UNIVERSIDAD
NACIONAL
DE COLOMBIA

Identificación de genes asociados a la producción de withanólidos en Physalideae

Santiago Perez Moreno

Universidad Nacional de Colombia Facultad de Ciencias, Departamento de Biología

Bogotá, Colombia 2024

UNIVERSIDAD
NACIONAL
DE COLOMBIA

Identification of Genes Associated with the Production of Withanolides in Physalideae

Santiago Perez Moreno

Universidad Nacional de Colombia Facultad de Ciencias, Departamento de Biología

Bogotá, Colombia 2024

Identificación de genes asociados a la producciónde withanólidos en Physalideae

Santiago Perez Moreno

Tesis o trabajo de investigación presentada como requisito parcial para optar al título de:

Magister en Ciencias Biología

Director (a):

Ph.D. en Biología, Federico Roda Fornaguera

Línea de Investigación:

Genómica Evolutiva

Grupo de Investigación:

Genómica Evolutiva del Metabolismo Secundario (GEME)

Universidad Nacional de Colombia Facultad de Ciencias,Departamento de Biología

Bogotá , Colombia 2024

A mi Mantis y Tuoh, por su amor incondicional, su sacrificio y dedicación para darme la mejor educación posible, gracias por ser mis guías y por enseñarme los valores que me han llevado hasta aquí.

A la Abuelita Te, el Bibito Bose, la Abuelita Bibi y el Abuelito Jose por sus sabios consejos y sus historias que han nutrido mi mente y mi corazón. Algunos de estos legados no fueron entregados directamente a mi, pero la esencia de sus seres siempre permanece y permanecerá en los que los amamos.

A mi hermano, por estar conmigo desde el principio y ser mi gemelo pequeño, gracias por las risas, las bromas y las peleas.

A mi Aniaceooo, quiero dar gracias por ayudarme a ser la persona que soy. Por creer en mi, por mostrarme que siempre hay una mejor manera de ser. Por las cuachabubbles aventuras que hemos tenido y tendremos, y sobre todo, por nuestro amor. Te amo

Y a mis amigos, por ser mi fuente de energía, por acompañarme en momentos felices y difíciles, y por brindarme su apoyo incondicional, en el baile y en la vida

Esta tesis es el resultado de mucho esfuerzo y dedicación, pero no habría sido posible sin ustedes. Espero que esta dedicatoria sea una pequeña muestra de mi gratitud hacia cada uno de ustedes.

Con todo mi amor y agradecimiento,

Declaración de obra original

Yo declaro lo siguiente:

He leído el Acuerdo 035 de 2003 del Consejo Académico de la Universidad Nacional. «Reglamento sobre propiedad intelectual» y la Normatividad Nacional relacionada al respeto de los derechos de autor. Esta disertación representa mi trabajo original, exceptodonde he reconocido las ideas, las palabras, o materiales de otros autores.

Cuando he presentado ideas o palabras de otros autores en esta disertación, he realizado su respectivo reconocimiento aplicando correctamente los esquemas de citas y referencias bibliográficas en el estilo requerido.

He obtenido el permiso del autor o editor para incluir cualquier material con derechos de autor (por ejemplo, tablas, figuras, instrumentos de encuesta o grandes porciones detexto).

Por último, he sometido esta disertación a la herramienta de integridad académica, definida por la universidad.

Santiago Perez Moreno

Fecha 5/10/2023

Agradecimientos

El autor expresa su gratitud al Dr. Federico Roda Fornaguera por su guía y apoyo en el desarrollo de este trabajo. También se agradece al candidato a PhD Pablo Andrés Pérez Mesa, a la candidata a PhD Gina Paola Sierra Reyes y a la Magister Sofía Rojas Contreras por su colaboración en las extracciones de RNA y metabolitos, así como en la identificación taxonómica de las plantas y el uso de software para la identificación de compuestos. Asimismo, se agradece a la PhD Rocío Deanna y a la PhD Stacey Smith por su contribución con especies asociadas al estudio y guía en la realización del presente trabajo. Además, se agradece a la PhD Eva Knock por su investigación en withanolidos, que sirvió como guía para el desarrollo de este trabajo y por compartir sus conocimientos en el área. También se agradece al grupo de investigación de Genómica del Metabolismo Especializado (GEME) de los grupos Tandem de Max Planck por su apoyo económico y por dar oportunidades de surgir a los investigadores de Colombia. Asimismo, se agradece al agrónomo Sixto Javier Sandoval Suárez y a la profesora María Cecilia Delgado Niño por su colaboración en el cuidado de los especímenes y la asignación de finanzas y asuntos jurídicos, respectivamente. Finalmente, el autor expresa su agradecimiento a su familia y allegados por su apoyo fundamental en el inicio, desarrollo y término de este trabajo.

Resumen

Identificación de genes asociados a la producción de withanólidos en *Physalideae*

Esta tesis explora la biosíntesis y regulación de los withanolidos en la familia Solanaceae, específicamente en la tribu Physalideae, arrojando luz sobre los complejos procesos bioquímicos y genéticos que conforman su diversidad y evolución. El estudio analiza cómo se sintetizan los withanolidos, destacando la importancia de procesos bioquímicos específicos y la interacción de rutas metabólicas clave. Se enfatiza el papel crucial de ciertos procesos biológicos y genes en la modulación de la síntesis de estos compuestos, resaltando la significativa coexpresión de genes vinculados a funciones esenciales en la planta.

El papel de las enzimas citocromo P450, genes específicos y clústeres biosintéticos como el propuesto WBC se examinan en profundidad, proporcionando una comprensión más profunda de la regulación genética en la producción de withanolidos. La investigación también revela interesantes interacciones entre diferentes vías metabólicas y muestra cómo la variación genética podría influir en la presencia o ausencia de ciertos withanolidos entre las especies de Solanaceae.

Finalmente, la tesis ofrece una mirada a la regulación y los factores de transcripción involucrados en la producción de withanolidos, subrayando el descubrimiento de patrones de expresión genética que podrían ser fundamentales para futuras investigaciones en biotecnología vegetal y desarrollo farmacológico. Este trabajo representa un avance significativo en el conocimiento de la fitoquímica y la biología evolutiva de las plantas, destacando la complejidad y el dinamismo del mundo vegetal.

Palabras clave: Withanolidos, Biosíntesis de withanolidos, Genómica vegetal, Solanaceae, Metabolitos secundarios, Evolución metabólica.

Abstract

This thesis explores the biosynthesis and regulation of withanolidos in the Solanaceae family, specifically in the Physalideae tribe, shedding light on the complex biochemicaland genetic processes that shape their diversity and evolution. The study examines howwithanolidos are synthesized, highlighting the importance of specific biochemical processes and the interaction of key metabolic pathways.

The crucial role of biologicalprocesses and genes, particularly cytochrome P450 enzymes, specific genes, and proposed biosynthetic clusters like WBC, is examined in depth, providing a deeper understanding of genetic regulation in withanolide production. The research also revealsintriguing interactions between different metabolic pathways and shows how genetic variation might influence the presence or absence of certain withanolidos among Solanaceae species.

Finally, the thesis offers insights into the regulation and transcription factors involved in withanolide production, emphasizing the discovery of gene expression patterns that could be key for future research in plant biotechnology and pharmacological development. This work represents a significant advancement in the understanding of plant phytochemistry and evolutionary biology, highlighting the complexity and dynamism of the plant world.

Keywords: Withanolides, Withanolide Biosynthesis, Plant Genomics, Solanaceae, Secondary Metabolites, Metabolic Evolution.

Contenido

Pág

1. Capítulo 1. Metodología.....26
1.1 Búsqueda Bibliográfica y Análisis de Datos……..……..27
1.2 Metabolomica……28
1.3 Analisis de Datos Metabolomicos…….29
1.4 Selección de Plantulas y Preparación.....31
1.5 Analisis Transcriptomicos y Filogeneticos 33
1.6 Cumplimiento Ético e Institucional……...39

2. Capítulo 2: Resultados 40
2.1 Resultados Metabolomicos 40
2.2 Analisis de Transcriptomas y Expresion Genica.....50
2.3 Analisis Filogenetico y Clasificacion de Especies 55
2.4 Integracion de datos y correlaciones.....53

3. Capítulo 3:* Discusión *73
3.1 Selección de Plantulas y Preparacion.....73
3.2 Metabolómica……74
3.3 Analisis Transcriptomicos y Filogeneticos 76
3.4 Integración de Metabolómica, Transcriptómica y Filogenética 84

4. Capítulo 4: Conclusiones 114

5. Capítulo 5: Recomendaciones 116

6. Capítulo 6: Bibliografia.....118

7. Capítulo 7: Anexos.....128

Lista de Figuras

Figura 1: Vía MEP y MVA con compuestos, y genes asociados.....22
Figura 2: Determinación de presencia de genes del cluster en especies de Solanaceae.....25
Figura 3. Espectro de masas MS -MS de Withaferin A en Physalis_peruviana..... 41
Figura 4. Espectro de Masas MS -MS de 1,3,27-trihydroxy-(20S,22R)-3- methoxywitha-4,6,24-trienolide en Physalis peruviana……..... 41
Figura 5: Figura 5. Espectro de masas MS -MS de Baimantuoluoside B en Withania somnífera ……..42
Figura 6. Espectro de masas MS -MS de 12-Deoxywithastramonolide en Cuatresia riparia…….43
Figura 7: Espectro de masas MS -MS de 27 hydroxywithanone Cuatresia riparia…….....44
Figura 8. Distribución de Withanolidos en Especies de Solanaceae…..... 47
Figura 9. Distribución de Withanolidos en Especies de Solanaceae adicionales..... 48
Figura 10. Análisis de diferenciación metabólica entre clados de plantas mediante t- score y t-score ortogonal..... 49
Figura 11. Filogenia de los parálogos de DWF1 en los genomas de Solanaceae……..... 54
Figura 12. Filogenia de los parálogos de DWF1 en los transcriptomas de Solanaceae..... 55
Figura 13. Filogenia de ASTRAL con los genes de copia única..... 58
Figura 14. Mapa de calor de correlación de expresión génica entre especies..... 57
Figura 15. Histograma de Normalización de EdgeR Withanólidos 60
Figura 16. Gráfico de Dispersión de Variabilidad Biológica (BCV) vs. Promedio Log CPM……61
Figura 17. Determinación de la independencia de escala y conectividad media para diferentes umbrales suaves (potencia)……65
Figura 18. Cluster Dendogram asociado a 'Log_cpm_filtered' y 'cpm_filtered TPM'..... 66
Figura 19. Determinación de la independencia de escala y conectividad media para diferentes umbrales suaves (potencia)..... 67
Figura 20. Comparación de la independencia de escala y la conectividad media en redes de co-expresión génica basadas en datos 'log_cpm_filtered' y ' cpm_filtered_TPM'……..... 68
Figura 21. Análisis de Expresión Génica para Biosíntesis de Withanolidos en Especies de Solanaceae……..... 70

Lista de tablas

Pág.

Tabla 1: Genes asociados al cluster biosintético de Knoch et al, 2018 y número de copias dentro de tres ejemplares de la tribu Physalideae 24
Tabla 2: Estadísticas de lecturas crudas y filtradas por Trimmomatic.....52
Tabla 3. Estadísticas de ensamblajes..... 53

Introducción

En la vasta y diversa familia Solanaceae, conocida por albergar miembros notables como el tomate (*Solanum lycopersicum*), la papa (*Solanum tuberosum*), el pimiento y chile (*Capsicum spp.*), la berenjena (*Solanum melongena*), y plantas con interesantes propiedades medicinales como el tabaco (*Nicotiana tabacum*), se encuentra la tribu Physalideae (Xu & Chang, 2017). Esta tribu destaca por su extraordinaria diversidad morfológica y bioquímica. Con aproximadamente 300 especies distribuidas en 29 géneros, esta tribu no solo es notable por su riqueza en formas y funciones, sino tambiénpor su significativo papel en la medicina, tanto tradicional como moderna (Deanna *et al*, 2019).

Ashwagandha, conocida como *Withania somnífera*, ha sido utilizada durante siglos en la medicina ayurvédica para una variedad de propósitos y es uno de los miembros más estudiados de la tribu (Chang *et al*. 2007) (Chandra, 2013). Se le atribuyen propiedades adaptogénicas, que ayudan al cuerpo a resistir los efectos del estrés y a promover un equilibrio general del sistema. Además, se utiliza para tratar condiciones como la debilidad, el agotamiento, la fatiga y la disminución de la libido. Los estudios modernos han respaldado varias de estas aplicaciones tradicionales, encontrando evidencia de que los extractos de ashwagandha y sus withanólidos pueden tener efectos antiinflamatorios, antioxidantes, ansiolíticos, y potencialmente antitumorales (Sukanya *et al*, 2010).

En el contexto de sus propiedades defensivas, es crucial explorar su impacto en organismos como bacterias e insectos. Los withanólidos han demostrado poseer propiedades antibacterianas significativas, lo que sugiere un mecanismo de defensa en las plantas contra patógenos bacterianos (Misico et al, 2011). Estos compuestos interfieren con las funciones bacterianas vitales, afectando principalmente a bacterias Gram-positivas. Se ha observado que los withanólidos alteran la permeabilidad de la membrana bacteriana, lo que conduce a la desestabilización y eventual ruptura de esta. Este mecanismo de acción podría estar asociado con la interrupción de procesos esenciales como la síntesis de proteínas y el metabolismo celular (Lobatto *et al*, 2023).

En el caso de los insectos, los withanólidos actúan como potentes agentes de disuasión, interfiriendo en su desarrollo y supervivencia. Estos compuestos ejercen un efecto tóxico en varios insectos, incluyendo aquellos que son plagas de cultivos importantes. La toxicidad de los withanólidos contra los insectos se debe a su capacidad para desequilibrar procesos fisiológicos clave como la regulación hormonal, el metabolismo y la función del sistema nervioso. En algunos insectos, los withanólidos pueden inhibir la alimentación y el crecimiento, lo que resulta en una disminución significativa de su población (Barthel *et al*, 2016; Huang *et al*., 2020).

Por otro lado, en la compleja tribu Physalideae, la clasificación interna ha revelado la existencia de tres subtribus distintas: Physalidinae, Iochrominae y Withaninae. Cada una de estas subtribus representa un conjunto de linajes que, aunque comparten la producción de withanólidos, presentan distintas rutas evolutivas y adaptaciones ecológicas (Fan et al., 2020). La subtribu Physalidinae es conocida por su diversidad de metabolitos secundarios, lo que sugiere una especialización en sus rutas biosintéticas (Deanna et al., 2019). Iochrominae, por otro lado, ha sido objeto de una reestructuración taxonómica, lo que indica una diversificación continua de los linajes existentes. Withaninae, aunque no es la única subtribu que no es monofilética, fue recientemente ampliada para incluir otros géneros (e.g.,

Tuberowithania, Wang et al., 2020) y refleja la dinámica de clasificación y evolución dentro de la tribu. Sin embargo, aún hay incertidumbre en las relaciones filogenéticas, especialmente dentro de Physalidinae, señalando la necesidad de investigación adicional para dilucidar completamente la historia evolutiva de estos grupos.

Como se menciona en el anterior párrafo, Las subtribus de *Physalideae* se distinguen por la producción de distintos tipos de withanólidos. La subtribu *Physalidinae* producephysalins junto con otros withanólidos que mantienen un esqueleto sin modificar(Misico et al., 2011). Iochrominae también genera compuestos con un esqueleto sin modificar, pero se distingue por la producción de ancistrins y withaphysalins (Huang etal., 2020). Withaninae, por su parte, comparte la producción de ancistrins con *Iochrominae*, pero no los withaphysalins o physalins característicos de las otras subtribus. Esto refleja la diversidad bioquímica entre las subtribus y su posible adaptación a distintas funciones ecológicas y biológicas.

En cuanto a los aspectos evolutivos asociados con los withanólidos en la tribu *Physalideae*, la investigación se ha enfocado principalmente en su biosíntesis y diversificación estructural más que en su historia evolutiva específica (Fan et al., 2020).Sin embargo, algunos estudios proporcionan información valiosa aun no asociada a loswithanolidos. Por ejemplo, se ha observado que la diversificación de la subtribu *Physalidinae* en la tribu *Physalideae* está marcada por notables cambios en su cariotipo (Deanna et al, 2018). Con un número cromosómico base de x = 12, la mayoría de las especies de esta subtribu exhiben cariotipos asimétricos. Se han identificado varios eventos de poliploidía y un caso de aneuploidía en *Quincula* (X = 11). Estos cambios cariotípicos, incluyendo variaciones en cromosomas telocéntricos y submetacéntricos, son cruciales para distinguir entre géneros dentro de la subtribu y reflejan su proceso evolutivo.

Continuando con *Iochrominae*, se resalta la complejidad de la clasificación basada en morfología y se sugiere que eventos de hibridación, junto con la selección mediada por polinizadores, podrían haber contribuido a la diversificación en esta subtribu (Deanna et al, 2019). Además, se indica que la subtribu Withaninae, aunque no tan diversa en número de especies como Physalidinae, presenta una mayor diversidad floral (Misico etal, 2011). Estos elementos subrayan la importancia de las interacciones ecológicas y lavariabilidad ambiental en la evolución de la subtribu.

Considerando la parte fisiológica, es importante observar que la biosíntesis de los withanólidos en la tribu *Physalideae* es un fenómeno que destaca por su regulación específica, tanto espacial como temporal, dentro de los tejidos de las plantas (Agarwal et al., 2017; Gupta et al., 2015). Esta regulación indica que la producción de estos compuestos no se distribuye de manera uniforme en todos los tejidos vegetales ni durante todas las fases del desarrollo de la planta. Por ejemplo, las hojas y las raíces suelen presentar concentraciones más elevadas de withanólidos, especialmente durante etapas específicas del ciclo de vida, como la floración y la fructificación. Además, factores ambientales como la luz, temperatura, disponibilidad de agua y nutrientes también influyen significativamente en la síntesis de estos compuestos. Estrés abiótico, como sequía o exposición a patógenos, puede aumentar la producción de withanólidos como mecanismo de defensa (Agarwal et al., 2017; Gupta et al., 2015). Este fenómeno ha sido documentado en estudios que muestran que condiciones como la luz y la temperatura afectan la concentración de withanólidos en *Withania somnifera* (Dhar et al., 2013; Zhang et al., 2011). La precisión en la regulación refleja una estrategia optimizada por parte de la planta para el uso de sus recursos en la síntesis de estos metabolitos secundarios.

Por otro lado, se ha observado una distinción clara en la presencia de grupos específicos de withanólidos en diferentes géneros. El género Physalis se caracteriza por una amplia gama de withanólidos, incluidos physalins, neophysalins, cyclophysalins y withaphysalins, todos con esqueletos sin modificar, reflejando una diversidad química significativa (Huang et al., 2019). Por otro lado, los géneros Witheringia y Brachistus comparten una asociación con physalins que también presentan esqueletos sin modificar, lo que sugiere un perfil de withanólidos más restringido. Cabe mencionar que el género Margaranthus, anteriormente considerado separado, hace mucho que ya es considerado parte de Physalis, específicamente como Physalis solanacea.

En cuanto a Acnistus, este género se distingue por contener tanto ancistrins como withaphysalins, junto con un esqueleto sin modificar, indicando una combinación única de metabolitos secundarios (Alfonso & Kapetanidis, 1994). Sin embargo, cabe mencionar que Acnistus es ahora considerado parte del género Iochroma, específicamente como Iochroma arborescens (Shawn, 2018). Los géneros Iochroma y Eriolarynx, aunque ambos presentan withaphysalins, difieren en la estructura de sus esqueletos, con Iochroma mostrando un esqueleto sin modificar (Alfonso & Kapetanidis, 1994) y Eriolarynx un tipo ligeramente diferente (Misico et al., 2011). Vassobia y Dunalia exhiben una mezcla de ancistrins y withaphysalins con esqueletos sin modificar (Erazo et al., 2008; Espinoza et al., 2012).

Withania y Tubocapsicum, mientras tanto, se asocian con estructuras de esqueletos sin modificar y ancistrins, respectivamente. En el caso de Deprea y las dos secciones del género Datura, se destaca la presencia de subtrifloralactones y withametelins, además de esqueletos sin modificar, señalando una especialización en su ruta biosintética de withanolidos (Gupta et al., 2015; Sharma et al., 2020). Este patrón específico en la biosíntesis de withanolidos sugiere un proceso evolutivo diferenciado en estas plantas, enfocado en la producción de compuestos con estructuras particulares y funciones biológicas únicas.

Finalmente, Jaborosa muestra una complejidad química con compuestos que incluyen un anillo aromático A y espiroanoid-γ-lactonas, además de un esqueleto sin modificar, lo que indica una química única de withanolidos en este género (Lobatto et al., 2023). Por otro lado, Nicandra y Salpichroea comparten la característica de tener un anillo aromático D en sus withanolidos, también con esqueletos sin modificar, reflejando similitudes en su biosíntesis de withanolidos (Misico et al., 2011). Schraderanthus se distingue por la presencia de physalins, una clase diferenciada de withanolidos (Fang etal., 2009). Esta diferenciación de withanolidos en los géneros mencionados indica una variabilidad significativa y posiblemente refleja adaptaciones evolutivas específicas dentro de la subfamilia Solanoideae.

Si bien los withanólidos son muy diversos en la tribu *Physalideae* y muestran especificidad para ciertas subtribus, géneros y especies, su estudio enfrenta una brecha significativa en nuestro entendimiento de cómo se producen diferentes tipos de estos metabolitos. Esta diversidad se extiende más allá de la familia Solanaceae, detectándoseen varias otras familias de plantas, lo que subraya la complejidad y la diversidad de estoscompuestos y su posible importancia en una variedad de contextos biológicos y farmacológicos. Por ejemplo, los withanólidos se encuentran en los rizomas de especiesde *Tacca* J.R. Forst. & G. Forst. (familia *Dioscoreaceae*), en las partes aéreas de *Sennasiamea* (familia *Leguminosae*), en la corteza de *Eucalyptus globulus* (familia *Myrtaceae*) y en especies de Ajuga L. (familia *Lamiaceae*) (Agarwal et al. 2018; Changet al. 2007).

Dentro del marco de mi investigación, es crucial identificar los genes asociados con la producción de withanólidos, ya que esto podría permitir una manipulación genética dirigida para aumentar la producción de estos compuestos valiosos. Entre los genes identificados se encuentra la 3-Hidroxi-3-metilglutaril coenzima A reductasa (HMGR), que es el paso limitante de la tasa de producción y un objetivo común para las drogas reductoras del colesterol como las estatinas (Bharadwaj et al., 2021). Los estudios han demostrado una correlación positiva entre los altos niveles de transcripción de HMGR y la acumulación

óptima de withanólidos en el tejido radicular de W. somnifera, sugiriendo una función esencial de HMGR en la biosíntesis de precursores clave como el IPP y DMAPP (Chaurasiya et al., 2012).

Otro gen crucial es la squalene synthase (WsSQS), que cataliza la formación deescualeno, un precursor vital en la biosíntesis de fitoesteroles y withanólidos (Razdan et al., 2017). Además, la squalene epoxidase (WsSQE) es responsable de la epoxidación del escualeno, un paso esencial en la vía biosintética del esterol (Sharma et al., 2020). Los oxidosqualene cyclases (OSCs) como WsOSC/LS, WsOSC/BS y WsOSC/CSjuegan un papel importante en la formación de diversos esqueletos triterpenoides, que son intermediarios cruciales para la diversidad estructural de los withanólidos (Shilpashree et al., 2022). La expresión de estos genes es modulada en respuesta a estímulos abióticos y bióticos, lo que indica la existencia de elementos reguladores en sus regiones promotoras que podrían ser claves en la regulación de la biosíntesis de withanólidos.

Sin embargo, muy pocos estudios han correlacionado los perfiles metabólicos y genómicos de múltiples especies en condiciones controladas (Schirmer, 2014). Esto impide las comparaciones directas entre especies, dado el efecto grande que tiene el ambiente sobre losgenes que se expresan y los metabolitos asociados a estos. Dichas comparaciones pueden darse desde un punto de investigación metabolómico y transcriptómico, siendo importantes estos campos, ya que al hablar de metabolitos secundarios es posible identificar cambios enla estructura de estos que se correlacionena expresión de genes antes no conocidos. Todo esto se ve impulsado por el hecho de que algunos géneros dentro del grupo han sido poco estudiados a un nivel genómico (i.e., Cuatresia, Quincula, Chamaesaracha).

Además, nuestro análisis genómico está intrínsecamente conectado con la filiación filogenética de las especies de Physalideae (Altenhoff et al., 2019). Al correlacionar las variaciones genómicas con la historia evolutiva de estas especies, buscamos descubrir cómo los eventos evolutivos han influido en la diversificación de los genes de Withanolidos (Fan et al., 2020). Este enfoque permite no solo identificar genes específicos y sus variantes, sino también entender cómo la evolución ha modelado las complejas redes genéticas que subyacen a la producción de estos valiosos metabolitos (Nützmann et al., 2018).

La justificación de nuestro estudio genómico radica en abordar la brecha de conocimiento sobre la diversidad de withanólidos en la tribu Physalideae, particularmente en relación con los genes que los producen (Bharadwaj et al., 2021). A través de la genómica, buscamos dilucidar los mecanismos que regulan la producción de estos compuestos, trazando el origen de su diversificación bioquímica (Gupta et al., 2015) y su distribución específica entre especies (Sharma et al., 2020). Este enfoque es esencial para desarrollar estrategias de mejora genética y biotecnológica que potencien las aplicaciones farmacológicas de los withanólidos y fortalezcan las defensas naturales de las plantas.

Planteamiento del problema

A pesar de la amplia diversidad de especies en la tribu Physalideae, la mayoría de los estudios sobre la producción de withanólidos se han centrado en la especie Withania somnifera, lo que ha generado un conocimiento limitado sobre la biosíntesis de estos metabolitos en otras especies de la tribu con respecto a sus genes (Gupta et al., 2015). Nuestra investigación busca llenar este vacío, enfocándose en la diversidad genética y metabólica de los withanólidos en la tribu Physalideae y explorando las bases genéticasde su variabilidad. Al mismo tiempo, intentamos aportar claridad filogenética dentro dela tribu Physalideae, evidenciando que ciertos géneros, como Witheringia, no son monofiléticos al encontrarse en diferentes subtribus, como Physalinae y Withaninae (Deanna et al., 2019).

