parámetros “Leading:3 y Trailing:3”. El ensamblado *de novo* de los transcriptomas se realizó mediante el ensamblador Trinity (v.2.9.1).

3.3.2. Anotación funcional de los transcriptomas

La anotación funcional de los transcriptomas se realizó mediante el Blast2GO. Los transcritos ensamblados fueron alineados mediante un tblastx con la base de datos del NCBI. Se usó un BLAST E-value: $1.0e^{-3}$ y por cada secuencia se realizaron 20 alineamientos. Posteriormente, se realizó un mapeo de las secuencias y la anotación funcional en los niveles de ontología génica (OG) Proceso biológico (PB), Función Molecular (FM) y Componente Celular (CC). Las secuencias que no fueron anotadas en OG se identificaron en las bases de datos InterPro, UniProt y TAIR.

3.3.3. Análisis de expresión diferencial

Se identificaron los pares de genes correspondientes entre individuos de ambas poblaciones de *Galphimia* spp., para esto se usó el programa de formación de cluster a partir de ensamblados *de novo*, Corset (v.1.09). Los transcritos resultantes de este agrupamiento fueron verificados con la anotación funcional. Se realizó una cuantificación de abundancia de los transcritos ensamblados en cada transcriptoma mediante el programa Kallisto (v.0.46.0.4). Luego, se construyó la matriz de conteo en R (v.4.0.4 para Windows), con el uso de las librerías plyr, dplyr y stats. Se filtraron los genes con conteos bajos en las librerías mediante el *CPM Filter*, sobre una base de conteos por millón (CPM). Para el análisis de expresión diferencial se empleó la

librería edgeR. Las muestras en la matriz de conteo fueron normalizadas sobre una base de transcritos por millón (TPM). Se comparó la expresión de los transcritos de la población productora contra los de la población no productora y el control. Se consideraron transcritos expresados diferencialmente aquellos con una tasa de falsos positivos (FDR) <0.01 . Los transcritos expresados diferencialmente se clasificaron en regulados positivamente cuando el logFold Change (logFC) >2 y negativamente cuando el logFC <2 . Los transcritos con mayores valores de expresión diferencial en cada muestra fueron visualizados en un mapa de calor que se construyó con la librería pheatmap.

3.4. Identificación de los posibles genes candidatos relacionados con la biosíntesis de galfiminas.

Con los datos obtenidos de la anotación funcional de los transcriptomas secuenciados y los resultados la cuantificación de abundancia de los transcritos de cada muestra, se identificaron genes involucrados en la síntesis de terpenos y posibles genes candidatos para la síntesis de galfiminas. Se construyó una red de co-expresión para identificar las interacciones entre los transcritos de estas rutas. Se usaron las librerías de R, GENIE3, igraph, RCy3 y Rgraphviz. La red de co-expresión se calculó en forma de una matriz de adyacencia ponderada, utilizando conjuntos de árboles de regresión. No se utilizaron genes reguladores candidatos. El método del árbol fue Random Forests. El número de árboles en cada objetivo fue de 1000. Los resultados se visualizaron en Cytoscape (v.3.8.2).

3.5. Validación de los datos del transcriptoma mediante PCR cuantitativa

A partir del RNA extraído de las hojas de las plantas de ambas poblaciones se obtuvo cDNA mediante el PrimerScript RT Reagent Kit with gDNA Eraser (Takara Bio USA, Inc.). Se mezclaron 3.75 μL de RNA de cada población con 5 μL de mRQ Buffer y 1.25 μL de la enzima mRQ. La mezcla se incubó 1 h a 37°C, seguido de 5 min a 85°C para inactivar la enzima. La concentración y pureza del cDNA se evaluaron en el espectrofotómetro Nano-Drop 2000 (NanoDrop Technologies, Technologies, Wilmington, DE).

De acuerdo con los datos del análisis transcriptómico de la población productora de galfiminas, la secuencia de nucleótidos de 10 posibles genes candidatos relacionados con la síntesis de galfiminas (CYP82C4, P450 TBP, PLAC8, 5'-AMP-activated protein kinase, serine/threonine-protein kinase AtPK2/AtPK19, P450, CYP71A24, CYP71B34, CYP72A, CYP81D) se usaron para el diseño de primers (Cuadro 1) mediante el programa *Primer3 plus* ya que representaron los miembros de la familia de monooxigenasas que mostraron los mayores conteos genéticos.

Cuadro 1. Secuencias de primers sintetizados para qPCR

Descripción del gen	Símbolo del gen	Secuencias de primers 5'-3'	Longitud del producto de PCR (bp)
CYP82C4	CYP82C4	F 5' TCATTTGGTGGGCTAAAAGC 3' R 5' ACGAGTCCGTTAATGGTTGC 3'	191
P450 TBP	P450 TBP	F 5' AGCTTCGTCGCAAGTGAAAT 3' R 5' TAGGACGCCTGCGTTATCTT 3'	189
PLAC8	PLAC8	F 5' ACTGTTGATAACCCCGATGG 3' R 5' TCTCAGGAAGTCCGAAGTGG 3'	170
5'-AMP-activated protein kinase	AMPA	F 5' CCAAGGAAAAGGTTTCACA 3' R 5' GGCAAAAGCAATGGCTAAGA 3'	229
serine/threonine-protein kinase AtPK2/AtPK19	Ser/Thr	F 5' TTTCAACTGGACCACAAAGG 3' R 5' CGTTCGGGAAAGTCTACCAA 3'	164
P450	P450	F 5' CCTTGAGGTCATGGCTGAAT 3' R 5' GCTTCTCTCCAAAGGCACAC 3'	232
CYP71A24	CYP71A24	F 5' CAAACCGGCCTAAATCAAAG 3' R 5' GCTCTTGTTCACTTTCTCAATC 3'	200
CYP71B34	CYP71B34	F 5' CCGTGAGAGAGGCCATTAAC 3' R 5' GTGCGAAAGATTTGCGTTCT 3'	160
CYP72A	CYP72A	F 5' TTGTTGGCTTTTCGAGGAAT 3' R 5' AAGCATCAGGAGTGGCAAAC 3'	167
CYP81D1	CYP81D1	F 5' TCGGAGGATTGGACTACGAC 3' R 5' TTCCGCCATAACATTTCTCC 3'	219

Se realizó una PCR cuantitativa mediante el uso del kit TB Green® Advantage® qPCR Premix (Takara Bio USA, Inc.). Se desarrolló el protocolo propuesto por el fabricante y la corrida en el termociclador StepOne™ Real-Time PCR System (Applied Biosystems, USA). Se programaron las siguientes condiciones de amplificación: activación/desnaturalización a 95°C durante 10 segundos, desnaturalización a 95°C por 5 segundos, extensión a 60°C por 30 segundos. Posteriormente, la curva de Melting en un rango de 56-95°C. Se usaron tres réplicas biológicas con tres replicas técnicas para cada muestra. La expresión relativa se

cuantificó por el método de $2^{-\Delta\Delta Ct}$ y se usó como gen normalizador de la expresión el U6 (Tang et al., 2019).

4. Resultados y discusión

4.1. Obtención de plantas de *Galphimia* spp.

A los 14 días posteriores a la colocación de las semillas en la cámara húmeda, se encontraron germinadas más del 40% de estas para ambas poblaciones. Se observó el desarrollo de las hojas cotiledonares. La morfología de la raíz fue variable, algunas fueron rectas y otras presentaron curvaturas (Figura 9 a, b). En todas las raíces se observaron abundantes pelos adsorbentes, lo cual indicó que la humedad relativa a la que estuvieron expuestas las semillas fue adecuada.

A los 30 días posteriores a la germinación se observó la elongación del tallo a partir de las hojas cotiledonares así como la aparición de las primeras hojas verdaderas (Figura 9 c, d). En las hojas verdaderas se distinguió la nervadura primaria y secundaria. Las yemas axilares adquirieron una coloración roja. Durante el periodo de 2-4 meses posteriores a la germinación fue evidente la elongación del tallo, así como el aumento del tamaño y el número de hojas (Figura 9 e, f), lo cual pudo estar relacionado con la transferencia de las plantas a macetas de 2 L de capacidad y el cambio a condiciones naturales, donde la luz solar pudo influir directamente en la actividad fotosintética.

Durante los 7-9 meses posteriores a la germinación, se observó un gran desarrollo de las ramas laterales y se incrementó el número de hojas por plantas (Figura 9 g-j).

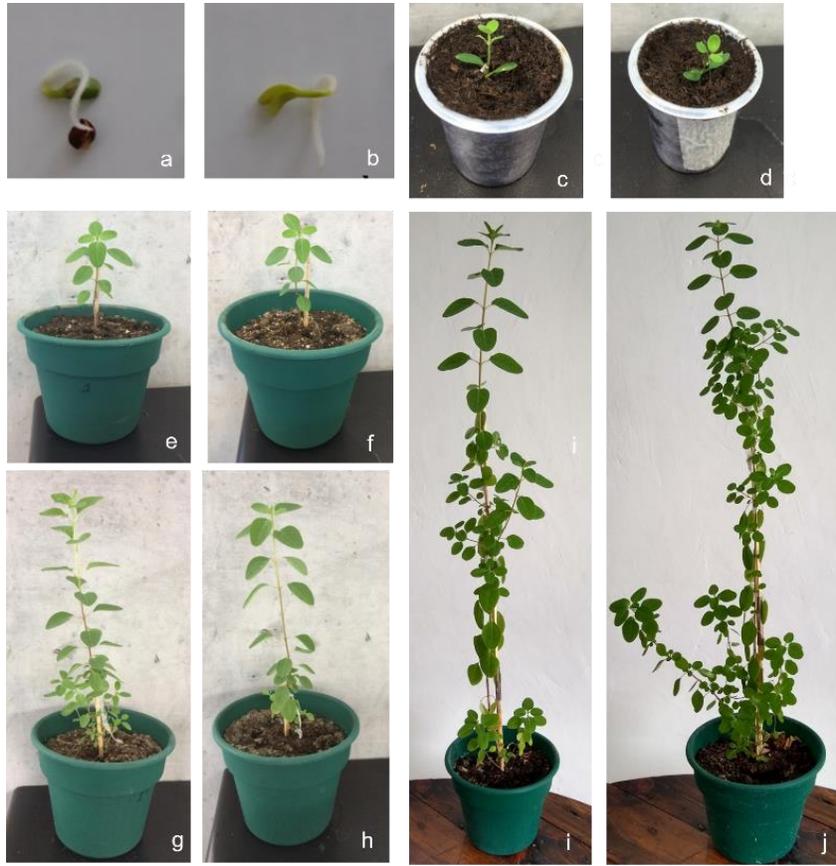


Figura 9. Germinación y crecimiento de plantas de dos poblaciones de *Galphimia* spp. en condiciones semicontroladas de temperatura (25°C) y humedad relativa (60%). Población productora de galpfiminas (a, c, e, g, i), población no productora de galpfiminas (b, d, f, h, j). 14 días después de la germinación (a, b). 1 mes después de la germinación (c, d). 2-4 meses después de la germinación (e, f). 7-9 meses posteriores a la germinación (g, h).

A los 14 días posteriores a la germinación, se observaron cambios significativos en las longitudes de los órganos vegetativos. En las plántulas PG, el área de las hojas cotiledonares fue menor con respecto a la de las plántulas NPG, pues existieron diferencias significativas entre los valores medios calculados (Cuadro 2). De modo similar, la longitud de las raíces de las plantas NPG fue mayor.

En la mayoría de las plántulas se observó la caída de las hojas cotiledonares. El área foliar fue diferente entre las plántulas de las dos poblaciones (Cuadro 2). Los mayores valores se calcularon en las plántulas NPG. Por otra parte, las plántulas PG presentaron la mayor altura (Cuadro 2). En el periodo de 2-4 meses, se observó

en ambas poblaciones un aumento del área foliar y la altura. En las plantas PG, la altura fue mayor.

Durante los 7-9 meses posteriores a la germinación, el área foliar no experimentó cambios significativos. Las plantas PG presentaron una mayor altura.

Cuadro 2. Desarrollo de plantas de dos poblaciones de *Galphimia* spp. desde los 14 días hasta los 15 meses posteriores a la germinación.

Tiempo posterior a la germinación	Métricas	Población productora de galfiminas	Población no productora de galfiminas
14 días	Área de hojas cotiledonares (cm ²)	0.0820b	0.1200a
	Longitud de las raíces (cm)	1.4070b	1.6730a
1 mes	Área foliar (cm ²)	0.7490b	0.7770a
	Altura (cm)	2.4780a	2.0440b
2 meses	Área foliar (cm ²)	8.608	8.810
	Altura (cm)	11.25a	8.693b
4 meses	Área foliar (cm ²)	12.657	12.858
	Altura (cm)	40.986	40.710
7 meses	Altura (cm)	63.44a	47.42b
	Número de ramas laterales	3	4
9 meses	Altura (cm)	71.90a	57.89b
	Número de ramas laterales	6b	11a
12 meses	Altura (cm)	100a	86b
	Total de inflorescencias	50	48
15 meses	Altura (cm)	130	130
	Total de frutos	140a	125b

Durante el cultivo de plántulas de *Galphimia* spp., la germinación de las semillas de las dos poblaciones fue adecuada. En la cámara húmeda que se estableció para este estadio fue posible mantener valores de humedad relativa superiores al 50% y regular la temperatura de acuerdo con los requerimientos de las plántulas. Además, se garantizó que la luz incidiera de modo uniforme sobre las semillas, lo que evitó el estiramiento de estas una vez que comenzara la germinación (Friesen, 2020).

En los primeros estadios de desarrollo de las plántulas se observaron diferencias en el área foliar de las dos poblaciones evaluadas. Durante este periodo se produce una división del contenido de carbono en la planta que justifica la variación del crecimiento entre los diferentes órganos vegetativos. El carbono que se fija durante la fotosíntesis se divide entre la síntesis de sacarosa que se usa rápidamente y exporta hacia las zonas donde ocurren otros procesos metabólicos como la

respiración, el crecimiento, la producción de nuevas hojas y raíces y el exudado de sustancias (Weraduwage et al., 2015). Si se tiene en cuenta esta división del contenido de carbono, se justifica que la planta NPG presentara mayores valores de área foliar y menor altura y las plantas PG mayor altura y menor área foliar. Sin embargo, en los últimos estadios de desarrollo, no se observó un crecimiento significativo de las hojas, así como diferencias en cuanto al área foliar de las plantas evaluadas. Esto podría estar relacionado con mecanismos de control que establece la planta adulta para evitar un potencial hídrico muy negativo. Generalmente, las hojas muestran un control estomático más fuerte para reducir la pérdida de agua, con el objetivo de disminuir la transpiración, la fotosíntesis y el crecimiento (N. Liu et al., 2019).

La altura de las plantas de *Galphimia* spp. fue variable entre las dos poblaciones durante su desarrollo. Este carácter depende del balance hídrico, el transporte de carbohidratos, la captación de la luz (Liu et al., 2019) y, principalmente, de la fertilidad del suelo (Santini et al., 2017). Las diferencias en cuanto a la altura en las dos poblaciones analizadas pudieron deberse a la relación que existió entre el contenido de nutrientes presentes en el suelo y las necesidades nutricionales de cada individuo en particular.

El desarrollo de las ramas laterales fue diferente en las poblaciones estudiadas. Estas se originan a partir de los meristemos del brote lateral que generalmente comienzan en el ápice del brote principal. Los meristemos laterales pueden convertirse en ramas, donde la iniciación y el crecimiento es un proceso único y continuo, o formar pocas hojas y entrar en un período de latencia (Couder, 2017). Las variaciones relacionadas al número y tamaño de ramas laterales en las plantas de ambas poblaciones pudieron estar relacionada con la fluctuación de los niveles de auxina (fitohormona con efectos pleiotrópicos sobre el desarrollo de la planta que se sintetiza principalmente en las hojas jóvenes del brote). Los diferentes estadios de crecimiento evaluados pudieron estar sujetos al proceso de dominancia apical en el cual elevados niveles de auxina aumentaron el crecimiento del ápice principal e inhibieron el desarrollo de las ramas laterales (Kebrom, 2017).

La aparición de inflorescencias se produjo de modo similar en las plantas de las dos poblaciones estudiadas. Sin embargo, el número de frutos fue mayor en las plantas PG. La aparición de frutos y su maduración está relacionado a la producción de reguladores del crecimiento vegetal como etileno, ácido abscísico, ácido jasmónico, giberelinas, brasinosteroides, ácido salicílico y melatonina (Fenn & Giovannoni, 2021). Estos reguladores pudieran haber estado en menor concentración en las plantas NPG e inhibir la aparición de un mayor número de frutos.

A los 17 meses posteriores a la germinación, las plantas de ambas poblaciones presentaron caracteres fenotípicos similares a los descritos por Anderson (2007) para el género *Galphimia* (Figura 10). Se caracterizaron por presentar un gran número de raíces adventicias, de 5-6 cm de longitud y abundantes pelos adsorbentes. Tenían porte de arbustos de 1.4 m de altura, con aproximadamente 15 ramas laterales en las plantas PG (Figura 10a) y 10 en las plantas NPG (Figura 10d). Las hojas presentaron forma elíptica con venas secundarias prominentes y el área foliar fue de 13.75 cm² (Figura 10b, e). Se observaron inflorescencias amarillas dispuestas en racimos terminales (Figura 10c, f).

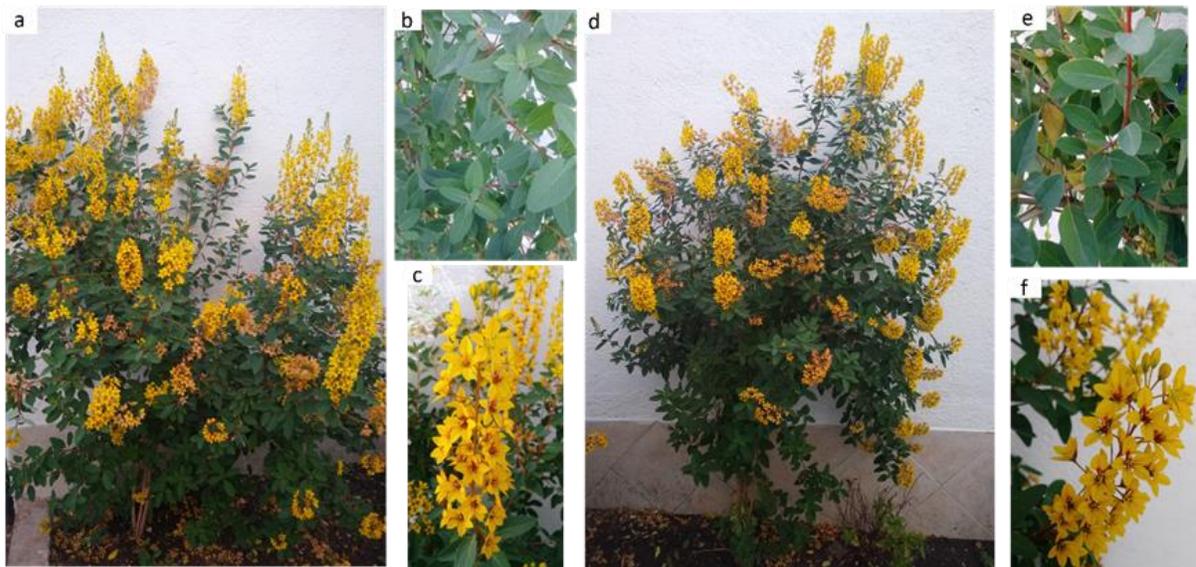


Figura 10. Plantas de dos poblaciones de *Galphimia* spp. cultivadas durante 17 meses. a, b, c (población productora de galfiminas), d, e, f (población no productora de galfiminas)

El cultivo de estas plantas en condiciones semicontroladas de temperatura y humedad, permitió su protección de la incidencia de factores intrínsecos o extrínsecos que les podrían ocasionar estrés durante su desarrollo. El control que se estableció para cultivarlas bajo estas condiciones posibilitó la obtención de un mayor número de individuos, así como una mayor pureza del componente activo que se deseaba aislar, debido a que todas las plantas crecieron en igualdad de condiciones. De modo general, todas las plantas de *Galphimia* spp. poseían los caracteres fenotípicos descritos para el género en publicaciones anteriores (Anderson, 2007). Es por ello que se consideraron individuos viables para la toma de muestras a utilizar en el análisis transcriptómico, objeto de estudio en este trabajo.

4.2. Determinación de la presencia de galfiminas en el extracto metanólico de *Galphimia* spp.

4.2.1. Obtención de los extractos

Se obtuvieron extractos metanólicos de las dos poblaciones de *Galphimia* spp. a una concentración de 50 mg/mL.

4.2.2. Cromatografía en capa fina

En el análisis por TLC se detectó la presencia de galfiminas en los extractos de las plantas PG, pues se observó la coloración violeta, característica de estos compuestos (Figura 11) bajo la aplicación del agente cromógeno ácido sulfúrico/vainillina seguido de calentamiento. En los extractos de las plantas NPG, no se detectó la presencia de galfiminas, ni tampoco en las raíces de las plantas PG (Figura 11). En todos los extractos de hojas analizados se observaron diferencias en cuanto a la aparición de bandas de otra coloración que correspondían a compuestos de baja polaridad.

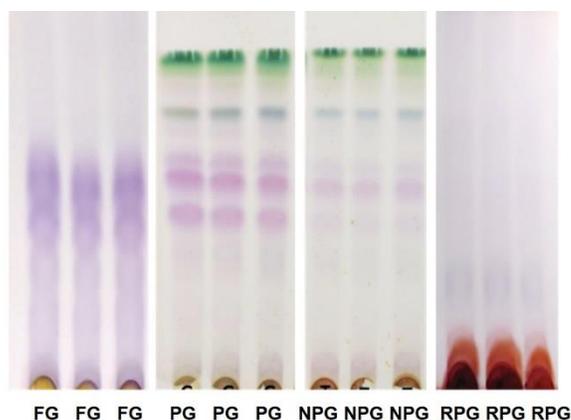


Figura 11. Separación de compuestos en extractos metanólicos de *Galphimia* spp. mediante TLC. FG (fracción de galfiminas), PG (plantas productoras de galfiminas), NPG (plantas no productoras de galfiminas), RPG (raíces de plantas productoras)

El análisis mediante TLC permitió la separación de las galfiminas del resto de los metabolitos presentes en los extractos metanólicos evaluados. El perfil químico obtenido mediante este análisis por TLC estableció diferencias entre las poblaciones de *Galphimia* spp. analizadas. La presencia de bandas coloreadas, que se observaron con el mismo R_f en el cromatograma de individuos de la misma población, indicó la consistencia de este perfil químico. En los cromatogramas de ambas poblaciones, se observaron bandas de coloración azul que han sido descritas en otros estudios de diferentes poblaciones de *Galphimia* spp. (Sharma et al., 2012b; Gesto-Borroto et al., 2019; Balderas et al., 2020). Estos compuestos azules no han sido relacionados con las actividades ansiolítica y sedante.

4.2.3. Cromatografía líquida de alta eficiencia

En el cromatograma correspondiente a la fracción de galfiminas puras se observaron picos en los tiempos de elución 5.5 y 6.5 min (Figura 12). Según lo descrito por Sharma et al. (2012a), estos se enumeran como II y V y se corresponden el II con las galfiminas B y F y el V con las galfiminas E y G. En correspondencia con lo identificado en la fracción de galfiminas puras, en el extracto perteneciente a la población productora se identificaron las galfiminas B,

F, E y G. En población no productora, y en las raíces de las plantas productoras no se detectaron picos que indiquen la presencia de galfiminas.

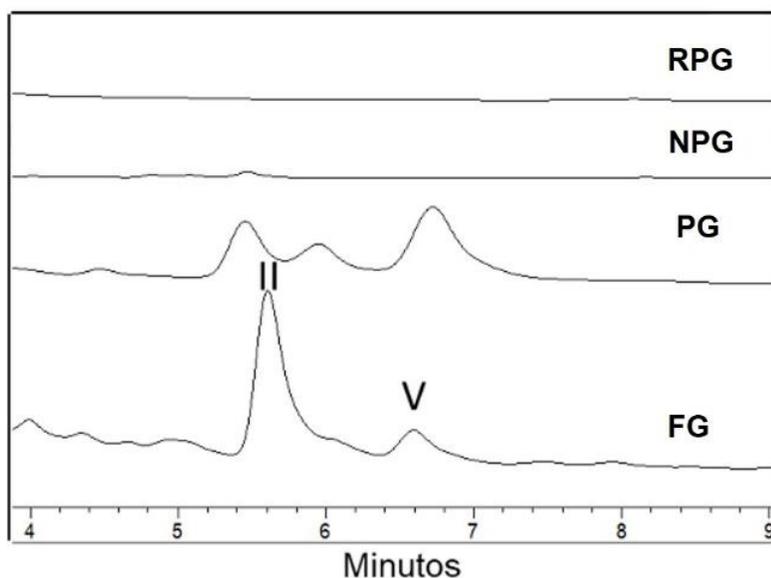


Figura 12. Separación de galfiminas por CLAE en extractos metanólicos de dos poblaciones de *Galphimia* spp. FG (fracción de galfiminas), PG (plantas productoras de galfiminas), NPG (plantas no productoras de galfiminas), RPG (raíces de la población productora).

El análisis mediante CLAE confirmó la presencia de galfiminas solo en la población de Guanajuato. La abundancia de este compuesto observada en el cromatograma resultó diferente a la obtenida en estudios anteriores como los de Cardoso-Taketa *et al.* (2008), Sharma *et al.* (2012a) y Balderas *et al.* (2020), en los cuales se analizó la misma población en condiciones experimentales similares. Estas diferencias pudieron deberse a la variación en la producción de triterpenos por mecanismos de control genético, así como a factores ambientales que pudieron influir sobre estas plantas como temperatura, humedad e iluminación. A diferencia de los estudios de Cardoso-Taketa *et al.* (2008), Sharma *et al.* (2012a) y Balderas *et al.* (2020), las plantas a partir de las cuales se obtuvieron los extractos en este trabajo crecieron en condiciones semicontroladas de temperatura y humedad.