Además de su importancia farmacológica y su diversidad en la tribu Physalideae, es fundamental entender la biosíntesis de los withanólidos, explicada aquí con más detalle.La biosíntesis de los withanólidos puede dividirse en tres etapas. La primera etapa implica la creación de pirofosfato de isopentenilo (IPP) a través de dos vías: la vía del ácido mevalónico (MVA) en el citosol (Chaurasiya et al., 2012) y la vía del fosfato de metileritritol (MEP) en los plastidios. La segunda etapa involucra la biosíntesis del colesterol 24-metileno, que es una molécula central para la biosíntesis de varios componentes de los withanólidos (Misra et al., 2008; Figura 1). Finalmente, la tercera etapa implica la formación de diferentes withanólidos (Sharma et al., 2020). Lasprimeras dos etapas son compartidas con múltiples metabolitos defensivos, incluyendosaponinas, glicoalcaloides esteroidales y brasinólidos.

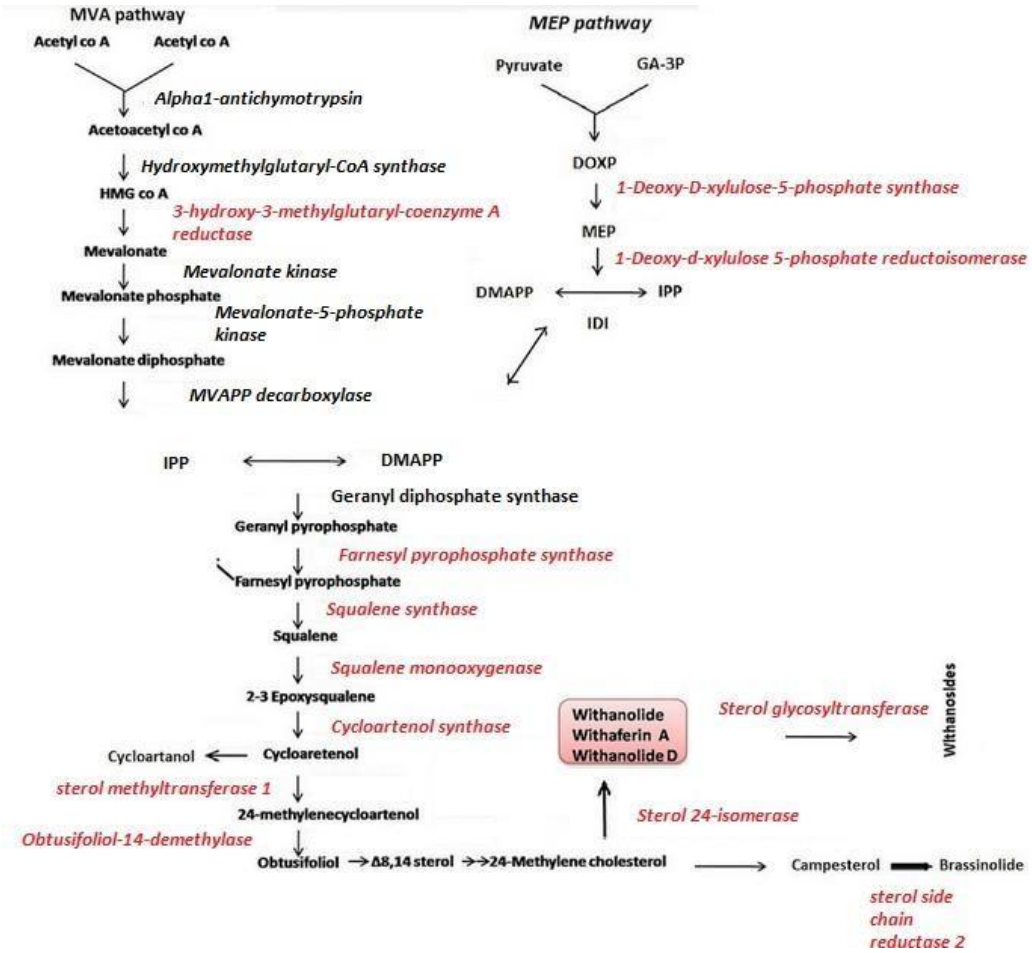


Figura 1. Vía MEP y MVA con compuestos, y genes asociados, a la transformación de estos (Modificado de Pandey et al, 2018) Los genes en rojo tienen impacto directo en laggeneración de withanólidos y al final de la cadena se observa la acción del esterol reductasa 2; generando brasinolidos, otros compuestos asociados al metabolismo secundario en Solanaceae.

Es importante destacar que, aunque los genes que median las dos primeras etapas son relativamente conservados entre las especies de Solanaceae (Knoch et al., 2018), hasta ahora se ha logrado identificar solo unos pocos genes involucrados en la etapa final de la biosíntesis de estos compuestos. Entre ellos, se encuentra el gen que codifica la enzima Δ24-isomerasa de esteroles (24-ISO), esencial para el primer paso comprometido de esta ruta (Knoch et al.,

2018), y algunos genes del citocromo-P450 (CYP450) relacionados con reacciones subsiguientes (Sharma et al., 2020).

El uso de técnicas de transcriptómica y genómica comparativa para confirmar el clúster biosintético de withanólidos (WBC), centrado en el gen 24-ISO, esencial en la biosíntesis de estos compuestos (Knoch et al., 2018), ha revelado que, rodeando al gen 24-ISO en los genomas de especies productoras, se encuentran citocromos P450 (CYP450s) y dioxygenasas dependientes de 2-oxoglutarato (2ODDs). Estos genes podrían intervenir en el metabolismo del 24-metildesmosterol (Gupta et al., 2015; Misra et al., 2008). Este hallazgo sugiere que un conjunto específico de genes asociados podría estar regulando las etapas finales de la biosíntesis de withanólidos. Aunque la función exacta de estos genes aún no ha sido validada funcionalmente, la creciente disponibilidad de secuencias genómicas completas de Solanaceae proporciona una oportunidad invaluable para identificar y evaluar genes candidatos. La correlación entre la presencia, el número de copias y la expresión de estos genes con la variabilidad de los niveles de withanólidos en diferentes especies es un área prometedora para futuras investigaciones. Estudios adicionales, utilizando técnicas avanzadas de edición genética y análisis funcional, podrían confirmar la participación de estos genes en la biosíntesis y abrir nuevas vías para la manipulación genética dirigida (Fan et al., 2020).

El hallazgo de posibles clústeres genéticos en la biosíntesis de withanólidos, como se observa en otras especies de Solanaceae (Fan et al., 2020), subraya su relevancia desde una perspectiva evolutiva y fisiológica. La agrupación de genes relacionados metabólicamente sugiere un mecanismo de co-regulación y co-evolución (Nützmann etal., 2018), facilitando la producción eficiente de moléculas complejas como los withanólidos. Esta organización genética puede reflejar la adaptación de las plantas a suentorno, respondiendo a presiones selectivas para mecanismos de defensa eficientes contra plagas o enfermedades.

Profundizando la idea anterior, la co-regulación en clústeres genéticos implica que genes relacionados se controlan juntos, lo que facilita respuestas coordinadas a estímulos ambientales y de desarrollo. Por ejemplo, la activación simultánea de genes en un clústerpuede optimizar la producción de un metabolito específico (Nützmann et al., 2018). La co-evolución, por su parte, se refiere a cómo estos clústeres de genes evolucionan conjuntamente en respuesta a presiones ambientales y adaptativas (Smit y Lichman, 2022). Al estar físicamente cercanos, estos genes pueden heredarse juntos, permitiendo la evolución sincronizada de trayectorias metabólicas complejas, lo que es crucial parala adaptación y supervivencia de la planta en diversos entornos.

Análisis preliminares, realizados por el grupo GEME, han permitido observar que genescomo la 24-ISO, P450 y 2OGD se expresan únicamente en especies productoras de withanólidos lo cual sugiere que están involucradas en la producción de estos metabolitos.

Tabla 1. Genes asociados al cluster biosintético de Knoch et al, 2018 y número de copias dentro de tres ejemplares de la tribu Physalideae.

Function	Physalis grisea	Physalis floridana	Iochroma cyaneum
24ISO	2	2	2
2OGD	6	11	8
ASAT4 like	1	1	1
P450	8	9	9
Dehydrogenase	2	4	2
Sulf1	2	2	1
VQBinding	1	0	0

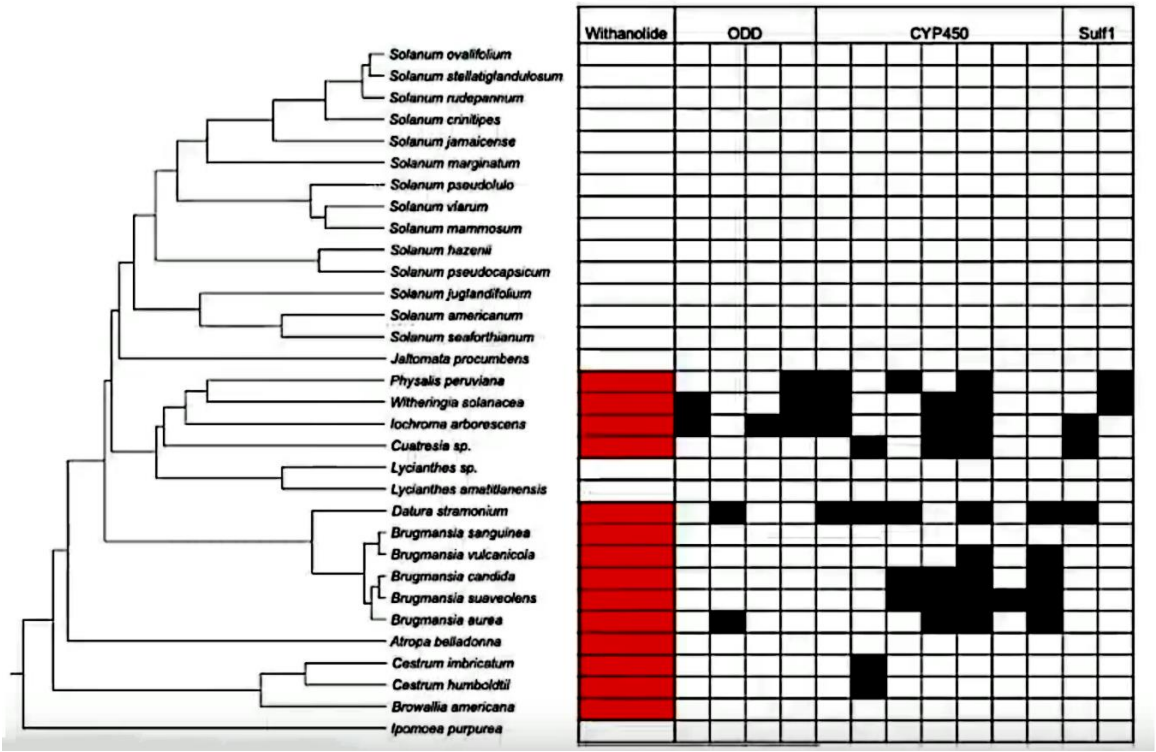


Figura 2. Determinación de presencia de genes del cluster en especies de Solanaceae.

Resumiendo, se ha llevado a cabo un análisis genómico exhaustivo previo con el objetivo de comprender en profundidad la diversidad y función de los withanólidos. Se ha centrado en las especies de Solanaceae, explorando la presencia, el número de copias y la distribución de genes relacionados con los withanólidos. Este enfoque permitió revelar cómo la evolución genética ha influido en la biosíntesis de estos compuestos. Además, los datos genómicos se correlacionaron con patrones de expresión y biosíntesis de withanólidos, lo que brinda una visión de las estrategias evolutivas que han permitido a la tribu Physalideae desarrollar una amplia variedad de withanólidos altamente efectivos.

Nuestros resultados preliminares respaldan fuertemente la importancia de un clúster biosintético de withanólidos (WBC) en la diversificación de estos compuestos. Se han identificado varios genes cuya expresión está fuertemente correlacionada con los nivelesde withanólidos y que se encuentran exclusivamente en el WBC de las especiesproductoras de withanólidos. Estos genes candidatos representan herramientas valiosaspara comprender la evolución, los roles biológicos y las posibles aplicaciones de los withanólidos en la tribu Physalideae.

El propósito esencial de esta tesis es elucidar los genes que confieren especificidad a los variados withanólidos dentro de la tribu Physalideae, destacando las relaciones filogenéticas entre las subtribus Physalidinae, Iochrominae y Withaninae. Así, este estudio genómico evolutivo tiene como fin profundizar en la comprensión de cómo se han originado los genes para la producción de withanólidos dentro de Physalideae y revelar la historia evolutiva y adaptativa detrás de la compleja biosíntesis de estos fascinantes compuestos.

1. Capítulo 1. Metodología

La metodología adoptada en este estudio se centró primordialmente en la caracterización metabolómica, estableciendo una base sólida para la comprensión integral de las relaciones entre metabolitos, transcriptomas y filogenética en la tribu Physalideae y el género Lycianthes. La variación metabólica, identificada mediante técnicas avanzadas de metabolómica, no solo destaca la diversidad química en estas especies, sino que también actúa como un punto de partida crucial para explorar las complejidades genéticas y evolutivas. Al correlacionar la producción de withanolidos y otros metabolitos con la expresión de genes específicos, obtenemos una visión detallada de las regiones genómicas implicadas en su síntesis. Este enfoque no solo mejora nuestra comprensión de los

mecanismos bioquímicos subyacentes, sino que también nos permite rastrear la evolución de estos compuestos defensivos y terapéuticos dentro del grupo. La inclusión de un marco filogenético enriquece esta exploración, permitiéndonos discernir cómo la segregación genética y la variación metabólica están interconectadas en el contexto evolutivo. Este párrafo introductorio establece el enfoque prioritario en la metabolómica y prepara al lector para la estructura no-cronologica de la sección, donde los métodos metabolómicos son discutidos primero, seguidos por análisis complementarios en transcriptómica y filogenética, reflejando así la integración y la interdependencia de estas disciplinas en nuestro estudio.

1.1 Búsqueda Bibliográfica y Análisis de Datos

En el marco de mi tesis, se llevó a cabo una búsqueda bibliográfica exhaustiva utilizando técnicas de búsqueda booleana para identificar literatura relevante sobre las especies asociadas a la tribu *Physalideae* y el género *Lycianthes*. El uso de consultas booleanas refinadas permite una identificación más efectiva de estudios relevantes, reduciendo el número total de documentos recuperados mientras aumenta el número de documentos relevantes identificados (Alharbi & Stevenson, 2020). La guía de la Universidad de Dundee sobre técnicas de búsqueda booleana también enfatiza la importancia de estas técnicas para obtener resultados de búsqueda más precisos y relevantes, lo cual fue esencial para asegurar la relevancia y calidad de la información obtenida en mi estudio (University of Dundee LibGuides, 2023).

1.2 Metabolomica

Para la extracción de metabolitos, se utilizaron 50 mg de tejido de plántulas completas de *Physalideae*, previamente macerado en nitrógeno líquido. Las plántulas tenían aproximadamente 3 semanas de edad y estaban en la etapa de la cuarta hoja verdadera. El tejido se dispuso en un tubo Eppendorf de 2 ml. A cada tubo se añadió 750 µl del buffer de extracción, que consistía en metanol al 100% y ribitol como estándar interno. Luego, se aplicó vortex para mezclar las muestras, manteniéndolas constantemente en hielo. Posteriormente, las muestras se agitaron durante 15 minutos a 70 °C, abriendo los tubos después de 2 minutos para liberar la presión. El sobrenadante de 700 µl se transfirió a un nuevo tubo Eppendorf de 2 ml y se añadieron 375 µl de cloroformo, seguido de otra ronda de vortex. Luego, se agregaron 750 µl de agua y se aplicó vortex nuevamente. Las muestras se centrifugaron a 12,000 rpm durante 10 minutos y se tomó el sobrenadante en volúmenes de 250, 250 y 500 µl, transfiriéndolos a tubos Eppendorf de 1.5 ml. Las muestras se almacenaron para su posterior envío y procesamiento mediante HPLC acoplado a masas (Protocolo grupo GEME).

Para los análisis metabolómicos, las fases polares secas al vacío se resuspendieron en 300 µL de metanol al 50% de agua. Después de la sonicación de las muestras durante 10 minutos en un baño de sonicador refrigerado por hielo, los tubos se centrifugaron durante 15 minutos a máxima velocidad (>12.000 rpm). El sobrenadante (100 µL), sin ningún sedimento del fondo del tubo, se transfirió a tubos de LC para su análisis. Las muestras se introdujeron en un equipo UPLC-LC-MS. El sistema UPLC está equipado con una columna de fase inversa HSS T3 C18 (100 × 2,1 mm de diámetro interno, 1,8 µm de tamaño de partícula; Waters) que funciona a una temperatura de 40 °C. Las fases móviles consisten en ácido fórmico al 0,1% en agua (disolvente A) y ácido fórmico al 0,1% en acetonitrilo (disolvente B). El caudal de la fase móvil es de 400 µL/min, y se cargan 2 µL de muestra por inyección. El instrumento UPLC se conecta a un Exactive Orbitrap- focus (Thermo Fisher Scientific) a través de una

fuente de electrospray calentado (Thermo Fisher Scientific).

Los espectros se registraron utilizando el modo de detección de iones positivos y negativos de barrido completo con todas las fragmentaciones de iones (MS2), cubriendo un rango de masas de m/z 100 a 1.500. La resolución se fijó en 70.000, yel tiempo máximo de exploración se fijó en 250 ms. El gas de cubierta se fijó en unvalor de 60 mientras que el gas auxiliar se fijó en 35. La temperatura del capilar de transferencia se fijó en 150°C, mientras que la temperatura del calentador se ajustó a 300°C. La tensión de pulverización se fijó en 3 kV, con una tensión capilar y una tensión de espumadera de 25 V y 15 V, respectivamente. Se registraron los espectros de MS desde los minutos 0 a 19 del gradientede UPLC. El procesamiento de los cromatogramas, la detección de picos y la integración se realizó con RefinerMS (versión 5.3; GeneData).

1.3 Analisis de Datos Metabolomicos

Los picos fueron anotados utilizando bases de datos personales mediante el análisis de masas y patrones de fragmentación (Villas-Bas et al., 2023). Los valores de masas (m/z) obtenidos se utilizaron para su análisis en MetaboAnalyst (Pang et al., 2022). Para asegurar la precisión y validez de los análisis estadísticos, los valores fueron normalizados mediante transformaciones basadas en logaritmos, adición y auto-escalado de datos (Dunn et al., 2011). Se llevaron a cabo dos tipos de análisis estadísticos específicos: 1) Se realizó una regresión de mínimos cuadrados parciales, un método supervisado de análisis multivariado que relaciona variables como especies y tipos de muestra con asociaciones simples o múltiples (metabolitos). 2) Para identificar los metabolitos con diferencias significativas entre especies, se utilizó la prueba t de Student, un análisis paramétrico múltiple, estableciendo un valor de p FDR ≤ 0.05 como criterio para la significancia estadística.

Los datos de metabolómica alineados en los modos positivo y negativo se procesaron utilizando la herramienta Metabolomic Search (Dhanasekaran et al., 2015), accesible en Metabolomic-search. Esta herramienta, diseñada para facilitar elanálisis y la visualización de datos complejos de metabolómica, se complementa conel uso de MetaboAnalyst (Pang et al., 2022). Para iniciar el procesamiento, se subieron los archivos de datos al formato CSV a la plataforma, asegurándose de quelos mismos siguieran un formato específico requerido por el sistema, tal como se ilustra en la interfaz de usuario de la herramienta.

La configuración general de la herramienta permite la selección del archivo a analizar y la personalización de los parámetros de importación, como la presencia de encabezados, el estilo decimal, el carácter de cotización y el carácter separador. Además, se configura la fila de inicio para la lectura de datos. En la sección de configuración de columnas, se establecen los nombres de las filas y se asignan los números de columna para los valores m/z, RT (tiempo de retención) y las muestras,que son críticos para el análisis posterior.

Una vez cargados y configurados los datos, la herramienta procesa la información yutiliza algoritmos de agrupamiento para determinar los clusters. Estos clusters, asociados con los compuestos presentes en las muestras, se determinan en función de sus perfiles de masas y tiempos de retención. El algoritmo de clustering agrupa los compuestos con perfiles similares, facilitando la identificación de metabolitos relacionados o que comparten rutas bioquímicas similares, un paso crucial en la exploración de datos y la generación de hipótesis en estudios metabolómicos(Dhanasekaran et al., 2015).

Posteriormente, se realizan análisis detallados de los datos de masas (m/z) y patrones de fragmentación de los withanolidos que han sido reportados en la literatura científica por medio de la herramienta Xcalibur. Xcalibur es un sistema de software sofisticado diseñado para el análisis y la interpretación de datos de espectrometría de masas (Thermo Fisher Scientific, 2023). Es ampliamente utilizado para la identificación y cuantificación de compuestos en muestras complejas y es particularmente útil en la metabolómica, donde la identificación precisa de moléculas es esencial.

Para el análisis, los datos de m/z y los patrones de fragmentación de los withanolidos reportados se introdujeron en Xcalibur, donde se emplearon sus capacidades de procesamiento de datos para comparar los espectros de masas teóricos con los obtenidos de las muestras experimentales. Este proceso de validación es crucial paraconfirmar la presencia de compuestos específicos en las muestras bajo estudio.

1.4 Selección de Plantulas y Preparación

Dentro de la colección de semillas del grupo GEME, se realiza una selección de semillas con el objetivo de incluir representantes de la mayor cantidad posible de géneros pertenecientes a la tribu Physalideae (Bharadwaj et al., 2021). En este estudio, se han seleccionado 15 semillas por taxon de un total de 37 taxones, reflejando la riqueza y variedad de esta importante tribu (Gupta et al., 2015).

Las semillas utilizadas en este estudio provienen de dos fuentes principales: algunas fueron recolectadas directamente por miembros del grupo GEME, mientras que otras fueron donadas por la Dra. Stacey Smith de la Universidad de Boulder en Colorado y por el Dr. Franco Chiarini de la Universidad de Córdoba en Argentina. Para el estudio de la tribu Physalideae, se seleccionó un conjunto diverso de taxones, lo que es crucial para comprender la diversidad genética y la evolución de los metabolitos secundarios en estas plantas (Nützmann, Scazzocchio, y Osbourn, 2018).

Los taxones elegidos para esta investigación incluyen Cuatresia riparia, Darcyanthus spruceanus, Dunalia solanacea, Dunalia spathulata, Dunalia spinosa, Eriolarynx australis, Eriolarynx fasciculata, Iochroma arborescens, Iochroma edule, Iochroma gesnerioides, Iochroma lorense, Iochroma sp., Iochroma sp. nov, Nicandra physaloides, Alkekengi officinaron, Physalis angulata, Physalis cinerascens, Physalis ixocarpa, Physalis peruviana, Physalis sp., Quincula lobata, Trozelia umbellata, Vassobia breviflora, Witheringia cf. killipiana, Witheringia solanacea, Withania somnifera, y Cuatresia cuspidata.

Adicionalmente, se incluyeron tres individuos por especie del género *Lycianthes*, como *Lycianthes amatitlanensis*, *Lycianthes aff. pauciflora*, *Lycianthes inaequilatera*, *Lycianthes medusocalyx*, *Lycianthes pauciflora* y *Lycianthes sp.*, como grupo externo (outgroup). Esta elección se basó en el hecho de que las especies de *Lycianthes* no tienen reportada producción de withanólidos, lo que permite una comparación evolutiva y bioquímica esencial para entender la biosíntesis y diversificación de withanólidos dentro de la tribu *Physalideae*. La inclusión de este grupo externo es crucial para aportar un contexto filogenético y metabólico al estudio, permitiendo así identificar características y vías evolutivas únicas asociadas con la producción de withanólidos en los taxones seleccionados.

Para aumentar la tasa de germinación de las semillas, se sumergieron en una solución de ácido giberélico a una concentración de 500 ppm. Posteriormente, se colocaron en un horno a una temperatura de 37 °C durante 24 horas y luego se germinaron en cajas de Petri que contenían papel absorbente y una solución de nitrato de potasio a una concentración de 1000 ppm, debido a su efecto fertilizante. Las cajas de Petri se colocaron en una cámara de germinación con una temperatura constante de 22 °C y un fotoperiodo de 16 horas de luz y 8 horas de oscuridad. Las plántulas germinadas se mantuvieron en estas condiciones hasta que alcanzaron su cuarta hoja verdadera.

Dada la regulación específica, tanto espacial como temporal, de la biosíntesis de withanólidos en la tribu Physalideae, se propendió por el uso de plántulas, ya que permiten una observación detallada de cómo estos compuestos se distribuyen y se concentran en diferentes tejidos (Agarwal et al., 2017). Esta aproximación es crucial para comprender la estrategia de la planta en la síntesis de estos metabolitos secundarios y cómo varían en concentración en partes jóvenes de la planta, reflejando una variabilidad inherente a la fisiología de la planta (Gupta et al., 2015). Por ejemplo, la sobreexpresión de ciertas enzimas clave en plántulas puede influir significativamente en las concentraciones de withanólidos en tejidos específicos (Sharma et al., 2020).

1.5 Analisis Transcriptomicos y Filogeneticos

Para la extracción de ARN, se maceró toda la plántula en nitrógeno líquido, lo cual es crucial para evitar la degradación del material genético debido al calor generado durante la maceración. El nitrógeno líquido asegura que el ARN se mantenga intacto al inactivar rápidamente las enzimas degradativas. Luego se agregó 1 ml de TRizol por cada 50-100 mg de tejido. El TRizol permitió estabilizar el ARN, ya que es un agente caotrópico, y a su vez desnaturalizó las ARNasas que podían dañarlo. La cantidad de muestra no debía superar el 10% del volumen de TRizol usado; una cantidad insuficiente de TRizol podía ocasionar contaminación con ADN en el ARN (Suaréz-Baron y Pabón-Mora, 2014). Se homogenizó con vórtex e incubó la muestra durante 5 minutos a temperatura ambiente con el fin de permitir la disociación del complejo de nucleoproteínas.

Posteriormente se adicionó 0.2 ml de Cloroformo por cada 1 ml de TRizol; lo cual permitió desarrollar una fase orgánica, una intermedia y una acuosa, donde el ARNquedó concentrado en la fase acuosa. Se mezcló vigorosamente por inversión durante 15 minutos. Se incubó 2-3 minutos a temperatura ambiente. Luego, se utilizóla centrifugadora a 11000 rpm durante 15 minutos a 4°C. El sobrenadante se eliminópor completo y se lavó el pellet de ARN con 1 ml de etanol al 75%, para precipitar cualquier proteína que pudiera contaminar la muestra. Se centrifugó de nuevo, peroen esta ocasión a 7500 x g durante 10 minutos a 4°C. El etanol se decantó y el precipitado de ARN se secó durante 5-10 minutos en hielo, teniendo cuidado de nosecar el pellet completamente. Al final se re-suspendió en 200 µl de agua DEPC y se almacenó a -80°C. La calidad del ARN se evaluó por medio de NanoDrop One/One y electroforesis en gel de agarosa al 2%. La concentración del ARN se cuantificó con Qubit 4 Fluorometer.

Los ARNs limpios, extraídos de plántulas completas de la tribu *Physalideae*, fueron enviados a Novogene (California) para su secuenciación. Las plántulas utilizadas tenían aproximadamente 3 semanas de edad y estaban en la etapa de la cuarta hoja verdadera. La calidad del ARN se determinó mediante electroforesis y el ensayo Qubit RNA IQ Assay. Para la electroforesis, se utilizó un gel de agarosa al 1% preparado con TBE (Tris-Borato-EDTA) y teñido con SYBR Green para visualizar el ARN. Se cargaron 5 µL de cada muestra junto con un marcador de peso molecular de ARN (RNA ladder) y se corrieron a 100 V durante 30 minutos. Los resultados de la electroforesis se incluyen en los anexos.

El Qubit RNA IQ Assay se utilizó para evaluar la integridad y cantidad del ARN, asegurando que las muestras cumplieran con los estándares de calidad para la secuenciación, con un IQ (Índice de Calidad) mínimo de 7. Las librerías de secuenciación fueron preparadas mediante enriquecimiento poli-A y se utilizaron insertos de 250-300 pb. Estas bibliotecas fueron secuenciadas con el sistema de secuenciación Illumina NovaSeq™ 6000 S4 ("paired end", longitud de lectura = 150 bp). Se obtuvo un mínimo de 6 Gb de secuencia bruta para todas las muestras.

Después de haber recibido las secuencias de los transcriptomas, se procedió a revisarla calidad de los transcritos utilizando FastQ. Aquellas secuencias con valores de calidad inferiores a 28 fueron sometidas al programa Trimmomatic (Bolger, A. M., Lohse, M. & Usadel, B., 2014), ya que esto indicaba una probabilidad de error mayora 1 en 1000 bases. El corte de las secuencias se realizó utilizando las siguientes especificaciones: se empleó un enfoque de PE (paired end) para cortar las secuenciastanto en el archivo forward como en el archivo reverse. Para asociar los valores proporcionados por Fastq, se leyó el archivo en formato Phred33, el cual hace referencia al sistema de codificación ASCII utilizado para determinar la calidad del transcrito (Edgar, R. C., 2004). Se sumó 33 a la calidad determinada y se asignó uncódigo para su identificación posterior.

Posteriormente, se introdujeron dos archivos de entrada en el software, uno con losreads forward y otro con los reads reverse. Se obtuvieron las secuencias que estabanemparejadas, superaban el valor de calidad asociado de 28 y tenían una longitud mínima de contig. Durante este proceso, también se verificó la calidad en los extremos 5' y 3' de la secuencia, además de verificar si los adaptadores de Illumina habían sido eliminados del read (Haas, B. J. et al., 2013).

Para llevar a cabo el ensamblaje de los transcritos, después de mejorar su calidad mediante Trimmomatic, se utilizó Trinity (Haas, B. J. et al., 2013). Trinity permite generar ensamblajes de novo cuando no se dispone de un genoma de referencia, utilizando gráficos de Bruijn sobre el ARN particionado. Los gráficos de Bruijn songrafos dirigidos que miden las relaciones entre diferentes partes de un sistema, lo que permite representar las superposiciones entre secuencias y obtener el conjunto completo del constructo en el transcriptoma. Se utilizó este constructo para generar alineamientos y determinar la expresión utilizando el programa Salmon-1.6.0. La cuantificación de la expresión génica se realizó mediante el mapeo de los reads de cada taxón a su respectivo transcriptoma (Patro, R. et al., 2017).

Antes de realizar la cuantificación, se redujo la redundancia debida a isoformas utilizando Perl. Luego, se identificaron las regiones codificantes de los transcritos mediante TransDecoder, donde se buscaban los marcos abiertos de lectura. Se utilizóOrthofinder para encontrar ortólogos entre transcriptomas de diferentes especies. Lassecuencias no redundantes de CDS se usaron como entrada. Orthofinder se inicializóutilizando alineamientos múltiples (MSA), y se asoció una carpeta a cada especie aanalizar. Se verificó que el árbol de especies generado fuera acorde a la filogenia másaceptada para el grupo (Deanna et al., 2019). En caso contrario, se generó un árbol de referencia en formato Newick para la correcta búsqueda de los genes ortólogos (Emms & Kelly, 2019).

OrthoFinder se utilizó también para clasificar genes en grupos ortólogos y parálogos, facilitando así la identificación de genes que comparten un ancestro común en diferentes especies (Emms, D. M. & Kelly, S., 2019). Este análisis nos permitió agrupar genes en 'Hierarchical Orthologous Groups' (HOGs), lo que nos ayudó a entender la evolución y la función de estos genes a través de diferentes linajes (Kriventseva, E. V. et al., 2019). Cada HOG, identificado por un código único, representa un conjunto de genes ortólogos que presumiblemente se originaron de ungen ancestral en un punto específico de la evolución. La identificación de estos grupos a través de OrthoFinder nos proporcionó una base sólida para

explorar la historia evolutiva y las relaciones genéticas entre las especies analizadas (Altenhoff,A. M. et al., 2019).

Adicionalmente, utilizamos la herramienta Salmon para cuantificar la abundancia de transcritos de estos HOGs en nuestros datos de secuenciación de ARN (RNA-Seq) (Patro, R. et al., 2017). Salmon, conocida por su rapidez y precisión, nos permitió estimar la expresión de transcritos a nivel de genes y HOGs, utilizando un enfoque de mapeo cuasi (quasi-mapping) para la cuantificación. Este enfoque de mapeo cuasi no requiere una alineación completa de las lecturas de secuenciación al genoma de referencia, lo que reduce significativamente el tiempo de procesamiento y mejora la eficiencia. Este método nos proporcionó información valiosa sobre la expresión génica de los HOGs identificados, lo que nos ayudó a correlacionar la presencia y la abundancia de estos genes con funciones biológicas específicas y su relevancia en el contexto de nuestro estudio. La combinación de OrthoFinder (Emms, D. M. & Kelly, S., 2019) y Salmon fue esencial para comprender no solo la presencia de genes conservados en diferentes especies, sino también su nivel de expresión y potencial implicación en procesos biológicos.

En el análisis de la expresión génica, es fundamental disponer de métodos precisos para normalizar los datos y obtener estimaciones fiables. En nuestro estudio, se ha implementado un enfoque avanzado para calcular estimaciones normalizadas de la expresión acumulativa en diferentes grupos ortólogos. Específicamente, hemoscalculado las estimaciones para cada Grupo Ortólogo (OG), Grupo Ortólogo Adaptado (AOG) y Transcritos por Millón por 10,000 Grupo Ortólogo (TPM10KOG). Este método está adaptado de una metodología previamente establecida en la literatura científica (Robinson, M. D., McCarthy, D. J. & Smyth, G.K., 2010; Zhao, Y. et al., 2021).

Para el cálculo de la expresión adaptada en los Grupos Ortólogos Adaptados (AOG),se empleó la siguiente fórmula:

$$\mathrm{AOG} = \left(\frac{\text{Suma de lecturas por OG} \times 10^3}{\text{longitud efectiva total de CDSs en OG}} \right)$$

Donde la 'Suma de lecturas por OG' representa el número total de lecturas de secuenciación que corresponden a un grupo ortólogo específico. Este valor es luegomultiplicado por 10^3, sirviendo como un factor de normalización para ajustar la escala de los datos. Por otro lado, 'longitud efectiva total de CDSs en OG' se refierea la longitud total de las Secuencias de Codificación (CDS) dentro de un grupo ortólogo. Este enfoque proporciona una medida más precisa y ajustada de laexpresión génica, teniendo en cuenta tanto la abundancia de lecturas como la longitudde las secuencias codificantes involucradas.

Se llevó a cabo un proceso de anotación genómica utilizando la herramienta BLASTn para las secuencias de ARN obtenidas. Para comenzar, se prepararon las secuencias de interés en formato FASTA, asegurándose de que cada archivo contuviera un encabezado precedido por el símbolo '>' seguido de la secuencia de nucleótidos. Esta preparación meticulosa es crucial para garantizar la precisión de los resultados de BLAST. La base de datos de nucleótidos de NCBI (nt) fue seleccionada para la búsqueda debido a su exhaustividad y relevancia para secuencias de ADN.

Para automatizar el proceso de anotación, se integró BLASTn en un script de Python utilizando el paquete Biopython. Este script leyó las secuencias desde los archivos FASTA, ejecutó BLASTn contra la base de datos seleccionada y almacenó los resultados en formato XML. Específicamente, la función NCBIWWW.qblast de Biopython se utilizó para ejecutar BLASTn, lo que permitió la rápida y eficiente identificación de regiones de similitud en las secuencias de ARN. Los resultados obtenidos proporcionaron información valiosa sobre la

función, estructura y evolución de las secuencias analizadas (BLAST: Basic Local Alignment Search Tool, n.d.).

Posteriormente, se buscó identificar aquellos genes cuya expresión estuviera asociada con los niveles de los diferentes withanólidos encontrados en las especies muestreadas. Para ello, se utilizó el paquete WGCNA en la plataforma R (Langfelder & Horvath, 2008). Se aplicó la función de correlación de Pearson (cor.test) para determinar la significancia de las correlaciones. Los valores de p resultantes fueron corregidos utilizando la tasa de falsos descubrimientos mediante la opción de corrección fdr correction de la función p.adjust en el software R. Se utilizaron datos filtrados y normalizados de expresión de genes obtenidos con EdgeR, así como los valores reportados de withanólidos a través de MetaboAnalyst como archivos de entrada.

1.6 Cumplimiento Ético e Institucional

El presente documento fue elaborado con base en los lineamientos establecidos en la Resolución 001 de 2011 de la Vicerrectoría Académica de la Universidad Nacional de Colombia, “por la cual se establecen los procedimientos para la publicación de lastesis de maestría y doctorado de los estudiantes de la Universidad Nacional de Colombia en el Repositorio Institucional UN”.

2. Capítulo 2: Resultados

2.1 Resultados Metabolómicos:

El objetivo del análisis metabolómico en el presente estudio es caracterizar y comparar los perfiles metabólicos de los especímenes de la tribu Physalidae y compararlos con miembros del grupo más cercano, Lycianthes, el cual se considerano productor de Withanolidos (Misico et al, 2011). Lo anterior, con el propósito deidentificar metabolitos y vías metabólicas asociadas a la producción de los metabolitos secundarios mencionados.

El análisis exhaustivo de la literatura científica y bases de datos especializadas permitió identificar 237 compuestos catalogados como withanolidos. De estos, se encontró que 91 compuestos tienen información detallada sobre sus patrones de fragmentación. Se evidenciaron 21 subgrupos de withanolidos asociados a las especies estudiadas y 135 compuestos reportados en las especies presentes, y algunasde los mismos géneros.

A través de este meticuloso proceso de revisión, se logró establecer una lista detallada de los withanólidos que han sido documentados en cada una de las especies estudiadas, proporcionando así una visión completa y actualizada de la composiciónquímica de estos compuestos en la tribu Physalidae.

En nuestro estudio, analizamos los perfiles metabolómicos de diversas especies de Physalideae para identificar y cuantificar los withanólidos presentes. Un compuesto notable identificado fue el Withaferin A en **Physalis peruviana**, cuyo espectro de masas se muestra en la Figura 3. Este espectro de masas fue obtenido mediante un análisis de MS-MS, proporcionando una identificación precisa del compuesto en cuestión.

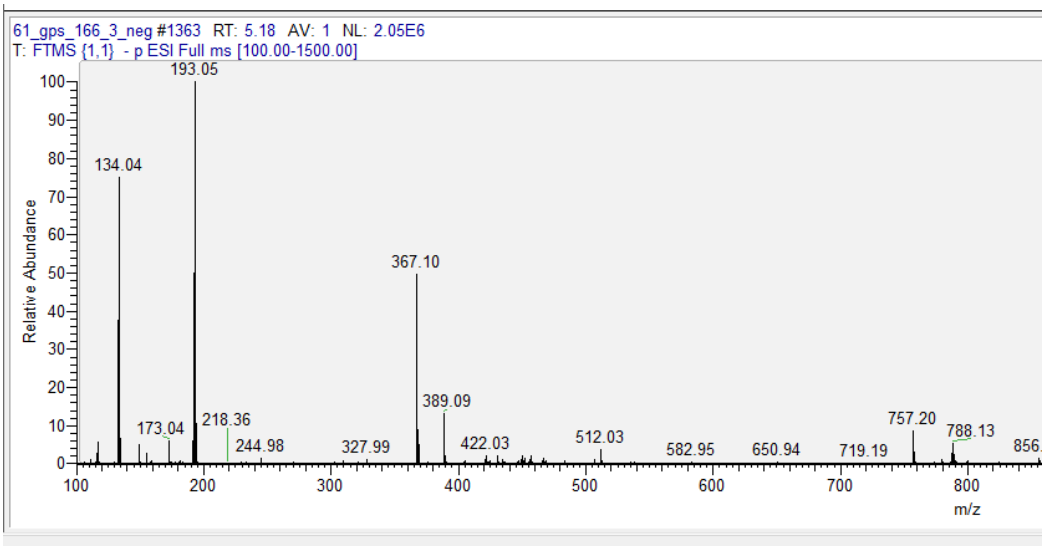


Figura 3. Espectro de masas MS -MS de Withaferin A en Physalis_peruviana

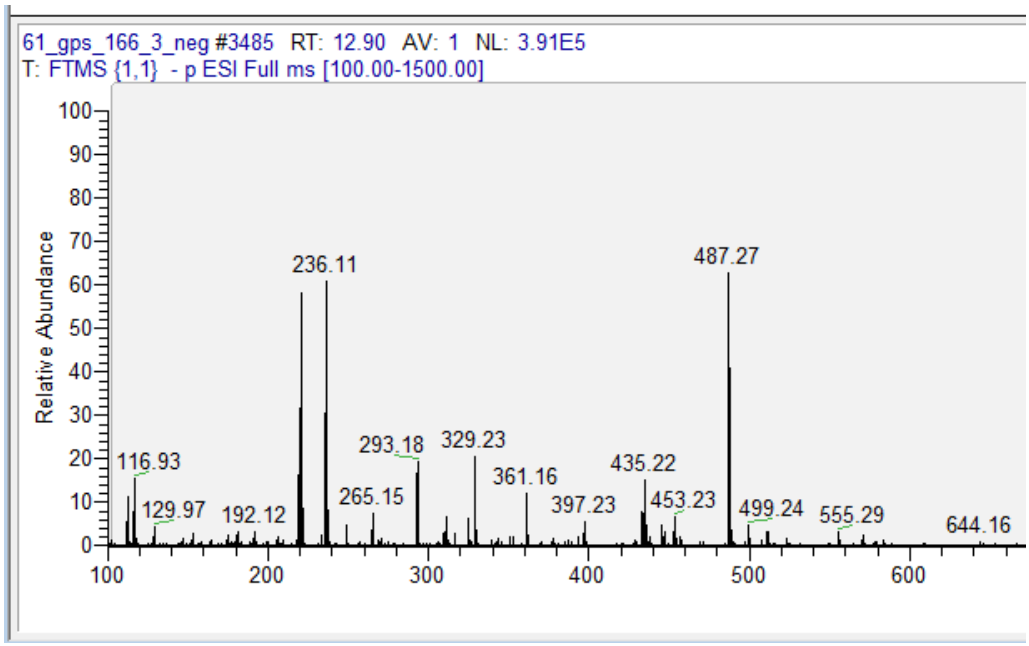


Figura 4. Espectro de Masas MS -MS de 1,3,27-trihydroxy-(20S,22R)-3- methoxywitha-4,6,24-trienolide en Physalis peruviana

Además del Withaferin A, identificamos otro compuesto significativo, el 1,3,27-trihydroxy-(20S,22R)-3-methoxywitha-4,6,24-trienolide, en *Physalis peruviana*. El espectro de masas de este compuesto se muestra en la **Figura 4**. Este análisis también se realizó mediante MS-MS, permitiendo una caracterización detallada del metabolito.

Otro compuesto identificado en nuestro análisis fue el Baimantuoluoside B en *Withania somnífera*. El espectro de masas de este compuesto se muestra en la **Figura 5**. Este análisis fue crucial para corroborar la presencia y estructura del compuesto.

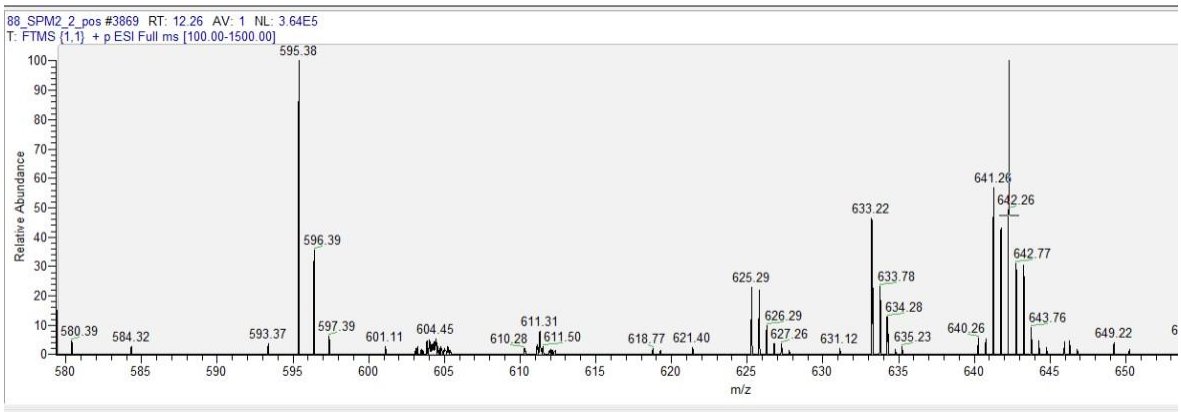


Figura 5. Espectro de masas MS -MS de Baimantuoluoside B en *Withania somnífera*

Otro compuesto significativo identificado en nuestro estudio fue el 12-Deoxywithastramonolide en Cuatresia riparia. Este compuesto se detectó mediante análisis de MS-MS, proporcionando una caracterización detallada del metabolito en cuestión. La Figura 6 muestra el espectro de masas correspondiente, destacando los picos relevantes que permiten su identificación.

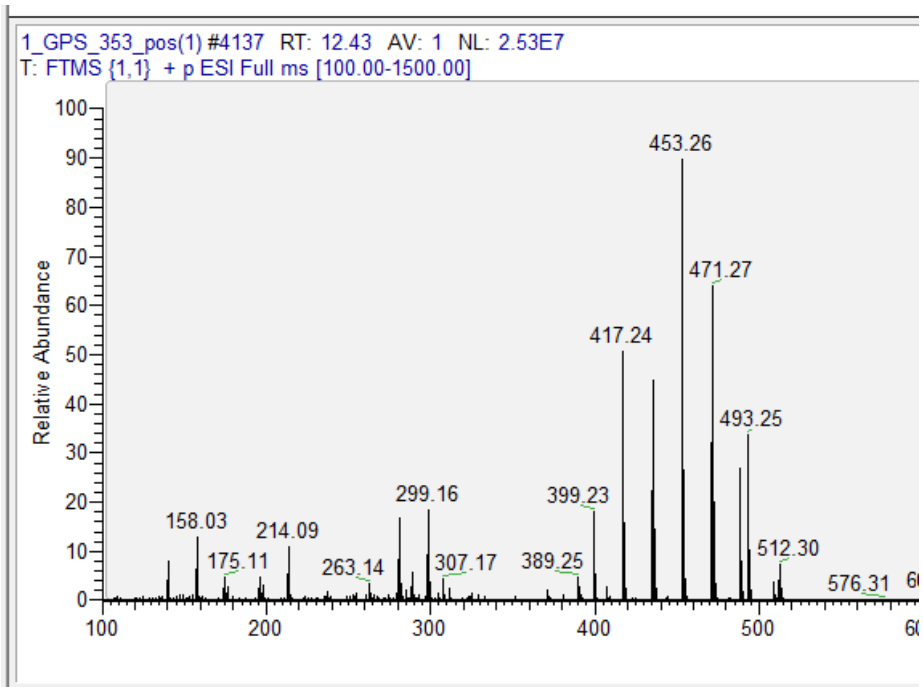


Figura 6. Espectro de masas MS -MS de 12-Deoxywithastramonolide en *Cuatresia riparia*

Finalmente, se identificó el **27-hydroxywithanone** en *Cuatresia riparia*. Este compuesto se detectó mediante análisis de MS-MS, proporcionando una caracterización detallada del metabolito en cuestión. La **Figura 7** muestra el espectro de masas correspondiente, destacando los picos relevantes que permiten su identificación.

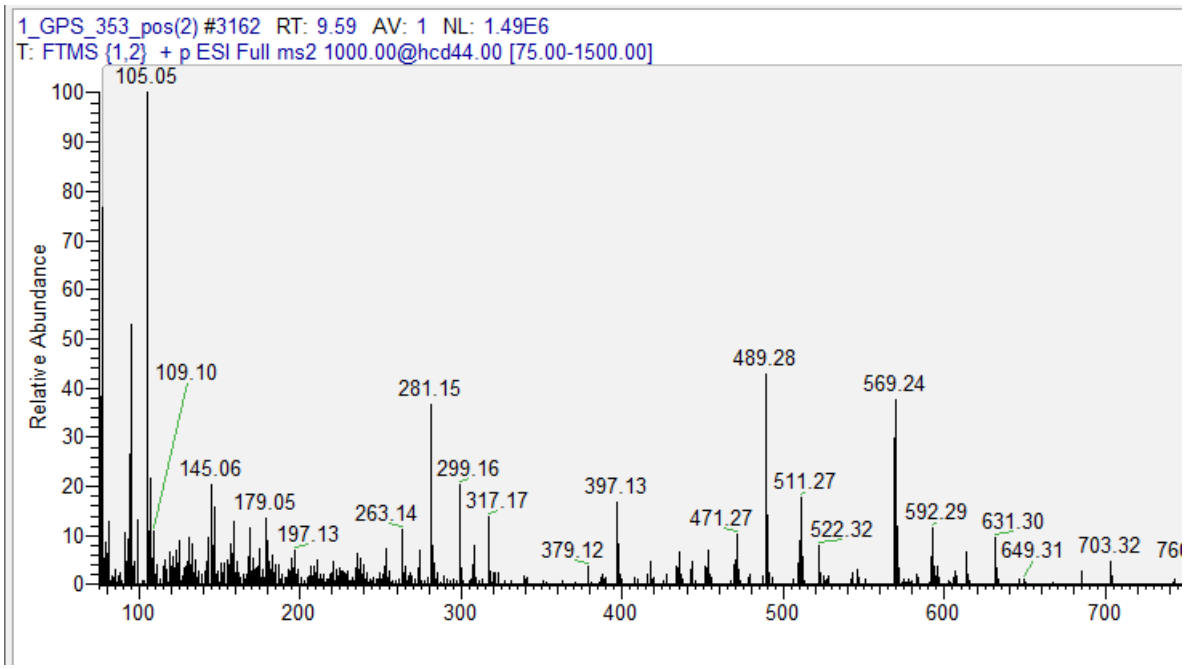


Figura 7. Espectro de masas MS -MS de 27 hydroxywithanone *Cuatresia riparia*

Withania somnífera: 5b,6b-EpoxyWithanolidos: Viscosalactona B (Misra et al., 2008). Other D5-Withanolidos: Withaferina A (Misico et al., 2011), Withanolido D(Misico et al., 2011), y 2,3-dihydrowithaferina A (Tong et al, 2011). C-14, C-17, y C-20 Hydroxylated Withanolidos and Related Compounds: Physagulin D,Withanoside IV, Withanoside V y Withanoside X (Tong et al., 2011). 1a,3b-Dihydroxy-D5-Withanolidos and Related Compounds: 12-Deoxywithastramonolide (Tong et al., 2011). 5a-Hydroxy-6a,7a-EpoxyWithanolidos and Related Compounds: Withanone (Tong et al., 2011).

Dunalia spinosa: 5b,6b-EpoxyWithanolidos: 4β, 5β, 6β)-3-amino-4,27-dihydroxy- 5,6-epoxy-1-oxo-witha-24-enolide (Espinoza et al., 2012). Withaferina A (Erazo etal., 2008)

Vassobia breviflora: 5b,6b-EpoxyWithanolido: Withaferin A (Samadi, 2015)

Iochroma gesnerioides: 5b,6b-EpoxyWithanolidos: Withaferin A (Alfonso & Kapetanidis, 1994). Withacnistine (RAlfonso & Kapetanidis, 1994).

Alkekengi officinaron: C-14, C-17, y C-20 Hydroxylated Withanolidos y Compuestos Relacionados: Physagulin J y Physagulin B (Li et al., 2008). 13,14-seco-16,24-cycloergostane: 5a,6a-dihydroxy physalin D, Physalin B, Physalin D, Physalin A, Physalin L, Physalin X, 4,7-didehydroneo-physalin B y Withangulatin A (Li et al., 2014). Withaphysalins y Withanolidos Relacionados: Withaphysalin N, Withaphysalin U y Withaphysalin B (Li et al., 2014).

Physalis angulata: Other D5-Withanolidos: Pubesenolide y Withaminimin (He et al., 2007). C-14, C-17, y C-20 Hydroxylated Withanolidos and Compuestos Relacionados: Physagulins L–O (He et al., 2007). 13,14-seco-16,24-cycloergostane:Withangulatin I (He et al., 2007). Withaphysalins y Withanolidos Relacionados: Withangulatin A, Withaphysalin Y y Withaphysalin Z (He et al., 2007).

Physalis cinerascens: 5b,6b-EpoxyWithanolidos: 4β-hydroxyWithanolido E (He et al., 2007). Other D5-Withanolidos: Withanolido S (He et al., 2007). C-14, C-17, y C-20 Hydroxylated Withanolidos and Related Compounds: 24,25- dihydroWithanolido S, Cinerolide y Physalactone (He et al., 2007). Aromatic Ring- A Withanolidos and 19-HydroxyWithanolidos: Cinerolide (He et al., 2007).

Nicandra physaloides: Aromatic Ring-A Withanolidos and 19- HydroxyWithanolidos: Salpichrolide A (He et al., 2007). Ixocarpalactones and Perulactones: Ixocarpalactone L y Ixocarpalactone A (He et al., 2007). OtherHydroxylated Withanolidos: 24α,25β-dihydroxy-nicandrenone-2, 17-(1α/1β-methylpropanone)-nicandrenone y 6β,7α dihydroxynicandrenone 10 (He et al., 2007).

Physalis ixocarpa: 5a-Hydroxy-6a,7a-EpoxyWithanolidos and Related Compounds: Ixocarpanolide y Withasomniferol B (He et al., 2007). 13,14-seco- 16,24-cycloergostane: Physalin F, Physalin D y Physalin B (He et al., 2007).