Mediante la separación cromatográfica se determinó la presencia de galfiminas en las plantas de la población productora. De esta forma se demostró que la síntesis

de este metabolito secundario no está estrictamente condicionada por factores ambientales, debido a que las plantas producen el compuesto fuera de su hábitat natural. No obstante, se observaron variaciones en la abundancia de galfiminas entre las plantas provenientes de su hábitat natural en las cual fue mayor el contenido de este metabolito (Cardoso-Taketa et al., 2008; Sharma et al., 2012a) y las obtenidas en condiciones semicontroladas de temperatura y humedad,

4.2.4. Secuenciación de transcriptomas de *Galphimia* spp. mediante RNAseq con el uso de Illumina Next Seq 550

4.2.5. Extracción de RNA

Se obtuvieron seis muestras de RNA de 50 uL cada una. Dos correspondientes a hojas de plantas PG, dos correspondientes a hojas de plantas NPG y dos de raíces de plantas PG.

4.2.6. Concentración, calidad e integridad del RNA extraído

Los análisis espectrofotométricos y fluorométricos mostraron que se obtuvieron muestras con concentraciones de RNA y calidades adecuadas (Cuadro 3), pues los valores de absorbancia en las diferentes longitudes de onda del espectro fueron superiores a 1.70. En las muestras de hojas se obtuvieron los mayores valores de absorbancia y concentración de RNA con respecto a las muestras de raíces. La mayor concentración de RNA se obtuvo en las hojas de las plantas PG.

Cuadro 3. Análisis espectrofotométrico y fluorométrico de muestras de RNA de hojas y raíces de plantas de *Galphimia* spp.

Muestras	Espectrofotómetro		Fluorómetro c(µg/mL)
	Abs 260/280	Abs 260/230	
Hojas de la planta productora 1	1.91	1.78	1430
Hojas de la planta productora 2	2.08	2.09	1850
Hojas de la planta no productora 1	2.00	1.84	1320
Hojas de la planta no productora 2	2.00	1.77	962
Raíces de la planta productora 1	1.85	1.85	267
Raíces de la planta productora 1	1.79	1.70	291

En la electroforesis en gel de agarosa se observaron bandas correspondientes con el RNA 28S y 18S según lo reportado por el marcador de peso molecular (Figura

13). Las bandas fueron de elevada intensidad en las muestras de hojas de ambas poblaciones, lo cual es un indicador de la integridad del RNA. Sin embargo, en las muestras de raíces solamente se observaron bandas de muy baja intensidad que se corresponden con el RNA 28S.

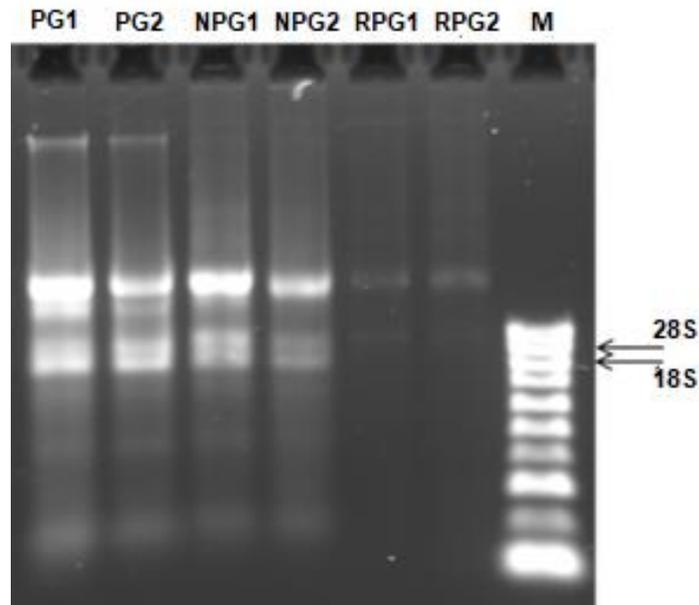


Figura 13. Electroforesis en gel de agarosa del RNA extraído de hojas y raíces de plantas de dos poblaciones de *Galphimia* spp. PG1, PG2 (Hojas de plantas productoras de galfiminas), NPG1, NPG2 (Hojas de plantas no productoras de galfiminas), RPG1, RPG2 (Raíces de plantas productoras de galfiminas), M (marcador de peso molecular).

Mediante la electroforesis capilar se comprobó que el RNA extraído de la mayoría de las muestras constaba de buena calidad, debido a que los valores de RQN fueron superiores a 7 (Figura 14). Los picos correspondientes al RNA 18S y 28S se observaron en la mayoría de los electroferogramas con una definición adecuada, lo que indica que no existió degradación. Sin embargo, en la muestra 1 de raíz se observaron picos adicionales a los RNA 18S y 28S, así como un menor valor de RQN, lo cual es un indicador de degradación y baja integridad.

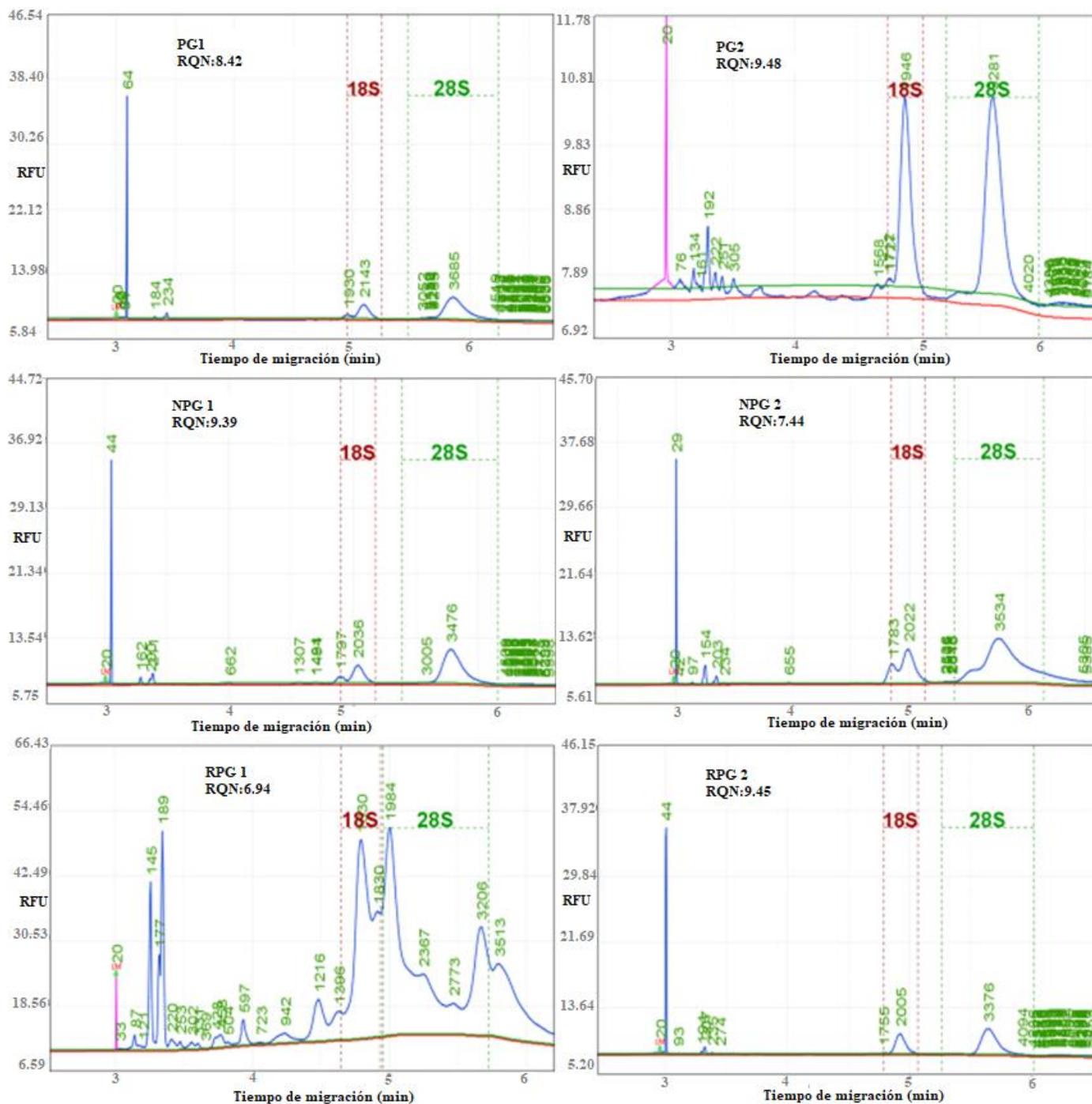


Figura 14. Electroferogramas de muestras de RNA de hojas y raíces de *Galphimia* spp. PG1, PG2 (hojas de plantas productoras de galfiminas), NPG1, NPG2 (hojas de plantas no productoras de galfiminas), RPG1, RPG (raíces de plantas productoras de galfiminas). RQN (Número de calidad de RNA).

La extracción de ácidos nucleicos constituye un punto crítico en la construcción de librerías para secuenciación. En las muestras de ambas poblaciones de *Galphimia* spp. las concentraciones de RNA obtenidas fueron adecuadas para la construcción de librerías según el protocolo TruSeq Stranded Total RNA de Illumina. La calidad e integridad de las muestras de hojas fue elevada, según lo indicado por los valores de RQN y absorbancia. Sin embargo, en las muestras de raíces la calidad fue menor, lo cual sugiere contaminación del RNA con otros metabolitos como polisacáridos, polifenoles o proteínas (Chang et al., 2016).

La pureza del RNA extraído puede depender de la cantidad de metabolitos secundarios producidos por la planta (Yockteng et al., 2013). En el caso de las especies con un elevado contenido de polisacáridos y polifenoles, la extracción de RNA de alta calidad puede resultar difícil debido a que las propiedades físicas y químicas de los polisacáridos y polifenoles son similares a las del RNA, por lo que la co-precipitación de estos se produce fácilmente y con ello la calidad del RNA se ve afectada (Liu et al., 2018). Además, este tipo de contaminación puede ocurrir con frecuencia cuando se emplean kits de extracción, pues los carbohidratos de la muestra pueden establecer interacciones hidrofóbicas con las membranas de separación, la cual posee los grupos oligo dT necesarios para la captura de las secuencias terminales polyA del RNA que se va a extraer (Sandoval Pineda et al., 2017).

4.2.7. Obtención de librerías

Se obtuvieron librerías de cDNA con la calidad requerida, según lo reportado por el protocolo TruSeq Stranded Total RNA de Illumina. El análisis de fluorescencia en el bioanalizador mostró que los tamaños de los fragmentos de cDNA fueron adecuados, pues el valor de unidades de fluorescencia relativa (RFU) correspondientes a estos fragmentos, se encontró representado por el pico de mayor tamaño en los electroferogramas (Figura 15). El tamaño de los fragmentos de cDNA fue homogéneo para todas las muestras, lo cual indica que el proceso de purificación fue adecuado. La concentración de cDNA fluctuó entre 7 y 9 ng/ μ l en

las muestras de hojas. De las librerías obtenidas a partir de RNA de raíces se obtuvieron concentraciones de cDNA menores.

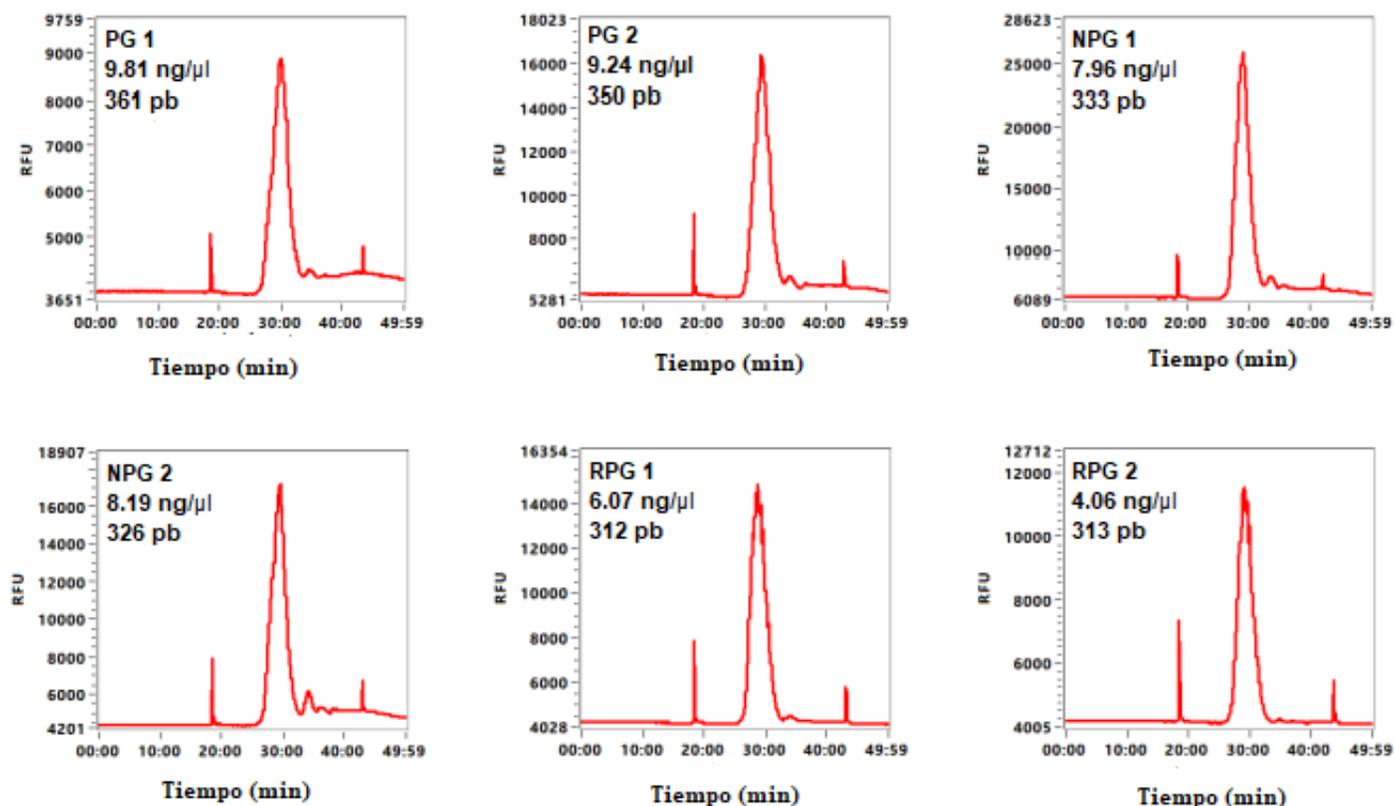


Figura 15. Electroferogramas de librerías de cDNA de *Galphimia* spp. PG1, PG2 (Hojas de plantas productoras de galfiminas), NPG1, NPG2 (Hojas de plantas no productoras de galfiminas), RPG1, RPG2 (Raíces de plantas productoras de galfiminas).

Aunque existieron diferencias entre las concentraciones y calidad de las muestras de RNA, se obtuvieron librerías de cDNA con tamaños de fragmentos homogéneos. En la obtención de este resultado influyeron factores como: el trabajo con una concentración de RNA inicial igual para todas las muestras y una correcta manipulación de las perlas magnéticas, específicamente durante la resuspensión y el secado de estas. Además, el manejo adecuado de los reactivos, así como el cuidado en los cambios de temperaturas a las que fueron sometidas las muestras durante todo el proceso. De esta forma cuando se evaluó la calidad de las librerías

en el bioanalizador no se observó contaminación con dímeros de primers, sobre amplificación de estas o la presencia de picos adicionales en los electroferogramas.

4.2.8. Secuenciación de librerías

Durante la secuenciación de las librerías el llamado de bases resultó exitoso pues la intensidad de aparición de las bases en cada ciclo fue homogénea. Además, se obtuvo un Q30 de 97.2, por lo que la probabilidad de obtener un error fue mínima. Se obtuvieron lecturas de extremos pareados en formato fastq.

Durante la secuenciación de las librerías la densidad homogénea de las muestras, así como la tasa de error mínima que se obtuvo durante el llamado de bases fueron indicadores de la calidad de este proceso. La obtención de lecturas de extremos pareados permitió la generación de datos de secuencias alineables de alta calidad, facilitó la detección de reordenamientos genómicos y elementos de secuencias repetitivas, fusiones de genes y transcripciones novedosas. Además, una alineación de lecturas más precisas y la capacidad de detectar variantes de inserción-delección. Este tipo de lecturas es recomendable cuando no existe un genoma de referencia de la especie debido a que se conserva información sobre la direccionalidad de la transcripción (www.illumina.com). La programación de 76 x 2 ciclos de secuenciación resultó de utilidad para la posterior obtención de un perfil de genes.

4.3. Análisis de los transcriptomas de *Galphimia* spp. mediante el uso de softwares especializados

4.3.1. Análisis de la calidad de las lecturas, depuración y ensamblado *de novo* de los transcriptomas

Después del análisis en el FastQC se comprobó que la calidad de bases por secuencia fue óptima, pues todas las lecturas correspondientes a cada una de las librerías secuenciadas se encontraban en la zona verde del gráfico. Sin embargo, algunas de las métricas evaluadas por este software como el contenido de bases por secuencias, el porcentaje de Guanina/Citocina y los niveles de duplicación presentaron variaciones. Después de la depuración de las lecturas con el Trimmomatic, estas métricas alcanzaron los valores estándares.

En el genoma de las plantas las secuencias repetidas pueden representar del 90-95% del DNA nuclear. Algunos DNA repetitivos codifican para el RNA ribosómico (ARNr). Otros no codifican las proteínas requeridas para el funcionamiento de la planta, pero si otras proteínas necesarias para su propia propagación en el genoma. Estas secuencias son importantes además para la estabilización y mantenimiento de la estructura del cromosoma pues participan en el reconocimiento y la disyunción cromosómica durante la mitosis y la meiosis (Shcherban, 2015). Las variaciones en los niveles de duplicación y el número de secuencias repetidas halladas en lecturas obtenidas pueden deberse además a que durante la secuenciación de transcriptomas los resultados dependen de los niveles de expresión de los transcritos que se obtienen, los cuales están sujetos a procesos endógenos o exógenos que estén afectando a la planta cuando se realiza la extracción de RNA (University Michigan State, 2020).

Luego del ensamblado de los transcriptomas se obtuvieron en las muestras de hojas aproximadamente 9000 transcritos y en las de raíces 5000 (Cuadro 4). El número de alineamientos obtenidos entre las lecturas crudas de secuenciación y los transcriptomas fue un indicador de la calidad del ensamblado *de novo*. Los porcentajes de alineamientos únicos fueron superiores al 50%. Las muestras correspondientes al control mostraron el mayor y el menor valor de alineamientos únicos. Sin embargo, en todas las muestras evaluadas las tasas de alineamientos generales fueron superiores al 97%. La tasa mayor de alineamiento correspondió a una de las muestras de la planta PG. Por otra parte, el tamaño promedio de los contigs que cubren la mitad del ensamblado (N50) osciló entre 264 y 375. Esta métrica fue menor en los controles y mayor en las plantas PG.

Cuadro 4. Medidas de calidad del ensamblado *de novo* de los transcriptomas de plantas de *Galphimia* spp.

Muestras	Total de lecturas	Número de contigs	Alineamientos únicos (%)	Tasa de alineamiento general	N50
Hojas de planta productora de galfiminas 1	30550783	9023	83.76	99.05	366
Hojas de planta productora de galfiminas 2	45139986	8915	79.02	97.89	375
Hojas de planta no productora de galfiminas 1	15130006	8667	89.35	98.05	332
Hojas de planta no productora de galfiminas 2	53443674	8896	92.58	98.23	355
Raíces de plantas productoras de galfiminas 1	20547069	5004	95.36	98.84	291
Raíces de plantas productoras de galfiminas 2	16305468	4667	50.36	98.78	264

El ensamblado de los transcriptomas de *Galphimia* spp. constó de excelente calidad debido a que las tasas de mapeo general entre los transcriptomas y las lecturas obtenidas de la secuenciación fueron elevadas en todas las muestras analizadas. Estos resultados fueron similares a los obtenidos con el transcriptoma de *Cucumis sativus* donde la tasa de mapeo fue de 96-98% (Slavokhotova et al., 2021). Sin embargo, al analizar el transcriptoma de *Camellia sinensis* y las lecturas de secuenciación, se obtuvieron tasas de mapeo de 79-82%. No obstante a este resultado, el transcriptoma ensamblado de *Camellia sinensis* proporcionó resultados valiosos de anotación funcional, así como la identificación de genes de interés (Liu et al., 2020).

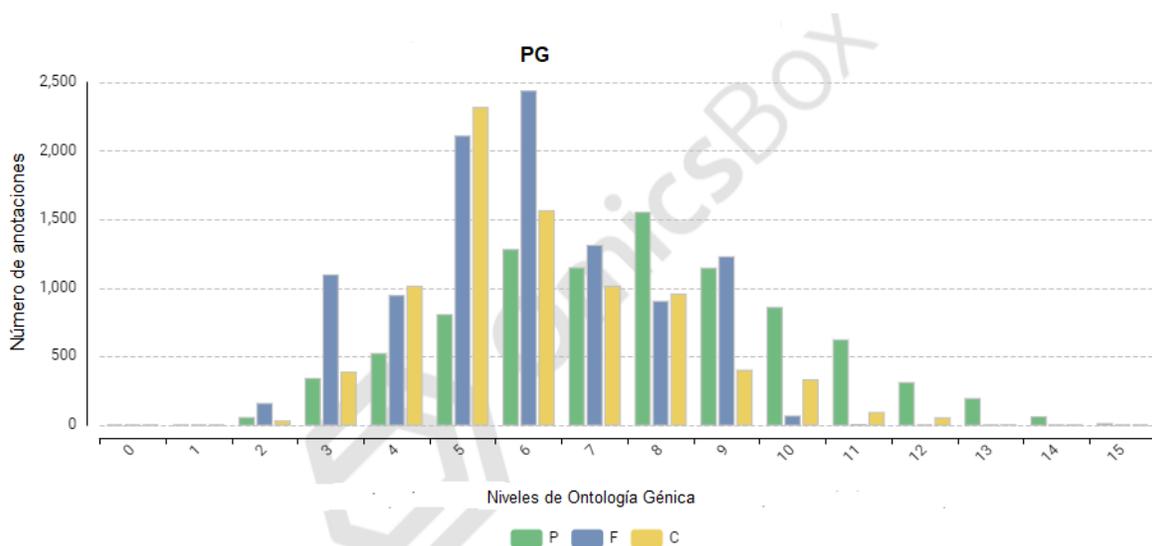
Al ensamblar los datos de un transcriptoma, generalmente se espera obtener numerosos gráficos individuales sin relación entre ellos, los cuales representan la complejidad transcripcional en loci no superpuestos. El ensamblador Trinity usado en este estudio divide los datos de secuenciación en estos gráficos individuales, y luego procesa cada gráfico de forma independiente para extraer isoformas y separar transcripciones derivadas de genes parálogos. Su efectividad se debe a que consta de tres componentes: *Inchworm* que elimina los *k-mer* erróneos de las secuencias de lectura y luego utiliza un método de superposición basado en extensiones codiciosas para ensamblar lecturas en contigs. *Chrysalis* que agrupa los contigs y

construye un gráfico de *Bruijn* para cada grupo y *Butterfly* que recorre los gráficos para construir transcritos (Grabherr et al., 2013).

El N50 en todos los transcriptomas ensamblados fue bajo, lo cual resultó diferente a otros estudios de transcriptómica en plantas como los de *Camellia sinensis* (Liu et al., 2016), *Andrographis paniculata* (Cherukupalli et al., 2016) y *Euphorbia jolkini* (Roy et al., 2020), los cuales fueron ensamblados con Trinity y se obtuvieron valores de N50 de 1123, 1880, 1460 respectivamente. Erróneamente se plantea que un valor de N50 elevado, es un indicador de buena calidad del ensamblado porque se superponen una mayor cantidad de lecturas para formar contigs más largos. A diferencia del ensamblaje del genoma, el transcriptoma incluye transcritos con longitudes variadas. Los contigs más largos pueden generar productos génicos adecuados, pero también podrían representar contigs quiméricos o sobre ensamblados por lo que las métricas basadas en la longitud no siempre son útiles como medidas de precisión (Neil y Emrich, 2013).

4.3.2. Anotación funcional de los transcriptomas

En todos los transcriptomas, más del 90% de las secuencias fueron alineadas por Blast. En los niveles de OG se anotaron 5900 secuencias en la planta PG, 3600 en la planta NPG y 1500 en el control. En la población PG y el control la mayoría de las transcritos se agruparon en el nivel FM y CC (Figura 16). En la población NPG la mayoría de los transcritos se agruparon en FM, seguido de PB.