Physalis peruviana: 5b,6b-EpoxyWithanolidos: Withanolido E (Lan et al, 2009), withaperuvin C (Sang-ngern et al, 2016), withaperuvin N (Sang-ngern et al, 2016) y4b-hydroxyWithanolido E (Sang-ngern et al, 2016). Other D5-Withanolidos: Withanolido S (Lan et al, 2009) y Withanolido C (Lan et al, 2009). C-14, C-17, andC-20 Hydroxylated Withanolidos and Related Compounds: phyperunolide A (Lan et al, 2009), phyperunolide B (Lan et al, 2009), phyperunolide C (Lan et al, 2009), phyperunolide D (Lan et al, 2009), phyperunolide E (Lan et al, 2009), phyperunolideF (Lan et al, 2009), Physalactone (Lan et al, 2009), physaperuvins I (Sang-ngern etal, 2016) y physaperuvins J (Sang-ngern et al, 2016). Ixocarpalactones and Perulactones: perulactone C (Sang-ngern et al, 2016), perulactone D (Fang et al , 2009) y perulactone B (Fang et al, 2009). Physalolacton: Physalolacton (Lan et al, 2009). Other 18-NorWithanolidos: Withaphysanolide (Lan et al, 2009).

Witheringia solanácea: 13,14-seco-16,24-cycloergostane: Physalin B, Physalin F yPhysalin D (Yang et al, 2021).

La investigación de los 237 compuestos, por medio de su m/z y RT ha revelado hallazgos notables en cuanto a la presencia y caracterización de withanolidos en lasmuestras estudiadas. Utilizando la herramienta Xcalibur para el análisis de datos, selogró confirmar la presencia de 7 withanolidos únicos. Esto fue después del uso de 24 perfiles positivos y 51 negativos en este estudio. Se lograron identificar tres withanolidos de perfiles negativos, los cuales son: 1,3,27-trihydroxy-(20S,22R)-3- methoxywitha-4,6,24-trienolide (NC_14929), Withaferin A (NC_14173) y Baimantuoluoside B (NC_ 21329). Por otro lado, en los perfiles positivos, se lograronasociar cuatro withanolidos identificados como: 12-deoxywithastramonolide (PC_ 25971), 27- hydroxywithanona (PC_ 26525), Withanoside V (PC_05301) y Withanoside VII (PC_05352) (Figura 3).

Adicionalmente, se observó que los 7 withanolidos pertecen a los siguientes subgrupos: 5b,6b-Epoxywithanolidos, C-12 and C-21 Oxygenated Withanolidos, 5a- Hydroxy-6a,7a-Epoxywithanolidos and Related Compound y 1a,3b-Dihydroxy-D5- Withanolidos and Related Compounds; Mientras que 1,3 ,27-trihydroxy-(20S,22R)-3-methoxywitha- 4,6,24-trienolide no pudo ser clasificado en ningún grupo. Estos compuestos no se detectaron en las especies del género *Lycianthes* y presentaron una distribución irregular en las especies de *Physalideae*; dos withanolidos se expresaron predominantemente en el género *Physalis*, dos mostraron mayor expresión en *Withania*, dos fueron más prevalentes en *Cuatresia*, y uno destacó en *Iochroma*. (Figura 8).

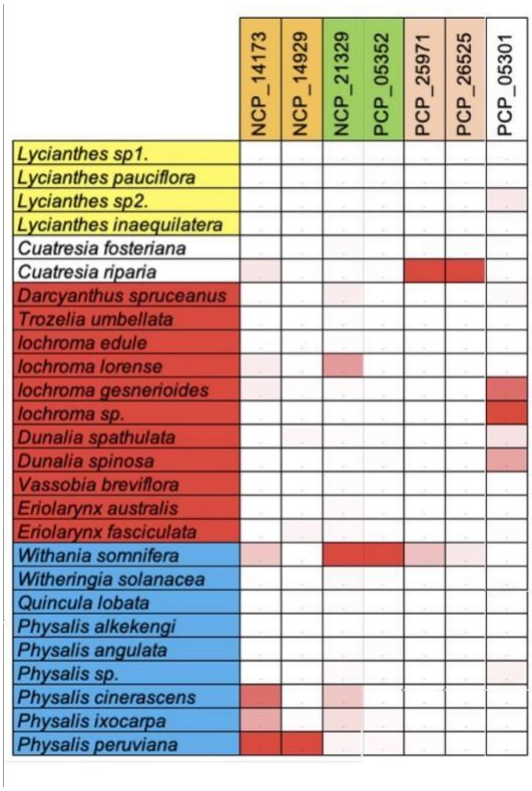


Figura 8. Distribución de Withanolidos en Especies de Solanaceae. La figura muestra la distribución de withanólidos en diferentes especies de la familia *Solanaceae*. La intensidad del color rojo indica la concentración de withanólidos, donde un rojo más intenso representa una mayor concentración. Las especies están categorizadas en diferentes colores de fondo para destacar distintas subtribus y un género: amarillo para *Lycianthes*, rojo para *Iochrominae*, azul para *Physalinae* y blanco para Cuatresia. Los códigos en la parte superior (NCP y PCP) representan identificadores específicos de los compuestos analizados.

Se realizo paralelamente, el analisis de los metabolomas de 59 especies de la familia *Solanaceae* para comprobar la presencia de withanolidos en diferentes clados. Descubrimos que tres de estos withanolidos no solo eran predominantes en los miembros de la tribu *Physalideae*, sino que también se encontraban en los géneros *Datura* y *Cestrum*, así como en la sección torva del género *Solanum* (Figura 9).

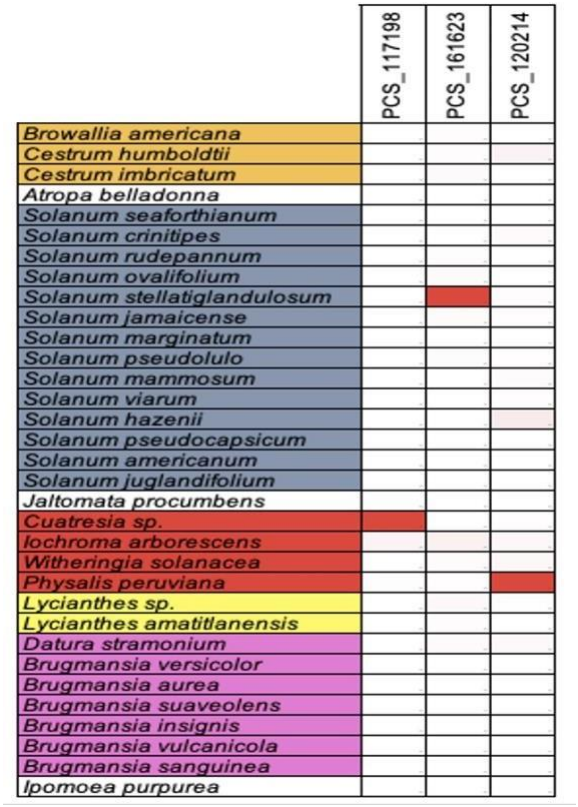


Figura 9. Distribución de Withanolidos en Especies de *Solanaceae* adicionales. La figura muestra la distribución de withanólidos en diferentes especies adicionales de la familia *Solanaceae*. La intensidad del color rojo indica la concentración de withanólidos, donde un rojo más intenso representa una mayor concentración. Las especies están categorizadas en diferentes colores de fondo para destacar distintos clados filogenéticos.

Se llevó a cabo la comparación de perfiles metabólicos entre Physalideae y Lycianthes, por medio de MetaboAnalyst, con el objetivo de identificar diferencias en las rutas metabólicas asociadas con la evolución y diversidad de estas especies. Para realizar este análisis, se emplearon métodos estadísticos, incluyendo t-score y t- score ortogonal, para evaluar la significancia y relevancia biológica de las diferencias en la abundancia de metabolitos.

Los resultados obtenidos revelan una clara diferenciación en términos de perfiles metabólicos entre los dos clados de plantas analizados. En el gráfico de t-score y t- score ortogonal, se observa una marcada separación entre los dos grupos, lo que indica que la abundancia de metabolitos es distintiva entre estas especies pertenecientes a diferentes clados (Figura 10) (Langfelder & Horvath, 2008).

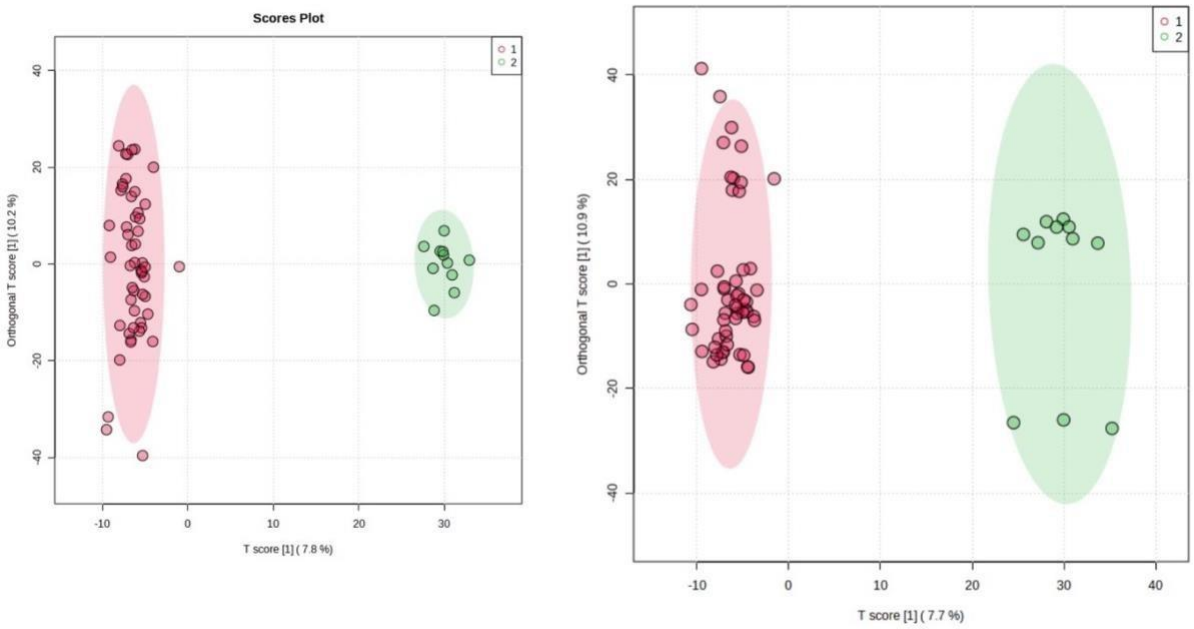


Figura 10. Análisis de diferenciación metabólica entre clados de plantas mediante t- score y t-score ortogonal (rojo representa Physalideae y verde *Lycianthes*).

Los metabolitos ubicados en las regiones más alejadas del gráfico, con valores extremos de t-score y t-score ortogonal, son aquellos que contribuyen significativamente a la diferenciación entre los dos clados de plantas (Huang et al., 2020; Manohar et al., 2020).. Estos metabolitos candidatos podrían tener roles importantes en la determinación de las características únicas y las vías metabólicas específicas asociadas con cada clado (Mehta et al., 2021).

2.2 Análisis de Transcriptomas y Expresión Génica:

En el curso de nuestra investigación, se obtuvieron inicialmente 296 semillas correspondientes a 36 taxones distintos dentro de la tribu *Physalideae* para su estudioy análisis. Sin embargo, nos enfrentamos a un desafío significativo en la fase de germinación, ya que no todas las semillas lograron germinar satisfactoriamente. A pesar de sembrar las semillas en más de tres ocasiones, solo pudimos estudiar efectivamente 200 semillas, pertenecientes a 25 taxones. Esta limitación en la germinación de semillas restringió el alcance de nuestra investigación a un subconjunto más pequeño de lo planeado inicialmente, pero aun abarcando las principales subdivisiones de *Physalideae*, asegurando una representación amplia y variada de la diversidad de la tribu (Deanna et al, 2019). Este desafío subraya la complejidad de trabajar con material biológico y la importancia de considerar la viabilidad de las semillas en estudios de este tipo. Los datos obtenidos, que incluyenun mínimo de 6Gb de secuencia bruta por muestra, garantizan una profundidad de análisis adecuada, reflejando alta calidad y fiabilidad en los resultados obtenidos (Chow et al, 2014).

Las especies que constituyeron el enfoque de este estudio incluyen *Darcyanthus spruceanus*, *Dunalia spinosa*, *Eriolarynx australis*, *Eriolarynx fasciculata*, *Iochroma edule*, *Iochroma gesnerioides*, *Iochroma lorense*, *Iochroma sp. nov*,*Lycianthes aff. pauciflora*, *Lycianthes inaequilatera*, *Lycianthes pauciflora*, *Lycianthes sp.*, *Alkekengi officinaron*, *Physalis angulata*, *Physalis cinerascens*, *Physalis ixocarpa*, *Physalis peruviana*, *Physalis sp.*, *Quincula lobata*, *Trozelia umbellata*, *Vassobia breviflora*, *Withania somnifera* y *Witheringia*

solanacea.

En el proceso de secuenciación de nuestro estudio, hemos obtenido un promedio de 190,365,630 lecturas por muestra. Es importante destacar que se observó una calidad ligeramente menor en el extremo reverse de los transcritos, aunque esta calidad se mantuvo siempre dentro de parámetros aceptables (Anexo 1). Durante la revisión delos datos, tomamos la decisión de conservar únicamente aquellas secuencias que superaran los 140 pares de bases. Esta elección resultó en la eliminación de más del50% de la información inicial. Este fenómeno sugiere una alta redundancia en los datos de secuenciación, posiblemente atribuible a la amplificación y secuenciación múltiple de fragmentos de RNA, lo que podría conducir a una sobre-representaciónde ciertas secuencias (Conesa et al, 2016).

Además, en cuanto al análisis de calidad de los datos de secuenciación, se observó que la distribución del contenido de GC por secuencia se encuentra dentro de los parámetros adecuados y cumple con las expectativas previas (Anexo 2). Estehallazgo se confirmó mediante un análisis exhaustivo con FastQC, que no solo demostró una distribución equilibrada del contenido de GC entre las secuencias, sinotambién un bajo nivel de bases ambiguas en todas las posiciones de las lecturas (Babraham Bioinformatics - FastQC a Quality Control Tool for High Throughput Sequence Data, n.d.). Este resultado subraya la coherencia del proceso de secuenciación y la ausencia de factores significativos que pudieran comprometer lacalidad de los datos obtenidos.

Tabla 2. **Estadísticas de lecturas crudas y filtradas por Trimmomatic.** Esta tabla presenta las estadísticas de calidad de las secuencias de ARN antes y después del proceso de filtrado con Trimmomatic. Incluye el número total de secuencias, las secuencias filtradas, la longitud de las secuencias y el contenido de GC. Los datos se dividen en lecturas de extremo 3' y extremo 5', comparando los datos iniciales con los filtrados para evaluar la efectividad del proceso de limpieza.

Muestra	Datos iniciales 3'	Datos filtrados 3'
Tipo de Archivo	Llamado de bases convencional	Llamado de bases convencional
Codificación	Sanger / Illumina 1.9	Sanger / Illumina 1.9
Total Secuencias	190185450	6789321
SecuenciasFiltradas	0	183396129
Longitud de las secuencias	150	140-150
%GC	43	43
Muestra	Datos iniciales 5'	Datos filtrados 5’
Tipo de Archivo	Llamado de bases convencional	Llamado de bases convencional
Codificación	Sanger / Illumina 1.9	Sanger / Illumina 1.9
Total Secuencias	190545810	6789321
SecuenciasFiltradas	0	183756489
Longitud de las secuencias	150	140-150
%GC	43	43

Tabla 3. Estadísticas de ensamblajes. Esta tabla muestra las estadísticas de ensamblaje de secuencias de ARN, incluyendo el número total de contigs (N), contigs mayores a 1000 bp (N1000), el número de contigs hasta N50 (N:N50), y varias medidas de longitud de contigs, como la longitud mínima, media, promedio, máxima, y la suma total de las longitudes de los contigs ensamblados. Estos datos ayudan a evaluar la calidad y completitud del ensamblaje.

ITEM	PROMEDIO
N	189735
N1000	56519
N:N50	18627
Longitud mínima	1000
Longitud media	1673
Longitud promedio	1928
N50	2042.625
Longitud máxima	13229.3333
Suma de longitudes	110098333

En nuestro análisis, utilizamos los ensamblajes DeNovo de los transcriptomas comoinsumo para el software Salmon, de los cuales se generaron índices. Salmon utilizaeste índice para alinear de manera eficiente las lecturas de secuenciación y estimar los niveles de abundancia de los transcritos. La herramienta identifico un total de 921,627 genes. En promedio, la longitud de los genes fue de aproximadamente 902.05 bases, mientras que la longitud efectiva, que ajusta la longitud de los genes en función de la probabilidad de muestreo, fue en promedio de 620.72 bases. Además, el nivel promedio de expresión génica, medido como Transcripts Per Million (TPM), fue de 37.25, lo cual proporciona una valiosa perspectiva

sobre la abundancia relativa de estos transcritos en el conjunto de datos. Por último, el número promedio de lecturas que mapean a estos genes, un indicador clave de la expresión génica, fue de 332.45 (Patro et al, 2017).

La longitud máxima de los genes alcanzó los 12,039 nucleótidos, mientras que la longitud efectiva, tuvo un pico de 11,752.88. En cuanto al TPM se observó un valormáximo 9,182.62. El número máximo de lecturas mapeadas a un gen específico fuede 64,163.52.

Por otro lado, los valores mínimos fueron: La longitud más corta registrada fue de 258 nucleótidos, y la longitud efectiva más baja fue 34.94. En términos de TPM, algunos genes mostraron una expresión mínima o nula, con un valor más bajo de 0,lo mismo que para el número de lecturas, donde el mínimo registrado también fue 0.

Regresando a los análisis de ortología, para analizar la expresión genética y las secuencias en un conjunto de especies sin genoma de referencia. Identificamos un total de 64,647 grupos de ortólogos en nuestras muestras, a los cuales pudimos asignar anotaciones funcionales a 35,558, incluyendo 60 grupos que contenían ortólogos de genes implicados en la biosíntesis de withanolidos. Los resultados obtenidos en esta tesis revelan un análisis exhaustivo de un conjunto de 19914 HOGs(Grupos Ortogonales Jerárquicos), de los cuales 15458 fueron exitosamente anotados.

Además, se identificaron 19618 HOGs que contaban con un ID de entrez, lo que permitió una caracterización más detallada de estos grupos. Asimismo, se encontróque 3958 HOGs disponían de una anotación específica mediante palabras clave proporcionadas por la herramienta DAVID. De un total de 19,683 HOGs (Grupos Ortólogos Homólogos), 4,333 de ellos (aproximadamente el 22.01%) no tuvieron anotación en DAVID, mientras que 15,350 HOGs (aproximadamente el 77.99%) sífueron anotados en DAVID. Además, 1,282 HOGs (aproximadamente el 6.51%) fueron identificados dentro de un pathway de metabolitos secundarios, y 116 HOGs(aproximadamente el 0.59%) fueron anotados en un cluster de metabolitos secundarios.

Contrastamos la expresión génica entre tres clados filogenéticos y el grupo externo Lycianthes, que no sintetiza withanólidos, frente a Iochrominae y Physalineae-Withaninae, que sí lo hacen. Utilizamos análisis de expresión diferencial con EdgeR para identificar grupos de ortólogos con expresión significativamente distinta (FDR < 0.05) entre los clados. A estos grupos los denominamos ortólogos diferencialmente expresados (ODEs).

2.3 Análisis Filogenético y Clasificación de Especies

Para rastrear la evolución de la vía biosintética de los withanolidos, analizamos la presencia, el número de copias y la distribución de genes de withanolidos en los genomas de 25 especies de *Solanaceae*, incluyendo 5 especies conocidas por producir withanolidos (*Atropa belladonna*, *Datura inoxia*, *Datura stramonium*, *Iochroma cyaneum*, *Physalis pruinosa*, *Physalis grisea*), así como miembros de los géneros *Capsicum* y *Solanum*. Además, utilizamos datos transcriptómicos de 59 especies para determinar si estos genes se expresan en diferentes clados de Solanaceae (Figuras 6 y 7).

El gen 24-ISO se distribuye de manera irregular entre diferentes especies. Se encuentra en el genoma de todos los géneros productores de withanólidos conocidos, como *Physalis*, *Iochroma*, *Datura* y *Atropa*, así como en *Capsicum* y miembros de la sección *Leptostemonum* de *Solanum*. En los géneros que producen withanólidos, este gen aparece en dos copias, mientras que en *Capsicum* y en la sección *Leptostemonum* de *Solanum* solo se presenta una copia. La Figura 11 muestra la filogenia de los parálogos de DWF1 en los genomas de *Solanaceae*, destacando la presencia del gen 24-ISO en los diferentes géneros mencionados.

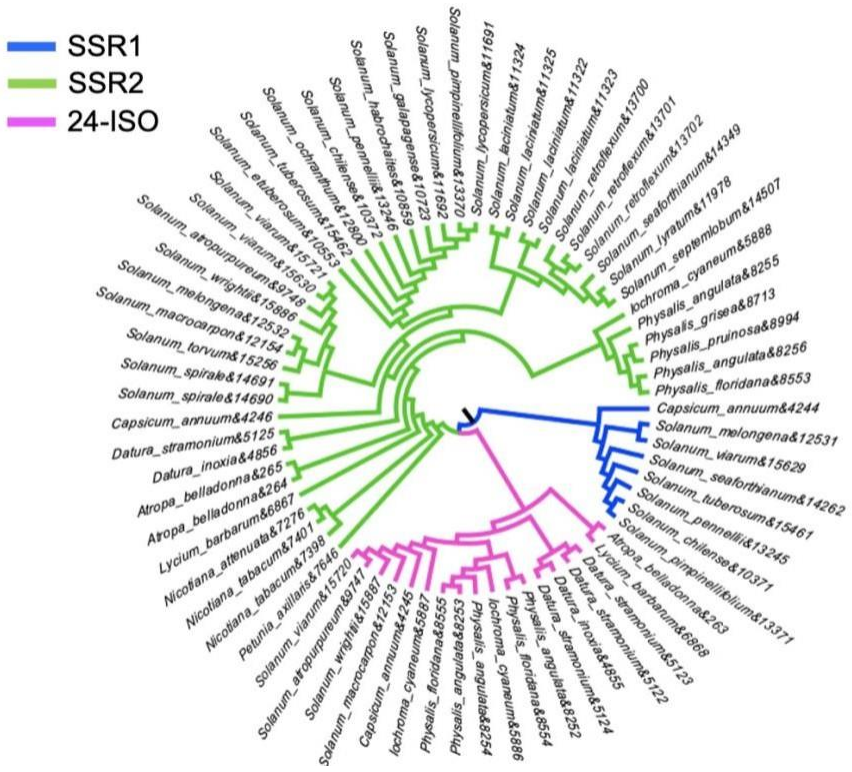


Figura 11. Filogenia de los parálogos de DWF1 en los genomas de Solanaceae. Los colores en la Figura 6 representan los diferentes parágolos de DWF1 en los genomas de Solanaceae: SSR1 está representado en azul, SSR2 en verde, y 24-ISO en rosa.

Confirmamos la expresión de dos copias de 24-ISO en hojas de miembros de los géneros *Physalis*, *Iochroma*, *Witheringia*, *Brugmansia*, *Datura* y *Solandra*. Sin embargo, no se observó expresión de 24-ISO en hojas de *Solanum*. También analizamos dos parálogos cercanos de 24-ISO, los genes SSR1 y SSR2 involucradosen la biosíntesis de Brassinolide y SGA respectivamente. Ambos genes estánpresentes en los genomas de todos los clados de Solanaceae. No obstante, SSR1 se expresa en hojas de todos los clados, mientras que SSR2 muestra preferencia por expresarse en hojas de especies de *Solanum* productoras de SGA (Figura 12).

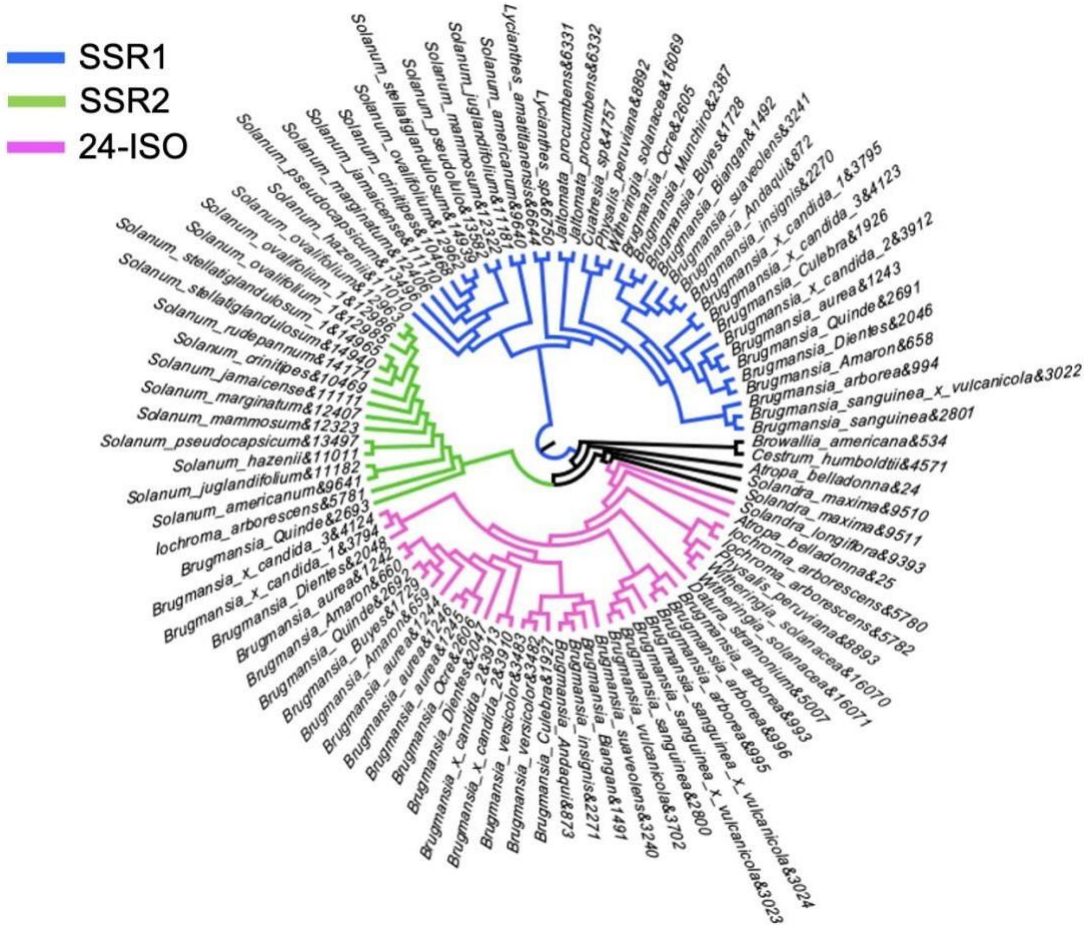


Figura 12. Filogenia de los parálogos de DWF1 en los transcriptomas de Solanaceae.