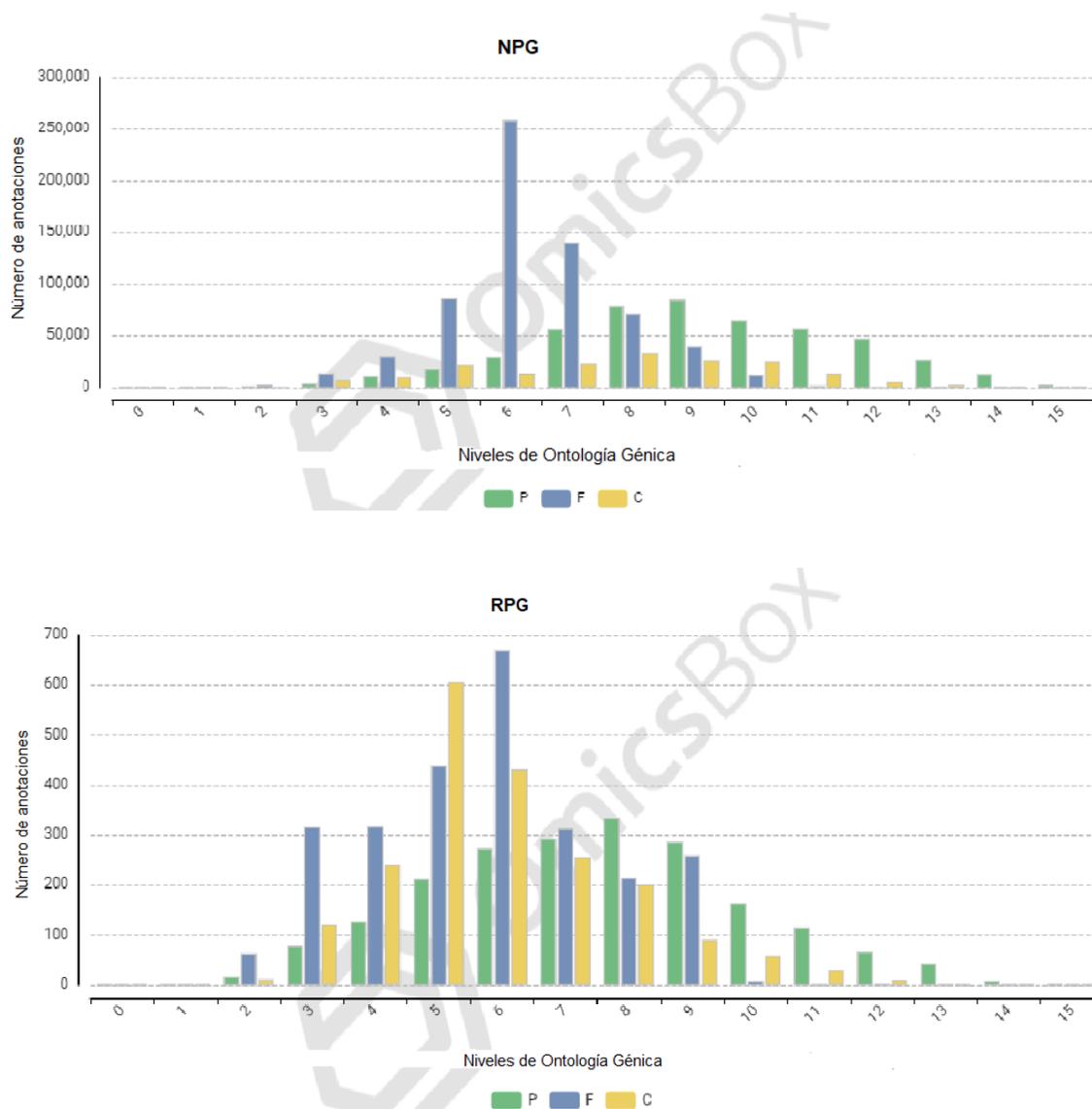


Figura 16. Niveles de distribución de ontología génica de plantas de dos poblaciones de *Galphimia* spp. P (Proceso celular), F (Función molecular), C (Componente celular). PG (hojas de planta productora de galfiminas), NPG (hojas de plantas no productoras de galfiminas), RPG (raíces de planta productora de galfiminas).

En el nivel PB la mayor cantidad de transcritos identificados se asociaron con procesos celulares, así como procesos y regulación biológicos (Figura 17). Con respecto al nivel FM, la mayoría de las transcritos anotados se relacionaron actividad de catálisis, unión a otras moléculas y transporte. En CC los transcritos de

las tres muestras se asociaron a entidades anatómicas celulares y complejos asociados a proteínas.

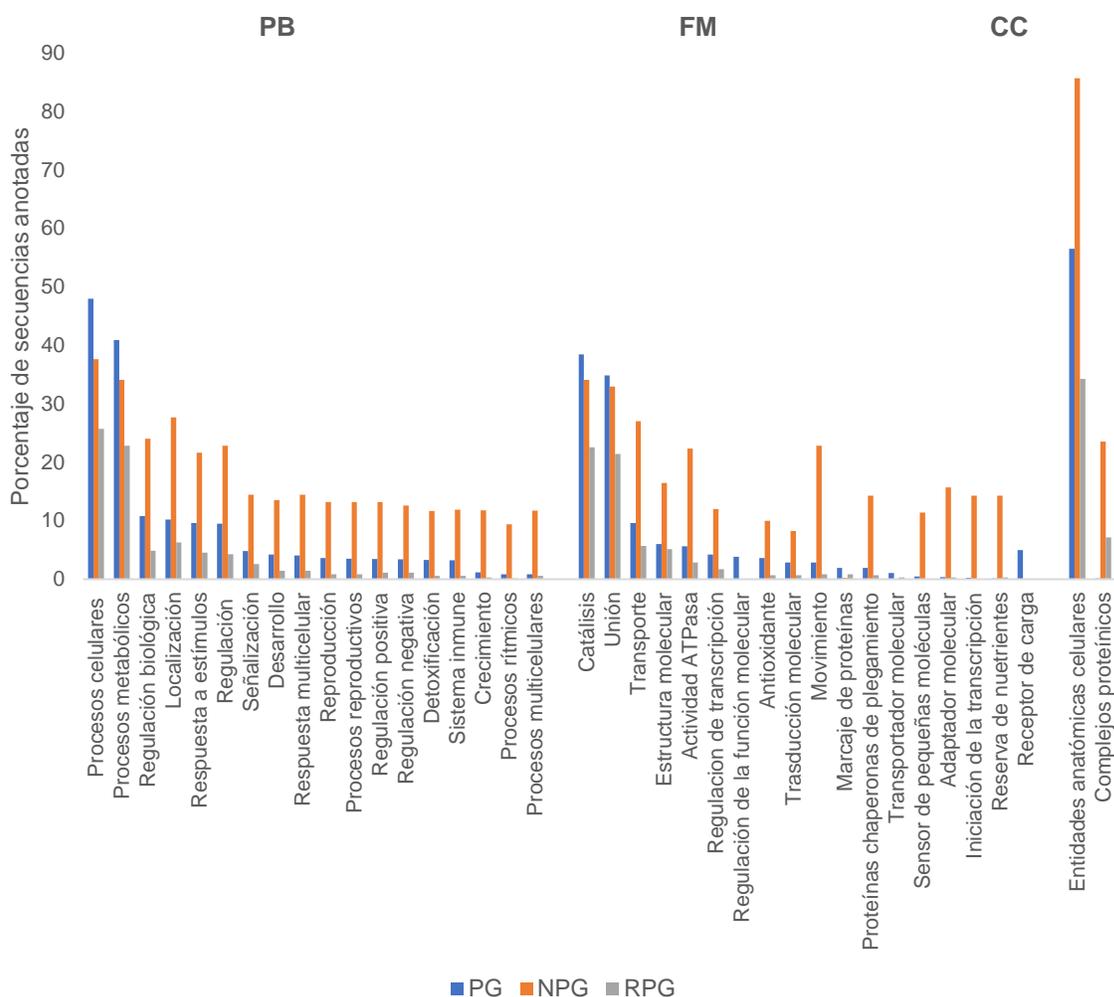


Figura 17. Anotación de secuencias de dos poblaciones de *Galphimia* spp. en los niveles de ontología génica Procesos Biológicos (PB), Función Molecular (FM) y Componente Celular (CC). PG (hojas de planta productora de galfiminas), NPG (hojas de plantas no productoras de galfiminas), RPG (raíces de plantas productoras de galfiminas).

De acuerdo con el código de distribución de enzimas, en la población PG se identificaron la mayor parte de los transcritos asociados a actividad transferasa, hidrolasa y oxidorreductasa (Figura 18). En la población NPG, fue mayor el número de transcritos con actividad oxidorreductasa, transferasa y translocasa. En el

control, los transcritos se relacionaron principalmente con actividad hidrolasa, transferasa y translocasa.

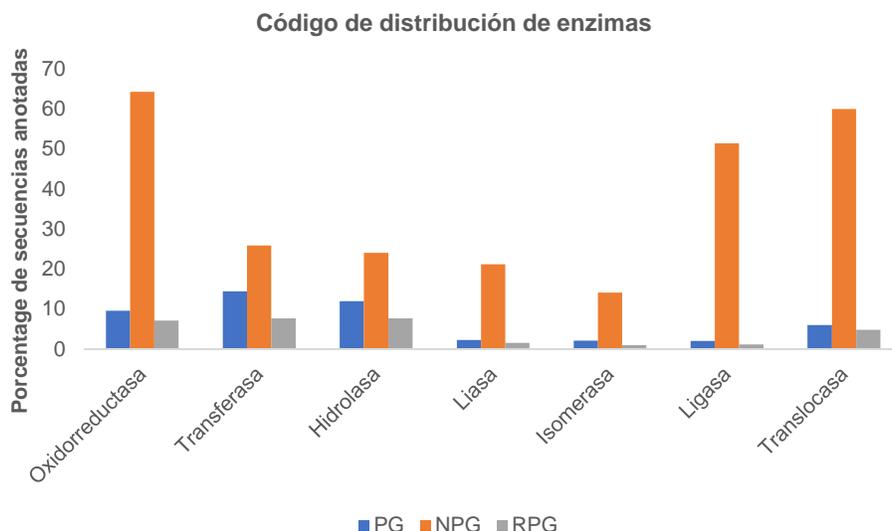


Figura 18. Código de distribución de enzimas de secuencias de dos poblaciones de *Galphimia* spp. PG (hojas de planta productora de galfiminas), NPG (hojas de plantas no productoras de galfiminas), RPG (raíces de plantas productoras de galfiminas).

La anotación en términos de OG proporciona una indicación preliminar de la naturaleza de un producto génico. En este estudio, la mayoría de los transcritos se registraron en los niveles de ontología génica FM y CC, lo que proporcionó una aproximación a las funciones elementales en las que participan los transcritos, así como las partes de las células o el entorno extracelular en el que se encuentran. Si se tiene en cuenta que las galfiminas son triterpenoides de estructura pentacíclica y que la síntesis de estos metabolitos incluye oxidación y adición de grupos funcionales, resulta de utilidad que la mayoría de los transcritos identificados se encuentren asociados con actividades transferasa y oxidorreductasa pues la posibilidad de hallar enzimas relacionadas con su ruta de síntesis es mayor.

4.3.3. Expresión diferencial

Al comparar la población PG con la población NPG, se obtuvieron 680 transcritos expresados diferencialmente. De estos, 313 con regulación positiva y 367 con

regulación negativa. En el gráfico del volcán, se observó que tanto los transcritos regulados positivamente (representados por puntos rojos) como los regulados negativamente (representados por puntos verdes) tienen un logFC de gran magnitud, debido a que los puntos están separados a la derecha o la izquierda del valor 0. Además, algunos de los transcritos con expresión diferencial mostraron un alto nivel de significación estadística, pues varios puntos se localizaron en la parte superior con valores de $-\log_{10}(\text{PValue})$ mayores que 100 (Figura 19a). En la comparación de plantas de la población PG y el control, los transcritos regulados positivamente presentaron una mayor significación estadística (Figura 19b).

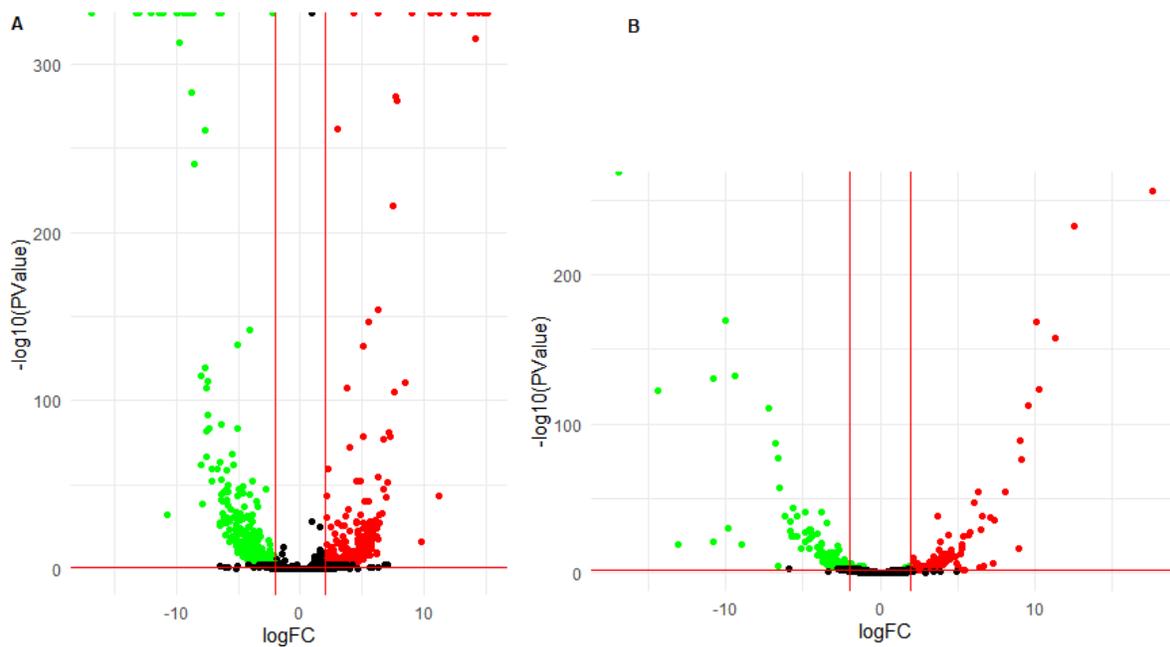


Figura 19. Expresión diferencial de transcritos de plantas de dos poblaciones de *Galphimia* spp. A (comparación entre hojas de plantas productoras y no productoras de galfiminas). B (comparación entre hojas y raíces de las plantas productoras de galfiminas). Puntos negros (transcritos sin expresión significativa), puntos verdes (transcritos regulados negativamente), puntos rojos (transcritos regulados positivamente).

En la población PG se observaron un mayor número de transcritos expresados diferencialmente con regulación positiva con respecto a la población NPG y el control (Figura 20). Esto sugiere una mayor actividad metabólica de estas plantas. En la comparación de la población PG y la NPG, los transcritos con mayores valores de expresión en las plantas PG fueron proteínas hipotéticas, proteínas ribosomales, transferasas, endonucleasas y carboxipeptidasas. En las plantas NPG, los

transcritos con mayor expresión codificaron para receptores mitocondriales, proteínas ribosomales, enzimas con actividad sintasa y poligalacturonasas

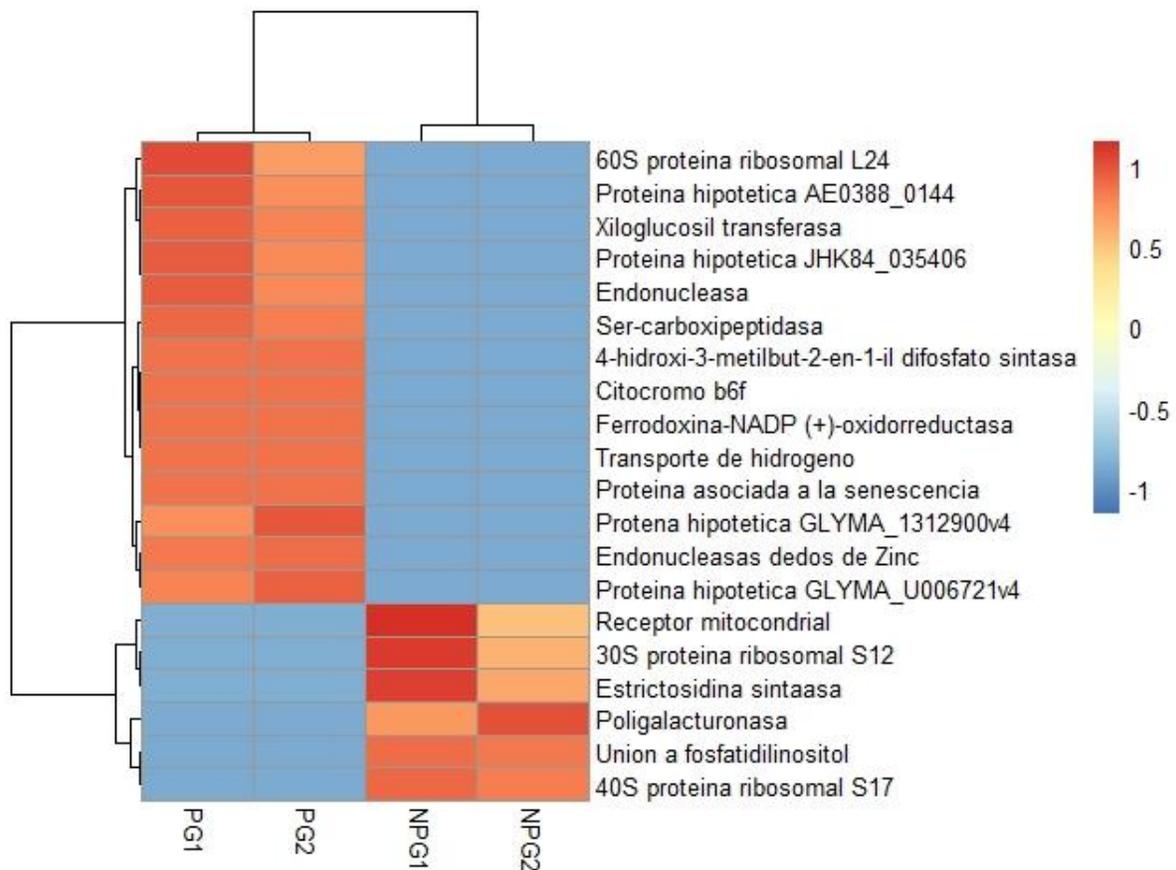


Figura 20. Comparación de transcritos con mayores valores de expresión diferencial en dos poblaciones de *Galphimia* spp. PG1, PG2 (hojas de plantas productoras de galfiminas). NPG1, NPG2 (hojas de plantas no productoras de galfiminas). Los mayores valores están representados en color rojos, los menores en azul.

En la comparación entre la población productora y el control, los transcritos con mayores valores de expresión en las plantas productoras fueron proteínas ribosomales, asociadas a senescencia y enzimas con actividad endonucleasa, polimerasa y deshidrogenasa (Figura 21). En el control los transcritos con mayores valores de expresión codifican para receptores, acuaporinas, chaperonas, endonucleasas y metiltransferasas. Se observó un mayor número de transcritos con regulación positiva en las muestras de hojas con respecto a las raíces.

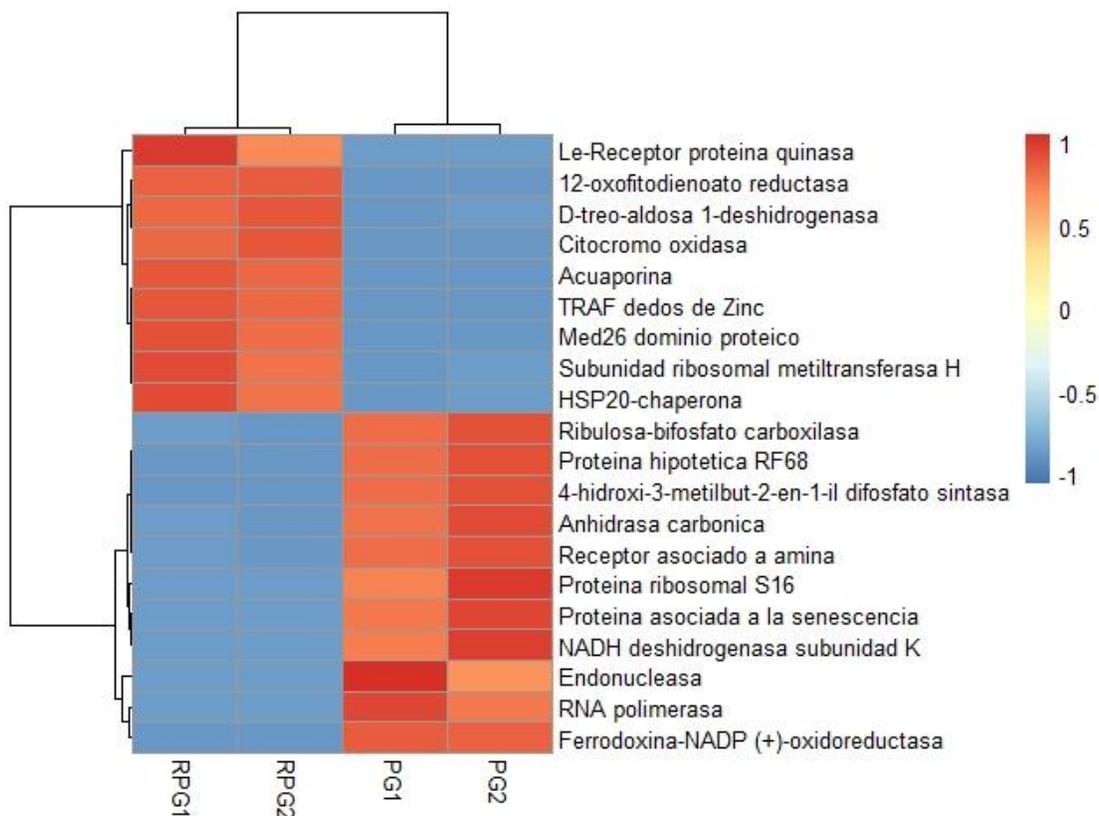


Figura 21. Comparación de transcritos con mayores valores de expresión diferencial en hojas y raíces de la población productora de galfiminas. PG1, PG2 (hojas de plantas productoras de galfiminas). RPG1, RPG2 (raíces de plantas productoras de galfiminas).

De acuerdo con los transcritos identificados que poseían los mayores valores de expresión diferencial en las hojas de las plantas PG y NPG, se puede afirmar que durante la toma de muestras en este órgano existía una elevada actividad fotosintética, respiratoria, de recambio celular, así como de defensa frente a factores externos. Entre los productos génicos identificados, las proteínas ribosomales como S12 han sido relacionadas con plastidios (Wolfe et al., 1994), la L24 con mecanismos de defensa (Park et al., 2001) y síntesis de miRNA y la S17 con procesos de ubiquitinación (Xiong et al., 2021). Las enzimas con actividad endonucleasas pudieran estar relacionadas con procesos como la muerte celular programada a partir de la degeneración del endospermo, la diferenciación de

elementos traqueales, la senescencia y respuestas hipersensibles de la planta (Sugiyama et al., 2000). Resulta pertinente señalar que en la planta PG, la enzima 4-hidroxi-3-metilbut-2-en-1-il difosfato sintasa, que interviene en la síntesis de terpenos (Gutbrod et al., 2019), mostró elevados valores de expresión. En la planta NPG se expresó considerablemente la estrictosidina sintasa que está relacionada con la síntesis de triterpenos índole alcaloides en tejidos meristemáticos activos (Zou et al., 2017).

La comparación entre las hojas y raíces de las plantas PG validó la diferencia funcional entre estos dos órganos. En las hojas se observó una mayor expresión en genes que codifican proteínas relacionadas con fotosíntesis, respiración celular, senescencia, recambio celular y síntesis de RNA. Sin embargo, en las raíces los genes con mayor expresión están relacionados con transporte de agua, síntesis y recambio celular.

4.4. Identificación de los posibles genes candidatos relacionados con la biosíntesis de galfiminas.

En base en los resultados de la anotación funcional, se identificaron 44 transcritos asociados con la síntesis de terpenos en la población PG (Cuadro 5). De estos, dos transcritos que codifican para 3-hidroxi-3-metilglutaril-CoA reductasa, la enzima reguladora de la vía del ácido mevalónico, así como 13 transcritos que codifican enzimas de la vía del fosfato de metileritritol. Además, se identificaron tres transcritos que codifican enzimas asociadas a la síntesis de isoprenoides, uno a la síntesis de terpenos y 25 a la síntesis de triterpenos.

En la población NPG, se identificaron 35 transcritos relacionados con la síntesis de terpenos. Entre estos, uno que codifica para la enzima reguladora de la vía del ácido mevalónico. Relacionado con la vía del metileritritol fosfato, se anotaron 10 transcritos. Además, cinco relacionados con intermediarios en la síntesis de isoprenoides, tres con la síntesis de terpenos y 16 con la síntesis de triterpenos. En el control, se anotó un transcripto que codifica para 3-hidroxi-3-metilglutaril-CoA reductasa, cinco para terpeno sintasa y uno para P450.

En las muestras de ambas poblaciones, se identificaron transcritos que codifican enzimas de la familia del citocromo P450 (CYP) que han sido asociadas a la síntesis de terpenos. En el control solo se identificó un transcripto que codifica para la familia CYP.

Cuadro 5. Anotación y conteo génico de transcritos relacionados con la síntesis de terpenos en dos poblaciones de *Galphimia* spp.