Los colores en la Figura 12 representan los diferentes parágolos de DWF1 en los transcriptomas de Solanaceae: SSR1 está representado en azul, SSR2 en verde, y 24-ISO en rosa.

Mediante el uso de Astral pro2 y 250 grupos de ortólogos presentes en todas las muestras, establecimos las relaciones filogenéticas entre las especies estudiadas. Losresultados, que concuerdan con estudios filogenéticos previos, clasifican a las especies en cuatro clados: el género *Lycianthes*, *Cuatresia riparia*, la subtribu *Iochromineae* y los miembros de las subtribus *Physalineae* y *Withaninae* (Figura 13).

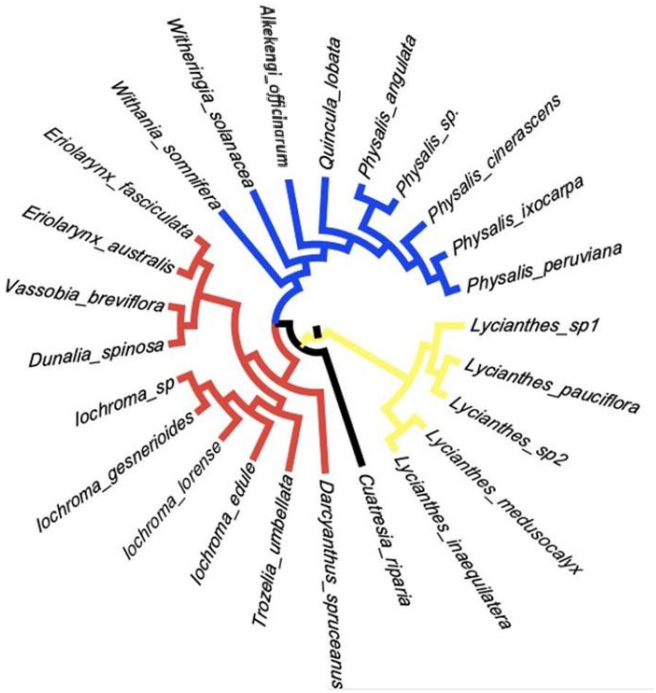


Figura 13. Filogenia de ASTRAL con los genes de copia única. El árbol filogenético muestra las relaciones entre especies de Physalideae usando genes de copia única, que están presentes solo una vez en el genoma, facilitando la interpretación filogenética. Los colores representan diferentes géneros: azul para Physalis, rojo para Iochroma y relacionados, amarillo para Lycianthes (grupo externo), y negro para géneros diversos como Cuatresia y Darcyanthus.

La filogenia obtenida en nuestro análisis muestra coherencia con la reportada en el artículo de Deanna et al. 2019. El análisis filogenético realizado mediante Orthofinder revela que

Quincula lobata se encuentra como una especie externa en relación al clado del género Physalis, lo cual sugiere que Physalis es monofilético. Además, se observa que Witheringia somnifera y Alkekengi officinarum están agrupadas fuera del clado principal de Physalis, y que Iochroma, Dunalia y Vassobia forman un clado distinto. Esta disposición refleja una robustez en términos de la proximidad de sus miembros, coincidiendo con lo reportado en la literatura y apoyando la consistencia de nuestras observaciones filogenéticas.

Además, para explorar las relaciones entre las muestras y visualizar el patrón de expresión génica entre especies, se generó un dendrograma jerárquico a partir de la matriz de correlación (Figura 9). Este análisis corroboró los cuatro clados identificados en el análisis filogenético. El grupo Physalidinae se agrupa en un clado en la parte inferior izquierda de la figura, mientras que Iochrominae forma un clado distinto en la parte superior derecha. Las especies del género Lycianthes están estrechamente relacionadas y se agrupan en el centro de la figura. Por otro lado, Cuatresia se encuentra como un clado aislado en la parte inferior de la Figura 9.

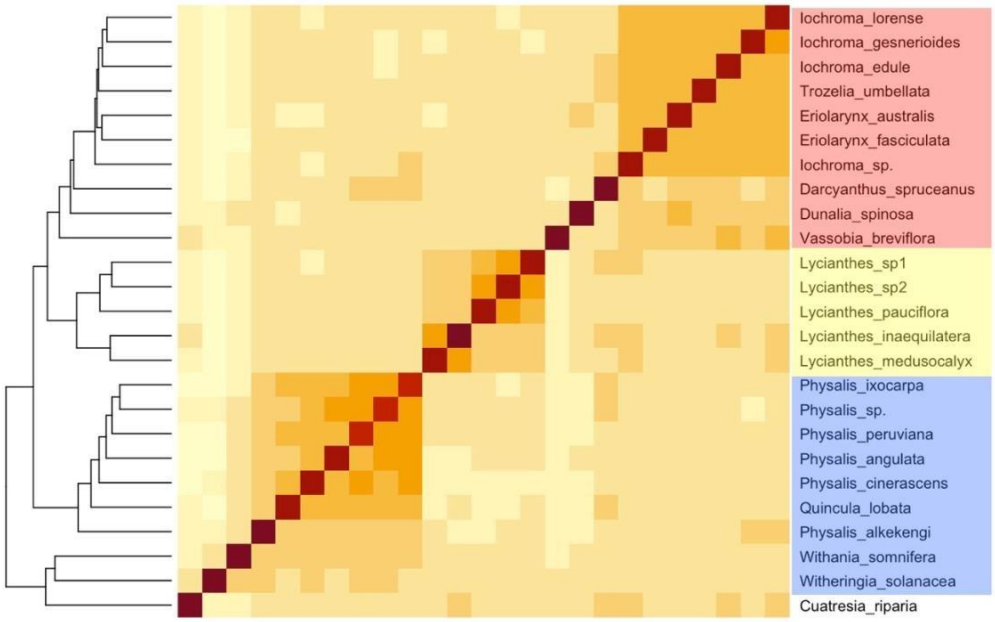


Figura 14. Mapa de calor de correlación de expresión génica entre especies

2.4 Integración de Datos y Correlaciones Metabolómicas, Transcriptómicas y Filogenéticas:

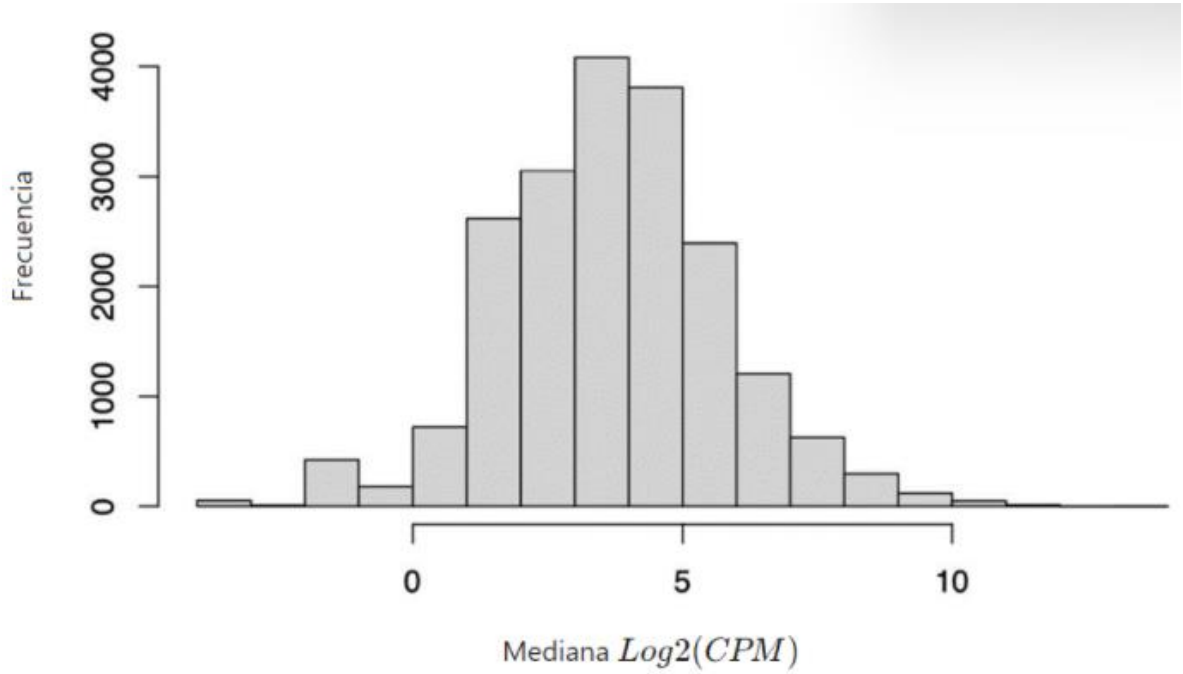


Figura 15. Histograma de Normalización de EdgeR

Para el análisis de expresión génica diferencial, empleamos el paquete EdgeR con la matriz de recuentos de expresión normalizados, utilizando la estadística AOG como entrada. Se filtraron los grupos ortólogos (OGs) con baja expresión aplicando un umbral de -3 en log2CPM. Los recuentos de genes se normalizaron mediante el método de medias recortadas de valores M (TMM), y la varianza de los recuentos de lectura fue modelada siguiendo una distribución de Poisson. Las matrices filtradas y normalizadas resultantes contenían expresiones de 19,683 OGs (Figura 10).

En la determinación de la variabilidad biológica de los genes, se observó que disminuye a medida que aumenta la expresión promedio de los mismos. Esto se evidencia por la aglomeración de puntos negros que se extienden verticalmente a bajos niveles de expresión promedio y que se condensan hacia un nivel más bajo de variabilidad a medida que la expresión promedio incrementa. La línea de tendencia azul (Trend) refleja esta relación decreciente, mostrando cómo la dispersión estimada disminuye a medida que aumenta el logaritmo de los recuentos por millón (CPM) (Figura 11).

El gráfico también presenta dos líneas horizontales, una roja (Common) y una azul(Trend), que representan dos formas de estimar la dispersión. La línea roja horizontalmuestra un nivel común de dispersión que no cambia con la expresión promedio delgen, mientras que la línea azul se ajusta a los datos, reflejando la tendencia real observada en la dispersión a diferentes niveles de expresión. La discrepancia entre estas dos líneas sugiere que un modelo de dispersión que varía con el nivel de expresión (el modelo "Tagwise") es más representativo de los datos que un modelode dispersión común.

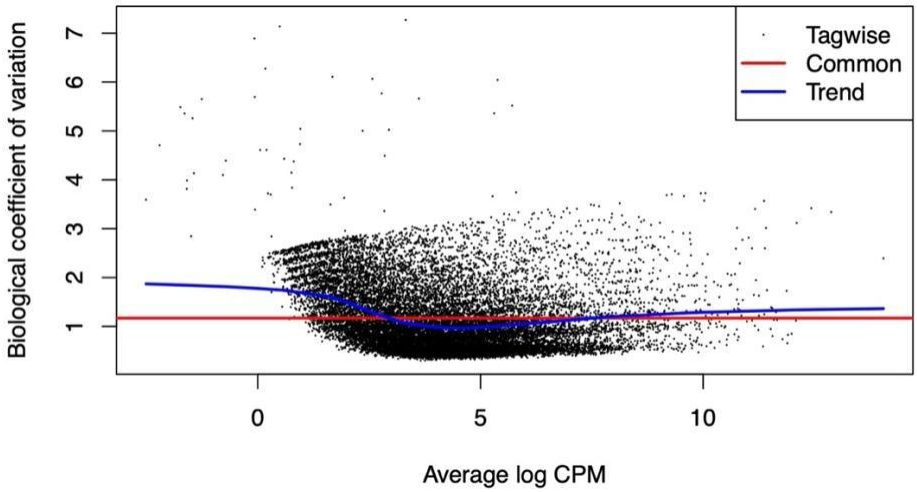


Figura 16. Gráfico de Dispersión de Variabilidad Biológica (BCV) vs. Promedio Log CPM

En el análisis comparativo realizado mediante EdgeR en el grupo de subtribus de *Physalideae*, se identificaron en total 19,654 orthogrupos en el estudio, de los cuales714 orthogrupos presentaron comparaciones significativas en al menos una de las dos tribus de *Physalideae* .

Centrándonos específicamente en la subtribu *Iochrominae*, se encontraron 156 ortogrupos significativos. De estos, cuatro se encuentran previamente anotados con una función relacionada con el metabolismo secundario, perteneciendo a la ruta de los withanolidos y alcaloides tropanos.

Por otro lado, en la subtribu *Physalineae* se identificaron 244 ortogrupos significativos. De ellos, 10 ortogrupos ya contaban con anotaciones previas relacionadas con el metabolismo secundario. Estos son cinco withanolidos, dos acil-azucares, dos glicoalacaloides esteroidales y un alcaloide tropano . Con respecto a *Lycianthes*, se encontraron 89 ortogrupos significativos, donde 3 de ellos estaban anotados en su orden como: CYP450, GAME6 y GAME11.

Por último, en la tribu Physalideae, se encontraron un total de 284 ortogrupos con comparaciones significativas. Entre ellos, 7 ortogrupos tenían anotaciones previas asociadas al metabolismo secundario. Estos genes anotados eran: CAT2, SQE, UGT, esqualenemonooxigenasa, ITTS1, citocromo P450 y 2ODD1.

En el análisis comparativo entre Physalineae y Iochrominae, se identificaron 314 orthogrupos que mostraron significancia estadística en ambas subtribus. De manerainteresante, dentro de este conjunto de 30 orthogrupos, se encontraron 3 de ellos queestaban previamente anotados como genes implicados en el metabolismo secundario. Estos genes anotados fueron identificados como SPDS, ASAT3 eICMEL2.

En el transcurso del uso de la herramienta EdgeR, se identificaron un total de 21 orthogrupos con genes potencialmente candidatos, con un alto grado de relevancia en relación con la producción y variabilidad de withanólidos en la tribu Physalideae.Estos orthogrupos fueron seleccionados meticulosamente, basándonos en criterios rigurosos de significancia estadística, con un False Discovery Rate (FDR) menor a 0.1, lo que asegura la robustez de nuestros hallazgos. Además, estos orthogrupos yacontaban con anotaciones relacionadas con diversos grupos de metabolitos secundarios, tales como glicoalcaloides esteroidales, acilazucares y alcaloides tropanos, lo que sugiere su papel multifuncional en la biosíntesis de compuestos secundarios. Es relevante destacar que estos orthogrupos candidatos también mostraron significancia en una o más comparaciones, respaldando aún más su importancia en el contexto de la síntesis de withanólidos.

FPPS1 (Farnesyl Pyrophosphate Synthase 1): FPPS1 es una enzima involucrada en la biosíntesis de isoprenoides, que son precursores clave en la síntesis de withanólidos. La síntesis de isoprenoides es un paso esencial en la ruta de biosíntesisde withanólidos.

P450reductase (Cytochrome P450 Reductase): P450reductase es una enzima que puede estar implicada en la activación de otros genes de la familia P450 para llevara cabo reacciones de oxidación. Esto podría resultar en cambios estructurales en loswithanólidos al introducir o modificar grupos funcionales en la molécula, como hidroxilos, epóxidos o cetonas; cambios pueden alterar la actividad biológica o las propiedades químicas de los withanólidos.

PgCYP88C8 (Cytochrome P450 88C8) y CYP749B10 (Cytochrome P450 749B10): Los citocromos P450, como PgCYP88C8 y CYP749B10, son conocidos por catalizar reacciones de hidroxilación y oxidación en compuestos orgánicos. Estas reacciones pueden dar lugar a cambios estructurales en los withanólidos, como la adición de grupos hidroxilo en posiciones específicas de la molécula.

ITTS1 (Isopentenyl-diphosphate Delta-isomerase 1): ITTS1 está relacionado con la isomerización de isopentenil difosfato (IPP), un precursor de los isoprenoides, que son precursores importantes para la producción de withanólidos.

Squalene Monooxygenase: Esta enzima podría estar involucrada en la síntesis de squalene, que es un paso importante en la producción de esteroides, incluyendo loswithanólidos que son un tipo de esteroides.

PgCYP88C10 (Cytochrome P450 88C10), Pg2ODD1 (2-Oxoglutarate-dependent Dioxygenase 1), PgCYP87G1_2 (Cytochrome P450 87G1_2), PgCYP88C7_1 (Cytochrome P450 88C7_1), CYP71B10: Estos genes pertenecen a la familia de citocromo P450 y, como se mencionó anteriormente, los citocromos P450 a menudo participan en la síntesis de metabolitos secundarios, incluyendo withanólidos.

FPPS1, DXS2, CYP749B10, TR, AAT, GAME17, UGT, PDS: Estos genes pueden estar relacionados con la producción de withanólidos al influir en la biosíntesis de isoprenoides, la modificación de sustratos o la conjugación de compuestos, todos procesos clave en la síntesis de withanólidos.

SSR2 (Site-specific Recombinase 2): SSR2, como recombinasa, puede estar involucrado en la reordenación de secuencias de ADN que controlan la expresión degenes relacionados con la síntesis de withanólidos. Cambios en la regulación génicapueden alterar la disponibilidad de precursores o la actividad de enzimas clave en laruta de biosíntesis, lo que podría tener un impacto en la producción y la estructura de los withanólidos.

Por otro lado, en los análisis de WGCNA, Los datos normalizados y filtrados utilizando el método de Transcripts Per Million (TPM) muestran una mejora significativa en la estructura de la red de co-expresión en comparación con los datosnormalizados por conteo por millón (CPM) seguido de una transformación logarítmica. Esto se evidencia en los gráficos de independencia de escala, donde losvalores de independencia de escala para los datos filtrados TPM están más cerca de1, lo que sugiere una robustez en la estructura de la red y una adhesión

al modelo dered libre de escala. Además, la conectividad media de la red para los datos filtrados

TPM es generalmente más baja en comparación con los datos log_cpm_filtered.

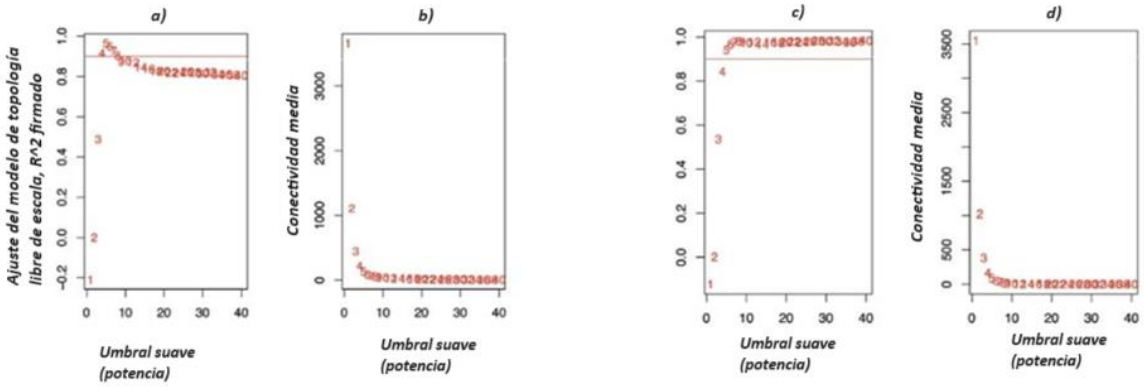


Figura 17. Determinación de la independencia de escala y conectividad media para diferentes umbrales suaves (potencia). En los paneles a) y b) se presentan los datos sin filtrar (Unfiltered_log_tpm), mientras que en los paneles c) y d) se muestran los datos filtrados (Filtered_log_tpm). Los gráficos a) y c) muestran la independencia de escala (Scale Free Topology Model Fit, signed R^2) en función del umbral suave (potencia), y los gráficos b) y d) muestran la conectividad media (Mean Connectivity) en función del umbral suave (potencia).

En el análisis realizado mediante el paquete WGCNA (Weighted Gene Co- expression Network Analysis), se observó una resolución de módulos más clara y definida en el Dendrograma de Clústeres asociado a los datos 'cpm_filtered TPM' encomparación con aquel derivado de 'Log_cpm_filtered'. La calidad de lasegmentación de módulos es un aspecto crítico en la interpretación de redes de co- expresión, ya que los módulos bien resueltos reflejan grupos de genes con patronesde co-expresión altamente correlacionados, lo que a menudo está relacionado con procesos biológicos compartidos o respuestas coordinadas a condiciones fisiológicas. La mejora en la resolución de módulos observada sugiere que la normalización 'cpm_filtered TPM', que ajusta los recuentos de expresión de genes no solo

por la profundidad de secuenciación sino también por la longitud del gen, proporciona un marco más fiable para discernir las relaciones biológicas intrínsecas.Por tanto, el enfoque 'cpm_filtered TPM' demuestra ser superior para detectar la estructura modular en los datos de expresión génica, lo que es esencial para la identificación de biomarcadores potenciales o dianas terapéuticas.

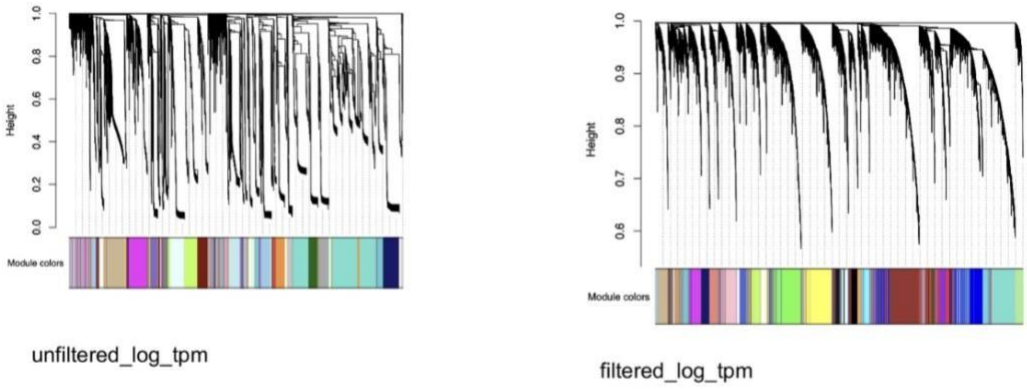


Figura 18. Cluster Dendogram asociado a 'Log_cpm_filtered' y 'cpm_filtered TPM'. Esta figura muestra dendrogramas de agrupamiento jerárquico antes (izquierda) y después (derecha) del filtrado de datos de expresión génica, usando TPM y CPM transformados logarítmicamente. Los colores en la base representan diferentes módulos de genes co-expresados (Anexo 3).

En "Unfiltered_log_tpm" en comparación con los "Filtered_log_tpm", se observan diferencias claras en los resultados que sugieren una mejora en la calidad de la red de co-expresión después de aplicar el filtrado.

Para los datos "Unfiltered_log_tpm", el gráfico de independencia de escala muestra una variabilidad considerable en la adhesión al modelo de red libre de escala, ya quelos valores de R cuadrado oscilan significativamente a medida que aumenta el soft threshold (potencia). En contraste, el gráfico correspondiente a los datos "Filtered_log_tpm" exhibe valores de R

cuadrado más estables y consistentemente altos a medida que aumenta el soft threshold, lo que indica que el filtrado ha mejorado la característica de la red libre de escala.

En lo que respecta a la conectividad media, los datos "Unfiltered_log_tpm" presentan una gran variabilidad en la conectividad a través de los diferentes valoresde soft threshold, mientras que los datos "Filtered_log_tpm" mantienen una conectividad media relativamente baja y consistente en todos los valores de soft threshold evaluados.

El proceso de filtrado aplicado a los datos de expresión génica, representados aquí como 'Filtered_log_tpm', resultó en una mejora significativa en la estructura de la red de co-expresión. Los gráficos de independencia de escala para los datos filtrados muestran un incremento en la consistencia de los valores de R cuadrado, lo que indica una robustez mejorada de la red y una conformidad más precisa con un modelo de red libre de escala. Por otro lado, los gráficos de conectividad media revelan que el filtrado ha conducido a una disminución en la variabilidad de la conectividad, lo cual es indicativo de una red con módulos de co-expresión más definidos.

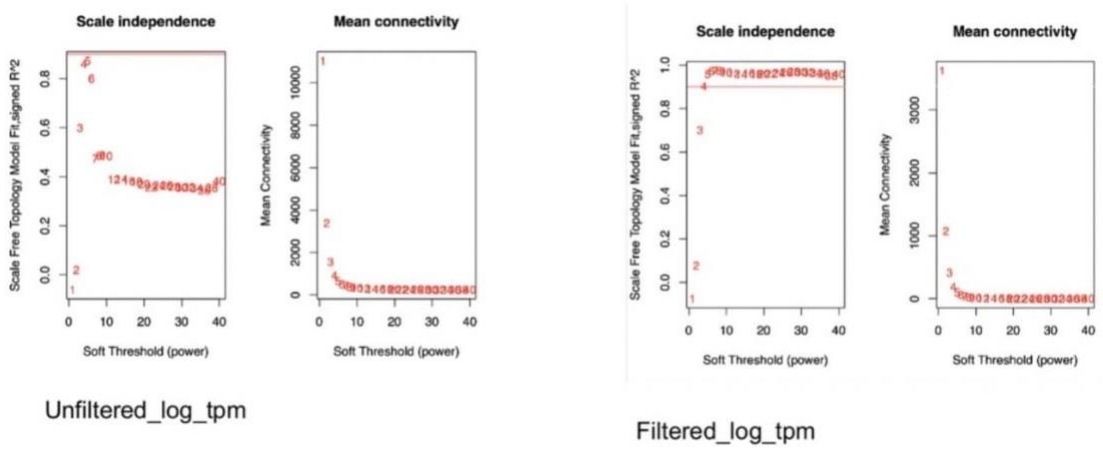


Figura 19. Determinación de la independencia de escala y conectividad media para diferentes umbrales suaves (potencia). Los paneles a) y b) muestran los datos sin filtrar (Unfiltered_log_tpm), mientras que los paneles c) y d) presentan los datos filtrados (Filtered_log_tpm). Los gráficos a) y c) ilustran la independencia de escala (Ajuste del modelo de topología libre de escala, R^2 firmado) en función del umbral suave (potencia), y

los gráficos b) y d) muestran la conectividad media (Conectividad media) en función del umbral suave (potencia).