Transcritos anotados	Población productora de galphiminas		Población no productora de galphiminas		Control	
	Número de anotaciones	Conteos génicos	Número de anotaciones	Conteos génicos	Número de anotaciones	Conteos génicos
3-hidroxi-3-metilglutaril-CoA reductasa	2	13	1	8	1	6
1-deoxi-D-xilulosa-5-fosfato sintasa	4	15	4	29	-	-
4-difosfocitidil-2-C-metil-D-eritritol quinasa	4	18	1	10	-	-
4-hidroxi-3-metilbut-2-en-1-il difosfato sintasa	5	19	5	16	-	-
isopentenil difosfato isomerasa	1	52	1	17	-	-
farnesil transferasa	1	6	3	4	-	-
geranil difosfato sintasa	1	14	1	4	-	-
terpeno sintasa	1	347	3	253	5	78
escualeno sintasa	1	7	-	-	-	-
escualeno epoxidasa	1	16	-	-	-	-
beta-amirina sintasa	2	23	4	14	-	-
lanosterol sintasa	7	18	6	23	-	-
P450	2	304.444	-	-	1	4
P450 TBP	19	417.183	-	-	-	-
CYP71A24	1	5	2	3	-	-
CYP71A26	-	-	1	14	-	-
CYP71B10	-	-	1	12	-	-
CYP71B34	1	9	-	-	-	-
CYP71D11	-	-	-	-	1	15
CYP72A	2	10	-	-	-	-
CYP72A15	-	-	1	7	-	-
CYP81D1	1	18	-	-	-	-
CYP82C4	1	231	-	-	-	-
CYP82G1	1	41	1	129	-	-
PLAC8	2	12	-	-	-	-
5'-AMP- proteína quinasa activada	1	30	-	-	-	-
serina/treonina-proteína quinasa AtPK2/AtPK19	1	6	-	-	-	-

*-ausencia del transcripto

La población PG presentó un mayor número y diversidad de transcritos relacionados con la síntesis de terpenos con altos valores de conteos génicos en comparación con la población NPG. Esto sugiere una mayor actividad metabólica en la vía de síntesis de terpenos en estas plantas. En ambas poblaciones analizadas existen transcritos que codifican las vías del ácido mevalónico y del fosfato de metileritritol, así como precursores en la biosíntesis de isoprenoides, terpenos y triterpenos. Sin embargo, en la población PG, el número y diversidad de enzimas identificadas fue mayor. Debido a que los terpenos son un grupo de metabolitos secundarios muy representados en las plantas (Tholl, 2015), es común encontrarlos en las dos poblaciones estudiadas. Según los resultados de este estudio, la población PG presenta una mayor expresión de las enzimas de esta vía y como consecuencia un mayor contenido de estos metabolitos. Esto coincide con los resultados del estudio de Sharma et al. (2012b), donde se analizó por TLC el perfil químico de las poblaciones de *Galphimia* spp utilizadas en este trabajo. Se observaron bandas de coloración violeta que indicaban la presencia de galfiminas en la población productora, así como otras bandas que indicaban la existencia de otros terpenos en la población NPG.

Las enzimas con mayores conteos génicos en la población NPG, fueron terpeno sintasa y CYP82G1. Tanto terpeno sintasa como CYP82G1 están muy representadas en todas las especies de plantas, debido a que diversifican los precursores isoprenoides en una gran cantidad de mono y sesquiterpenos (Booth et al., 2017) y participan en la síntesis de homoterpenos volátiles, respectivamente. Estos homoterpenos se consideran entre los compuestos vegetales más comunes que actúan como parte de los mecanismos de defensa contra los herbívoros (Lee et al., 2010).

En la población PG las enzimas con mayores conteos génicos fueron P450 TBP, P450 y terpeno sintasa. Se ha estudiado que durante la síntesis de triterpenos, P450 cataliza un gran número de oxidaciones regio y estereoespecíficas de los intermediarios de esta ruta (Ghosh, 2017), lo cual ha sido estudiado en varias especies de plantas (Malhotra y Franke, 2022). En hojas de *Nicotiana benthamiana* se realizó un análisis funcional de CYP51 y se demostró que esta enzima es capaz

de catalizar la hidroxilación y la epoxidación del triterpeno simple β -amirina para obtener 12,13 β -epoxi-3 β , 16 β -dihidroxi-oleanano (12, 13 β -epoxi-16 β -hidroxi- β -amirina) (Geisler et al., 2013). Además, se aislaron y caracterizaron tres OSC; TwOSC1, TwOSC2 y TwOSC3 en *Tripterygium wilfordii*. TwOSC1 y TwOSC3 fueron multiproductos de la friedelina sintasa y estaban involucradas en la biosíntesis de celastrol. TwOSC2 era una β -amirina sintasa. Estos autores propusieron una vía biogénica para el celastrol en la que la síntesis de hidroxifriedelanos estaba mediada por P450 (Zhou et al., 2019). Por otro parte, en *Aralia elata*, CYP716A295 y CYP716A296 se identificaron como los genes candidatos con mayor probabilidad de estar asociados con la síntesis de ácido oleanólico (Cheng et al., 2020).

Las galfiminas son triterpenoides modificados, clasificados como nor-secofriedelanos, que se derivan de la friedelina (Han et al., 2019). La vía biosintética para la formación de la friedelina, friedelano pentacíclico oxigenado en C₃, se inicia a través de la protonación de oxidoscualeno, ciclización, numerosos reordenamientos y, finalmente, la desprotonación (Figura 22).

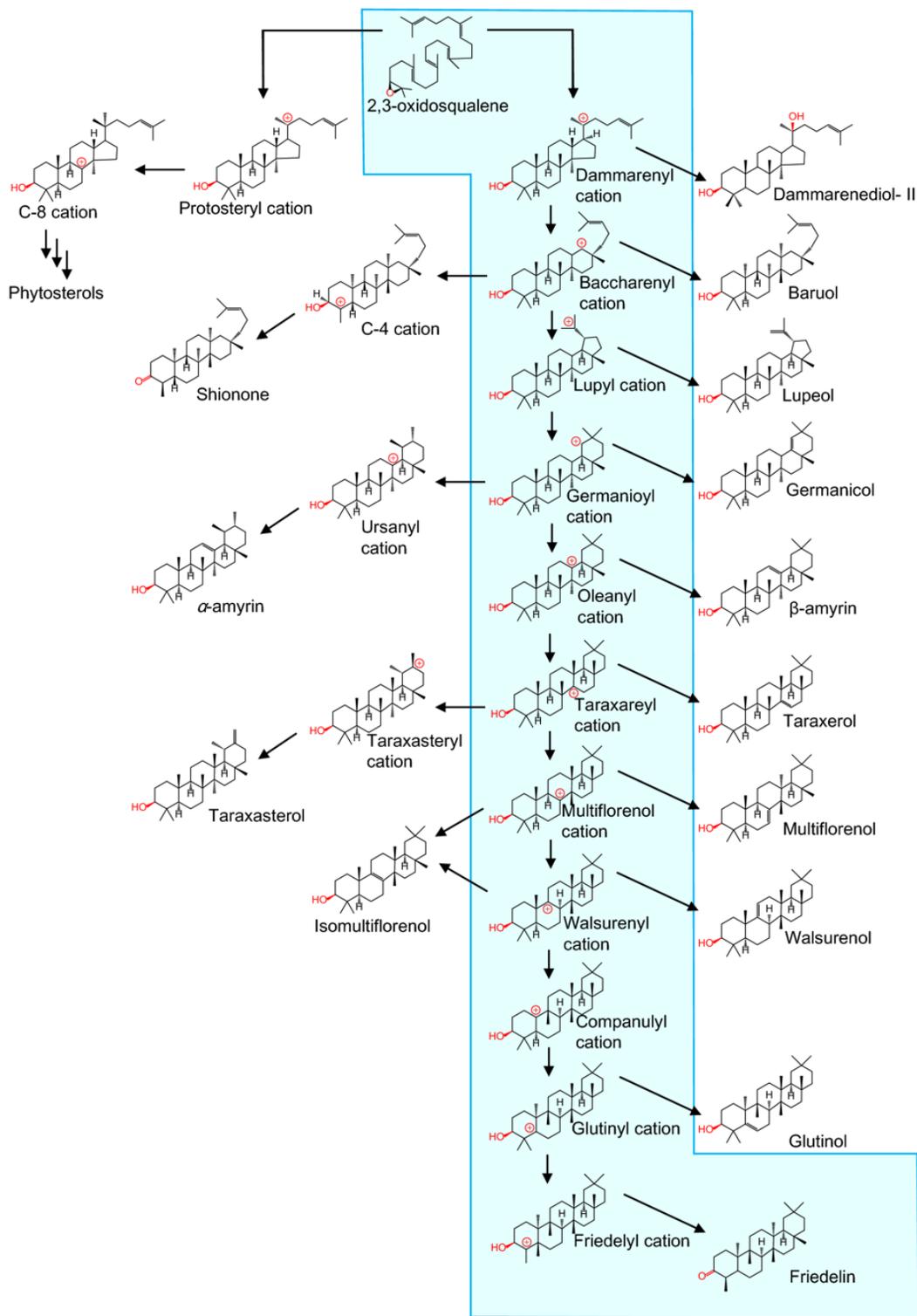


Figura 22. Ruta biosintética de la friedelina pentacíclica a través de la protonación del oxidosqualeno y la formación de cationes intermediarios por reacciones de ciclización, reordenamientos y desprotonación. Tomado de Han et al. (2019).

Enseguida, mediante la participación del citocromo P450, la mayoría de las modificaciones estructurales de la friedelina se producen en el anillo A debido a la facilidad de transformación de sus enlaces, pues el grupo funcional C₃-ceto es el único químicamente activo en la molécula (Das et al., 2018). Los arreglos en los intermediarios entre la friedelina y los norsecofriedelanos implican reacciones de oxidación para la ruptura del anillo A en el enlace C₃-C₄, que resulta en la generación de los dos grupos funcionales necesarios para la formación de la lactona de siete miembros característica de las galfiminas.

Si se consideran los planteamientos de los trabajos sobre el papel de la familia del citocromo P450 en la síntesis de triterpenos y el conteo de genes obtenidos de los transcritos que codifican a miembros de la familia P450 en la población PG (Cuadro 5), así como la ausencia de muchos de estos en la población NPG, estas enzimas podrían considerarse posibles candidatas para la síntesis de galfiminas, como se sugiere en la Figura 23.

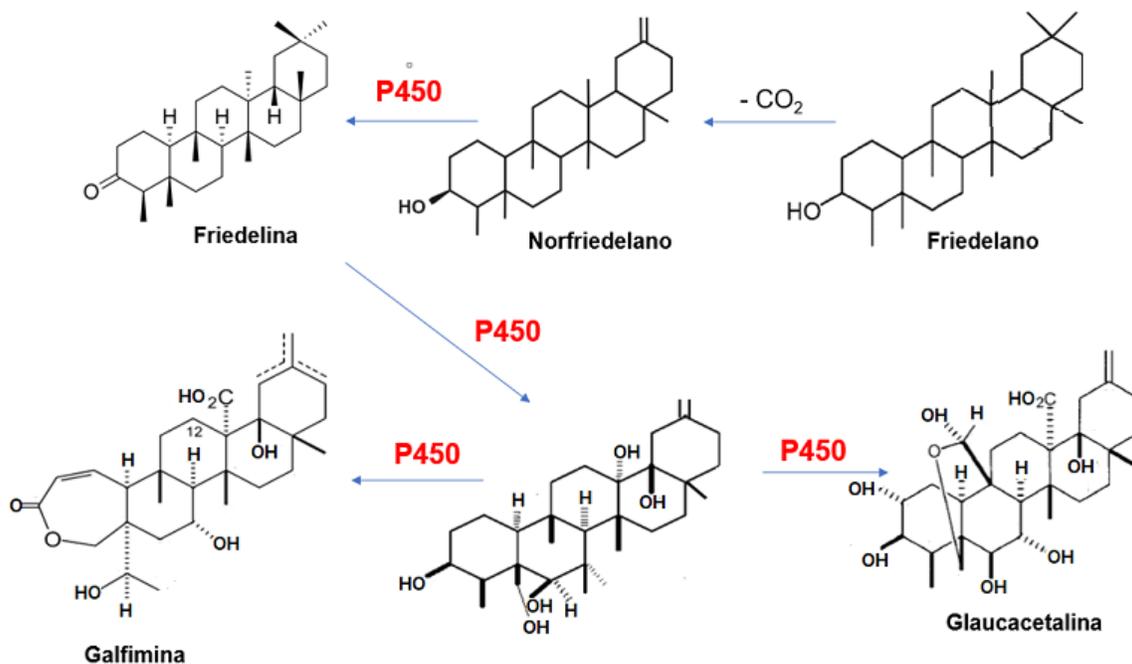


Figura 23. Propuesta de ruta de síntesis de las galfiminas mediada por el citocromo P450 en plantas de *Galphimia* spp.

En esta propuesta de ruta de síntesis de las galfiminas, el citocromo P450 cataliza las reacciones mediante las que se sintetizan los nor-friedelanos (como la friedelina) y los nor-secofriedelanos. P450 es una enzima que luego de ser biosintetizada puede sufrir varias modificaciones postraduccionales mediadas por fosforilación, la quinasa dependiente de AMP (sustrato: Ser, Thr), la quinasa dependiente de fosfolípido activada por Ca^{2+} (sustrato: Arg)], la ubiquitinación, óxido nítrico y glicosilación (Aguiar et al., 2005). Estas modificaciones postraduccionales pudieran ser la causa de que la enzima esté presente en varias de las reacciones de síntesis de los intermediarios en la ruta de las galfiminas. El sitio catalítico de todos los miembros del citocromo P450 se encuentra altamente conservado a lo largo de todo el árbol filogenético de la vida, las modificaciones se encuentran principalmente fuera de esta porción enzimática. Al punto de que una planta y las células hepáticas humanas pueden presentar un citocromo P450 con un 80% de similitud en la secuencia de aminoácidos (Uno et al 2022). Aun cuando su sitio catalítico se encuentre altamente conservado es posible que las enzimas participantes en cada reacción de la ruta de síntesis de las galfiminas posean diferencias a causas de estas modificaciones.

En la red de coexpresión, se observaron 22 nodos (Figura 24). De estos, IPPI (isopentenil difosfato isomerasa), HMGCR (3-hidroxi-3-metilglutaril-CoA reductasa) y FT (farnesil transferasa) mostraron el mayor número de interacciones con otros nodos. La expresión de la mayoría de los miembros de la familia del P450 se encuentran estrechamente relacionada. Las redes de co-expresión son un nuevo recurso para comprender las vías convergentes y sus relaciones, el cual se ha utilizado para identificar grupos de genes que se encuentren relacionados funcionalmente o bajo el mismo patrón de regulación (Hartl, 2021). En este estudio, las enzimas seleccionadas como posibles candidatas para la vía de síntesis de las galfiminas interactuaron estrechamente entre sí, así como con otras enzimas involucradas en etapas previas a la síntesis de triterpenos. Estos resultados sugieren que estos genes participan en la misma ruta metabólica, o están regulados de forma similar. Resulta pertinente resaltar la interacción de la 3-hidroxi-3-metilglutaril-CoA reductasa, con un número considerable de las enzimas

identificadas que están asociadas a la síntesis de triterpenos policíclicos. Esta es la enzima reguladora de la ruta del ácido mevalónico, la cual se ha probado que se encuentra directamente relacionada con la síntesis de triterpenos (Itoh et al., 2003). La 3-hidroxi-3-metilglutaril-CoA reductasa es el centro regulador del metabolismo de los terpenos en el citosol, por lo que modula la producción de los compuestos obtenidos, así como la actividad de los productos génicos implicados en estas reacciones.

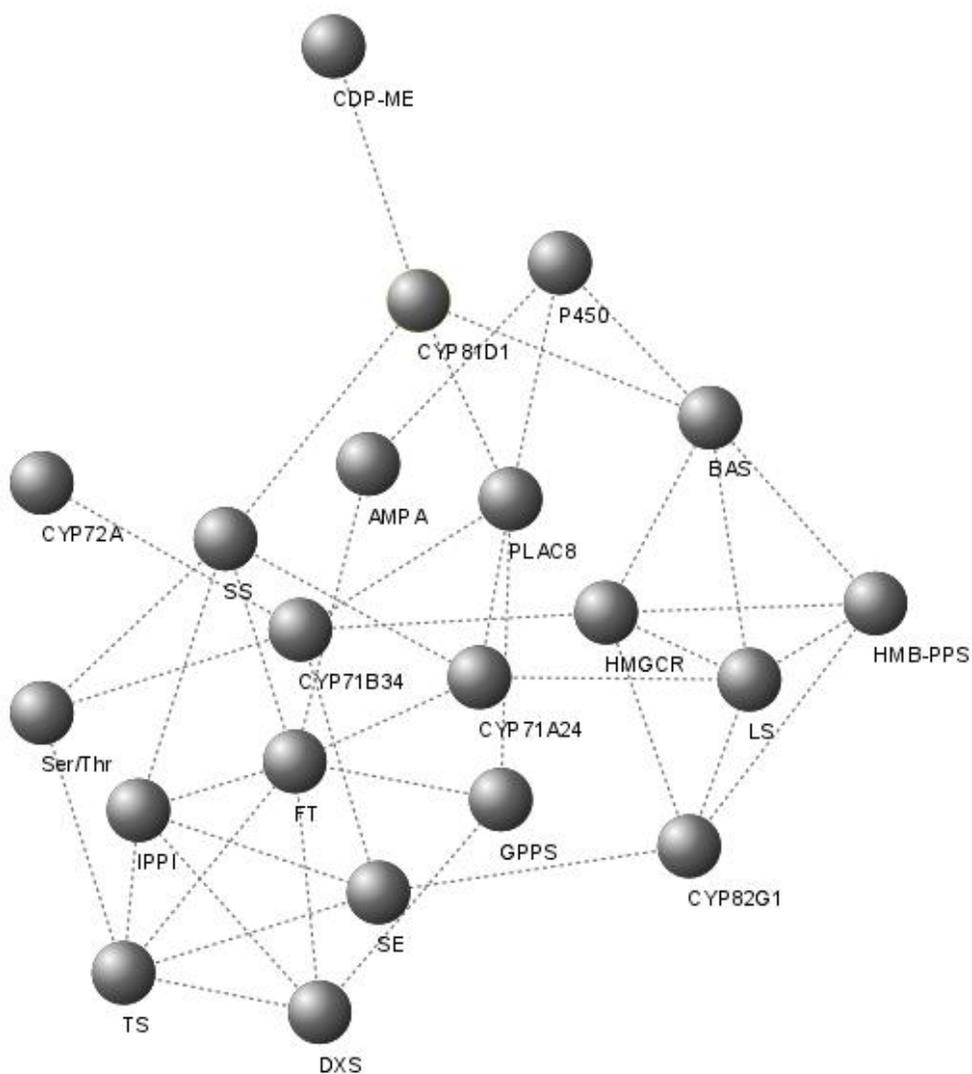


Figura 24. Red de co-expresión de transcritos relacionados con la ruta de síntesis de triterpenos y miembros de la familia del citocromo P450 en las poblaciones productoras y no productoras de

galfiminas. CDP-ME (4-difosfocitidil-2-C-metil-D-eritritol quinasa), TS (terpeno sintasa), GPPS (geranil difosfato sintasa), BAS (beta-amirina sintasa), DXS (1-deoxi-D-xilulosa-5-fosfato sintasa), LS (lanosterol sintasa), HMB-PPS (4-hidroxi-3-metilbut-2-en-1-il difosfato sintasa), HMGCGR (3-hidroxi-3-metilglutaril-CoA reductasa), IPP1 (isopentenil difosfato isomerasa), SE (escualeno epoxidasa), SS (escualeno sintasa), FT (farnesil transferasa), AMPA (5'-AMP- proteína quinasa activada), Ser/Thr (serina/treonina-proteína quinasa AtPK2/AtPK19).

4.5. Validación de los datos del transcriptoma mediante una PCR cuantitativa

Cuando se realizó la transcripción inversa a partir del RNA extraído de las plantas pertenecientes a las dos poblaciones de *Galphimia* spp. se obtuvieron concentraciones de cDNA superiores a las del RNA de entrada (Cuadro 6). Las mayores concentraciones de cDNA corresponden a las plantas PG, lo cual puede deberse a que la concentración del RNA de entrada fue mayor que el de las plantas NPG. Para la realización de la PCR cuantitativa se ajustaron las concentraciones en un rango de 94-120 ng/μL.

Cuadro 6. Transcripción inversa de muestras de hojas de las poblaciones productoras y no productoras de galfiminas de *Galphimia* spp.

Muestra	Concentración de RNA (ng/μL)	Concentración de cDNA (ng/μL)	Concentración final de cDNA (ng/μL)
Planta productora de galfiminas, réplica 1	1469.3	3039.6	100
Planta no productora de galfiminas, réplica 1	767	3099.9	109.3
Planta productora de galfiminas, réplica 2	875.8	2639.1	94
Planta no productora de galfiminas, réplica 2	485	1998.3	96
Planta productora de galfiminas, réplica 3	2030.6	3481.4	110.8
Planta no productora de galfiminas, réplica 3	970.3	2861	119.4

Los resultados del estudio de transcriptómica fueron validados mediante un análisis por PCR cuantitativa. La expresión de los posibles genes candidatos relacionados con la síntesis de galfiminas fue mayor en la población PG, pues los valores de umbral de ciclo (CT) fueron inferiores con respecto a la población NPG (Figura 25). Esto significa que en las muestras de las plantas NPG fue necesario un mayor número de ciclos de amplificación para obtener la señal de expresión de los genes.

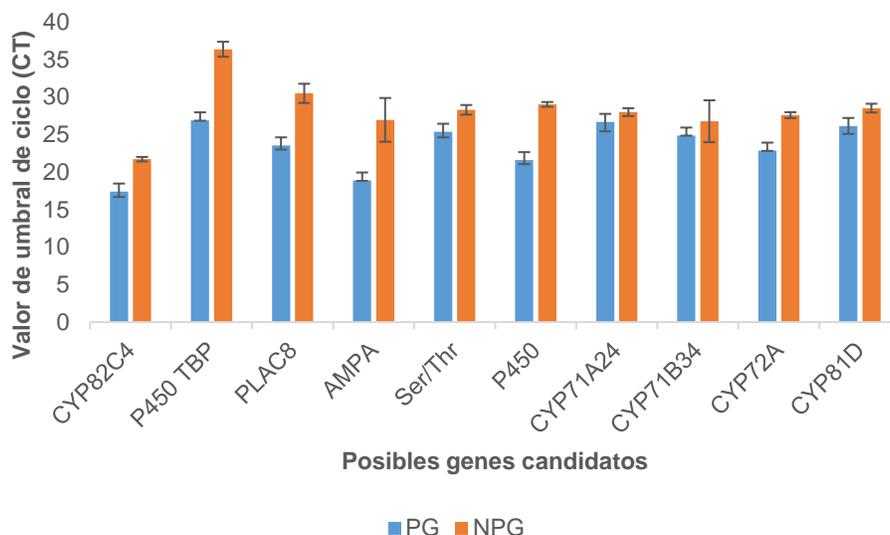


Figura 25. Valores de umbral de ciclo (CT) de muestras de hojas de dos poblaciones de *Galphimia* spp. calculados por una PCR cuantitativa. PG (plantas productoras de galfiminas), NPG (plantas no productoras de galfiminas). AMPA (5'-AMP- proteína quinasa activada), Ser/Thr (serina/treonina-proteína quinasa AtPK2/AtPK19)

Los niveles de expresión de los genes evaluados fueron similares en el análisis de expresión diferencial (Cuadro 5) y en la PCR cuantitativa (Figura 26). Se observaron diferencias significativas entre los valores de Fold Change de ambas poblaciones (Figura 23). En la población PG los genes con mayor expresión relativa fueron P450 TBP, 5'-AMP-proteína quinasa activada y P450. Sin embargo, los genes CYP71B34 y CYP71A24 no se expresaron, lo cual pudo deberse a que están relacionados con la biosíntesis de sesquiterpenos (Liu et al., 2020), la cual pudo no estar activa durante la colecta de las hojas de las plantas usadas en este trabajo. Sin embargo, en la población NPG solo se obtuvieron valores de expresión relativa en los genes CYP71B34, CYP71A24 y CYP81D. De modo general, se observó una mayor expresión de los genes evaluados en la población PG.

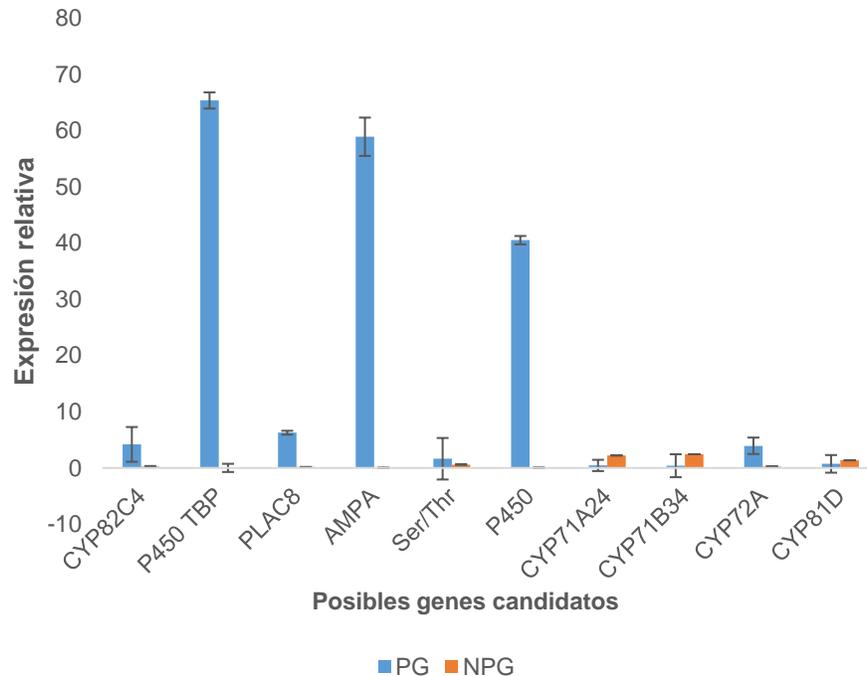


Figura 26. Expresión relativa de posibles genes candidatos en hojas de plantas productoras de galfiminas (PG) y no productoras (NPG) calculada por el método $2^{-\Delta\Delta Ct}$ en una PCR cuantitativa en tiempo real. Eje Y representa los valores de Fold change entre PG y NPG. AMPA (5'-AMP- proteína quinasa activada), Ser/Thr (serina/treonina-proteína quinasa AtPK2/AtPK19)

La transcripción inversa combinada con la PCR cuantitativa en tiempo real constituye una metodología de gran utilidad para cuantificar la expresión de genes. De modo particular, la cuantificación relativa describe los cambios en la expresión de un conjunto de genes seleccionados contra uno o más genes de referencia. Esta es una técnica sensible, específica y con un amplio rango de cuantificación dinámica. Para su ejecución se puede garantizar el equipamiento necesario y los suministros son de bajo costo económico. Además, se puede realizar cuando se tiene una información genómica limitada, su dificultad técnica es baja y las estadísticas que se usan son simples. Los resultados de expresión relativa obtenidos de la PCR cuantitativa en este trabajo validaron el uso de la tecnología de RNA seq para la identificación de genes con una expresión estable en plantas. El valor de los resultados obtenidos al usar esta técnica molecular reside en que su límite de detección y cuantificación es más alto que las ómicas de modo general (Harshitha y Arunraj, 2021).