En datos 'log_cpm_filtered' y 'cpm_filtered_TPM'. La comparación se centró en dosaspectos críticos de la topología de la red: la independencia de escala y la conectividad media. Se encontró que la red derivada de los datos 'cpm_filtered_TPM' exhibía una mayor resolución en la estructura modular. Esto seevidenció por una curva de independencia de escala más estable y una conectividadmedia más consistente, lo que implica una mayor robustez en la red y una adhesiónmás precisa a una topología libre de escala. Estos factores son indicativos de una redcon módulos bien definidos, que son fundamentales para el desciframiento de complejas interacciones biológicas. En consecuencia, los datos 'cpm_filtered_TPM' demostraron ser más eficaces para la identificación de módulos de co-expresión biológicamente significativos, lo que subraya la importancia de una selecciónccuidadosa de métodos de normalización y filtrado en estudios de co-expresión génica.

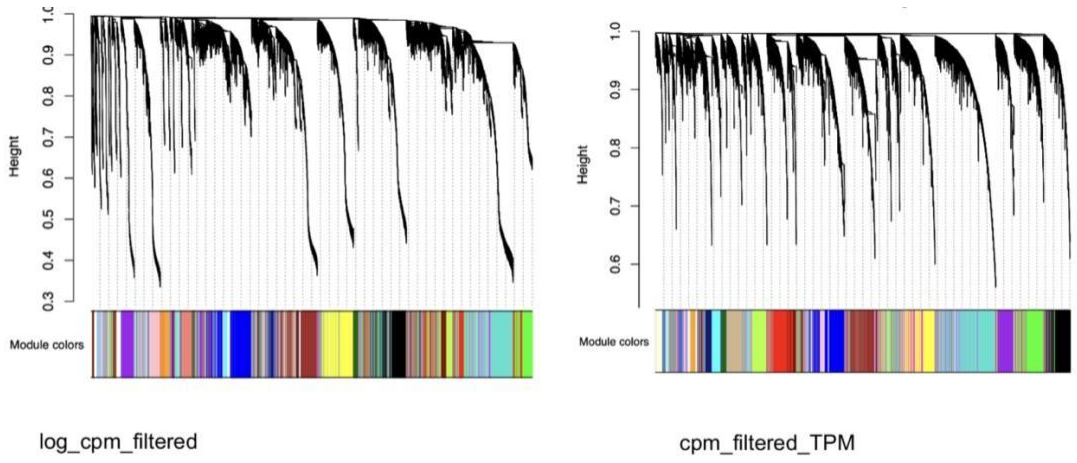


Figura 20. Comparación de la independencia de escala y la conectividad media en redes de co-expresión génica basadas en datos 'log_cpm_filtered' y ' cpm_filtered_TPM'

La herramienta WGCNA permitió determinar patrones de coexpresión genética. En *Lycianthes* se notó una notable ausencia de expresión en ciertos orthogrupos relevantes para el metabolismo secundario de plantas, marcando diferencias expresivas frente a grupos productores de withanolidos.

Por otro lado, empleamos el análisis de redes de coexpresión con WGCNA para detectar módulos de genes relacionados con la variabilidad de withanolidos entre especies. Iniciamos buscando genes que se expresan conjuntamente con el gen 24- ISO, esencial en la síntesis de withanolidos. Dentro del mismo módulo de coexpresión de 24-ISO, encontramos 17 genes conocidos por su rol en la biosíntesisde withanolidos, así como veintisiete grupos ortólogos implicados en el metabolismo de glicoalcaloides esteroides, incluyendo factores de transcripción como GAME9, ERF y BHLH. El análisis revela variaciones expresivas entre especies. Mientras que *Lycianthes* exhibe mínima actividad génica relacionada, en *Physalis* se aprecian altos niveles de expresión (Figura 21).

Identificamos 802 genes con expresión diferencial, ricos en funciones de oxigenasas-oxidoreductasas y CYP450. Nueve de ellos relacionados con withanolidos, como ACCT y CYP71B10, tenían una expresión más alta en *Physalideae* que en *Lycianthes* y variaban entre subtribus de *Physalideae*. Nuestro análisis de correlación gen-especie reveló patrones que vinculan ciertos genes con grupos taxonómicos específicos. Por ejemplo, ALC está presente en *Iochrominae*, mientras que FPPS1 se halla en todas las especies de *Physalis*,sugiriendo una asociación genética con estos taxones. TR, compartido entre *Physalis*, *Iochroma* y *Eriolarynx*, demuestra una correlación más amplia. Genes como PgCYP88C10 y PgCYP88C7 se detectan en *Iochrominae* y *Physalineae*, indicando una proximidad evolutiva. PgCYP87G1_2 se distribuye en toda la tribu *Physalideae*, mientras que DXS2 y Pg2ODD1, presentes incluso en *Lycianthes*, parecen desempeñar roles fundamentales en múltiples especies (Figura 21).

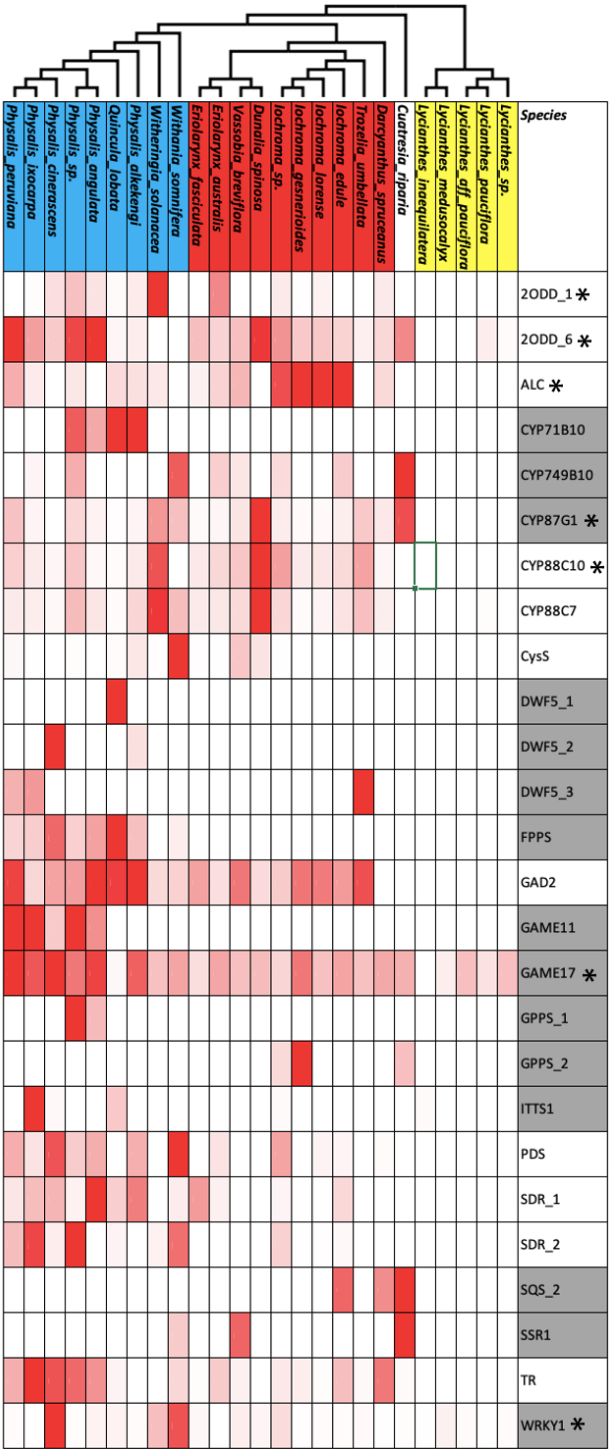


Figura 21. Análisis de Expresión Génica para Biosíntesis de Withanolidos en Especies de Solanaceae

En el análisis de enriquecimiento, el único clúster que se identificó con una significancia mayor a 0.1 (ajustada utilizando el método de Benjamini-Hochberg) fue el observado en la comparación entre *Physalineae* e *Iochrominae*, denominado 'UP_KW_LIGAND,' el cual está asociado con la actividad Heme. El método de Benjamini-Hochberg se utilizó para controlar la tasa de falsos descubrimientos, ajustando los valores de p obtenidos en el análisis. Dentro de los genes relacionados con este clúster, se destacan los siguientes: L-ascorbate peroxidase 5 (LOC101264583), Cytochrome P450 711A1 (LOC107838975), Cytochrome P450 71AU50 (LOC107872290), Cytochrome P450 CYP736A12 (LOC101247609), Cytochrome P450 CYP736A12-like (LOC101244275), Cytochrome P450 CYP749A22-like (LOC101260816), Ethylene receptor NR (NR), Geraniol 8-hydroxylase-like (LOC101262043), Peroxidase 70-like (LOC112940153) y Premnaspirodiene oxygenase-like (LOC101262745).

Empleando el análisis de redes de coexpresión ponderada de genes (WGCNA, por sus siglas en inglés), calculamos las correlaciones de Pearson para la expresión de ciertos grupos de genes ortólogos (OGs) en relación con la concentración de site Withanolidos específicos (Figura 10). Posteriormente, identificamos OGs denominados ExWs, que mostraban correlaciones significativas ($p < 0.05$) con cualquiera de los Withanolidos estudiados. Como paso final, nos enfocamos en encontrar módulos de coexpresión que presentaran una alta proporción de ExWs, resultando en la identificación de dos módulos distintos.

Se identificaron un total de 39 módulos. Dentro de estos, los módulos 'greenyellow'y 'lightcyan' destacaron por tener el mayor número de correlaciones positivas en relación con la variable Count_Pos_A, con cuatro correlaciones positivas cada uno.Por otro lado, el módulo 'steelblue' mostró una preeminencia en la variable Count_Pos_Logcpm, registrando el mayor número de correlaciones positivas con untotal de cinco.

En lo que respecta a las correlaciones negativas, el módulo 'black' fue el más destacado para la variable Count_Neg_A, con tres correlaciones negativas. De manera similar, el módulo 'skyblue3' mostró la mayor cantidad de correlaciones negativas en relación con Count_Neg_Logcpm, contabilizando seis en total.

Capitulo 3: Discusion

3.1 Selección de Plantulas y Preparacion:

En nuestro estudio sobre la tribu Physalideae, la selección de semillas y su posterior germinación desempeñaron un papel fundamental. Se obtuvieron 296 semillas de 36 taxones distintos, representando las subtribus Physalidinae, Withaninae, y Iochrominae, así como Lycianthes como grupo externo. Esta selección buscaba abarcar una representación extensa de la diversidad en Physalideae, una estrategia similar a la empleada por autores como Yang et al. (2021) en su estudio de Witheringia solanacea. Sin embargo, enfrentamos desafíos en la germinación, lo cual es común en estudios botánicos, como se observa en la investigación de He et al. (2007). A pesar de los intentos repetidos, solo 200 semillas de 25 taxones germinaron exitosamente, lo que limitó, pero no disminuyó significativamente, la amplitud de nuestra investigación. Esta variabilidad en la germinación puede atribuirse a factores como la viabilidad de las semillas y las condiciones de almacenamiento previas, aspectos destacados en estudios como los de Lan et al. (2009).

Para mejorar la tasa de germinación, se implementaron técnicas especializadas, incluyendo la inmersión de semillas en una solución de ácido giberélico y el uso deun fotoperiodo controlado, como se describe en la metodología. Esta técnica, basadaen la eficacia de los tratamientos con ácido giberélico reportada por Sang-ngern et al. (2016), fue crucial para optimizar la germinación. A pesar de estos esfuerzos, lagerminación fue menor a la esperada, lo que resalta la importancia de considerar lavariabilidad biológica y las condiciones de almacenamiento de las semillas en futuras investigaciones.

La decisión de utilizar plántulas para estudiar la distribución y concentración de metabolitos secundarios en Physalideae se basó en la importancia de examinar tejidos jóvenes de plantas, donde la biosíntesis de withanólidos es más activa. Esta elección metodológica sigue la línea de razonamiento de Misra et al. (2008), quienesdestacaron la importancia de examinar tejidos específicos para un análisis detalladode metabolitos. Las plántulas ofrecen una ventana única a los procesos metabólicos tempranos, permitiendo una observación detallada de la distribución y concentraciónde withanólidos en diferentes tejidos. Esta estrategia es coherente con la literatura existente, que indica una regulación espacial y temporal específica en la biosíntesisde estos compuestos, como lo demuestran Misico et al. (2011) en su estudio sobre Withania somnifera.

3.2 Metabolómica:

La eficiencia de la extracción de metabolitos se validó mediante la comparación con estándares internos y la observación de la consistencia en los perfiles metabolómicos obtenidos. El uso de ribitol como estándar interno asegura la precisión en la cuantificación y comparabilidad de los datos (Roessner et al., 2001, 'Metabolic Profiling Allows Comprehensive Phenotyping of Genetically or Environmentally Modified Plant Systems'). La extracción se realizó sumergiendo el tejido vegetal en un solvente adecuado, seguido de la homogenización y centrifugación para separar los metabolitos del material celular. Adicionalmente, al no estar presente naturalmente en las muestras, el ribitol evita cualquier interferencia en los resultados.

La identificación de compuestos en diferentes subgrupos de withanólidos destaca la diversidad y complejidad de los perfiles metabólicos en las especies de *Physalidae* estudiadas. Utilizando la metodología de Xcalibur, se logró no solo identificar estoscompuestos, sino también resaltar la ausencia de withanólidos en el género *Lycianthes*, lo cual refuerza la validez de nuestros hallazgos. Los resultados muestran una clara diferenciación en los perfiles metabólicos entre los dos clados deplantas analizados, evidenciada en el gráfico de t-score y t-score ortogonal. Está marcada separación indica que la abundancia de metabolitos es distintiva entre especies de diferentes clados.

Por otro lado, estas diferencias pueden atribuirse a variaciones en la expresión de enzimas clave en la biosíntesis de metabolitos, como los withanolidos (Sharma et al., 2020), o a diferencias en las metodologías de análisis de perfiles metabólicos (Dunn et al., 2011). Por ejemplo, la elección entre técnicas como la espectrometría de masas (MS) y la cromatografía líquida de alto rendimiento (UPLC) puede tener un impacto significativo. Mientras que MS ofrece mayor sensibilidad y especificidad en la detección de metabolitos, UPLC se destaca en la separación de compuestos con estructuras similares, como se discute en "Comparative Metabolomics in Plant Biology" de Smith et al. (2018). Por lo tanto, la selección deHPLC resulta adecuada, permitiendo una identificación más precisa y detallada de los withanolidos presentes en las especies estudiadas.

Adicionalmente, la variabilidad en la composición química y la actividad biológicade los metabolitos entre distintas especies (Lobatto et al., 2023) también podría influir en estas diferencias observadas. Esta variabilidad es particularmente relevante en el contexto de los withanolidos, donde diferencias sutiles en la estructura química pueden traducirse en cambios significativos en su actividad biológica. Por ejemplo, pequeñas modificaciones en los grupos funcionales o en la conformación estereoquímica de estos compuestos pueden alterar su afinidad por determinados receptores biológicos o su capacidad para interactuar con

enzimas específicas, posiblemente explicando las diferencias entre las especies en el t-scoreortogonal (Mehta et al., 2021; Tong et al., 2011; He et al., 2007).

La distribución irregular de withanólidos en las especies de *Physalidae*, observada en este estudio, sugiere una diversificación metabólica significativa en esta tribu. Tal variabilidad en la producción de estos metabolitos secundarios refleja posibles adaptaciones ecológicas o presiones evolutivas que han moldeado su biosíntesis (Misra *et al*. 2008). Esta hipótesis está respaldada por investigaciones previas, que han identificado y caracterizado withanólidos específicos en *Withania somnifera*, uno de los miembros clave de Physalidaae.

La revisión detallada de la literatura y bases de datos especializadas culminó en la identificación de 237 withanólidos, con información específica de fragmentación disponible para 91 de estos.

3.3 Analisis Transcriptomicos y Filogeneticos:

La eficacia de nuestra búsqueda bibliográfica se ve reflejada en la identificación de21 subgrupos de withanolidos asociados a las especies estudiadas, junto con 135 compuestos reportados en las especies presentes. Sin embargo, la falta de información detallada sobre más de la mitad de los withanolidos encontrados subraya una limitación común en la investigación de metabolitos secundarios en plantas, como también se observa en otros estudios similares. Por ejemplo, estudioscomo los de Suárez-Baron y Pabón-Mora (2014) han enfrentado desafíos similares en la caracterización completa de compuestos secundarios en plantas.

Esta situación no es única de los withanólidos, ya que otros compuestos fitoquímicos también sufren de una falta de información detallada, lo que demuestra la necesidad de métodos más avanzados y exhaustivos en la investigación fitoquímica. La presente investigación destaca la importancia de continuar explorando y caracterizando estos compuestos para comprender mejor su diversidad, biosíntesis y potencial en diversas aplicaciones.

En este contexto, la mejora de la calidad de los datos de secuenciación es crucial para obtener resultados más precisos y fiables. El uso de Trimmomatic para la mejora de la calidad de las lecturas de secuenciación es una decisión estratégica clave en el estudio. Trimmomatic es una herramienta ampliamente reconocida y utilizada para el recorte y filtrado de datos de secuenciación de alta producción, especialmente útil para eliminar secuencias de baja calidad y adaptadores. Según Bolger, Lohse, y Usadel (2014), Trimmomatic se destaca por su eficiencia en la limpieza de lecturas y su flexibilidad para adaptarse a diferentes tipos de datos y necesidades experimentales. La decisión de conservar secuencias que superan los 140 pares de bases, resultando en la eliminación de más del 50% de los datos, sugiere una rigurosa calidad en el procesamiento de datos. Esto es fundamental para reducir la redundancia y la sobre-representación de secuencias, comúnmente observada en datos de secuenciación masiva.

Por otro lado, la elección de Trinity es una herramienta poderosa para el ensamblajede novo de transcriptomas, especialmente en ausencia de un genoma de referencia.Haas et al. (2013) resaltan la habilidad de Trinity para reconstruir con precisión transcriptomas complejos, aprovechando su algoritmo basado en gráficos de Bruijnpara manejar eficientemente la diversidad y complejidad de las secuencias de ARN. El uso de Trinity facilita una reconstrucción detallada y precisa de los transcritos, crucial para estudios de expresión génica y análisis funcionales.

En relación con la contigüidad y continuidad de las secuencias ensambladas, los valores de N50 obtenidos indican una calidad alta. Un N50 elevado sugiere que unaparte considerable de la longitud total de las secuencias está concentrada en secuencias más largas, lo cual es beneficioso para proporcionar información más completa sobre regiones genómicas de interés. La proximidad entre las longitudes media y promedio de las secuencias sugiere una distribución simétrica o aproximadamente normal, lo que indica una representación equilibrada de secuencias cortas y largas en el conjunto de datos. Esta distribución equilibrada es ventajosa para diversos análisis posteriores, ya que garantiza una longitud promediorepresentativa para las secuencias.

El rango entre las longitudes mínimas y máximas de las secuencias revela la presencia de una variedad de secuencias, tanto cortas como largas, lo que añade flexibilidad en el análisis y permite una exploración exhaustiva del material genético. Un N50 alto, en comparación con otras tecnologías de secuenciación, implica que los consensos (contigs) más largos son más representativos del conjunto total de bases, indicando una calidad de ensamblaje superior. Además, una suma más alta delas longitudes de las secuencias sugiere una cobertura más completa de los genomas o regiones objetivo, proporcionando un conjunto de datos más extenso para el análisis.

Finalmente, al comparar nuestros datos con el ensamblaje generado por Garzón-Martínez et al. (2012) y las comparaciones realizadas por Zarate (2013), observamosque nuestro N50 y la longitud media de contigs superan los obtenidos en organismos como *H. symbiolongicarpus*, *Acropora millepora*, *Chuniphyes multidentata* y *Physalis peruviana*, entre otros. Además, en nuestro proceso de filtración de secuencias, logramos filtrar aproximadamente el 97% de los datos, en comparacióncon un máximo del 87% en los estudios mencionados.

La cuantificación de la expresión génica utilizando Salmon-1.6.0 también demuestrauna elección acertada en términos de software. Según Patro, Duggal, Love, Irizarry, y Kingsford (2017), Salmon ofrece una cuantificación rápida y precisa de la expresión génica, aprovechando algoritmos que permiten un mapeo eficiente de lecturas a transcritos. Esto es vital para obtener estimaciones precisas de abundancia de transcritos, un aspecto crítico en el análisis de datos de secuenciación.

En comparación con otros softwares, como HTSeq o featureCounts, que también se utilizan para cuantificar la expresión génica, Salmon ofrece ventajas significativas en términos de velocidad y eficiencia, especialmente en conjuntos de datos grandes y complejos. Según Patro et al. (2017), Salmon no solo es más rápido, sino que también es capaz de corregir sesgos inherentes en los datos de secuenciación, como la composición de GC y la longitud del fragmento, lo que resulta en estimaciones más precisas de la abundancia de transcritos. Estas correcciones permiten obtener datos más fiables y precisos, mejorando la calidad general del análisis de expresión génica.

En nuestro análisis, utilizamos los ensamblajes DeNovo de los transcriptomas como insumo para el software Salmon, de los cuales se generaron índices. Salmon utiliza este índice para alinear de manera eficiente las lecturas de secuenciación y estimar los niveles de abundancia de los transcritos. La herramienta identificó un total de 921,627 genes, el cual podría considerarse un número considerablemente alto. Por ejemplo, en estudios de plantas como el árbol de helecho *Sphaeropteris brunoniana*, la cantidad de genes expresados suele ser mucho menor. En un estudio reciente, se identificaron 11,625, 14,391 y 10,099 unigenes en las raíces, raquis y pinnas de *S. brunoniana*, respectivamente (Peng et al., 2024). Sin embargo, la diferencia puede atribuirse a la variabilidad entre especies, así como a la profundidad y metodología de secuenciación utilizada. Un número elevado de genes identificados puede reflejar la complejidad del transcriptoma de la especie

estudiada. En plantas, especialmente en aquellas que no son modelos comunes en investigación, esto puede indicar una diversidad genética significativa y una riqueza de vías metabólicas y procesos biológicos. Por ejemplo, en estudios de transcriptoma de plantas no modelo, se han identificado una amplia gama de genes que sugieren complejidad en su regulación y función genética (Peng et al., 2024).

En promedio, la longitud de los genes fue de aproximadamente 902.05 bases, mientras que la longitud efectiva, que ajusta la longitud de los genes en función de la probabilidad de muestreo, fue en promedio de 620.72 bases. Además, el nivel promedio de expresión génica, medido como Transcripts Per Million (TPM), fue de37.25, lo cual proporciona una valiosa perspectiva sobre la abundancia relativa de estos transcritos en el conjunto de datos. Aunque no hay estudios directamente comparables, la expresión génica y su regulación varían ampliamente en las plantas,dependiendo de factores como el estrés ambiental, por lo que asegurar condiciones iguales para todas las muestras permitió reducir el efecto de esa variable (Gao et al, 2022).

El número promedio de lecturas mapeadas fue de 332.45, reflejando una expresión génica significativa. Este elevado recuento podría deberse a la duplicación de genes,un fenómeno que induce variaciones en la expresión génica (Das & Bansal, 2019). Específicamente, la duplicación y evolución del gen 24-ISO es crucial en la producción de withanolidos en ciertas especies de plantas, como se observa en Physalis, Iochroma, Datura y Atropa, así como en Capsicum y algunos Solanum. Interesantemente, el gen 24-ISO no se encuentra en todas las especies de Solanum, incluyendo aquellas que dan origen al tomate y la patata, y se han identificado versiones truncadas de este gen en Capsicum annuum y Solanum melongena (Schwander et al., 2014).

Un promedio elevado de lecturas mapeadas a los genes indica una expresión génica activa y significativa. Esto es crucial para entender los patrones de expresión en las plantas estudiadas y puede revelar procesos biológicos importantes en curso, dondeal investigar los agentes externos a las plantas podría identificar procesos adicionales que se asocien a la producción de withanolidos tanto en cantidad como en variedad. Se conoce asociaciones con endófitos en Withania somnífera, donde Los endofitos como *Bacillus amyloliquefaciens* y *Pseudomonas fluorescens* no solo ayudaron a laplanta a tolerar el estrés biótico causado por Alternaria alternata, sino que también incrementaron la expresión de genes clave en la ruta biosintética de los withanolidos(Mishra et al., 2018).

En el análisis de RNA-seq de nuestro estudio, se observaron valores mínimos destacables. La longitud más corta de un gen fue de 258 nucleótidos, y la menor longitud efectiva registrada fue de 34.94 nucleótidos. Estos valores sugieren la presencia de genes con regiones codificantes pequeñas o posiblemente fragmentos de genes más grandes, lo cual es relevante para la comprensión de la estructura genómica y la regulación génica (Agarwal et al., 2017). Además, algunos genes mostraron una expresión mínima o nula, con valores de TPM inferiores a 0 y un número de lecturas también en 0. Este hallazgo es consistente con variaciones en laexpresión génica reportadas en otros estudios, lo que subraya la complejidad del transcriptoma de las plantas (Alfonso & Kapetanidis, 1994; Alharbi & Stevenson, 2020).

La presencia de genes cortos puede tener implicaciones biológicas interesantes, como la codificación de péptidos pequeños con funciones específicas o fragmentos de RNA que participan en la regulación de la expresión génica (Altenhoff et al., 2019; Bharadwaj et al., 2021). Sin embargo, cabe destacar que se identificaron copias truncadas del gen 24-ISO en *Capsicum annuum* y *Solanum melongena*, lo que, al identificar si se asocian a estos fragmentos cortos, añade una

capa adicionala la comprensión de la distribución y función de este gen en distintas especies, abriendo puertas al análisis de mecanismos de regulación génica post-transcripcional. (Emms & Kelly, 2019; Ellegren & Galtier, 2016.

Por otro lado, la elección de OrthoFinder para clasificar genes en grupos ortólogos y parálogos resulta particularmente acertada para el análisis de expresión genética ysecuencias en un conjunto de especies sin genoma de referencia. OrthoFinder, desarrollado por Emms y Kelly (2015), se ha establecido como una herramienta eficaz para identificar grupos ortólogos a través de diferentes especies, proporcionando una perspectiva evolutiva y funcional valiosa. A diferencia de otrasherramientas, OrthoFinder no solo identifica ortólogos, sino que también clasifica los genes en "Hierarchical Orthologous Groups" (HOGs), lo que facilita una comprensión más profunda de la evolución y función de estos genes en diferentes linajes. Este enfoque es especialmente útil en estudios, donde se analizan múltiplesespecies sin un genoma de referencia claro. La capacidad de OrthoFinder para manejar grandes conjuntos de datos y proporcionar resultados precisos y detalladoses fundamental para el análisis de relaciones evolutivas complejas (Kriventseva et al., 2019).