En plantas la identificación de nuevos genes resulta complejo, debido a la disponibilidad limitada de genomas de referencia. La estrategia más viable y económica para la identificación genética es la transcriptómica. El surgimiento de la NGS, así como de plataformas entre las que destaca Illumina ha constituido un avance significativo en la secuenciación de ácidos nucleicos. Con esta tecnología es posible la secuenciación de millones de lecturas en paralelo, lo que reduce notablemente el esfuerzo y el tiempo. Además, aumenta la calidad de los datos obtenidos en el secuenciador.

En este trabajo, el análisis de seis transcriptomas de dos poblaciones diferentes de *Galphimia* spp. permitió la exploración y el análisis de un conjunto de genes con expresión diferencial, lo cual resultó útil en la identificación de productos génicos de interés, así como para la caracterización de su estructura. El uso de la plataforma Illumina Next seq 550 posibilitó la obtención de lecturas de RNAseq que fueron de utilidad para la identificación de posibles genes candidatos relacionados con la síntesis de galfiminas.

5. Conclusiones

- 1) Se obtuvieron plantas de *Galphimia* spp. con caracteres fenotípicos similares a los descritos para este género.
- 2) Se identificaron galfiminas en los extractos metanólicos de las plantas de la población Dr. Mora Guanajuato.
- 3) Las lecturas obtenidas de la secuenciación fueron de gran calidad, lo que influyó positivamente en el ensamblado *de novo* de los transcriptomas y la anotación funcional.
- 4) La mayoría de los transcritos se agruparon en el nivel de ontología génica función molecular y se relacionaron con actividad transferasa.
- 5) La población productora de galfiminas presentó el mayor número de transcritos expresados diferencialmente con regulación positiva.
- 6) El P450 identificado solamente en la población productora de galfiminas puede ser un posible gen candidato relacionado con la síntesis de estos nor-secotriterpenoides.
- 7) Los resultados del análisis transcriptómico se validaron molecularmente, los genes potencialmente relacionados con la síntesis de galfiminas presentaron mayores niveles de expresión en las plantas productoras.
- 8) La comparación de transcriptomas de una población productora de galfiminas y otra no productora posibilitó la identificación del P450 como posible gen candidato relacionado con la ruta de síntesis de las galfiminas.

6. Perspectivas

- 1) Secuenciar el genoma de *Galphimia* spp. lo que permitirá la obtención de un archivo de anotación de genes más completo para las especies de este género.
- 2) Obtener el proteoma y metaboloma de plantas de la población productora y no productora de galfiminas para validar el análisis transcriptómico en su totalidad.
- 3) Realizar análisis de acoplamiento molecular para explicar posibles interacciones de P450.
- 4) Silenciar el gen que codifica para el P450 en las plantas productoras de galfiminas con el objetivo de comprobar si esta enzima está directamente relacionada con la síntesis de galfiminas.
- 5) Elicitar plantas de la población productora de galfiminas para aumentar la síntesis de este metabolito.

Referencias

- Aguiar, M., Masse, R., & Gibbs, B. F. (2005). Regulation of Cytochrome P450 by Posttranslational Modification. *Drug Metabolism Reviews*, 37(2), 379-404. <https://doi.org/10.1081/dmr-46136>
- Aminfar, Z., Rabiei, B., Tohidfar, M., & Mirjalili, M. H. (2019). Identification of key genes involved in the biosynthesis of triterpenic acids in the mint family. *Scientific Reports*, 9(1), 1-15. <https://doi.org/10.1038/s41598-019-52090-z>
- Ortíz, A., Cardoso-Taketa, A., Rodríguez, M., Arellano, J., Hernández, G., & Villarreal, ML. (2010). Transformed cell suspension culture of *Galphimia glauca* producing sedative nor-friedelanes. *Planta Medica*, 76(4), 386-392. <https://doi.org/10.1055/s-0029-1186166>
- Anamika, K., Verma, S., Jere, A., & Desai, A. (2016). Transcriptomic profiling using Next Generation Sequencing - Advances, advantages, and challenges. In *Next Generation Sequencing - Advances, Applications and Challenges*. <https://doi.org/10.5772/61789>
- Anderson, C. (2007). Revision of *Galphimia* (Malpighiaceae). *Contributions University Michigan Herbarium*, 25, 1-82.
- Ateto, A. A. (2014). *Bioinformatics for Beginners, Genes, Genome, Molecular Evolution, Databases and Analytical Tools*.
- Ayturk, U. (2019). RNA-seq in Skeletal Biology. *Current Osteoporosis Reports*, 17(4), 178-185. <https://doi.org/10.1007/s11914-019-00517-x>
- Balderas, G. S., Alcalá, R. E., Ortíz-Caltempa, A., Cardoso-Taketa, A., & Villarreal, M. L. (2020). Variation in the production of sedative and anxiolytic compounds among *Galphimia* sp. populations grown in a greenhouse. *Revista Brasileira de Farmacognosia*, 30(1), 99-102. <https://doi.org/10.1007/s43450-020-00010-y>
- Bathe, U., & Tissier, A. (2019). Cytochrome P450 enzymes: A driving force of plant diterpene diversity. *Phytochemistry*, 161(October), 149-162. <https://doi.org/10.1016/j.phytochem.2018.12.003>
- Besser, J., Carleton, H. A., Gerner-smidt, P., Lindsey, R. L., & Trees, E. (2018). NGS and their application to the study and control of bacterial Infections. *Clinical Microbiology and Infection*, 24(4), 335-341. <https://doi.org/10.1016/j.cmi.2017.10.013>.Next-Generation
- Boncan, D. A. T., Tsang, S. S. K., Li, C., Lee, I. H. T., Lam, H. M., Chan, T. F., & Hui, J. H. L. (2020). Terpenes and terpenoids in plants: Interactions with environment and insects. *International Journal of Molecular Sciences*, 21(19), 1-19. <https://doi.org/10.3390/ijms21197382>

- Booth, J. K., Page, J. E., & Bohlmann, J. (2017). Terpene synthases from *Cannabis sativa*. *PLoS ONE*, 12(3), 1-20. <https://doi.org/10.1371/journal.pone.0173911>
- Campos, M. G., Toxqui, E., Tortoriello, J., Oropeza, M. V., Ponce, H., Vargas, M. H., & Montaña, L. M. (2001). *Galphimia glauca* organic fraction antagonizes LTD4-induced contraction in guinea pig airways. *Journal of Ethnopharmacology*, 74(1), 7-15. [https://doi.org/10.1016/S0378-8741\(00\)00322-6](https://doi.org/10.1016/S0378-8741(00)00322-6)
- Cardoso-Taketa, A. T., Pereda-Miranda, R., Young, H. C., Verpoorte, R., & Villarreal, M. L. (2008). Metabolic profiling of the Mexican anxiolytic and sedative plant *Galphimia glauca* using nuclear magnetic resonance spectroscopy and multivariate data analysis. *Planta Medica*, 74(10), 1295-1301. <https://doi.org/10.1055/s-2008-1074583>
- Cardoso Taketa, A. T., Lozada-Lechuga, J., Fragoso-Serrano, M., Villarreal, M. L., & Pereda-Miranda, R. (2004). Isolation of Nor-secofriedelanes from the sedative extracts of *Galphimia glauca*. *Journal of Natural Products*, 67(4), 644-649. <https://doi.org/10.1021/np0304666>
- Chang, E., Zhao, Y., Wei, Q., Shi, S., & Jiang, Z. (2016). Isolation of high-quality RNA from *Platycladus orientalis* and other Cupressaceae plants. *Electronic Journal of Biotechnology*, 23, 21-27. <https://doi.org/10.1016/j.ejbt.2016.08.003>
- Chen, S., Cai, X., & Luo, M. (2019). Mechanistic Aspects of Methyltransferases with Small-Molecule Metabolites and Natural Products as Substrates.
- Cheng, Y., Liu, H., Tong, X., Liu, Z., Zhang, X., Li, D., Jiang, X., & Yu, X. (2020). Identification and analysis of CYP450 and UGT supergene family members from the transcriptome of *Aralia elata* (Miq.) seem reveal candidate genes for triterpenoid saponin biosynthesis. *BMC Plant Biology*, 20(1), 1-17. <https://doi.org/10.1186/s12870-020-02411-6>
- Cherukupalli, N., Divate, M., Mittapelli, S. R., Khareedu, V. R., & Vudem, D. R. (2016). *De novo* assembly of leaf transcriptome in the medicinal plant *Andrographis paniculata*. *Frontiers in Plant Science*, 7(AUG2016), 1-13. <https://doi.org/10.3389/fpls.2016.01203>
- Das, J., Sarkar, A., & Ghosh, P. (2018). Friedelane triterpenoids: Transformations toward A-ring modifications including 2-: Homo derivatives. *New Journal of Chemistry*, 42(9), 6673-6688. <https://doi.org/10.1039/c8nj00009c>
- Del Rayo Camacho, M., Phillipson, J. D., Croft, S. L., Marley, D., Kirby, G. C., & Warhurst, D. C. (2002). Assessment of the antiprotozoal activity of *Galphimia glauca* and the isolation of new nor-secofriedelanes and nor-friedelanes. *Journal of Natural Products*, 65(10), 1457-1461. <https://doi.org/10.1021/np010419i>
- Dirihan, S., Terho, P., Helander, M., & Saikkonen, K. (2013). Efficient analysis of ploidy levels in plant

evolutionary ecology. *Caryologia*, 66(3), 251-256.
<https://doi.org/10.1080/00087114.2013.849414>

- Fenn, M. A., & Giovannoni, J. J. (2021). Phytohormones in fruit development and maturation. *Plant Journal*, 105(2), 446-458. <https://doi.org/10.1111/tpj.15112>
- Forestier, E., Romero-Segura, C., Pateraki, I., Centeno, E., Compagnon, V., Preiss, M., Berna, A., Boronat, A., Bach, T. J., Darnet, S., & Schaller, H. (2019). Distinct triterpene synthases in the laticifers of *Euphorbia lathyris*. *Scientific Reports*, 9(1), 1-12. <https://doi.org/10.1038/s41598-019-40905-y>
- Geisler, K., Hughes, R. K., Sainsbury, F., Lomonossoff, G. P., Rejzek, M., Fairhurst, S., Olsen, C. E., Motawia, M. S., Melton, R. E., Hemmings, A. M., Bak, S., & Osbourn, A. (2013). Biochemical analysis of a multifunctional cytochrome P450 (CYP51) enzyme required for synthesis of antimicrobial triterpenes in plants. *Proceedings of the National Academy of Sciences of the United States of America*, 110(35). <https://doi.org/10.1073/pnas.1309157110>
- Gesto-Borroto, R., Cardoso-Taketa, A., Yactayo-Chang, J. P., Medina-Jiménez, K., Hornung-Leoni, C., Lorence, A., & Villarreal, M. L. (2019). DNA barcoding and TLC as tools to properly identify natural populations of the Mexican medicinal species *Galphimia glauca* Cav. *PLoS ONE*, 14(5), 1-18. <https://doi.org/10.1371/journal.pone.0217313>
- Ghosh, S. (2017). Triterpene structural diversification by plant cytochrome P450 enzymes. *Frontiers in Plant Science*, 8(November), 1-15. <https://doi.org/10.3389/fpls.2017.01886>
- González-Cortazar, M., Herrera-Ruiz, M., Zamilpa, A., Jiménez-Ferrer, E., Marquina, S., Álvarez, L., & Tortoriello, J. (2014). Anti-inflammatory activity and chemical profile of *Galphimia glauca*. *Planta Medica*, 80(1), 90-96. <https://doi.org/10.1055/s-0033-1360150>
- González-Cortazar, M., Tortoriello, J., & Alvarez, L. (2005). Norsesquiterpenes as spasmolytics, advances of structure-activity relationships. *Planta Medica*, 71(8), 711-716. <https://doi.org/10.1055/s-2005-871224>
- Goyal, N.Y. (2014). Chemistry of terpenoids. *International Journal of Pharmaceutical Sciences Review and Research*, 27(2), 272-278.
- Grabherr, M. G., Haas, B. J., Moran Yassour, J., Joshua Z., Levin, D.A., Thompson, I. et al., & Friedman, A. R. (2013). Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology*, 29(7), 644-652. <https://doi.org/10.1038/nbt.1883>
- Gupta, N., Verma, V. K. (2019). Next-Generation Sequencing and its application: empowering in public health beyond reality. In P. Arora (Ed.), *Microbial Technology for the Welfare of Society. Microorganisms for Sustainability*. Springer, Singapore. https://doi.org/10.1007/978-981-13-8844-6_15
- Gutbrod, K., Romer, J., & Dörmann, P. (2019). Phytol metabolism in plants. *Progress in Lipid*

Research, 74(2018), 1-17. <https://doi.org/10.1016/j.plipres.2019.01.002>

- Dorsch, W., & Wagner, H. (1991). New antiasthmatic drugs from traditional medicine? *International Archives of Allergy and Immunology*, 94, 262-265. <https://doi.org/10.1159/000235378>
- Han, J. Y., Ahn, C. H., Adhikari, P. B., Kondeti, S., & Choi, Y. E. (2019). Functional characterization of an oxidosqualene cyclase (PdFRS) encoding a monofunctional friedelin synthase in *Populus davidiana*. *Planta*, 249(1), 95-111. <https://doi.org/10.1007/s00425-018-2985-8>
- Hansen, N. L., Miettinen, K., Zhao, Y., Ignea, C., Andreadelli, A., Raadam, M. H., Makris, A. M., Møller, B. L., Stærk, D., Bak, S., & Kampranis, S. C. (2020). Integrating pathway elucidation with yeast engineering to produce polypunonic acid the precursor of the anti-obesity agent celastrol. *Microbial Cell Factories*, 19(1), 1-18. <https://doi.org/10.1186/s12934-020-1284-9>
- Hari, R., & Parthasarathy, S. (2018). Next generation sequencing data analysis. In *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics* (Vols. 1-3). Elsevier Ltd. <https://doi.org/10.1016/B978-0-12-809633-8.20093-9>
- Harshitha, R., & Arunraj, D. R. (2021). Real-time quantitative PCR: A tool for absolute and relative quantification. *Biochemistry and Molecular Biology Education*, 49(5), 800-812. <https://doi.org/10.1002/bmb.21552>
- Hartl, C. L., Ramaswami, G., Pembroke, W. G., Muller, S., Pintacuda, G., Saha, A., & Geschwind, D. H. (2021). Coexpression network architecture reveals the brain-wide and multiregional basis of disease susceptibility. *Nature Neuroscience*, 24(9), 1313-1323. <https://doi.org/10.1038/s41593-021-00887-5>
- Herrera-Arellano, A., Jiménez-Ferrer, E., Zamilpa, A., Morales-Valdéz, M., García-Valencia, C. E., & Tortoriello García, J. (2007). Efficacy and tolerability of a standardized herbal product from *Galphimia glauca* on generalized anxiety disorder. A randomized, double-blind clinical trial controlled with lorazepam. *Planta Medica*, 73(8), 713-717. <https://doi.org/10.1055/s-2007-981539>
- Herrera-Arellano, A., Jiménez-Ferrer, J. E., Zamilpa, A., García-Alonso, G., Herrera-Alvarez, S., & Tortoriello, J. (2012). Therapeutic effectiveness of *Galphimia glauca* vs lorazepam in generalized anxiety disorder. A controlled 15-week clinical trial. *Planta Medica*, 78(14), 1529-1535. <https://doi.org/10.1055/s-0032-1315110>
- Herrera-Ruiz, M., Jiménez-Ferrer, J. E., De Lima, T. C. M., Avilés-Montes, D., Pérez-García, D., González-Cortazar, M., & Tortoriello, J. (2006). Anxiolytic and antidepressant-like activity of a standardized extract from *Galphimia glauca*. *Phytomedicine*, 13(12), 23-28. <https://doi.org/10.1016/j.phymed.2005.03.003>
- Itoh, D., Kawano, K., & Nabeta, K. (2003). Biosynthesis of chloroplastidic and extrachloroplastidic terpenoids in liverwort cultured cells: ¹³C serine as a probe of terpene biosynthesis via

- mevalonate and non-mevalonate pathways. *Journal of Natural Products*, 66(3), 332-336. <https://doi.org/10.1021/np0204141>
- Kebrom, T. H. (2017). A growing stem inhibits bud outgrowth-The overlooked theory of apical dominance. *Frontiers in Plant Science*, 8(October), 1–7. <https://doi.org/10.3389/fpls.2017.01874>
- Kellner, F., Kim, J., Clavijo, B. J., Hamilton, J. P., Childs, K. L., Vaillancourt, B., Cepela, J., Habermann, M., Steuernagel, B., Clissold, L., Mclay, K., Buell, C. R., & O'Connor, S. E. (2015). Genome-guided investigation of plant natural product biosynthesis. *Plant Journal*, 82(4), 680-692. <https://doi.org/10.1111/tpj.12827>
- Knoch, E., Sugawara, S., Mori, T., Poulsen, C., Fukushima, A., Harholt, J., Fujimoto, Y., Umemoto, N., & Saito, K. (2018). Third DWF1 paralog in Solanaceae, sterol $\Delta 24$ -isomerase, branches withanolide biosynthesis from the general phytosterol pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 115(34), E8096-E8103. <https://doi.org/10.1073/pnas.1807482115>
- Lee, S., Badiyan, S., Bevan, D. R., Herde, M., Gatz, C., & Tholl, D. (2010). Herbivore-induced and floral homoterpene volatiles are biosynthesized by a single P450 enzyme (CYP82G1) in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, 107(49), 21205–21210. <https://doi.org/10.1073/pnas.1009975107>
- Liu, L., Han, R., Yu, N., Zhang, W., Xing, L., Xie, D., & Peng, D. (2018). A method for extracting high-quality total RNA from plant rich in polysaccharides and polyphenols using *Dendrobium huoshanense*. *PLoS ONE*, 13(5), 1-9. <https://doi.org/10.1371/journal.pone.0196592>
- Liu, N., Buckley, T. N., He, X., Zhang, X., Zhang, C., Luo, Z., Wang, H., Sterling, N., & Guan, H. (2019). Improvement of a simplified process-based model for estimating transpiration under water-limited conditions. *Hydrological Processes*, 33(12), 1670-1685. <https://doi.org/10.1002/hyp.13430>
- Liu, S. C., Jin, J. Q., Ma, J. Q., Yao, M. Z., Ma, C. L., Li, C. F., Ding, Z. T., & Chen, L. (2016). Transcriptomic analysis of tea plant responding to drought stress and recovery. *PLoS ONE*, 11(1), 1-21. <https://doi.org/10.1371/journal.pone.0147306>
- Liu, X., Zhu, X., Wang, H., Liu, T., Cheng, J., & Jiang, H. (2020). Discovery and modification of cytochrome P450 for plant natural products biosynthesis. *Synthetic and Systems Biotechnology*, 5(3), 187-199. <https://doi.org/10.1016/j.synbio.2020.06.008>
- Liu, Y., Pang, D., Tian, Y., Li, Y., Jiang, H., Sun, Y., Xia, L., & Chen, L. (2020). Comparative transcriptomic analysis of the tea plant (*Camellia sinensis*) reveals key genes involved in pistil deletion. *Hereditas*, 157(1), 1-11. <https://doi.org/10.1186/s41065-020-00153-x>
- López de Heredia, U. (2016). Las técnicas de secuenciación masiva en el estudio de la diversidad

- biológica. *Munibe Ciencias Naturales*, 64, 7-31. <https://doi.org/10.21630/mcn.2016.64.07>
- Luchnikova, N. A., Grishko, V. V., & Ivshina, I. B. (2020). Biotransformation of oleanane and ursane triterpenic acids. *Molecules*, 25(23). <https://doi.org/10.3390/molecules2523526>
- Ma, C. H., Gao, Z. J., Zhang, J. J., Zhang, W., Shao, J. H., Hai, M. R., Chen, J. W., Yang, S. C., & Zhang, G. H. (2016). Candidate genes involved in the biosynthesis of triterpenoid saponins in *Platycodon grandiflorum* identified by transcriptome analysis. *Frontiers in Plant Science*, 7, 1-14. <https://doi.org/10.3389/fpls.2016.00673>
- Malhotra, K., & Franke, J. (2022). Cytochrome P450 monooxygenase-mediated tailoring of triterpenoids and steroids in plants. *Beilstein Journal of Organic Chemistry*, 18, 1289-1310. <https://doi.org/10.3762/bjoc.18.135>
- McDermaid, A., Monier, B., Zhao, J., Liu, B., & Ma, Q. (2019). Interpretation of differential gene expression results of RNA-seq data: Review and integration. *Briefings in Bioinformatics*, 20(6), 2044-2054. <https://doi.org/10.1093/bib/bby067>
- Mugford, S. T., Louveau, T., Melton, R., Qi, X., Bakht, S., Hill, L., Tsurushima, T., Honkanen, S., Rosser, S. J., Lomonosoff, G. P., & Osbourn, A. (2013). Modularity of plant metabolic gene clusters: A trio of linked genes that are collectively required for acylation of triterpenes in oat. *Plant Cell*, 25(3), 1078-1092. <https://doi.org/10.1105/tpc.113.110551>
- Nader, B. L. (2004). Transformación genética de la especie *Galphimia glauca* Cav. (Malpighiaceae) por *Agrobacterium rhizogenes* para la producción de nor-friedelanos. Universidad Autónoma del Estado de Morelos.
- Neil, S. T. O., & Emrich, S. J. (2013). Assessing *de novo* transcriptome assembly metrics for consistency utility. *BMC Genomics*, 14, 465.
- Ortega, A., Pastor-Palacios, G., Ortiz-Pastrana, N., Ávila-Cabezas, E., Toscano, R. A., Joseph-Nathan, P., Morales-Jiménez, J., & Bautista, E. (2020). Further galphimines from a new population of *Galphimia glauca*. *Phytochemistry*, 169(May 2019), 112180. <https://doi.org/10.1016/j.phytochem.2019.112180>
- Osuna, L., Pereda-Miranda, R., Tortoriello, J., & Villarreal, M. L. (1999). Production of the sedative triterpene galphimine b in *Galphimia glauca* tissue culture. *Planta Medica*, 65(2), 149-152. <https://doi.org/10.1055/s-1999-14057>
- R. A. Toscano, A. Ortega, E. Maldonado, R. Gaviño, X. L. and J. T. (1993). Structure of galphimine B. *Acta Crystallographica Section C*, 49(4), 774-776. <https://doi.org/10.1107/S0108270192006644>
- Raghavachari, N., y Garcia-Reyero, N. (2018). Overview of gene expression analysis: transcriptomics. In N. Raghavachari, N., y Garcia-Reyero (Ed.), *Gene Expression Analysis* (pp.

1-6). Humana Press.

- Rios, M. Y., Ortega, A., Domínguez, B., Déciga, M., & Rosa, V. de la. (2020). Glucacetalin E and galphimidin B from *Galphimia glauca* and their anxiolytic activity. *Journal of Ethnopharmacology*, 259(May), 112939. <https://doi.org/10.1016/j.jep.2020.112939>
- Romero-Cerecero, O., Islas-Garduño, A. L., Zamilpa, A., Pérez-García, M. D., & Tortoriello, J. (2018). Therapeutic effectiveness of *Galphimia glauca* in young people with social anxiety disorder: A pilot study. *Evidence-Based Complementary and Alternative Medicine*, 2018. <https://doi.org/10.1155/2018/1716939>
- Roy, N. S., Lee, I. hye, Kim, J. A., Ramekar, R. V., Park, K. C., Park, N. il, Yeo, J. H., Choi, I. Y., & Kim, S. (2020). *De novo* assembly and characterization of transcriptome in the medicinal plant *Euphorbia jolkini*. *Genes and Genomics*, 42(9), 1011-1021. <https://doi.org/10.1007/s13258-020-00957-1>
- Sandeep, Misra, R. C., Chanotiya, C. S., Mukhopadhyay, P., & Ghosh, S. (2019). Oxidosqualene cyclase and CYP716 enzymes contribute to triterpene structural diversity in the medicinal tree banaba. *New Phytologist*, 222(1), 408-424. <https://doi.org/10.1111/nph.15606>
- Sandoval Pineda, J. F., Ochoa Corona, F., & Torres Rojas, E. (2017). Evaluación de diferentes métodos de extracción de ARN a partir del hongo nativo *Xylaria* sp. *Revista Colombiana de Biotecnología*, 19(1), 42-52. <https://doi.org/10.15446/rev.colomb.biote.v19n1.57114>
- Sanfilippo, P., Miura, P., & Lai, E. C. (2017). Genome-wide profiling of the 3' ends of polyadenylated RNAs. *Methods*, 126, 86–94. <https://doi.org/10.1016/j.ymeth.2017.06.003>
- Santini, B. A., Hodgson, J. G., Thompson, K., Wilson, P. J., Band, S. R., Jones, G., Charles, M., Bogaard, A., Palmer, C., & Rees, M. (2017). The triangular seed mass-leaf area relationship holds for annual plants and is determined by habitat productivity. *Functional Ecology*, 31(9), 1770-1779. <https://doi.org/10.1111/1365-2435.12870>
- Sawai, S., & Saito, K. (2011). Triterpenoid biosynthesis and engineering in plants. *Frontiers in Plant Science*, 2(JUN), 1-8. <https://doi.org/10.3389/fpls.2011.00025>
- Seki, H., Tamura, K., & Muranaka, T. (2015). P450s and UGTs: Key Players in the Structural Diversity of Triterpenoid Saponins. *Plant and Cell Physiology*, 56(8), 1463-1471. <https://doi.org/10.1093/pcp/pcv062>
- Sharma, A., Folch, J. L., Cardoso-Taketa, A., Lorence, A., y Villarreal, M. L. (2012)b. DNA barcoding of the Mexican sedative and anxiolytic plant *Galphimia glauca*. *Journal of Ethnopharmacology*, 144(2), 371-378. <https://doi.org/10.1016/j.jep.2012.09.022>
- Sharma, A., Cardoso-Taketa, A., Choi, Y. H., Verpoorte, R., & Villarreal, M. L. (2012)a. A comparison on the metabolic profiling of the Mexican anxiolytic and sedative plant *Galphimia glauca* four years later. *Journal of Ethnopharmacology*, 141(3), 964-974.

<https://doi.org/10.1016/j.jep.2012.03.033>

- Shcherban, A. B. (2015). Repetitive DNA sequences in plant genomes. *Russian Journal of Genetics: Applied Research*, 5(3), 159-167. <https://doi.org/10.1134/S2079059715030168>
- Shetty, P. J., Amirtharaj, F., & Shaik, N. A. (2019). Essentials of Bioinformatics, Volume I. In *Essentials of Bioinformatics, Volume I*. <https://doi.org/10.1007/978-3-030-02634-9>
- Slatko, B. E., Gardner, A. F., & Ausubel, F. M. (2018). Overview of next generation sequencing technologies (and bioinformatics) in cancer. *Molecular Biology*, 122(1), 1–15. <https://doi.org/10.1002/cpmb.59.Overview>
- Slavokhotova, A., Korostyleva, T., Shelenkov, A., Pukhalskiy, V., Korottseva, I., Slezina, M., Istomina, E., & Odintsova, T. (2021). Transcriptomic analysis of genes involved in plant defense response to the cucumber green mottle mosaic virus infection. *Life*, 11(10). <https://doi.org/10.3390/life11101064>
- Soetaert, S. S. A., Van Neste, C. M. F., Vandewoestyne, M. L., Head, S. R., Goossens, A., Van Nieuwerburgh, F. C. W., & Deforce, D. L. D. (2013). Differential transcriptome analysis of glandular and filamentous trichomes in *Artemisia annua*. *BMC Plant Biology*, 13(1). <https://doi.org/10.1186/1471-2229-13-220>
- Song, M. C., Kim, E. J., Kim, E., Rathwell, K., Nam, S. J., & Yoon, Y. J. (2014). Microbial biosynthesis of medicinally important plant secondary metabolites. *Natural Product Reports*, 31(11), 1497-1509. <https://doi.org/10.1039/c4np00057a>
- Soto Sedano, J. C., y López Carrascal, C. E. (2014). RNA-seq: herramienta transcriptómica útil para el estudio de interacciones planta-patógeno. *Fitosanidad*, 16(2), 101–113.
- Souza-Moreira, T. M., Alves, T. B., Pinheiro, K. A., Felipe, L. G., De Lima, G. M. A., Watanabe, T. F., Barbosa, C. C., Santos, V. A. F. F. M., Lopes, N. P., Valentini, S. R., Guido, R. V. C., Furlan, M., & Zanelli, C. F. (2016). Friedelin synthase from *Maytenus ilicifolia*: Leucine 482 plays an essential role in the production of the most rearranged pentacyclic triterpene. *Scientific Reports*, 6(October), 1-13. <https://doi.org/10.1038/srep36858>
- Srivastava, A., George, J., & Karuturi, R. K. M. (2018). Transcriptome analysis. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, 1-3, 792-805. <https://doi.org/10.1016/B978-0-12-809633-8.20161-1>
- Sugiyama, M., Ito, J., Aoyagi, S., & Fukuda, H. (2000). Endonucleases. *Programmed Cell Death in Higher Plants*, 143-153. https://doi.org/10.1007/978-94-010-0934-8_11
- Tang, F., Chu, L., Shu, W., He, X., Wang, L., & Lu, M. (2019). Selection and validation of reference genes for quantitative expression analysis of miRNAs and mRNAs in Poplar. *Plant Methods*, 15(1), 1-15. <https://doi.org/10.1186/s13007-019-0420-1>
- Thimmappa, R., Geisler, K., Louveau, T., O'Maille, P., & Osbourn, A. (2014). Triterpene biosynthesis

- in plants. *Annual Review of Plant Biology*, 65(September), 225-257. <https://doi.org/10.1146/annurev-arplant-050312-120229>
- Tholl, D. (2015). *Biosynthesis and biological functions of terpenoids in plants. Biotechnology of isoprenoids.*
- Tortoriello, J., Ortega, A., Herrera-Ruíz, M., Trujillo, J., y Reyes-Vázquez, C. (. (1998). Galphimine-B modifies electrical activity of ventral tegmental area neurons in rats. *Planta Medica*, 64(4), 309-313. <https://doi.org/10.1055/s-2006-957440>
- Tortoriello, J., & Lozoya, X. (1992). Effect of *Galphimia glauca* methanolic extract on neuropharmacological tests. *Planta Medica*, 58(3), 234–236. <https://doi.org/10.1055/s-2006-961442>
- Türktaş, M., Kurtoglu, K. Y., Dorado, G., Zhang, B., Hernandez, P., & Ünver, T. (2015). Sequencing of plant genomes - A review. *Turkish Journal of Agriculture and Forestry*, 39(3), 361-376. <https://doi.org/10.3906/tar-1409-93>
- Unamba, C. I. N., Nag, A., & Sharma, R. K. (2015). Next generation sequencing technologies: The doorway to the unexplored genomics of non-model plants. *Frontiers in Plant Science*, 6(DEC). <https://doi.org/10.3389/fpls.2015.01074>
- University, M. S. (2020). *FastQC Tutorial & FAQ*. Research Technology Support Facility. <https://rtsf.natsci.msu.edu/genomics/tech-notes/fastqc-tutorial-and-faq/>
- Uno, Y., Ushirozako, G., Uehara, S., Murayama, N., Fujiki, Y., Kawaguchi, H., Tsukiyama-Kohara, K., & Yamazaki, H. (2022): Newly identified tree shrew cytochrome P450 2B6 (CYP2B6) and pig CYP2B6b are functional drug-metabolising enzymes. *Xenobiotica*, 3, 1-10 DOI: 10.1080/00498254.2022.2141153
- Upadhyay, S., Phukan, U. J., Mishra, S., & Shukla, R. K. (2014). *De novo* leaf and root transcriptome analysis identified novel genes involved in Steroidal sapogenin biosynthesis in *Asparagus racemosus*. *BMC Genomics*, 15(1), 1-13. <https://doi.org/10.1186/1471-2164-15-746>
- Wendt, K. U. (2005). Enzyme mechanisms for triterpene cyclization: New pieces of the puzzle. *Angewandte Chemie - International Edition*, 44(26), 3966-3971. <https://doi.org/10.1002/anie.200500804>
- Weraduwege, S. M., Chen, J., Anozie, F. C., Morales, A., Weise, S. E., & Sharkey, T. D. (2015). The relationship between leaf area growth and biomass accumulation in *Arabidopsis thaliana*. *Frontiers in Plant Science*, 6(APR), 1–21. <https://doi.org/10.3389/fpls.2015.00167>
- Xiong, W., Lan, T., & Mo, B. (2021). Extraribosomal Functions of Cytosolic Ribosomal Proteins in Plants. *Frontiers in Plant Science*, 12(April), 1-10. <https://doi.org/10.3389/fpls.2021.607157>
- Xue, L., He, Z., Bi, X., Xu, W., Wei, T., Wu, S., & Hu, S. (2019). Transcriptomic profiling reveals MEP pathway contributing to ginsenoside biosynthesis in *Panax ginseng*. *BMC Genomics*, 20(1), 1-

14. <https://doi.org/10.1186/s12864-019-5718-x>

- Y, C. (2017). The evolution of branching in and plants: between conservation and diversity. In M. G. Nuno de la Rosa L. (Ed.), *Evolutionary Developmental Biology*. Springer. https://doi.org/10.1007/978-3-319-33038-9_63-1
- Yasumoto, S., Seki, H., Shimizu, Y., Fukushima, E. O., & Muranaka, T. (2017). Functional characterization of CYP716 family P450 enzymes in triterpenoid biosynthesis in tomato. *Frontiers in Plant Science*, 8(JANUARY). <https://doi.org/10.3389/fpls.2017.00021>
- Yockteng, R., Almeida, A. M. R., Yee, S., Andre, T., Hill, C., & Specht, C. D. (2013). A method for extracting high-quality RNA from diverse plants for next-generation sequencing and gene expression analyses. *Applications in Plant Sciences*, 1(12), 1300070. <https://doi.org/10.3732/apps.1300070>
- Yun Lu, Yunfeng Luo, Jiawei Zhou, Tianyuan Hu, Lichan Tu, Yuru Tong, Ping Su, Yuan Liu, Jiadian Wang, Zhouqian Jiang, Xiaoyi Wu, Xiaochao Chen, Luqi Huang, W. G. (2021). Probing the functions of friedelane-type triterpene cyclases from four celastrol-producing plants. *Plant Journal*, 109(3), 555-567. <https://doi.org/https://doi.org/10.1111/tpj.15575>
- Zhang, Z., Liu, W., Ma, Z., Zhu, W., & Jia, L. (2019). Transcriptional characterization and response to defense elicitors of mevalonate pathway genes in cotton (*Gossypium arboreum* L.). *PeerJ*, 2019(11). <https://doi.org/10.7717/peerj.8123>
- Zhao, Z., Meng, F., Wang, W., Wang, Z., Zhang, C., & Jiang, T. (2017). Comprehensive RNA-seq transcriptomic profiling in the malignant progression of gliomas. *Scientific Data*, 4, 1-7. <https://doi.org/10.1038/sdata.2017.24>
- Zhou, J., Hu, T., Gao, L., Su, P., Zhang, Y., Zhao, Y., Chen, S., Tu, L., Song, Y., Wang, X., Huang, L., & Gao, W. (2019). Friedelane-type triterpene cyclase in celastrol biosynthesis from *Tripterygium wilfordii* and its application for triterpenes biosynthesis in yeast. *New Phytologist*, 223(2), 722–735. <https://doi.org/10.1111/nph.15809>
- Zou, T., Li, S., Liu, M., Wang, T., Xiao, Q., Chen, D., Li, Q., Liang, Y., Zhu, J., Liang, Y., Deng, Q., Wang, S., Zheng, A., Wang, L., & Li, P. (2017). An atypical strictosidine synthase, OsSTRL2, plays key roles in anther development and pollen wall formation in rice. *Scientific Reports*, 7(1), 1-13. <https://doi.org/10.1038/s41598-017-07064-4>

Anexos

Anexo 1. Artículo de investigación: Iglesias, D.; Donato Capote, M.d.; Méndez Tenorio, A.; Valdivia, A.V.; Gutiérrez-García, C.; Paul, S.; Iqbal, H.M.N.; Villarreal, M.L.; Sharma, A. Identification of putative candidate genes from *Galphimia* spp. encoding enzymes of the galphimines triterpenoids synthesis pathway with anxiolytic and sedative effects. *Plants*



Article

Identification of Putative Candidate Genes from *Galphimia* spp. Encoding Enzymes of the Galphimines Triterpenoids Synthesis Pathway with Anxiolytic and Sedative Effects

Dianella Iglesias ^{1,2}, Marcos de Donato Capote ², Alfonso Méndez Tenorio ³, Ana Victoria Valdivia ², Claudia Gutiérrez-García ², Sujay Paul ², Hafiz M. N. Iqbal ⁴, María Luisa Villarreal ^{1,*} and Ashutosh Sharma ^{2,*}

- ¹ Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos, Cuernavaca 62210, CP, Mexico; dianella.iglesiasrod@uaem.edu.mx
 - ² Centre of Bioengineering, NatProLab, School of Engineering and Sciences, Tecnológico de Monterrey, Querétaro 76130, CP, Mexico; mdedonate@tec.mx (M.d.D.C.); victoria.valdivia@tec.mx (A.V.V.); a00887520@itesm.mx (C.G.-G.); spaul@tec.mx (S.P.)
 - ³ Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Ciudad de Mexico 56565, CP, Mexico; amendezt@ipn.mx
 - ⁴ School of Engineering and Sciences, Tecnológico de Monterrey, Monterrey 64849, CP, Mexico; hafiz.iqbal@tec.mx
- * Correspondence: luisav@uaem.mx (M.L.V.); asharma@tec.mx (A.S.)



Citation: Iglesias, D.; Donato Capote, M.d.; Méndez Tenorio, A.; Valdivia, A.V.; Gutiérrez-García, C.; Paul, S.; Iqbal, H.M.N.; Villarreal, M.L.; Sharma, A. Identification of Putative Candidate Genes from *Galphimia* spp. Encoding Enzymes of the Galphimines Triterpenoids Synthesis Pathway with Anxiolytic and Sedative Effects. *Plants* **2022**, *11*, 1879. <https://doi.org/10.3390/plants11141879>

Academic Editors: Ivayla Dinecheva, Ilan Badjakov and Bistra Galunska

Received: 9 June 2022

Accepted: 5 July 2022

Published: 20 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: *Galphimia* spp. is popularly used in Mexican traditional medicine. Some populations of *Galphimia* exert anxiolytic and sedative effects due to the presence of the modified triterpenoids galphimines. However, the galphimine synthesis pathway has not yet been elucidated. Hence, in this study, a comparative transcriptome analysis between two contrasting populations of *Galphimia* spp., a galphimine-producer, and a non-galphimine-producer, is performed using RNA-Seq in the Illumina Next Seq 550 platform to identify putative candidate genes that encode enzymes of this metabolic pathway. Transcriptome functional annotation was performed using the Blast2GO in levels of gene ontology. For differential expression analysis, edgeR, pheatmap, and Genie3 library were used. To validate transcriptome data, qPCR was conducted. In producer and non-producer plants of both populations of *Galphimia* spp., most of the transcripts were grouped in the Molecular Function level of gene ontology. A total of 680 differentially expressed transcripts between producer and non-producer plants were detected. In galphimine-producer plants, a larger number of highly expressed transcripts related to acyclic and polycyclic terpene synthesis were identified. As putative candidate genes involved in the galphimine synthesis pathway, P450 family members and enzymes with kinase activity were identified.

Keywords: functional annotation; *Galphimia* spp.; transcriptome

1. Introduction

Galphimia spp. is a medicinal plant widely used in Mexico to treat various ailments, including fever, labor pains [1], diarrhea [2], and anxiety, due to its anti-inflammatory and anxiolytic properties [3,4]. The anxiolytic and sedative effects of the extract have made this plant the target of pharmacological [5–7], phytochemicals [3,8–10], biotechnological [6,11,12], metabolomic [13,14] and genetics studies [15,16]. Nevertheless, the effects of some populations of the plant, are due to the presence of the modified triterpenoid galphimines [17].

Different metabolomic studies have already been carried out to detect the presence of galphimines as well as their concentration among six Mexican populations of *Galphimia* spp. [13,14], from which only two (located in Jalpan, (Querétaro) and Dr. Mora

(Guanajuato)) were identified as galphimines producers. Moreover, to identify the individual populations of *Galphimia* a DNA barcode study was conducted by Sharma et al. [15], and the clades obtained showed clear differences between the galphimine-producer populations (Doctor Mora and Jalpan) and the other analyzed populations. In another study, conducted by Gesto-Borroto et al. [16], among nine populations of *Galphimia* spp. evaluated by DNA barcode, four galphimine-producing populations were identified, including Dr. Mora (Guanajuato) and Jalpan (Querétaro), corroborating the previous studies.

In several plants, stages of triterpene synthesis have been studied [18–20]; however, the galphimine-producing pathway in *Galphimia* spp., is still elusive. Nevertheless, a biogenetic proposal for its production has been published from the oxidation and decarboxylation of the Taraxasteryl cation [21], although it has not yet been demonstrated.

Regarding the biological properties of the galphimines, the hypothesis to be tested was the following: „If a transcriptomic study and a comparison are carried out in two populations of *Galphimia* spp.; a galphimine producer population (GP) with a non-galphimine producer population (NPG), the putative candidate genes, that encode the enzymes involved in the production pathway of these triterpenoids, can be identified. Hence, this paper aimed to analyze transcriptomes from a GP and NPG populations of *Galphimia* spp., to identify putative candidate genes that encode enzymes for the synthesis of these triterpenoids.

2. Results

2.1. Plant Material

Plants with phenotypic characteristics similar to those described for the *Galphimia* genus [22], were obtained from both populations GP and NPG. They were characterized by having several adventitious roots, 5–6 cm long, with abundant adsorbent hairs. They are 1.4 m tall shrubs, with approximately 15 lateral branches in the GP plants and 10 in the NPG plants. The leaves were elliptical, with prominent secondary veins and their area was 13.75 cm². Yellow inflorescences were observed arranged in terminal racemes.

2.2. Transcriptome Sequencing

RNA samples had a high quality with RNA integrity number (RIN) ≥ 8 . Around 56.6 Gb of clean data were obtained from transcriptome sequencing. Q30 percentages of clean data for all samples were 97.2%, and the Guanine/Cytosine (GC) were rated between 50 to 51%. The total bases per sample obtained were in a range between 2,704,225 and 675,148.

After FastQC analysis, it was observed that the quality of bases per sequence was optimal since all readings were located in the green area of the graphic. The fast QC analysis showed that the filtered reads had quality values $>Q30$, but the reads by sample GC content and duplication levels differed. After trimming the reads with Trimmomatic, these metrics reached the standard values. Post-assembly, around 9000 transcripts from leaf samples and 5000 transcripts from roots samples were obtained (Table 1).

Table 1. Transcripts were obtained after de novo assembly of transcriptomes.

Sample and Replica	Number of Transcripts
Leaves of galphimine producing-population	9023
Leaves of galphimine producing population, replica	8915
Leaves of non-galphimine producing-population	8667
Leaves of non-galphimine producing-population, replica	8896
Roots of galphimine producing-population	5004
Roots of galphimine producing population, replica	4667

2.3. Transcriptome Functional Annotation

The gene ontology analysis showed a different number of annotated genes: 5900 in GP plants, 3600 in the NGP ones, and 1500 in the controls. In GP plants and controls, most of the transcripts were grouped in gene ontology categories MF and CC. In the NGP plants, most of the transcripts were grouped mostly in the MF and BP categories. At the level of gene ontology BP, the largest number of transcripts were associated with cellular processes, biological processes, and biological regulation (Figure 1), while at the MF level, most of the transcripts annotated were related to catalysis, binding, and transporter. In the CC, transcripts of three samples are located as a cellular anatomical entity and protein-containing complex.

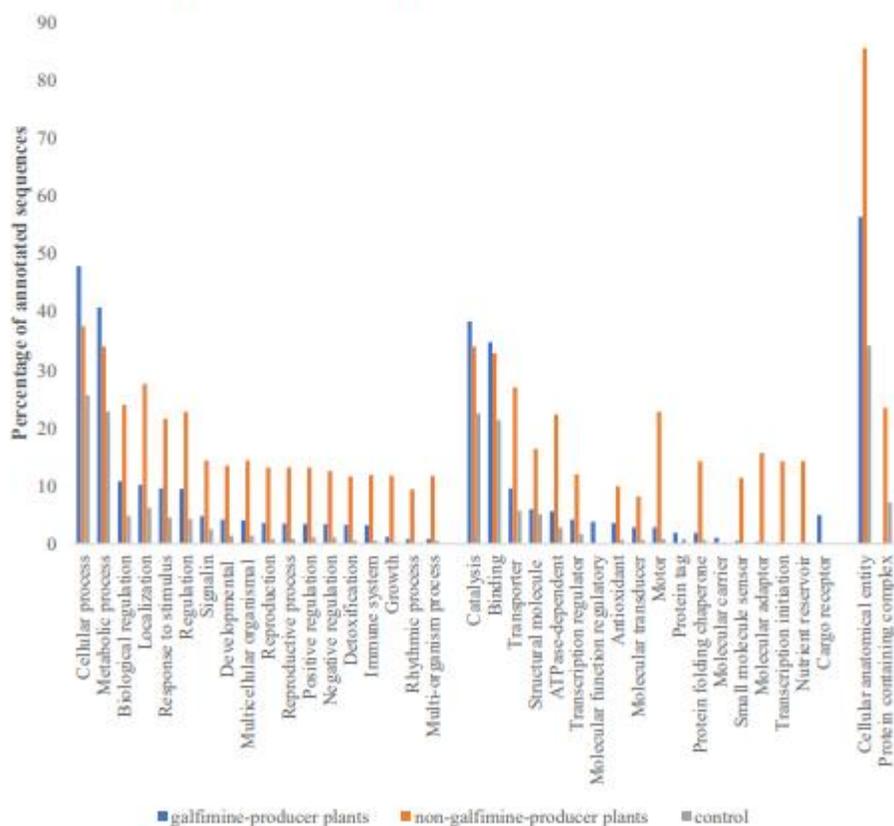


Figure 1. Percentage of annotated sequences of galphimine-producer plants (leaves samples, non-galphimine producer plants, (leaves samples) and control (roots samples) in gene ontology levels Biological Process (BP), Molecular Function (MF), and Cellular Component (CC).

According to the enzyme distribution code in the GP population, several transcripts were identified (Figure 2), such as 1300 with transferase activity, 1020 with hydrolase activity, and 810 with oxidoreductase activity; while in the NGP population transcripts were observed as follows: 2250 with oxidoreductase activity, 2220 with transferase activity, and 2050 with translocase activity. In control, 270 transcripts with hydrolase activity, 260 with transferase activity, and 250 with translocase activity were observed.

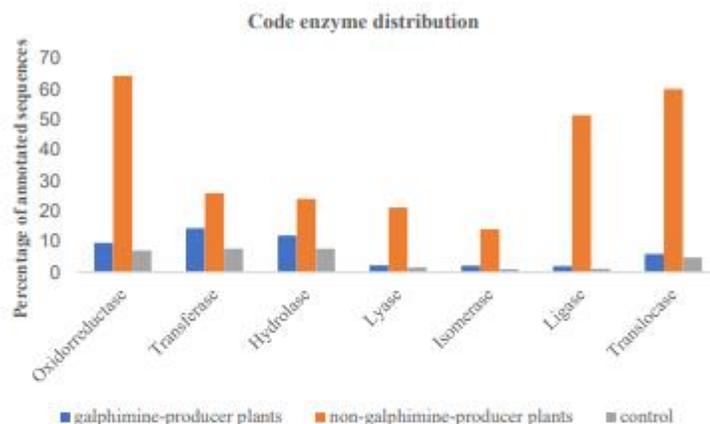


Figure 2. Enzyme distribution code of annotated sequences (percentage) in galphimine producer plants (leaves samples), non-galphimine producer plants (leaves samples), and control (roots samples).

2.4. Differential Expression Analysis

When comparing the GP population with the NGP population, a total of 680 transcripts with differential expressions were found. Of these, 313 were positively regulated and 367 negatively regulated. The volcano plot shows that both positively regulated (represented by red dots) and negatively regulated transcripts (represented by green dots) have a large-magnitude logFC due to the significant separation from the center, and a high level of statistical significance since the location of dots over 100 of log₁₀ (PValue). The greatest statistical significance was observed in positive regulated transcripts. (Figure 3a).

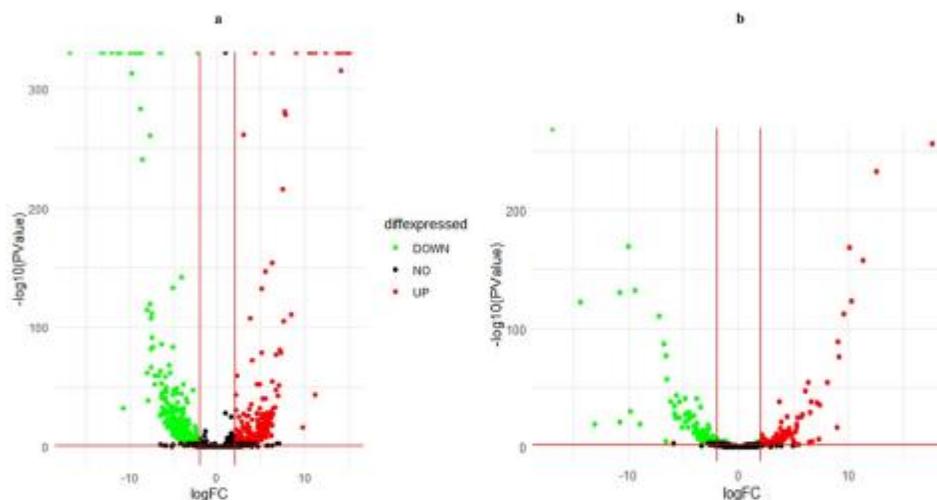


Figure 3. Volcano plot of differentially expressed genes (a) galphimine producer plants (leaves samples) versus non-galphimine producer plants (leaves samples) (b) galphimine producer plants (leaves samples) versus control (roots samples). X-axis and y-axis represent log₂ fold-change differences between the compared samples and statistical significance as the negative log of differentially expressed genes *p*-values, respectively. The significantly up-regulated and down-regulated genes are indicated with red and green dots, respectively, while non-significant genes are shown as black dots.