Otra estrategia aquí aplicada fue la decisión de mapear a HOGs en lugar de a genes individuales. Los HOGs proporcionan una visión más holística y evolutiva, yaque cada HOG representa un conjunto de genes ortólogos que se originaron de un gen ancestral común. Este enfoque es consistente con estudios recientes que han utilizado HOGs para explorar la evolución de genes a través de diferentes especies y linajes (Altenhoff et al., 2019). Al mapear a HOGs, el estudio se alinea con una tendencia creciente en la genómica comparativa que busca entender no solo la función de los genes, sino también su historia evolutiva.

Ahora, hablando de la identificación de 64,647 grupos de ortólogos y la anotación funcional de 35,558 de ellos, incluyendo ortólogos de genes implicados en la biosíntesis de withanolidos, es un hallazgo significativo. Esta proporción de genes anotados es notable y sugiere una cobertura extensa de la diversidad genética en tusmuestras. En comparación con otros estudios, esta cantidad de HOGs anotados es considerable y destaca la riqueza de información genética presente en tus muestras (Huerta-Cepas et al., 2019).

Profundizando esta idea, la proporción de genes anotados sugiere una cobertura extensiva de la diversidad genética en las muestras. Esto es importante porque una alta diversidad genética puede revelar una mayor cantidad de funciones biológicas y rutas metabólicas, lo que es crucial en estudios que buscan entender la complejidad de sistemas biológicos específicos (Ellegren y Galtier, 2016). La capacidad de anotar más del 50% de los grupos de ortólogos identificados demuestra la profundidad y la calidad del análisis genómico. Además, al utilizarse la herramienta DAVID, permitió tener un porcentaje de anotación de HOGs del 77,99%, lo cual evidencia su capacidad para proporcionar anotaciones funcionales detalladas (Huang da, Sherman, y Lempicki, 2009).

Para el análisis de expresión génica, se optó por utilizar datos de conteos de lecturas en lugar de valores normalizados como TPM, basándose en la mayor precisión y fiabilidad que ofrecen los conteos normalizados. Un estudio en modelos de xenoinjertos derivados de pacientes (PDX) sugiere que los conteos normalizados son más efectivos para agrupar muestras replicadas con precisión, mostrando un coeficiente de variación más bajo y valores de correlación intraclase más altos. Además, para abordar el desafío de la variabilidad en la longitud total de secuencia por ortogrupo entre especies, se utilizó el valor "A" en los análisis para corregir las diferencias, mejorando así la fiabilidad de los resultados. Este enfoque proporcionauna medida más directa de la abundancia de transcritos y es útil para

identificar patrones de coexpresión basados en la cantidad absoluta de expresión (Zhao *et al*, 2021).

3.4 Integración de Metabolómica, Transcriptómica y Filogenética:

En esta sección clave de la tesis, exploramos la integración sinérgica de la metabolómica, la transcriptómica y la filogenética para obtener una comprensión más profunda y completa de los sistemas biológicos estudiados. Abordamos cómo los perfiles de metabolitos y patrones de expresión génica, analizados en el contexto de las relaciones filogenéticas, revelan interacciones y adaptaciones complejas a nivel molecular y evolutivo. Esta integración multidisciplinaria no solo profundiza nuestra comprensión de los procesos biológicos subyacentes, sino que también abrenuevas avenidas en campos como la farmacología, la agronomía y la conservación, destacando la importancia de un enfoque holístico en la investigación biológica moderna.

Para empezar, realizamos una búsqueda de módulos que contienen genes cuya expresión está asociada al nivel de Withanolidos entre diferentes especies (Figura 8). Los genes se agrupan en módulos basados en la similitud de sus patrones de expresión. Estos módulos representan conjuntos de genes que muestran coexpresión, es decir, sus niveles de expresión varían de manera similar bajo diferentes condiciones experimentales. El dendrograma de clúster facilita la visualización de estos módulos, agrupando genes con perfiles de expresión similares.

En la subtribu Iochrominae, la integración de la información disponible en la literatura y los resultados experimentales recientes arroja una luz reveladora sobre la distribución y diversidad de los Withanolidos (Alfonso, D., & Kapetanidis, I. 1994). Estos metabolitos secundarios, conocidos por sus notables

propiedadesbiológicas y potencial terapéutico, son cruciales para comprender la bioquímica y laecología de las especies de Iochrominae (Chaurasiya et al, 2012).

La literatura existente identifica un espectro extenso de Withanolidos en Iochrominae, que incluyen 5b,6b-EpoxyWithanolidos, C-18 Hydroxylated Withanolidos, C-12 and C-21 Oxygenated Withanolidos, Other Hydroxylated Withanolidos, Acnistins, y Physalins and Withaphysalins. Contrastando con esta información, los análisis cuantitativos, desarrollados en el marco de este trabajo, revelan un perfil ligeramente diferente (Alfonso, D., & Kapetanidis, I. (1994)). Confirman la presencia de 5b,6b-EpoxyWithanolidos, coherente con la literatura, pero también destacan la aparición de 1a,3b-Dihydroxy-D5 Withanolidos, un grupode compuestos no reportados anteriormente (Chaurasiya et al, 2012).

Además, los resultados experimentales aportan una perspectiva más específica sobre la presencia de un withanolido en una especie dentro de la subtribu Iochrominae. Iochroma sp. nov, una especie recientemente descubierta, destaca en la producción de Withanoside V (Nützmann, Scazzocchio & Osbourn, 2018). En el estudio de larelación entre el Withanoside V y la expresión génica, con un enfoque particular enlas correlaciones positivas y la asociación con clústeres y el metabolismo secundario, se ha identificado el módulo 'brown' como el más representativo. Este módulo destaca por incluir genes cuya expresión se ve positivamente correlacionadacon el Withanoside V y que están relacionados con el metabolismo secundario o forman parte de clústeres específicos. A continuación, se describen algunos de los genes más notables dentro de este módulo y sus posibles roles biológicos.

El gen N0.HOG0000024, relacionado con la biosíntesis de alcaloides esteroideos, muestra una correlación positiva con el Withanoside V, lo que sugiere que este último puede influir en la defensa y adaptación de las plantas. El gen N0.HOG0017883, asociado a la lipasa SDP1-like y al metabolismo de esteroides,

también presenta una correlación positiva con el Withanoside V, indicando un posible impacto en el metabolismo lipídico y la regulación de esteroides en plantas.Por último, el gen N0.HOG0012318, importante en el metabolismo de la purina, muestra una correlación positiva con el Withanoside V, sugiriendo que puede afectar la síntesis de nucleótidos y la gestión de energía celular. Estos hallazgos destacan lainfluencia potencial del Withanoside V en varios procesos vitales de las plantas (Agarwal et al, 2017) (Sharma et al, 2020).

Estos genes, ubicados en el módulo 'brown', proporcionan una visión integral de cómo el Withanoside V podría estar influenciando positivamente la expresión génica en relación con el metabolismo secundario y clústeres específicos. El aumento en laexpresión de estos genes sugiere que el Withanoside V podría estar jugando un papel en la modulación de diversas vías biológicas, incluyendo la biosíntesis de alcaloidesesteroideos, el metabolismo lipídico y esteroideo, y el metabolismo de las purinas.

En el estudio de la relación entre el Withanoside V y la expresión génica, con un enfoque particular en las correlaciones negativas y la asociación con clústeres y el metabolismo secundario, se ha identificado el módulo 'turquoise' como el más representativo. Este módulo destaca por incluir genes cuya expresión se ve negativamente correlacionada con el Withanoside V y que están relacionados con elmetabolismo secundario o forman parte de clústeres específicos. A continuación, sedescriben algunos de los genes más notables dentro de este módulo y sus posibles roles biológicos.

El gen N0.HOG0009871, responsable de codificar la enzima geranylgeranyl diphosphate reductase en la vía de los terpenoides, muestra una correlación negativa con el Withanoside V, lo que sugiere que este último podría influir negativamente en la síntesis de terpenoides, afectando funciones biológicas clave como la defensacontra patógenos y la señalización celular. Aun no se conoce que

implicación puedatener el hecho de que una enzima del pathway de los terpenoides afecte a una sustancia generada por ese mismo camino, pero la interacción entre enzimas de la ruta de los terpenoides y los withanolidos no es solo una cuestión de síntesis bioquímica, sino que también tiene un impacto directo en las propiedades biológicas y terapéuticas de estos importantes metabolitos secundarios.

Por otro lado, el gen N0.HOG0010410, que codifica para la UDP-glycosyltransferase 13 en la vía de los flavonoides, también presenta una correlaciónnegativa con el Withanoside V, indicando un posible impacto en el metabolismo delos flavonoides, que son cruciales para la coloración de las plantas, la protección contra la radiación UV y el estrés oxidativo (Gupta et al, 2015). Withanoside V es un glucósido, lo que significa que es una molécula que incluye un grupo de azúcar glucosilado en su estructura. Esto es relevante ya que la presencia de glucósidos enlos withanolidos como el Withanoside V sugiere que la UDP-glycosyltransferase 13 podría estar implicada en su formación o modificación. La glucosilación es un proceso importante en la modificación y actividad biológica de muchos compuestosvegetales, incluidos los flavonoides y withanolidos. Por lo tanto, la correlación negativa con el Withanoside V y la UDP-glycosyltransferase 13 en la ruta de los flavonoides podría tener implicaciones significativas en el metabolismo y la función(PubChem, n.d.).

Estos genes, ubicados en el módulo 'turquoise', proporcionan una visión integral decómo el Withanoside V podría estar influenciando negativamente la expresión génica en relación con el metabolismo secundario y clústeres específicos. La disminución en la expresión de estos genes sugiere que el Withanoside V podría estar jugando un papel en la modulación de diversas vías biológicas, incluyendo la biosíntesis de terpenoides y flavonoides. Estos hallazgos abren nuevas perspectivaspara investigaciones futuras sobre los mecanismos moleculares y las rutas biológicasafectadas por este compuesto.

Sin embargo, la ausencia en los datos cuantitativos de compuestos como C-18 Hydroxylated Withanolidos, C-12 and C-21 Oxygenated Withanolidos, Other Hydroxylated Withanolidos, Acnistins, y Physalins and Withaphysalins, a pesar desu mención en la literatura, es intrigante (Misico et al, 2011. Esto plantea interrogantes sobre las técnicas de detección utilizadas o sobre la variabilidad en lasmuestras analizadas. Podría indicar que algunos Withanolidos están presentes en concentraciones por debajo del umbral de detección de los métodos actuales o que existen diferencias metodológicas y de muestreo entre los estudios literarios y experimentales.

Este análisis comparativo entre la literatura y los resultados experimentales resalta la necesidad de una mayor investigación y refinamiento en las técnicas de detecciónyanálisis. La comprensión más profunda de la distribución y diversidad de los Withanolidos en Iochrominae no solo es fundamental para la fitoquímica y la farmacología, sino también para entender mejor las adaptaciones evolutivas y las interacciones ecológicas de estas especies.

La exploración continuada y detallada de Iochrominae y sus Withanolidos no solo ampliará nuestro conocimiento sobre estos compuestos fascinantes, sino que también podría revelar aplicaciones terapéuticas innovadoras, contribuyendo significativamente al campo de la medicina natural y la farmacología.

Además, este estudio comparativo enfatiza la importancia de considerar tanto el contexto ecológico como la variabilidad genética de las especies al examinar la producción de metabolitos secundarios. La presencia o ausencia de ciertos Withanolidos en diferentes especies de Iochrominae podría reflejar adaptaciones a nichos ecológicos específicos o respuestas a presiones evolutivas (Misico et al, 2011). Por ejemplo, la presencia de Withanolidos únicos en Dunalia brachyacantha,como los dunawithanines G, H y F, y la diversidad de acnistins en Dunalia solanacea, resaltan cómo las diferencias en la química de productos

naturales pueden estar vinculadas a las adaptaciones específicas de cada especie.

La variabilidad observada en los perfiles de Withanolidos entre especies también plantea preguntas sobre los mecanismos de biosíntesis y regulación de estos compuestos en Iochrominae. La ausencia de Withanolidos en especies como Iochroma edule y Dunalia spathulata, a diferencia de otras especies de la subtribu, sugiere rutas metabólicas distintas o diferencias en la expresión génica que podríanestar influenciadas por factores ambientales, genéticos o evolutivos.

Estos hallazgos subrayan la necesidad de una aproximación integradora que combine estudios fitoquímicos, genéticos, ecológicos y evolutivos para comprender completamente la diversidad y el papel de los Withanolidos en Iochrominae. Tal enfoque no solo proporcionará una visión más profunda de la fitoquímica de estas especies, sino que también podría revelar nuevas vías para el desarrollo de fármacos y terapias basadas en Withanolidos, aprovechando su variedad de propiedades biológicas.

En conclusión, la subtribu Iochrominae ofrece un campo fértil para la investigación fitoquímica y farmacológica, con un potencial significativo para descubrimientos innovadores. La exploración continua de sus variados Withanolidos es esencial no solo para la ciencia de productos naturales y la medicina, sino también paracomprender mejor las complejas interacciones entre las plantas y su entorno, así como los procesos evolutivos que moldean la biodiversidad química en el mundo natural.

La utilización de Astral pro2 24 con 250 genes presentes en todas las muestras paradeterminar las relaciones filogenéticas entre las especies estudiadas condujo a la agrupación conjunta de miembros de las subtribus Physalineae y Withaninae. Este resultado nos lleva a presentar conjuntamente la información de ambas subtribus.

En el contexto de Physalineae, un análisis minucioso de la literatura existente y losdatos cuantitativos revela patrones intrigantes en la presencia y ausencia de Withanolidos. Se observa una notable ausencia de subgrupos específicos de Withanolidos, como 1a,3b-Dihydroxy-D5-Withanolidos, sugiriendo que estos compuestos podrían ser raros o no detectados en estudios previos. Sin embargo, la identificación de compuestos no reportados anteriormente, como el 1,3,27-trihydroxy-(20S,22R)-3-methoxywitha-4,6,24-trienolide, indica una posible diversidad química no explorada en Physalineae. La discrepancia entre los datos cuantitativos y las referencias literarias plantea preguntas sobre la precisión de las técnicas de detección y la variabilidad de las muestras (Misico et al, 2011).

Paralelamente, en Withaninae, la evaluación de la literatura y los datos cuantitativos muestra una variedad de Withanolidos, incluyendo 5b,6b-EpoxyWithanolidos y otros D5-Withanolidos, reflejando una significativa diversidad bioquímica. La presencia de compuestos como Withajardins y Aromatic Ring-A Withanolidos and 19-HydroxyWithanolidos destaca la complejidad de la composición química en Withaninae (Misico et al, 2011).

Así, la agrupación de Physalineae y Withaninae bajo un análisis filogenético común revela conexiones interesantes y diferencias significativas en la composición y diversidad de Withanolidos entre estas subtribus, sugiriendo líneas de investigaciónfuturas en su evolución y función biológica.

En el contexto de la subtribu Physalineae, un análisis detallado de la literatura existente y los datos cuantitativos revela patrones intrigantes en la presencia y ausencia de Withanolidos. A pesar de la riqueza de estudios fitoquímicos, se observauna notable ausencia de subgrupos específicos de Withanolidos, tales como 1a,3b- Dihydroxy-D5-Withanolidos y otros compuestos relacionados. Esta omisión sugiereque estos compuestos podrían ser raros o no detectados en los estudios previos realizados en esta subtribu. Paralelamente, la identificación de

compuestos no mencionados en la literatura, como el 1,3,27-trihydroxy-(20S,22R)-3- methoxywitha-4,6,24-trienolide, resalta la posibilidad de una diversidad química inexplorada en Physalineae. No obstante, la discrepancia entre los datoscuantitativos y las referencias literarias, especialmente en lo concerniente a Withanolidos como los C-18 Hydroxylated, plantea interrogantes sobre la precisión de las técnicas de detección actuales y la variabilidad de las muestras (Misico et al,2011).

En este escenario, Alkekengi officinaron emerge como un ejemplo emblemático. Esta especie, representativa de la subtribu Physalineae, ha ganado atención por su diversidad de Withanolidos, reflejando el potencial aún no totalmente explorado en la subtribu. Un compuesto particularmente destacado es el physagulin J, reportado por Li et al. en 2008. Este Withanolido, con su estructura química inusual, sugiere un vasto potencial para aplicaciones en el ámbito de la salud.

Profundizando en los Withanolidos específicos de Alkekengi officinaron, se encuentrancompuestos como physalin T, 11,16-cyclophysalin y 5a,6a-dihydroxy physalin D, identificados en 2014. Estos Withanolidos no solo amplían la comprensión de la diversidad química de la especie, sino que también abren nuevas posibilidades para la exploración de sus propiedades farmacológicas. Además, la presencia de withaphysalin N y withaminimin añade a la singularidad estructural y biológica de la especie.

La diversidad de Withanolidos en Alkekengi officinaron, incluyendo physalin B, withaphysalin U, y otros, refleja las observaciones generales sobre la variedad y el potencial farmacológico en la subtribu Physalineae. La complejidad de estos compuestos subraya la importancia de esta especie en la investigación de productosnaturales y su potencial en la medicina.

Continuando con esta línea de investigación, otro caso notable dentro de la subtribuPhysalineae es Physalis peruviana. Esta especie se ha distinguido por su variado espectro de Withanolidos, reflejando la riqueza fitoquímica inherente a esta subtribu.Investigaciones realizadas por Lan et al. en 2009 y Sang-ngern et al. en 2016, entreotros, han revelado la presencia de compuestos como el Withanolido E, 3-O- glycoside of coagulin R, phyperunolide A, y perulactone C.

La identificación de Physalolacton y Withaphysanolide (Lan et al, 2009), así como withaperuvin C y Withanolido S, amplía aún más la comprensión de la complejidad química de esta especie. La presencia de phyperunolide B, perulactone D (Fang et al, 2009), y withaperuvin N (Sang-ngern et al, 2016), junto con Withanolido C, phyperunolide C, y 4b-hydroxyWithanolido E, destaca la riqueza de Withanolidos en Physalis peruviana.

Además, el descubrimiento de phyperunolide D, phyperunolide E, phyperunolide F, Physalactone, physaperuvins I y J, 14a,17b,20-trihydroxyWithanolidos, 28-hydroxyWithanolido, 3-ethoxy Withanolido, y el deacetylated analogue of physapubenolide ofrece una visión más completa de la diversidad de Withanolidos presentes en esta especie (Misico et al, 2011).

La especie Physalis peruviana se ha identificado como una fuente significativa de Withaferin A, un compuesto conocido por sus propiedades anticancerígenas y antiinflamatorias. Esta especie también destaca en la producción de 1,3,27-trihydroxy-(20S,22R)-3-methoxywitha- 4,6,24-trienolide, reforzando su papel comouna planta de interés farmacológico (Misico et al, 2011).

En el estudio de la relación entre el Withaferin A y la expresión génica, con un enfoque particular en las correlaciones positivas y la asociación con clústeres y el metabolismo secundario, se ha identificado el módulo 'blue' como el más representativo. Este módulo destaca por incluir genes cuya expresión se ve positivamente correlacionada con el Withaferin A y que están relacionados con el

metabolismo secundario o forman parte de clústeres específicos. A continuación, se describen algunos de los genes más notables dentro de este módulo y sus posibles roles biológicos.

El gen N0.HOG0008905, que codifica para la L-ascorbate peroxidase 3 (LOC107859858), muestra una correlación positiva de 0.596461 con un p-valor de 0.002095. Este gen está asociado con las vías de los alcaloides esteroideos (SGAs).La correlación positiva sugiere que L-ascorbate peroxidase 3 podría estar influenciando procesos relacionados con la síntesis y el metabolismo de los alcaloides esteroideos, que son importantes en la respuesta de defensa de las plantas y en diversas funciones biológicas.

Otro gen relevante en el módulo 'blue' es el N0.HOG0022662, que codifica para la enzima pyrophosphate--fructose 6-phosphate 1-phosphotransferase, relacionada con el metabolismo secundario (SM). Presenta una correlación positiva de 0.564534 con un p-valor de 0.004055. Este gen juega un papel crucial en el metabolismo de los carbohidratos, lo que indica que el Withaferin A podría tener un impacto en rutas metabólicas fundamentales para la energía y la biomasa de la planta.

Además, el N0.HOG0011366, identificado como un factor de transcripción bHLH14-like (LOC101250434), tiene una correlación positiva de 0.558546 con un p-valor de 0.004557. Este gen también está relacionado con las vías de los alcaloides esteroideos (SGAs). La correlación positiva sugiere una posible influencia del Withaferin A en la regulación génica y en procesos de señalización celular asociados con estos alcaloides.

Por otro lado, en el estudio de la relación entre el Withaferin A y la expresión génica, con un enfoque particular en las correlaciones negativas y la asociación con clústeresy el metabolismo secundario, se ha identificado el módulo 'blue' como el más representativo. Este módulo destaca por incluir genes cuya expresión se ve negativamente correlacionada con el Withaferin A y que están relacionados con el metabolismo secundario o forman parte de clústeres específicos. A continuación, sedescriben algunos de los genes más notables dentro de este módulo y sus posibles roles biológicos.

El gen N0.HOG0015933, que codifica para una enzima bifuncional aspartato aminotransferase y glutamato-oxaloacetato transaminasa, muestra una correlación negativa de -0.542077 con un p-valor de 0.006213. Este gen está asociado con la vías de los Alcaloides Tropanos (TAs). La correlación negativa sugiere que el Withaferin A podría estar influenciando procesos relacionados con el metabolismo de estos alcaloides, los cuales son conocidos por sus roles en la defensa de la planta.Otro gen relevante en el módulo 'blue' es el N0.HOG0007010, que codifica para la proteína COX15, implicada en el ensamblaje del citocromo c oxidasa, y presenta una correlación negativa de -0.481010 con un p-valor de 0.017336. Este gen está relacionado con el metabolismo secundario (SM). La correlación negativa indica que el Withaferin A podría tener un impacto en la respiración celular y en procesosenergéticos de la planta.

Además, el N0.HOG0014179, identificado como CAAX prenyl protease 2 (LOC101259663), tiene una correlación negativa de -0.475249 con un p-valor de 0.018926. Este gen está involucrado en la vía de los terpenoides. La correlación negativa sugiere una posible influencia del Withaferin A en la síntesis y el metabolismo de terpenoides, componentes importantes en diversas funciones biológicas de las plantas, incluyendo la defensa y la señalización (Pryor et al, 2008).

En el estudio de la relación entre el compuesto 1,3,27-trihydroxy-(20S,22R)-3-methoxywitha-4,6,24-trienolide y la expresión génica, con un enfoque particular en las correlaciones positivas y la asociación con clústeres y el metabolismo secundario, se ha identificado el módulo 'blue' como el más representativo. Este módulo destaca por incluir genes cuya expresión se ve positivamente correlacionada con el compuesto y que están relacionados con el metabolismo secundario o forman parte de clústeres específicos. A continuación, se describen algunos de los genes más notables dentro de este módulo y sus posibles roles biológicos.

El gen N0.HOG0008355, implicado en la biosíntesis de withanolidos, muestra una correlación significativa con un withanolide específico, destacando su papel en su síntesis. Asimismo, el gen N0.HOG0009466, vinculado a las vías de los alcaloides esteroideos, y situado en el clúster SGA_chr12, exhibe una correlación positiva con uno de estos compuestos. La localización del gen N0.HOG0009466 en el clúster SGA_chr12, asociado a las vías de los alcaloides esteroideos (SGAs), y su correlación positiva con compuestos específicos, es un indicativo significativo de que genes involucrados en la biosíntesis de SGAs pueden compartirse o estar relacionados con la ruta de biosíntesis de los withanolidos. Esto sugiere una interconexión entre las vías metabólicas de los alcaloides esteroideos y los withanolidos, lo cual es relevante para entender la complejidad y la regulación de la producción de estos importantes metabolitos secundarios en las plantas (Fan et al 2020).

Además, el N0.HOG0022662, una enzima pyrophosphate--fructose 6-phosphate 1- phosphotransferase relacionada con el metabolismo secundario, también presenta una correlación positiva, sugiriendo un impacto en rutas metabólicas esenciales para la energía y biomasa de la planta. Estos genes, ubicados en el

módulo 'blue', proporcionan una visión integral de cómoel compuesto 1,3,27-trihydroxy-(20S,22R)-3-methoxywitha-4,6,24-trienolide podría estar influenciando positivamente la expresión génica en relación con el metabolismo secundario y clústeres específicos. El aumento en la expresión de estos genes sugiere que el compuesto podría estar jugando un papel en la modulación de diversas vías biológicas, incluyendo la biosíntesis de Withanolidos, alcaloides esteroideos y el metabolismo de los carbohidratos.

La diversidad de Withanolidos en Alkekengi officinaron, incluyendo physalin B, withaphysalin U, y otros, refleja las observaciones generales sobre la variedad y el potencial farmacológico en la subtribu Physalineae. La complejidad de estos compuestos subraya la importancia de esta especie en la investigación de productosnaturales y su potencial en la medicina.

La transición fluida de estas observaciones específicas a las generales demuestra larelevancia de la investigación continua en Alkekengi officinaron dentro del contexto más amplio de la subtribu Physalineae. La integración de estos hallazgos específicoscon el conocimiento general existente destaca la necesidad de métodos analíticos avanzados y una exploración más profunda para comprender completamente el perfil fitoquímico y el potencial terapéutico de los Withanolidos en esta subtribu.

En este escenario, Alkekengi officinaron emerge como un ejemplo emblemático. Esta especie, representativa de la subtribu Physalineae, ha ganado atención por su diversidad de Withanolidos, reflejando el potencial aún no totalmente explorado en la subtribu. Un compuesto particularmente destacado es el physagulin J, reportado por Li et al. en 2008. Este Withanolido, con su estructura química inusual, sugiere un vasto potencial para aplicaciones en el ámbito de la salud.

Profundizando en los Withanolidos específicos de Alkekengi officinaron, se encuentrancompuestos como physalin T, 11,16-cyclophysalin y 5a,6a-dihydroxy physalin D, identificados en 2014. Estos Withanolidos no solo amplían la comprensión de la diversidad química de la especie, sino que también abren nuevas posibilidades para la exploración de sus propiedades farmacológicas. Además, la presencia de withaphysalin N y withaminimin añade a la singularidad estructural y biológica de la especie (Misico et al, 2020).

La diversidad de Withanolidos en Alkekengi officinaron, incluyendo physalin B, withaphysalin U, y otros, refleja las observaciones generales sobre la variedad y el potencial farmacológico en la subtribu Physalineae. La complejidad de estos compuestos subraya la importancia de esta especie en la investigación de productosnaturales y su potencial en la medicina.