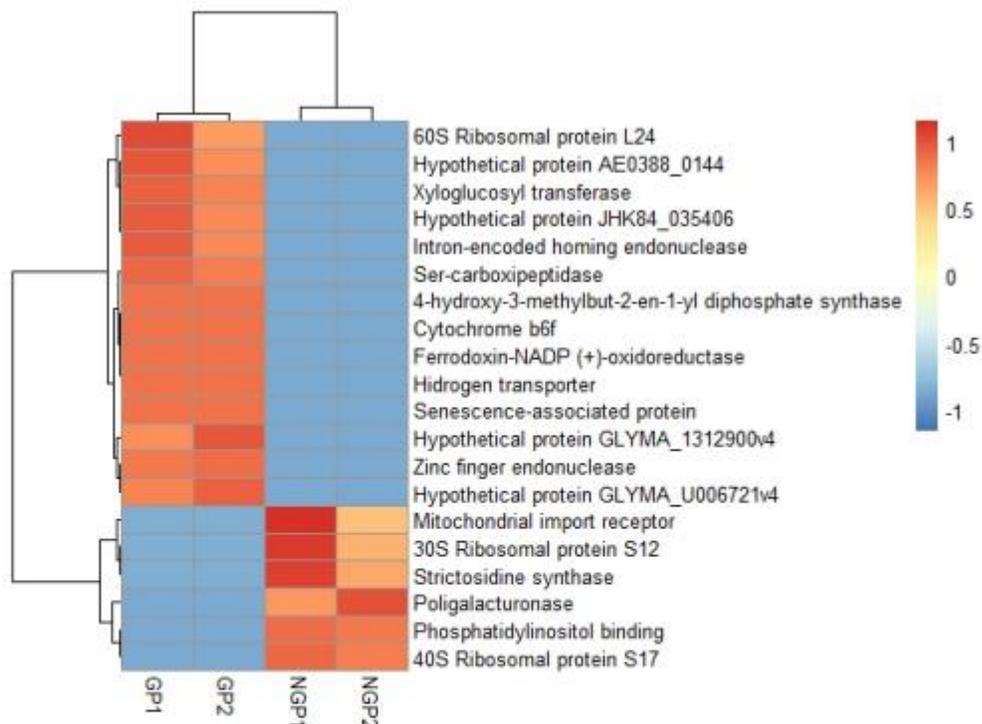


Figure 4. Comparison of transcripts with higher differential expression values between two populations of *Galphimia* spp. GP1, GP2 (leaves of galphimine producer plants). NGP1, NPG2 (leaves of non-galphimine producer plants). High values are represented in red color, the lowest values are represented in blue.

In the comparison between GP plants and controls, the transcripts with the highest expression values in the GP plants were ribosomal proteins, proteins associated with senescence, and enzymes with endonuclease, polymerase, and dehydrogenase activities (Figure 5). In controls, the transcripts with higher expression values code for receptors, aquaporins, chaperones, endonucleases, and methyltransferases.

A total of 203 transcripts were differentially regulated between GP plants and control. The number of negatively regulated transcripts ($\log_{2}FC < -2$) was 116, higher than positively regulated transcripts ($\log_{2}FC > 2$), which was 87 in number. The positively regulated transcripts presented a higher statistical significance (Figure 3b). In addition, there were changes in the $\log_{2}FC$ and high values of $-\log_{10}(P\text{Value})$, indicators of the statistical significance of some transcripts.

In the GP plants, a greater number of differentially expressed transcripts with positive regulation were observed as compared to the NGP and the control (Figure 4), suggesting a high metabolic activity in these plants. Moreover, the transcripts with the highest expression values in the GP plants were mostly hypothetical proteins, ribosomal proteins, transferases, endonucleases, and carboxypeptidases. However, in NGP plants, the transcripts with the highest expression encoded mitochondrial receptors, ribosomal proteins, enzymes with synthase activity, and polygalacturonases.

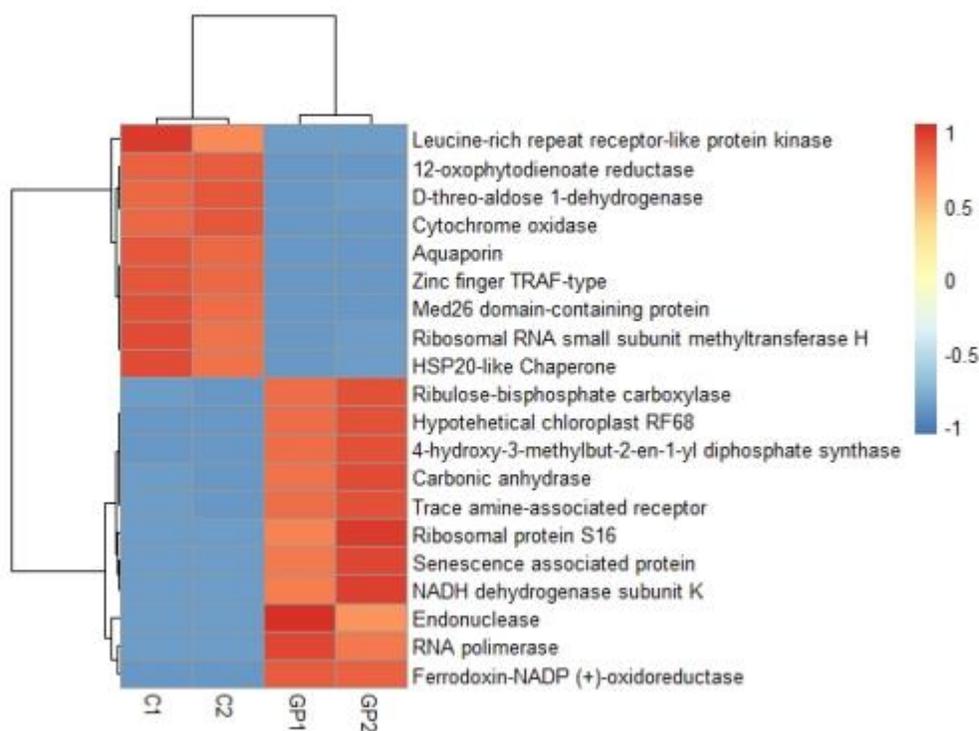


Figure 5. Comparison of transcripts with higher differential expression values between leaves and roots of galphimine producer plants. C1, C2 (roots). GP1, GP2 (leaves). High values are represented in red color, the lowest values are represented in blue.

2.5. Identification of Putative Candidate Genes That Encode Enzymes of the Galphimine Synthesis Pathway

Based on the results of functional annotation, 44 transcripts associated with the terpenes synthesis pathway were identified in GP plants (Table 2). Out of those, two transcripts encode for 3-hydroxy-3-methylglutaryl-CoA reductase (a regulatory enzyme of the mevalonic acid pathway), and 13 transcripts encode enzymes of the methylerythritol phosphate pathway. Among other transcripts, three encode for enzymes associated with intermediates in the isoprenoid biosynthesis, one related to terpenes synthesis, and 25 with triterpenes.

In the NGP plants, 35 transcripts were found associated with terpene synthesis, among which one encodes the regulatory enzyme of the mevalonic acid pathway, 10 associated with methylerythritol phosphate pathway, five with the intermediary in the isoprenoid biosynthesis, three transcripts related to terpenes synthesis and 16 with triterpenes (Table 2). In the control, the following transcripts were observed: one that encodes for 3-hydroxy-3-methylglutaryl-CoA reductase, one for P450, and five for terpene synthase.

The GP plants presented a greater number and diversity of transcripts related to the synthesis of terpenes with high values of gene counts compared to the NGP plants. This suggests a higher metabolic activity in the terpene synthesis pathway in these plants. P450 TBP, P450, and terpene synthase were the enzymes with the highest gene count. In the NGP plants, the enzymes with the highest gene counts were terpene synthase and CYP82G1.

Table 2. Annotation and gene counts of transcripts related to terpene synthesis of two populations of *Galphimia* spp.

Annotated Transcripts	Galphimine-Producer Population		Non-Galphimine-Producer Population		Control	
	Number of Annotations	Gen Counts	Number of Annotations	Gen Counts	Number of Annotations	Gen Counts
3-hydroxy-3-methylglutaryl-CoA reductase	2	13	1	8	1	6
1-deoxy-D-xylulose-5-phosphate synthase	4	15	4	29	-	-
4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	4	18	1	10	-	-
4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	5	19	5	16	-	-
isopentenyl diphosphate isomerase	1	52	1	17	-	-
farnesyl transferase	1	6	3	4	-	-
geranyl diphosphate synthase	1	14	1	4	-	-
terpene synthase	1	347	3	253	5	78
squalene synthase	1	7	-	-	-	-
squalene epoxidase	1	16	-	-	-	-
beta-amyrin synthase	2	23	4	14	-	-
lanosterol synthase	7	18	6	23	-	-
P450	2	304.444	-	-	1	4
P450 TBP	19	417.183	-	-	-	-
CYP71A24	1	5	2	3	-	-
CYP71A26	-	-	1	14	-	-
CYP71B10	-	-	1	12	-	-
CYP71B34	1	9	-	-	-	-
CYP71D11	-	-	-	-	1	15
CYP72A	2	10	-	-	-	-
CYP72A15	-	-	1	7	-	-
CYP81D1	1	18	-	-	-	-
CYP82C4	1	231	-	-	-	-
CYP82G1	1	41	1	129	-	-
PLAC8	2	12	-	-	-	-
5'-AMP-activated protein kinase	1	30	-	-	-	-
serine/threonine-protein kinase AtPK2/AtPK19	1	6	-	-	-	-

In the co-expression network, 22 nodes were observed (Figure 6), among which, HMGCR (3-hydroxy-3-methylglutaryl-CoA reductase), IPPI (isopentenyl diphosphate isomerase), and FT (farnesyl transferase) showed the highest number of interactions with other nodes. The expression of the majority of P450 family members was closely related.

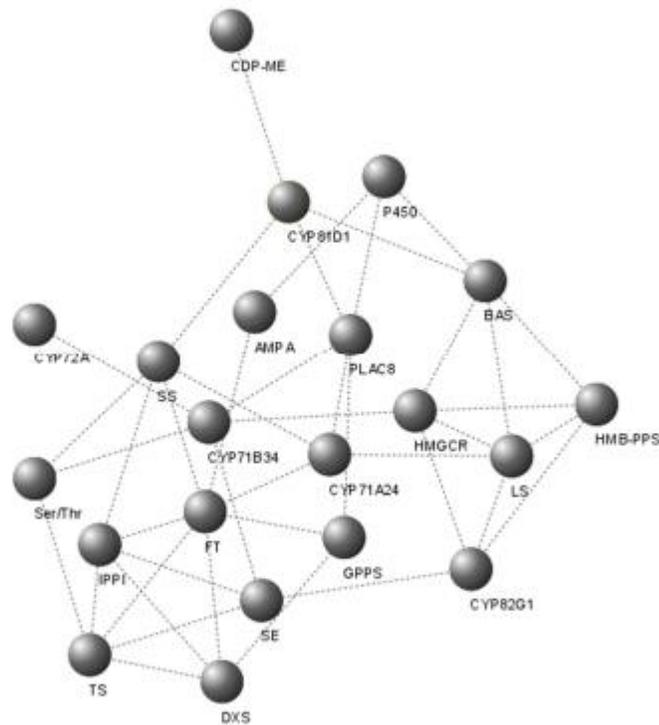


Figure 6. Co-expression network of transcripts related to the synthesis pathway of triterpenes and members of the cytochrome P450 family in leaves of the galphimine producer plants and non-galphimine producer plants. Transcripts are represented by spheres and dotted lines are their interaction. HMGCR (3-hydroxy-3-methylglutaryl-CoA reductase), DXS (1-deoxy-D-xylulose-5-phosphate synthase), CDP-ME (4-diphosphocytidyl-2-C-methyl-D-erythritol kinase), HMB-PPS (4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase), IPP isomerase (isopentenyl diphosphate isomerase), GPPS (geranyl diphosphate synthase), FT (farnesyl transferase), TS (terpene synthase), SS (squalene synthase), SE (squalene epoxidase), BAS (beta-amyrin synthase), LS (lanosterol synthase).

2.6. Validation of Transcriptome Data by qPCR

By qPCR analysis, the results of the transcriptomic study were validated. The expression levels of putative candidate genes related to galphimine synthesis were found to be similar in both differential gene expression analysis (Table 2) and qPCR (Figure 7). The expression was higher in GP plants because CT values were lower than in NGP plants (Figure 8) and significant differences between plants of both populations in fold change value were observed (Figure 6). In GP plants the genes with a higher level of relative expression were P450 TBP, 5'-AMP-activated protein kinase, and P450. Overall, the relative expression of these genes in NGP plants was low.

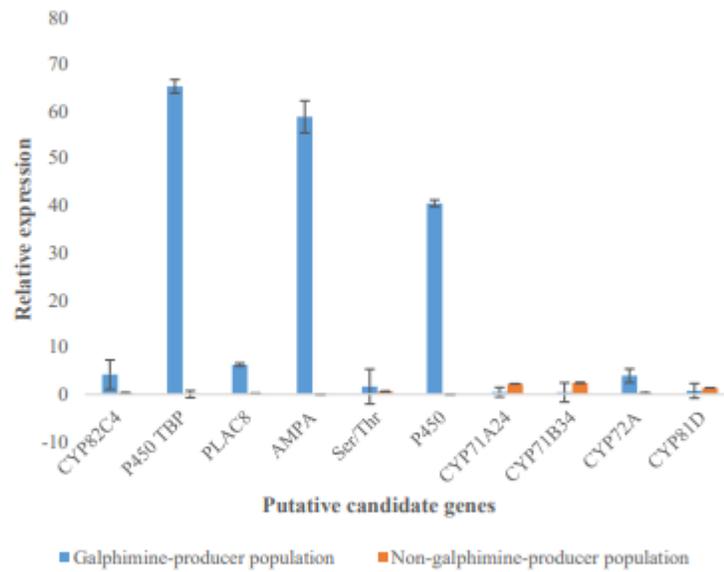


Figure 7. Relative expression of putative candidate genes in leaves of galphimine producer plants (GP) and non-galphimine producer plants (NGP) was calculated by the $2^{-\Delta\Delta Ct}$ method in quantitative PCR analysis. Y-axis represents the fold change value between GP and NGP. AMPA (5'-AMP-activated protein kinase), Ser/Thr (serine/threonine-protein kinase AtPK2/AtPK19).

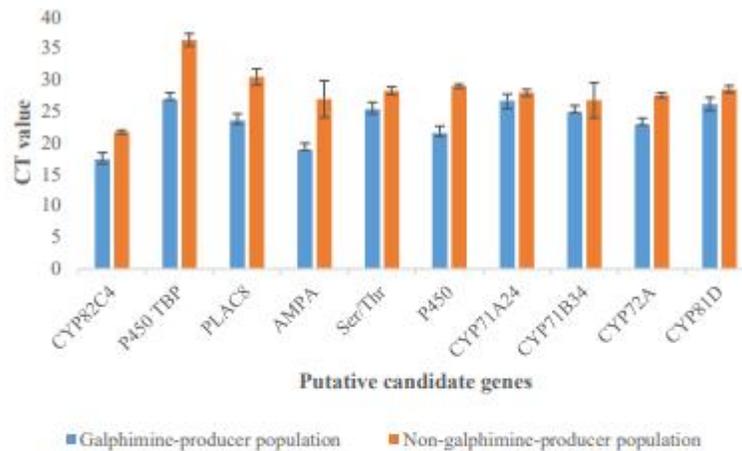


Figure 8. CT values of leaves of galphimine producer plants (GP) and non-galphimine producer plants (NGP) obtained by quantitative PCR analysis. AMPA (5'-AMP-activated protein kinase), Ser/Thr (serine/threonine-protein kinase AtPK2/AtPK19).

3. Discussion

Galphimia spp. is an important medicinal plant in Mexico, with several applications in traditional medicine; thus, the main investigations with this plant are based on its pharmacological effects, especially for the modified triterpenoids galphimines. However, genes that encode enzymes of the galphimines synthesis pathway are unknown. Nevertheless,

identifying new genes is a complex and great important process, especially when the genome is not available. The least complex and economical strategy is the use of omics tools such as transcriptome and digital gene expression profile analysis. These strategies allowed to explore and analyze genes with a differential expression which increases the chance of discovering new genes as well as the characterization of their structure. In this study, the Illumina Next seq 550 platform has been used to investigate the transcriptomes of *Galphimia* spp, which allowed the successful identification of several galphimine synthesis-related genes.

The GO annotation provides a preliminary indication of the nature of a gene product. In this study, most of the transcripts have been recorded at the MF and CC gene ontology levels, which provided an approximation to the elementary functions in which the transcripts participate as well as the parts of the cells or extracellular environment they're located in. Given that galphimines are triterpenoids of polycyclic structure and that the synthesis of these metabolites includes oxidation and transfer of functional groups, it is predicted that most of the transcripts identified are associated with transferase and oxidoreductase activities.

Triterpenes are one of the most important groups of secondary metabolites in plants. These are generally produced from the acyclic 30-carbon precursors 2,3-oxidosqualene in eukaryotes. The conversion of 2,3-oxidosqualene to various cyclic triterpenes by squalene cyclase is the first step in the diversification of the biosynthetic pathway of the triterpenes. Subsequently, the generated intermediaries undergo a large number of modifications regio and stereospecific structures catalyzed by cytochrome P450 (CYP), acetylases, and UDP glucosyltransferases that catalyze oxidation, methylation, acetylation, malonylation, and glycosylation, [23]. Plant CYPs transfer two electrons from cofactors NADPH and atmospheric oxygen to their substrates. Generally, plants require a CYP reductase to support electron transfer [24]. The oleanane, derived from β -amyrin, is one of the largest representatives of triterpene scaffolds pentacyclic. Approximately 55 CYPs act on scaffolds of pentacyclic triterpenes of plants. Most of them belong to members of the CYP716 family, although other families such as CYP51, CYP71, CYP72, CYP87, CYP88, and CYP93 also participate in modifications of pentacyclic triterpenes [23].

Members of the cytochrome P450 family are involved in the synthesis of triterpenes. Geisler et al. performed a functional analysis of CYP51 in *Nicotiana benthamiana* leaves and demonstrated that this enzyme can catalyze the hydroxylation and the epoxidation of simple triterpene β -amyrin to give 12,13 β -epoxy-3 β , 16 β -dihydroxy-oleanane (12, 13 β -epoxy-16 β -hydroxy- β -amyrin) [25].

Also, three oxidosqualene cyclases (OSC) TwOSC1, TwOSC2, and TwOSC3 were isolated and characterized from *Tripterygium wilfordii*. TwOSC1 and TwOSC3 were multiproduct friedelin synthases and were involved in celastrol biosynthesis, while TwOSC2 was a β -amirin synthase [26]. In that work, the authors proposed a biogenetic pathway for celastrol in which hydroxyfriedelanes synthesis was mediated by P450 [26]. On the other hand, in *Aralia elata*, CYP716A295 and CYP716A296 were identified as the candidate genes most likely associated with oleanolic acid synthesis [27].

Galphimines are modified triterpenoids, classified as Nor-secofriedelanes. These are derived from oleanane and later from friedelin. The arrangements in the intermediates between friedelin and norsescofriedelanes involve oxidation-reduction reactions for opening a ring and the C3-C4 bond break, resulting in the two functional groups required for the formation of the seven-membered lactone characteristic of galphimines. Considering the functions that have been reported for the P450 family in the synthesis of triterpenes and the gene count of the transcripts encoding members of the P450 family in the GP population (Table 2), as well as the absence of many of these in the NGP population, these enzymes could be considered potential candidates for galphimines synthesis, as suggested in Figure 9.

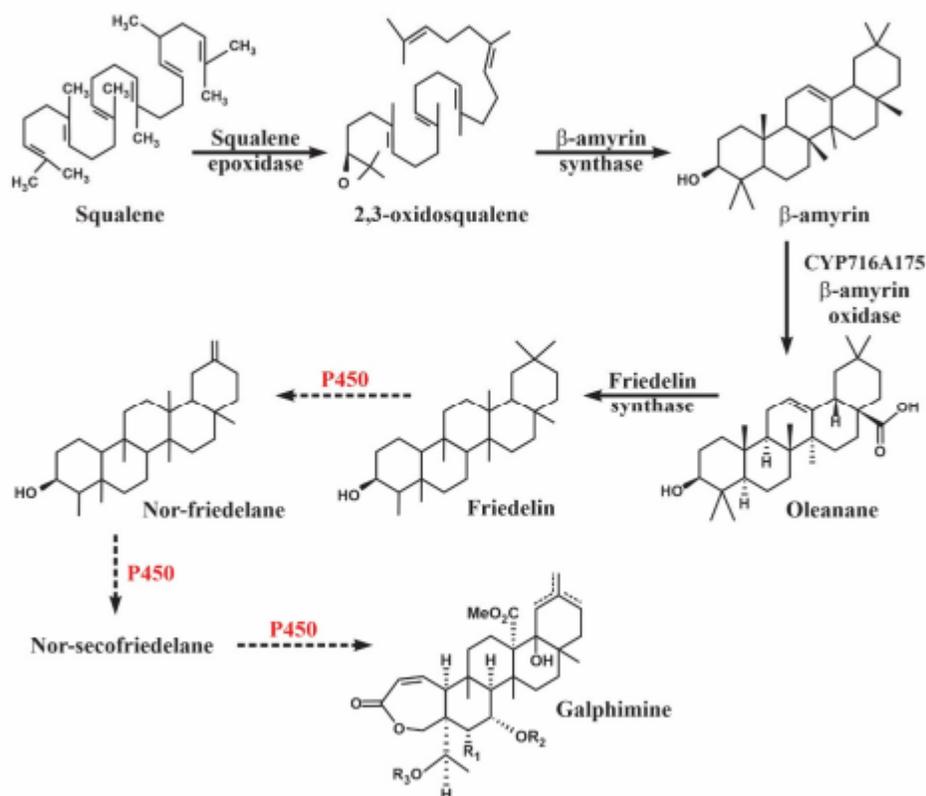


Figure 9. Proposed galphimines synthesis pathway mediated by cytochrome P450 in *Galphimia* spp. plants. Solid arrows indicate pathway reactions identified in previous work [18,28]. Dotted arrows indicate reactions proposed in this study.

In differential expression analysis, the GP plants presented a greater number of genes related to the synthesis of terpenes as compared to the NGP plants. In both analyzed populations, some transcripts encode the mevalonic acid and methylerythritol phosphate pathways as well as precursors of the biosynthesis of isoprenoids, terpenes, and triterpenes. However, in the GP plants, the number and diversity of identified enzymes were higher. Since terpenes are a group of secondary metabolites that are highly represented in plants [29], they are present in both studied populations. However, the GP plants presented a higher expression of the enzymes of this pathway and, as a consequence, a higher content of these metabolites. This corroborates with the results of the study by Sharma et al. where the chemical profile of the *Galphimia* spp., the populations used in this study was analyzed by thin-layer chromatography (TLC), suggesting the presence of galphimines in the GP population, while other terpenes were in the NGP population [15].

Among the enzymes with the highest levels of expression in both populations were terpene synthase and CYP82G1. This is expected because terpene synthase diversifies isoprenoid precursors into a large number of mono and sesquiterpenes [30]. CYP82G1 participates in the synthesis of volatile homoterpenes. These homoterpenes are considered among the most common plant compounds that act as part of defense mechanisms against herbivores [31].

Co-expression networks are a new resource for understanding convergent pathways and their relations. This resource has been used to address various biological questions [32].

In the present study, the enzymes selected as possible candidates for the galphimines synthesis pathway interacted closely with each other, as well as with other enzymes involved in stages before the triterpenes synthesis. It has shown the genes that participate in the same metabolic pathway or those are under the same regulatory mechanism. It is worth remarking on the interaction of 3-hydroxy-3-methylglutaryl-CoA reductase, with a considerable number of the identified enzymes that are associated with the synthesis of polycyclic triterpenes. This enzyme is a regulator of the mevalonic acid pathway, which seems to be the proper route for synthesizing triterpenes [33]. The 3-hydroxy-3-methylglutaryl-CoA reductase seems to control the core of the terpene's metabolism in the cytosol, regulating the production of the compound obtained, as well as the activity of the gene products involved in these reactions.

In this study, relative expression determined by qPCR corroborated the results from the transcriptomic analysis. Our results allow us to support that the genes selected as candidates involved in galphimine biosynthesis are present in high concentrations in the GP plants and are in low concentrations or absent in the NGP plants.

Given the limited availability of reference genomes in plants, the transcriptomic analysis of bioactive metabolites in producer and nonproducer plants is a valuable strategy to detect the genes involved in the synthesis pathways of the compounds of interest.

4. Material and Methods

4.1. Plant Material

The collection of plants and seeds was carried out manually under the permission of Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT) in a GP population, Doctor Mora (Guanajuato) (W 100° 19.22; N 21° 08.7; 2120 m altitude) and in a NGP population, Tepoztlán (Morelos) (W 99° 06.97 N 18° 59.35, 1700 m altitude). The samples were deposited in the HUMO Herbarium of the Sierra de Huatla Environmental Education and Research Center of the Universidad Autónoma del Estado de Morelos (UAEM) to be properly identified, voucher numbers 15189 and 15485, respectively. From 10 seeds collected from each population, *Galphimia* spp. seedlings were obtained and cultivated at 20 °C of temperature and 60% of humidity for 2 months and then in natural conditions for 15 months. The experiment was performed from September 2019 to January 2021 in the Tissue Culture Lab at the Bioengineering Center of Instituto Tecnológico y de Estudios Superiores de Monterrey, Querétaro, México. During the cultivation period of the plants, parameters related to their growth and development were evaluated, such as height, number of lateral branches, leaf area by ellipse equation, and appearance of inflorescences.

4.2. RNA Extraction, Library Preparation, and Transcriptome Sequencing

Total RNA was extracted using RNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) from the leaves of plants originating from both populations as well as from the roots of the plants of the producer population as a control. The quality and integrity of the total RNA were evaluated using a Qsep 100 Advance, Bioptic. RNA concentration was measured using a NanoDrop 2000 spectrophotometer (NanoDrop Technologies, Technologies, Wilmington, DE, USA) and a Qubit 4^{TD} fluorometer (Invitrogen, Thermo Fisher, Waltham, MA, USA).

Six RNA-Seq libraries (leaves samples from a GP population plant, leaves samples from the NGP population, root samples from the GP population, and one replica of each sample) were constructed using the Illumina TruSeq Stranded Total RNA protocol. Roots were used as a control since the production of galphimines is negligible in this organ. The quantity and quality of enriched dscDNA were assessed using a Qubit fluorometer (Thermo Fisher, Waltham, MA, USA) and Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The libraries were sequenced using the Illumina Next Seq 550 platform; for 2 × 76 sequencing, cycles to obtain paired-end readings.

Reads Quality Control and De Novo Transcriptome Assembly

The quality of the reads obtained by sequencing was analyzed using the FastQC software (v.0.11.8) and Trimmomatic (v.0.38) tool was used to trim the sequence reads. Afterward, de novo transcriptomes assembly was carried out, for each sample, by Trinity (v.2.9.1) using the default settings mode.

4.3. Transcriptome Functional Annotation

Transcriptome functional annotation was performed using the Blast2GO (v.4.1.9). The assembled transcripts were aligned by a tblastx with the NCBI database. A BLAST expectation value of 1.0×10^{-3} was applied, and for each sequence, 20 alignments were made. A sequence mapping was conducted, followed by functional annotation in levels of gene ontology Biological Process (BP), Molecular Function (MF), and Cellular Component (CC). Sequences that were not annotated by gene ontology were identified by InterPro, UniProt, and TAIR databases.

4.4. Differential Expression Analysis

To identify corresponding gene pairs between individuals in the two populations of *Galphimia* spp., Corset programs (v.1.09) (clustering de novo assembled transcripts) was used. After transcripts were compared and grouped, results were also verified by functional annotation. An abundance quantification of the transcripts assembled in each transcriptome was conducted using the Kallisto program (v.0.46.0.4). Then, the counting matrix was built in R (v.4.0.4) for Windows, with the use of the plyr, dplyr, and stats libraries. Genes with low library counts were filtered using the CPM Filter on a per-million count (CPM) basis. For the analysis of differential expression, the edgeR library was used. Samples in the counting matrix were normalized on a transcript-per-million (TPM) basis. Transcripts expression in the leaves of the GP plants was compared with transcripts of the NGP and with the control. The differentially expressed transcripts were considered to be those with a false positive rate (FDR) < 0.01 and classified as positively regulated when log fold Change (logFC) > 2 and negative when logFC < -2 . The transcripts with the highest differential expression values in each sample were visualized on a heatmap constructed with the pheatmap library.

4.5. Identification of Putative Genes Related to the Galphimines Synthetic Pathway

Data obtained from the functional annotation of transcriptomes and the results of the differential expression analysis were used to identify genes involved in the synthesis of terpenes and putative candidate genes for the synthesis of galphimines. To identify the interactions between the genes related to the synthesis of terpenes and the possible candidate genes for the synthesis of galphimines, a co-expression network was also built, using the libraries of R, GENIE3, igraph, RCy3, and Rgraphviz. The co-expression network was calculated in the form of a weighted adjacency matrix, using ensembles of regression trees. Candidate regulatory genes were not used. The tree method was Random Forests. The number of candidate regulators randomly selected at each tree node was the square root of the total number of candidate regulators. The number of trees in each target was 1000. The co-expression network was visualized by Cytoscape (v.3.8.2).

4.6. Validation of Transcriptome Data by qPCR

cDNA of GP plants and NGP plants was obtained using the PrimerScript RT Reagent Kit with gDNA Eraser (TaKaRa, San Jose, CA, USA). The concentration and purity of cDNA were evaluated by Nano-Drop 2000 spectrophotometer (NanoDrop Technologies, Technologies, Wilmington, DE). According to transcriptome data of GP plants, nucleotide sequences of ten putative candidate genes related to galphimines synthetic pathway (CYP82C4, P450 TBP, PLAC8, 5'-AMP-activated protein kinase, serine/threonine-protein kinase AtPK2/AtPK19, P450, CYP71A24, CYP71B34, CYP72A, CYP81D), were used for primers designing (Table 3) by Primer3 plus program. The quantitative PCR was conducted

using the TB Green[®] Advantage[®] qPCR Premix (Takara Bio USA, Inc., San Jose, CA, USA) following the manufacturer's protocol in a StepOne[™] Real-Time PCR System (Applied Biosystems, Waltham, MA, USA). The amplification conditions used were as follows: activation/denaturation at 95 °C for 10 s, 45 cycles of denaturation at 95 °C for 5 s, and annealing/extension at 60 °C for 30 sec followed by a melting curve analysis ranging from 56–95 °C. Three biological replicates with three technical replicates were used for each sample. The relative expression of the RNA was quantified by the $2^{-\Delta\Delta C_t}$ method using U6 as the reference gene [34].

Table 3. Sequences of synthesized primers for qPCR.

Gene Description	Gene Symbol	Primer Sequence 5'-3'	PCR Product Length (bp)
CYP82C4	CYP82C4	F 5'TCATTGGTGGGCTAAAAGC 3' R 5'ACGAGTCCGTTAATGGTTGC 3'	191
P450 TBP	P450 TBP	F 5'AGCTTCGTCGCAAGTAAAAT 3' R 5'TAGGACGCCTGCGTTATCTT 3'	189
PLAC8	PLAC8	F 5'ACTGTTGATAACCCCGATGG 3' R 5'TCTCAGGAAGTCCGAAGTGG 3'	170
5'-AMP-activated protein kinase	AMPA	F 5'CCCAAGGAAAAGGTTTACACA 3' R 5'GGCAAAGCAATGGCTAAGA 3'	229
serine/threonine-protein kinase AtPK2/AtPK19	Ser/Thr	F 5'TTTCAACTGGACCACAAAAGG 3' R 5'CGTTCGGGAAAGTCTACCAA 3'	164
P450	P450	F 5'CCTTGAGGTCATGGCTGAAT 3' R 5'GCTTCTCTCAAAGGCACAC 3'	232
CYP71A24	CYP71A24	F 5'CAAACCGGCTAAATCAAAG 3' R 5'GCTCTTGTTCATAACTTCTCAATC 3'	200
CYP71B34	CYP71B34	F 5'CCGTGAGAGAGGCCATTAAC 3' R 5'GTGCGAAAGATTTCGCTTCT 3'	160
CYP72A	CYP72A	F 5'TTGTTGGCTTTTCGAGGAAT 3' R 5'AAGCATCAGGAGTGGCAAAC 3'	167
CYP81D1	CYP81D1	F 5'TCGGAGGATTGGACTACGAC 3' R 5'TCCGCCATAACATTCTCC 3'	219

5. Conclusions

Transcriptome analysis of two populations of *Galphimia* spp. allowed the detection of differences between a galphimine producer GP plant and a non-galphimine producer NGP plant. The higher expression of P450 family members and kinases in GP plants and their well-known role in triterpene synthesis make them putative candidates to encode enzymes involved in the galphimine synthesis pathway.

Author Contributions: Conceptualization, A.S. and M.L.V.; Methodology, D.I., A.V.V., M.d.D.C. and C.G.-G.; Software, A.M.T. and D.I.; Validation, A.S., M.L.V., M.d.D.C., S.P., H.M.N.I. and D.I.; Formal Analysis, A.S., M.L.V., M.d.D.C., A.M.T. and D.I.; Investigation, D.I., Resources, A.S., Data Curation, A.S., M.L.V., M.d.D.C., A.M.T. and D.I.; Writing-Review & Editing, A.S., M.L.V., S.P., M.d.D.C. and D.I.; Visualization, D.I., Supervision, A.S. and M.L.V.; Project Administration, A.S.; Funding Acquisition, A.S. All authors have read and agreed to the published version of the manuscript.

Funding: The authors want to thank the Regional Department of Bioengineering, South Central Region of the Tecnológico de Monterrey, for the logistics and equipment to support the conduction of this work.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Thanks to CONACYT, Mexico, for the scholarship granted to carry out Dianella Iglesias' Ph.D. thesis (946795).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tortoriello, J.; Ortega, A.; Herrera-Ruiz, M.; Trujillo, J.; Reyes-Vázquez, C. Galphimine-B modifies electrical activity of ventral tegmental area neurons in rats. *Planta Med.* **1998**, *64*, 309–313. [\[CrossRef\]](#)
2. Camacho, M.D.R.; Phillipson, J.D.; Croft, S.L.; Marley, D.; Kirby, G.C.; Warhurst, D.C. Assessment of the Antiprotozoal Activity of *Galphimia glauca* and the Isolation of New Nor-secofriedelanes and Nor-friedelanes. *J. Nat. Prod.* **2002**, *65*, 1457–1461. [\[CrossRef\]](#)
3. González-Cortazar, M.; Herrera-Ruiz, M.; Zamilpa, A.; Jiménez-Ferrer, E.; Marquina, S.; Álvarez, L.; Tortoriello, J. Anti-inflammatory activity and chemical profile of *Galphimia glauca*. *Planta Med.* **2014**, *80*, 90–96. [\[PubMed\]](#)
4. Campos, M.G.; Toxqui, E.; Tortoriello, J.; Oropeza, M.V.; Ponce, H.; Vargas, M.H.; Montaña, L.M. *Galphimia glauca* organic fraction antagonizes LTD4-induced contraction in guinea pig airways. *J. Ethnopharmacol.* **2001**, *74*, 7–15. [\[CrossRef\]](#)
5. Dorsch, W.; Wagner, H. New antiasthmatic drugs from traditional medicine? *Int. Arch. Allergy Immunol.* **1991**, *94*, 262–265. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Osuna, L.; Pereda-Miranda, R.; Tortoriello, J.; Villarreal, M.L. Production of the sedative triterpene Galphimine B in *Galphimia glauca* tissue culture. *Planta Med.* **1999**, *65*, 149–152. [\[CrossRef\]](#)
7. Romero-Cerecero, O.; Islas-Garduño, A.L.; Zamilpa, A.; Pérez-García, M.D.; Tortoriello, J. Therapeutic effectiveness of *Galphimia glauca* in young people with social anxiety disorder: A pilot study. *Evid. Based Complement. Altern. Med.* **2018**. [\[CrossRef\]](#)
8. Taketa, A.T.C.; Lozada-Lechuga, J.; Fragoso-Serrano, M.; Villarreal, M.L.; Pereda-Miranda, R. Isolation of Nor-secofriedelanes from the Sedative Extracts of *Galphimia glauca*. *J. Nat. Prod.* **2004**, *67*, 644–649. [\[CrossRef\]](#)
9. González-Cortazar, M.; Tortoriello, J.; Alvarez, L. Nor-secofriedelanes as spasmolytics, advances of structure-activity relationships. *Planta Med.* **2005**, *71*, 711–716. [\[CrossRef\]](#)
10. Ortega, A.; Pastor-Palacios, G.; Ortiz-Pastrana, N.; Ávila-Cabezas, E.; Toscano, R.A.; Joseph-Nathan, P.; Morales-Jiménez, J.; Bautista, E. Further galphimines from a new population of *Galphimia glauca*. *Phytochemistry* **2019**, *169*, 112180. [\[CrossRef\]](#)
11. Náder, B.L.; Cardoso-Taketa, A.T.; Iturriaga, G.; Pereda-Miranda, R.; Villarreal, M.L. Genetic transformation of *Galphimia glauca* by *Agrobacterium rhizogenes* and the production of Norfriedelanes. *Planta Med.* **2004**, *70*, 1174–1179. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Ortiz, A.; Cardoso-Taketa, A.; Rodríguez, M.; Arellano, J.; Hernández, G.; Villarreal, M.L. Transformed cell suspension culture of *Galphimia glauca* producing sedative Nor-friedelanes. *Planta Med.* **2010**, *76*, 386–392. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Cardoso-Taketa, A.T.; Pereda-Miranda, R.; Choi, Y.H.; Verpoorte, R.; Villarreal, M.L. Metabolic profiling of the Mexican anxiolytic and sedative plant *Galphimia glauca* using Nuclear Magnetic Resonance spectroscopy and multivariate data analysis. *Planta Med.* **2008**, *74*, 1295–1301. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Sharma, A.; Cardoso-Taketa, A.; Choi, Y.H.; Verpoorte, R.; Villarreal, M.L. A comparison on the metabolic profiling of the Mexican anxiolytic and sedative plant *Galphimia glauca* four years later. *J. Ethnopharmacol.* **2012**, *141*, 964–974. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Sharma, A.; Folch, J.L.; Cardoso-Taketa, A.; Lorence, A.; Villarreal, M.L. DNA barcoding of the Mexican sedative and anxiolytic plant *Galphimia glauca*. *J. Ethnopharmacol.* **2012**, *144*, 371–378. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Gesto-Borroto, R.; Cardoso-Taketa, A.; Yactayo-Chang, J.P.; Medina-Jiménez, K.; Hornung-Leoni, C.; Lorence, A.; Villarreal, M.L. DNA barcoding and TLC as tools to properly identify natural populations of the Mexican medicinal species *Galphimia glauca* Cav. *PLoS ONE* **2019**, *14*, e0217313. [\[CrossRef\]](#)
17. Toscano, R.A.; Maldonado, E.; Lozoya, X.; Tortoriello, J.; Ortega, A.; Gaviño, R. Structure of galphimine B. *Acta Crystallogr. Sect. C Cryst. Struct. Commun.* **1993**, *49*, 774–776. [\[CrossRef\]](#)
18. Ma, C.H.; Gao, Z.J.; Zhang, J.J.; Zhang, W.; Shao, J.H.; Hai, M.R.; Chen, J.W.; Yang, S.C.; Zhang, G.H. Candidate genes involved in the biosynthesis of triterpenoid saponins in *Platycodon grandiflorum* identified by transcriptome analysis. *Front. Plant Sci.* **2016**, *7*, 1–14. [\[CrossRef\]](#)
19. Srivastava, G.; Sandeep; Garg, A.; Misra, R.C.; Chanotiya, C.S.; Ghosh, S. Transcriptome analysis and functional characterization of oxidosqualene cyclases of the arjuna triterpene saponin pathway. *Plant Sci.* **2019**, *292*, 110382. [\[CrossRef\]](#)
20. Lertphadungkit, P.; Qiao, X.; Sirikantaramas, S.; Satitpatipan, V.; Ye, M.; Bunsupa, S. De novo transcriptome analysis and identification of candidate genes associated with triterpenoid biosynthesis in *Trichosanthes cucumerina* L. *Plant Cell Rep.* **2021**, *40*, 1845–1858. [\[CrossRef\]](#)
21. Rios, M.Y.; Ortega, A.; Domínguez, B.; Déciga, M.; de la Rosa, V. Glaucacetalin E and galphimidin B from *Galphimia glauca* and their anxiolytic activity. *J. Ethnopharmacol.* **2020**, *259*, 112939. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Anderson, C. Revision of *Galphimia* (Malpighiaceae). *Contrib. Univ. Mich. Herb.* **2007**, *25*, 1–82.
23. Ghosh, S. Triterpene Structural Diversification by Plant Cytochrome P450 Enzymes. *Front. Plant Sci.* **2017**, *8*, 1886. [\[CrossRef\]](#)
24. Bathe, U.; Tissier, A. Cytochrome P450 enzymes. *Phytochemistry* **2019**, *161*, 149–162. [\[CrossRef\]](#)
25. Geisler, K.; Hughes, R.K.; Sainsbury, F.; Lomonosoff, G.P.; Rejzek, M.; Fairhurst, S.; Olsen, C.E.; Motawia, M.S.; Melton, R.E.; Hemmings, A.M.; et al. Biochemical analysis of a multifunctional cytochrome P450 (CYP51) enzyme required for synthesis of antimicrobial triterpenes in plants. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 3360–3367. [\[CrossRef\]](#)

26. Zhou, J.; Hu, T.; Gao, L.; Su, P.; Zhang, Y.; Zhao, Y.; Chen, S.; Tu, L.; Song, Y.; Wang, X.; et al. Friedelane-type triterpene cyclase in celastrol biosynthesis from *Tripterygium wilfordii* and its application for triterpenes biosynthesis in yeast. *New Phytol.* **2019**, *223*, 722–735. [[CrossRef](#)]
27. Cheng, Y.; Liu, H.; Tong, X.; Liu, Z.; Zhang, X.; Li, D.; Jiang, X.; Yu, X. Identification and analysis of CYP450 and UGT super-gene family members from the transcriptome of *Aralia elata* (Miq.) seem to reveal candidate genes for triterpenoid saponin biosynthesis. *BMC Plant Biol.* **2020**, *20*, 1–17. [[CrossRef](#)]
28. Souza-Moreira, T.M.; Alves, T.B.; Pinheiro, K.; Felipe, L.G.; De Lima, G.M.A.; Watanabe, T.F.; Barbosa, C.C.; Santos, V.A.F.F.M.; Lopes, N.P.; Valentini, S.R.; et al. Friedelin Synthase from *Maytenus ilicifolia*: Leucine 482 Plays an Essential Role in the Production of the Most Rearranged Pentacyclic Triterpene. *Sci. Rep.* **2016**, *6*, 1–3. [[CrossRef](#)]
29. Tholl, D. Biosynthesis and biological functions of terpenoids in plants. In *Biotechnology of Isoprenoids. Advances in Biochemical Engineering/Biotechnology*, 1st ed.; Chrader, J., Bohlmann, J., Eds.; Springer: Luxemburgo, 2015; Volume 148, pp. 63–106.
30. Booth, J.K.; Page, J.E.; Bohlmann, J. Terpene synthases from *Cannabis sativa*. *PLoS ONE* **2017**, *12*, e0173911. [[CrossRef](#)]
31. Lee, S.; Badieyan, S.; Bevan, D.R.; Herde, M.; Gatz, C.; Tholl, D. Herbivore-induced, and floral homoterpene volatiles are biosynthesized by a single P450 enzyme (CYP82G1) in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 21205–21210. [[CrossRef](#)]
32. Hartl, C.L.; Ramaswami, G.; Pembroke, W.G.; Muller, S.; Pintacuda, G.; Saha, A.; Parsana, P.; Battle, A.; Lage, K.; Geschwind, D.H. Coexpression network architecture reveals the brain-wide and multiregional basis of disease susceptibility. *Nat. Neurosci.* **2021**, *24*, 1313–1323. [[CrossRef](#)] [[PubMed](#)]
33. Itoh, D.; Kawano, A.K.; Nabeta, K. Biosynthesis of Chloroplastidic and Extrachloroplastidic Terpenoids in Liverwort Cultured Cells: ¹³C Serine as a Probe of Terpene Biosynthesis via Mevalonate and Non-mevalonate Pathways. *J. Nat. Prod.* **2003**, *66*, 332–336. [[CrossRef](#)] [[PubMed](#)]
34. Tang, F.; Chu, L.; Shu, W.; He, X.; Wang, L.; Lu, M. Selection and validation of reference genes for quantitative expression analysis of miRNAs and mRNAs in Poplar. *Plant Methods* **2019**, *15*, 35. [[CrossRef](#)] [[PubMed](#)]

Cuernavaca, Morelos, a 24 de noviembre del 2022.

VOTOS APROBATORIOS DE TESIS

Los integrantes de la Comisión Revisora de la tesis titulada: **"Comparación de transcriptomas de dos poblaciones de *Galphimia* spp. para identificar posibles genes candidatos que codifican enzimas de la ruta de síntesis de las galfiminas"**, que presenta la **C. Dianella Iglesias Rodríguez**, del Programa de Posgrado Doctorado en Ciencias Naturales, bajo la dirección de la DRA. MARÍA LUISA TERESA VILLARREAL ORTEGA y codirección del DR. ASHUTOSH SHARMA, han determinado que el documento reúne los requisitos académicos para su defensa oral en el examen de grado, por lo que emiten su **VOTO APROBATORIO**.

Comisión Revisora de tesis (firma electrónica)

Dra. María Luisa Teresa Villarreal Ortega (directora de tesis, CEIB UAEM)

Dr. Ashutosh Sharma (codirector de tesis, ITESM)

Dr. Alexandre Toshirrico Cardoso Taketa (CEIB UAEM)

Dr. Jorge Luis Folch Mallol (CEIB UAEM)

Dr. Alfonso Méndez Tenorio (IPN)

Dr. Rogelio Gregorio Pereda Miranda (UNAM)

Dr. Ramón Suárez Rodríguez (CEIB UAEM)



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

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

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

Sello electrónico

MARIA LUISA TERESA VILLARREAL ORTEGA | Fecha:2022-11-24 11:06:35 | Firmante
YJskou2DUURsAKev7Qlyk+HT566a67LSEMyJ3ql5J6ySZMRTa4Fs74gr1Nx0vXAlrdoDR50EYRfsljVydOLcSXu6T/VE84xyOZC++quUCDDW4fTYTfWrnViUNoComAINb3hE3xv
u1aYUwLrnyPS1OuQHa5qwK6ORGfK3OSciuvlCxFc808YAc3DwR81JcYlejqr+ZjgKTQZ3Gs+1WpXcjmXclIK7MiKUKlAfJnKjKCCMUjczTzVL0z5uHSEYAtld1jqBCh6opGq
m3Vh3DDDDGmNJO9ZHYbevppk7XwurSIY0pwwXarQHTNegJ09+2qjETwlsGuspgNbiA==

JORGE LUIS FOLCH MALLOL | Fecha:2022-11-24 11:19:25 | Firmante
CkQks8Lic1vnPj6kxvqNcdEj3g+NHxHCvzhIO7ONPB2nHOJiRwa8dJ1ITz+hbFrUAe/IsyJaj3Jt2xHzp8YfpFdPru+a5/mF9XIt4fqikSLRfSLBXx3KlVn1GLi6u57ctt1jDa2qmxF05BrVh
Y2reJp/+Tstj3fKaRUeSupQ3smDXvbs9Jd5jfhD77G9c+M3rVxaj+c20tUsaxwKNSjdDA0fEJLeF7Cxu9Zop2bXbBBmY3ueKdthk99f3zXaD Tms7j5/oHXm4u63PgUHXtMGUgCiPL
KwVshPYJM2TOcpev2cVJqonrSE7eCs3cmkQ2O236xq0suk78honbviDw==

RAMON SUAREZ RODRIGUEZ | Fecha:2022-11-24 11:32:31 | Firmante
BduPTFQaYezrwO2uANA3bjdVe6EKQk2AjEuxwiPS3OroaloFR+MnRmc65RCXoA4SfE4yonbWsJQdaN7/LO7xhvaWYvOxnwmtiaSBIL4sZZ1APlv3Dvu9mxizL968E/iinkSDD7jh
sistScO4VsbQQSS2BoyNjtlp0wR0EgNx/74loxiv8rfW6FYQNOZYVvncbDREYgxmH659wRZixQUyBSZCfm6lrnSlgBGSp0tZfttqLbOFh9grM6AGpivGh6cR5jcBEw/hXwQ+zk8g
Mgw170KIY9CFNfhtVkwUKA+SJnkQYw0+eiPCNICo4BsT4fpN0o+mR5gR9ThB8dg==

ROGELIO GREGORIO PEREDA MIRANDA | Fecha:2022-11-24 13:21:50 | Firmante
eCe90ytupzDvBlq+78R8O86sZTxL13cwyit/5Mu6/a/43YWXaO2o30cJhvWqQofuHTM7dL+wWwO1C8SHDIVU1fhu8uPvCiRgXrHbYH5gOUBXJ0PdBUbbJgEIK7cTb0pnjDpJ0p3K
juS+xr7Pjvh7f7MsGlrcFPTZXwdN23iZ3Ck8HJF0z17LBmiBaOPHTOonySG9i61BZAQ0wkERPwkeD8CMnxbwiz547j81uMvLacSqyYOG+/fUf/NFcv2AH6mjcB/AZfzHc4IEObH+V
ZuVMCOdlwVaJiCUCSNY13TqZRIFFtKgy5i/WnP/i1OWISs8MP09EqqRVzknjLow==

ALEXANDRE TOSHIRICO CARDOSO TAKETA | Fecha:2022-11-25 11:11:24 | Firmante
eLFZ4fpeOZlrY73JObdwUdyUMxJ9rvLXev+PRJmtyZoP37O9XwvmmE6zkgGpmkflw+/d/a58fwfR18bGhkl0uerDskatm2Y0gjcUdAw3jm9PoWy1zj36yrkW7IAbmhr/QGSYBNT40
aUV62GH94wj9CZDHTbRS1Hnlm2SD1z7U9Uubz8P7pXWgv0RaMgU56syASSDDMnpua59zxxujC8Uz7KqozRqF6UHQJ3i8ZwQv5JuiE8YPMqPTZZwb4TAnZq7R4h83ZmUk
ygEzvwWqGWwp5utesu9Hj+vD+OiwSMKD3nRHae1S9xpHseD4VzscbQHHeUSKsoc+X3Yr9SoDQ==

ASHUTOSH SHARMA | Fecha:2022-11-26 08:46:57 | Firmante
wr+w7dLYc45+M6k8Bqj2f/3ofCEeX7hU9oSWKpXvVpVUllsb2w7tq26x6Hn8NVStSMIKDoNgQIXEpxmiOqmBSq2CX6BunxbgT+x2HYwx/Mn+LwYSyTcPE4hemg2R7K/QJVFJ
UbWJKkWTsSo/acdpYS3w0hkuPHIUOb6nc3hpdBrOlnnYNsYnBYHBDVwmNYfmltzp0PgdgkleyZ3TE7j43FgoWePsc+CMlZK32UNNoxytl2/1aJsSVKfifVdch7ooZOPYKYARS
MPmnUxgRsiXHYwXaJp5SZUocqZpViiXqUvCWNjw7LbE83xN7UAENv0/qw08nPi3JTGd4A==

ALFONSO MÉNDEZ TENORIO | Fecha:2022-11-26 16:41:05 | Firmante
DFyVJ7qpn0DniekGP6PZCB39SZWxBmbR0PfiUbu0Yk8M89pydW3qx7hNCqKeg8HrKNEbQA4IHehA5v6QqJc9YTtXZTLcYxDgwJLxUYfd2VCLJjGHASdrDznVZ3PrTPU5iQ4
MtyrxY6N6xZbjz1G3kGFwHEf701FKmY7M1xHJGNFfBqfD3PhpbHOuXKp97sy1ol4H9VAUjRkXyVXaq2Mp6nj3gt7m4eq2mmF+9bhpZKqNoGialUPNhrXc8cnX0MbdKyLtQzCw
4IDJRit2WMrwiAt+bJOZwJcUYLLgyKvTBab03V/qWUrhAnjOk7jp6il0nWGLA0msdkJAyZQuQ==

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



L310xYhG

<https://efirma.uaem.mx/noRepudio/ifaV2Qv11VC7ni3Y3CjOsbPEKfQGUZam>