Retomando Withaninae, los datos cuantitativos revelan la presencia de C-12 and C-21 Oxygenated Withanolidos en Withaninae, un subgrupo no mencionado en la literatura. Este descubrimiento indica una diversidad potencialmente inexplorada deWithanolidos en Withaninae y sugiere que los procesos de biosíntesis de estos compuestos pueden ser más complejos y variados de lo previamente entendido.

Sin embargo, la ausencia de Withanolidos como los C-18 Hydroxylated Withanolidos, entre otros mencionados en la literatura, en los datos cuantitativos, sugiere posibles limitaciones en las técnicas de detección utilizadas o diferencias en las muestras analizadas. Esta discrepancia entre los datos literarios y experimentalessubraya la importancia de emplear métodos analíticos avanzados y realizar investigaciones más profundas para comprender completamente el perfil de Withanolidos en Withaninae.

En la literatura científica, se ha reportado la presencia de viscosalactone B en Withania somnifera, destacado por su estructura única y su potencial en la actividadantiinflamatoria y antitumoral, como indican los estudios de Misra et al.

(2008) y (Misico et al, 2011). Este hallazgo sugiere un camino prometedor para futuras investigaciones farmacológicas.

Glucosomniferanolide y physagulin D han sido reportados en Withania somnifera por Tong et al. (2011). Estos compuestos presentan interesantes estructuras con grupos epóxidos y glucósidos, lo que indica su potencial en la inhibición de la proliferación celular en ciertos tipos de cáncer. Los withanosides IV y V, también reportados en la misma fuente, ofrecen perspectivas en cuanto a su solubilidad y biodisponibilidad.

Tong et al. (2011) y Misico et al. (2011) han reportado la presencia de 12-deoxywithastramonolide y withaferin A en Withania somnifera. Estos compuestos son notables por sus estructuras y actividades biológicas, siendo withaferin A particularmente reconocido por su amplio espectro de actividades farmacológicas.

Withanone, Withanolido D y withanoside X han sido reportados en Withania somnifera, ampliando su perfil fitoquímico. Estos compuestos, con sus estructuras únicas, sugieren potencial en aplicaciones neuroprotectoras y antitumorales.

Los estudios han reportado withasomniferols A y B, 24,25-dihydrowithanoside VI y 2,3-dihydrowithaferin A en Withania somnifera. Estos compuestos añaden a la diversidad química de la planta y sugieren una variedad de actividades biológicas, desde antioxidantes hasta potenciales usos en neuroprotección.

El 17α-hydroxy derivative 2 ha sido reportado en Withania somnifera y representa un área de interés emergente. Su estructura y actividad biológica propuesta ameritan una investigación detallada para explorar nuevas perspectivas en el tratamiento de enfermedades crónicas y la prevención del envejecimiento celular.

En el marco de nuestros análisis, se observó que la especie Withania somnifera poseeconcentración de Baimantuoluoside B. Este hallazgo significativo el primer reportede Baimantuoluoside B en Withania somnífera.

El Módulo Azul, sobresaliente en mi análisis, se distingue por varias característicasclave que lo convierten en un foco de interés para la investigación en torno a la producción de baimantuoluoside B. La primera característica notable es el elevado número de genes correlacionados significativamente. Con 51 genes, este módulo nosolo es el que cuenta con la mayor cantidad de genes asociados, sino que también demuestra una fuerte conexión colectiva con la producción de baimantuoluoside B.Esta alta concentración de genes correlacionados proporciona un rico campo de estudio para entender cómo interactúan en conjunto para influir en la producción deeste compuesto.

En términos de diversidad metabólica, el Módulo Azul es igualmente intrigante. A pesar de la ausencia de pathways específicos, las anotaciones de metabolismo secundario incluyen términos clave como "Secondary Metabolism" y "Terpenoid". Esto indica que los genes dentro de este módulo podrían participar en una variedad de procesos metabólicos, algunos de los cuales podrían ser fundamentales en la producción de baimantuoluoside B. Esta diversidad sugiere un entorno metabólico complejo y multifacético, crucial para comprender la biosíntesis y regulación de metabolitos secundarios.

Además, la falta de pathways específicos identificados en este módulo se presenta como una oportunidad única para descubrimientos novedosos. Esta situación abre la posibilidad de explorar nuevas vías metabólicas o interacciones genéticas relacionadas con la producción de baimantuoluoside B, lo cual puede llevar a avances significativos en el campo de la biología molecular de las plantas.

Finalmente, la relevancia biológica y las aplicaciones prácticas de este módulo no pueden subestimarse. La significativa cantidad de genes y la diversidad en el metabolismo secundario ofrecen una visión más integral de los procesos biológicosque regulan la producción de baimantuoluoside B. El estudio detallado de este módulo podría descubrir objetivos genéticos clave para la manipulación y mejora dela producción de este compuesto, lo que tendría implicaciones directas en labiotecnología vegetal y la farmacología.

Dentro del Módulo Azul, caracterizado por su prominente correlación con la producción de baimantuoluoside B, se han identificado varios genes de interés. El gen más destacado es el identificado como HOG ID N0.HOG0006396, que exhibe la correlación más alta y una significancia estadística considerable. Este gen se anotacomo "separase (LOC101263947)", y su función sugiere un rol potencial en la regulación del ciclo celular, Siguiendo en relevancia, el gen HOG ID N0.HOG0010495, anotado como "midasin (LOC107853493)", también muestra una correlación significativa. Aunque no se identificaron anotaciones de metabolismo secundario específicas, su función en procesos nucleares podría indicar una influencia en la regulación genética asociada con la síntesis de baimantuoluoside B. Otro gen de interés es el HOG ID N0.HOG0023097, conocido como "auxin-responsive protein SAUR36-like (LOC109236871)", y etiquetado con palabrasclave relacionadas con hormonas y auxina. Este hallazgo sugiere una conexión entre las hormonas vegetales y la producción de baimantuoluoside B, lo que abre un interesante campo de estudio en la regulación hormonal de la síntesis de este metabolito.

En el contexto de mi investigación sobre los factores que inhiben la producción de baimantuoluoside B, el análisis de los módulos con correlación negativa significativa reveló varios candidatos. Entre estos, el Módulo Azul (Blue) resalta como el más prometedor para un estudio detallado, basándome en varios aspectos críticos.

Primero, el Módulo Azul se distingue por albergar el mayor número de genes correlacionados negativamente, con un total de 124 genes. Esta extensa agrupaciónde genes sugiere una fuerte influencia colectiva en la disminución de la producciónde baimantuoluoside B. La diversidad de estos genes proporciona una base amplia para explorar cómo las interacciones genéticas complejas podrían converger para influir en este proceso metabólico.

Además, la asociación del Módulo Azul con una variedad de pathways y funciones metabólicas, incluyendo "Secondary Metabolism (SM)", "Steroids", "Anthocyanins" y "Steroidal Glycoalkaloids (SGAs)", indica su participación en múltiples procesos metabólicos. Esta diversidad es particularmente relevante, ya quesugiere que los genes dentro de este módulo podrían estar ejerciendo un impacto significativo sobre distintas rutas que afectan la producción de baimantuoluoside B,ya sea directa o indirectamente.

La complejidad y la variedad de pathways y funciones metabólicas asociadas con elMódulo Azul también presentan una oportunidad única para descubrimientos novedosos. Este módulo podría albergar mecanismos de regulación y control metabólico previamente desconocidos, cuya comprensión podría ser fundamental para desentrañar las dinámicas moleculares que limitan la producción de baimantuoluoside B.

Por último, la investigación del Módulo Azul es crucial desde una perspectiva práctica. Comprender los factores que inhiben la producción de baimantuoluoside Bes esencial para desarrollar estrategias que puedan aumentar su síntesis. El estudio de este módulo podría revelar objetivos específicos para intervenciones genéticas ometabólicas, lo que tendría implicaciones directas en el mejoramiento de la producción y el rendimiento de baimantuoluoside B en aplicaciones biotecnológicasy farmacológicas.

Por lo tanto, basándome en estos argumentos, el Módulo Azul se perfila como un área de investigación crucial para entender y potencialmente manipular los factoresque restringen la producción de baimantuoluoside B. Su estudio no solo promete aportar conocimientos valiosos sobre la biología molecular de las plantas, sino también ofrece potenciales aplicaciones prácticas para mejorar la producción de esteimportante compuesto.

Dentro del Módulo Azul, he identificado un grupo de genes que muestran una correlación negativa significativa con la producción de baimantuoluoside B. Estos genes representan puntos clave para investigar cómo ciertos procesos biológicos podrían estar inhibiendo la producción de este compuesto. Entre ellos, he seleccionado cinco genes que destacan por su fuerte correlación negativa y su relevancia potencial, basándome en criterios científicos sólidos para su inclusión enmi análisis detallado.

El análisis de genes relevantes para la producción de baimantuoluoside B revela varios roles potenciales. El gen HOG ID N0.HOG0004330, una enzima 1,4-alfa-glucano-ramificadora, sugiere una influencia en la disponibilidad de precursores o en la regulación de vías metabólicas esenciales. Por otro lado, el HOG ID N0.HOG0013415, un probable factor de transcripción KAN2, destaca la importancia de la regulación genética en la síntesis de metabolitos secundarios. Finalmente, el HOG ID N0.HOG0003079, que codifica para la proteína ribosomal 40S S14-2, podría estar involucrado en la traducción de proteínas clave para la producción de baimantuoluoside B. Estos genes representan puntos de interés para futuras investigaciones en la modulación de la producción de metabolitos secundarios.

Actualmente, no existe información en la literatura científica que vincule específicamente algún Withanolido con *Cuatresia riparia*. Aunque hay un vacío en los datos sobre esta planta, sí se ha realizado un estudio sobre Cuatresia

colombiana,donde se identificaron varios Withanolidos. Este estudio, publicado en la revista "Steroids", informó sobre el aislamiento de dos nuevos Withanolidos, conocidos como withalongolide A 4-acetate y otro sin nombre específico detallado, además desiete Withanolidos conocidos. Aunque este estudio aporta datos valiosos sobre los Withanolidos en una especie del mismo género, no ofrece información directa sobre Cuatresia riparia, resaltando la necesidad de investigaciones específicas para esta especie.

En el marco de nuestro estudio metabolómico enfocado en Cuatresia riparia, hemosllevado a cabo un análisis exhaustivo utilizando la plataforma Xcalibur. Este análisis ha permitido la confirmación de que los perfiles cromatográficos obtenidos de Cuatresia riparia exhiben una correspondencia significativa con los patrones de 12-deoxywithastramonolide. A través de una rigurosa cuantificación, hemos determinado que Cuatresia riparia presenta la mayor concentración de este metabolito en comparación con las demás especies estudiadas. Este hallazgo aporta una evidencia contundente sobre la presencia y predominancia de 12-deoxywithastramonolide en Cuatresia riparia, lo cual representa un avance significativo en la comprensión de su perfil fitoquímico y potencial farmacológico.

Dentro del contexto de mi investigación sobre la producción de 12-deoxywithastramonolide, el análisis de módulos con correlación positiva significativa me llevó a identificar varios módulos de interés. De estos, he decidido centrar mi atención en el Módulo Marrón (Brown), basándome en criterios científicos sólidos y su relevancia potencial en la biosíntesis de este compuesto.

El Módulo Marrón se destaca por varias razones clave. En primer lugar, este móduloalberga la mayor cantidad de genes correlacionados positivamente con la producciónde 12-deoxywithastramonolide, sumando un total de 196 genes. Esta concentración significativa de genes sugiere una fuerte y compleja red de interacciones genéticas que podrían estar jugando un papel crucial en la regulación

y promoción de la síntesis de este metabolito.

Además, la diversidad de los pathways asociados a este módulo es particularmente notable. Incluye una amplia gama de rutas metabólicas, como "Secondary Metabolism (SM)", "Steroidal Glycoalkaloids (SGAs)", "Withanolido", "Anthocyanins", "Triterpenoid Saponins (TAs)" y "Flavonoids". Esta variedad sugiere que los genes en el Módulo Marrón están involucrados en múltiples aspectos del metabolismo secundario, lo que podría ser fundamental para comprender la producción de 12-deoxywithastramonolide. La presencia de múltiples rutas metabólicas refleja la complejidad y la riqueza de los procesos biológicos que podrían estar contribuyendo a la síntesis de este compuesto.

Las anotaciones de metabolismo secundario, como "Secondary Metabolism" y "Oxygenase oxydoreductase", también resaltan la relevancia del Módulo Marrón. Estas anotaciones indican que los genes dentro de este módulo pueden estar directamente involucrados en la regulación y ejecución de procesos bioquímicos que son críticos para la producción de metabolitos secundarios como el 12-deoxywithastramonolide.

En mi investigación sobre la producción de 12-deoxywithastramonolide, y tras un análisis exhaustivo de los genes en el Módulo Marrón, he identificado cinco genes de particular interés que están asociados con los pathways de "Withanolido" y "Withanolidos". La selección de estos genes se ha basado en su correlación significativa con la producción de 12-deoxywithastramonolide y su relevancia en losprocesos metabólicos pertinentes.

El primer gen en mi lista es el HOG ID N0.HOG0017228, que codifica para la "squalene synthase (SQS2)" y está asociado con el pathway de "Withanolido". Estegen es especialmente relevante debido a su función en la síntesis de escualeno, un precursor clave en la biosíntesis de esteroides y terpenoides, incluidos los Withanolidos. La correlación de bicorrelación de 0.500579 y un valor p de

0.012725hacen de este gen un candidato de estudio prominente para entender la producción de 12-deoxywithastramonolide.

Otro gen significativo es el HOG ID N0.HOG0006946, identificado como "probable 1-deoxy-D-xylulose-5-phosphate synthase" y relacionado con el pathway de "Withanolidos". Con una correlación de bicorrelación de 0.412531 y un valor p de 0.045142, este gen sugiere un papel en la vía del mevalonato independiente, crucialpara la formación de terpenoides, y por ende, en la biosíntesis de 12-deoxywithastramonolide.

El tercero es HOG ID N0.HOG0014747, que codifica para "amidophosphoribosyltransferase, chloroplastic". Este gen, asociado con el pathway de "SM" (Secondary Metabolism), muestra una correlación de bicorrelación de 0.752591 y un valor p de 0.000022. Su función en el metabolismo secundario lo convierte en un candidato intrigante para estudiar cómo ciertos procesos bioquímicos pueden estar promoviendo la síntesis de 12-deoxywithastramonolide.

Otro gen significativo es el HOG ID N0.HOG0011680, identificado como "oxalate--CoA ligase". Está relacionado con el pathway de "Phenylpropanoids" y tiene una correlación de bicorrelación de 0.748294 con un valor p de 0.000026. Dada su implicación en el metabolismo de los fenilpropanoides, un grupo de compuestos orgánicos conocidos por su importancia en el metabolismo vegetal, este gen es de gran relevancia para entender la red metabólica que subyace a la producción de 12-deoxywithastramonolide.

El quinto es el HOG ID N0.HOG0004644, aunque no tiene una función anotada en los datos, su correlación de bicorrelación es de 0.641420 con un valor p de 0.000730.Está asociado con el pathway de "Anthocyanins". A pesar de la falta de información específica sobre su función, su fuerte correlación y su asociación con los antocianinos, conocidos por su papel en el metabolismo secundario de las

plantas, sugieren que este gen podría desempeñar un rol importante en procesos metabólicosrelevantes para la producción de 12-deoxywithastramonolide.

El primer gen que capturó mi atención, al realizar el análisis de las correlaciones negativas fue el HOG ID N0.HOG0019914, que codifica para "cytochrome P450 736A (CYP736A)" y se asocia con el metabolismo secundario. Este gen mostró unacorrelación negativa significativa, lo que indica que podría estar involucrado en la modulación del metabolismo que afecta la producción de 12-deoxywithastramonolide. La función de los citocromos P450 en la oxidación de compuestos orgánicos es bien conocida, y su papel en este contexto podría ser fundamental para entender cómo ciertos procesos metabólicos inhiben la síntesis deeste metabolito secundario.

Otro gen de interés fue el HOG ID N0.HOG0010308, un gen no caracterizado asociado con el pathway de "Withanolidos". A pesar de la falta de información detallada sobre su función, la correlación negativa significativa de este gen con la producción de 12-deoxywithastramonolide sugiere que podría estar desempeñando un papel importante en la inhibición de su biosíntesis. Este hallazgo resalta la importancia de investigar genes no caracterizados, ya que podrían ofrecer nuevas perspectivas sobre los mecanismos reguladores en el metabolismo de las plantas.

Además, identifiqué el HOG ID N0.HOG0017054, que codifica para "agmatine deiminase" y está relacionado con el pathway de "Triterpenoid Saponins". Este gentambién exhibió una correlación negativa significativa, lo que indica su posible papel en procesos metabólicos que regulan negativamente la producción de 12-deoxywithastramonolide. La relación entre las saponinas triterpenoides y la síntesisde este compuesto es un área intrigante que merece una investigación más profunda.

El HOG ID N0.HOG0007403, identificado como "pyruvate decarboxylase 2" y asociado con el metabolismo secundario, también mostró una correlación negativa significativa. La decarboxilación del piruvato es un paso clave en varias rutas metabólicas, y su regulación podría tener un impacto considerable en la disponibilidad de precursores para la biosíntesis de 12-deoxywithastramonolide. Este hallazgo sugiere que los procesos metabólicos que involucran el piruvato podrían ser cruciales para la regulación de la producción de este compuesto.

Por último, el gen HOG ID N0.HOG0015102, que codifica para "F-box/kelch-repeat protein At1g80440" y está relacionado con el pathway de "Phenylpropanoids", también presentó una correlación negativa significativa. Las proteínas F-box están implicadas en la degradación de proteínas y podrían estar regulando factores clave en la biosíntesis de fenilpropanoides, lo que afecta indirectamente la producción de12-deoxywithastramonolide.

Por otro lado, especie Cuatresia riparia, tambien mostró niveles notables de 27-hydroxywithanona. Esta observación fue confirmada mediante análisis utilizando Xcalibur. Tomando los resultados de WGCNA, se dirigió el análisis hacia el módulode expresión genética denominado "pink", identificando genes con una correlación positiva significativa con la producción de este compuesto. Este enfoque permitió un estudio más concentrado en un conjunto de genes que posiblemente estén co-regulados o implicados en procesos biológicos comunes esenciales para la biosíntesis de 27-hydroxywithanona.

Dentro de este módulo, se identificaron cuatro genes en particular que eran de interés. Entre ellos, el gen N0.HOG0060368, etiquetado como DWF5-2 y asociado con el pathway de Withanolidos, mostró una correlación positiva significativa. La implicación de DWF5-2 en la biosíntesis o regulación de Withanolidos sugiere su participación directa en la producción de 27-hydroxywithanona.

Otro gen relevante fue el N0.HOG0005820, conocido como WRKY1, también anotado en el pathway de Withanolidos. Este gen podría estar involucrado en la regulación de la expresión génica o en la respuesta a señales relacionadas con la biosíntesis de 27-hydroxywithanona, dado el papel conocido de los factores de transcripción WRKY en la respuesta al estrés en las plantas y en la regulación metabólica.

Además, se examinó el gen N0.HOG0033224, etiquetado como P450reductase, vinculado con el pathway de Withanolido. Este gen podría participar en reacciones redox esenciales para la síntesis de 27-hydroxywithanona, considerando la importancia de las enzimas P450 reductasas en la biosíntesis de diversos metabolitossecundarios en las plantas.

La subtribu Lycianthes, al no presentar Withanolidos en ninguno de los análisis realizados, ofrece una perspectiva interesante en el estudio de estos metabolitos secundarios. La ausencia de Withanolidos en Lycianthes, tanto en la literatura científica como en los datos cuantitativos obtenidos, sugiere que esta subtribu no produce estos compuestos o los produce en niveles indetectables por los métodos deanálisis empleados.

Esta observación es particularmente relevante cuando se considera a Lycianthes como un grupo externo en el contexto del estudio de los Withanolidos. Tradicionalmente, los grupos externos se utilizan en la biología evolutiva para proporcionar un contexto comparativo. La ausencia de Withanolidos en Lycianthes podría indicar que la biosíntesis de estos compuestos es una característica que no secomparte con esta subtribu, lo cual es importante para entender las relaciones evolutivas y la distribución de estos metabolitos en las plantas.

La confirmación de que Lycianthes no produce Withanolidos refuerza la idea de quela presencia de estos compuestos en otras subtribus podría ser el resultado de procesos evolutivos específicos, adaptaciones ecológicas o diferencias en la expresión genética. Esto puede implicar que las rutas biosintéticas que conducen a la producción de Withanolidos están ausentes o significativamente alteradas en Lycianthes.

En este sentido el análisis de enriquecimiento de conjuntos de genes (GSEA, por sus siglas en inglés) demostró que estos módulos estaban enriquecidos en funciones metabólicas (metabolismo especializado, metabolismo de glicoxilato y dicarboxilato), compartimientos de organelos (plastidios y mitocondrias), funcionesreguladoras (factores de transcripción WRKY) e interacciones planta-patógeno. Por tanto, investigamos estos dos módulos en busca de OGs con anotaciones de metabolismo secundario. Descubrimos OGs involucrados en múltiples vías, incluyendo la biosíntesis de Withanolidos (21 OGs), alcaloides esteroideos glicosilados (SGAs, por sus siglas en inglés) (27 OGs) y Carotenoides (9 OGs).

De manera específica, las anotaciones identificadas en el grupo GEME para estos 21 OGS revisten una importancia particular para el análisis minucioso de las rutas metabólicas secundarias. Entre estas anotaciones, destacan las siguientes:

La anotación Withanolido$HMGS2 sugiere la presencia de la enzima 3-hidroxi-3- metilglutaril coenzima A sintasa 2 (HMGS2), crucial en la vía de biosíntesis del mevalonato, un precursor esencial para la formación de isoprenoides, moléculas clave en la síntesis de terpenoides y esteroides. La relevancia de esta vía en la síntesisde withanolidos se ve reflejada en la investigación de Chaurasiya et al. (2012), que demuestra la interacción significativa entre las vías MVA y MEP en la modulaciónde la biosíntesis de Withanolidos. Por otro lado, la anotación Withanolido$WRKY1,al indicar un factor de transcripción WRKY, podría tener

un rol en la regulación transcripcional en la síntesis de withanolidos, un aspecto respaldado por Agarwal etal. (2017), quienes destacaron la importancia de los factores de transcripción en la expresión génica específica para la biosíntesis de estos compuestos.

Además, las anotaciones como Withanolido$SMT y Withanolido$HYD1, que se refieren a la S-adenosil metionina transferasa y a una hidroxilasa respectivamente, podrían estar involucradas en pasos críticos de la biosíntesis de withanolidos, lo queincluye la metilación y modificación de sus estructuras, impactando en sus propiedades bioactivas. Estos hallazgos subrayan la complejidad y la interconexiónde las vías metabólicas en la producción de withanolidos y otros metabolitossecundarios, tal como se describe en el trabajo de Agarwal et al. (2017)

Para continuar, las anotaciones Withanolido$CYP71B10 y TAs$CPA también son clave. CYP71B10, un citocromo P450, puede desempeñar un rol en la biosíntesis dewithanolidos, especialmente en reacciones de hidroxilación, un paso crucial en la formación de estos compuestos. Por otro lado, la CPA, una enzima carboxipeptidasa,podría estar implicada en la síntesis de triperpenos ácidos, que son importantes en la interacción planta-ambiente y podrían influir en la generación de nuevos tipos de withanolidos. Estos aspectos resaltan la interacción compleja y clado-específica entre diferentes vías metabólicas en la producción de metabolitos secundarios, subrayando la importancia de la investigación en la biosíntesis y evolución de los withanolidos, como se menciona en estudios como los de Agarwal et al. (2017) y Chaurasiya et al. (2012.

Además, las anotaciones SGAs$GORKY y SGAs$GAME17 en el contexto de las saponinas glicosiladas (SGAs) sugieren genes relacionados con la síntesis de estos compuestos, que poseen propiedades antifúngicas y beneficios para las plantas. Aunque su relación directa con los withanolidos no es clara, su presencia destaca lacomplejidad metabólica y regulación compartida que podría influir en otras vías

metabólicas. En particular, la anotación SGAs$HMG-CoA_reductase, refiriéndose a la 3-hidroxi-3-metilglutaril coenzima A reductasa, subraya la interconexión entre las vías metabólicas y enzimas clave en la producción de diferentes categorías de metabolitos secundarios. Este enfoque integrado en el estudio de la biosíntesis de withanolidos y SGAs, como lo indican los trabajos de Agarwal et al. (2017) y Chaurasiya et al. (2012), es esencial para comprender la evolución y el desarrollo deestos metabolitos complejos en las plantas.

La biosíntesis del esqueleto de los Withanolidos requiere unidades de isopreno, como isopentenylnpyrophosphate (IPP) y dimethyl allyl pyrophosphate (DMAPP),producidas a través de la vía citosólica del ácido mevalónico (MVA) y la vía plastidial del fosfato de metileritritol (MEP) (Newman and Chappell, 1999). Se ha evidenciado que la isoprenogénesis es crucial para modular la biosíntesis de Withanolidos, con una interacción significativa entre las vías MVA y MEP (Chaurasiya et al., 2012). Esto se refleja en que los módulos de coexpresión de Withanolidos están enriquecidos en funciones relacionadas con el cloroplasto y la biosíntesis de carotenoides, incluyendo genes de carotenoides importantes cuya expresión correlaciona positivamente con la variación de Withanolidos, como la isomerasa de carotenoides (CRTISO2) y el fitoeno desaturasa (PDS).

Los módulos de withanolidos enriquecidos en funciones relacionadas con el cloroplasto y la biosíntesis de carotenoides refuerzan la importancia de la vía del MEP en la producción de estos metabolitos. La evolución de withanolidos está estrechamente ligada a la de los SGAs, compartiendo precursores comunes y mostrando compensaciones bioquímicas. La regulación negativamente correlacionada de genes de estas dos vías indica una compleja regulación clado-específica.

La variabilidad en la producción de withanolidos en las especies de Solanaceae destaca la complejidad de la red genética que controla estas vías metabólicas. Investigaciones como las de Schwander et al. (2014) e Itkin et al. (2013) demuestran que los cambios en los perfiles metabólicos están estrechamente vinculados a la expresión génica, evidenciando una regulación genética dinámica y compleja en la biosíntesis de estos compuestos.

El papel crucial de la ruta del MEP en la biosíntesis de withanolidos sugiere una interacción intrincada entre diversas vías metabólicas y la expresión genética. Este entendimiento es vital para comprender cómo se regulan y evolucionan estos metabolitos en Solanaceae. La correlación entre los niveles de withanolidos y la expresión génica, estudiada mediante análisis de redes de coexpresión y correlaciones de Pearson, subraya la existencia de redes genéticas complejas que sustentan la diversidad de withanolidos. Este enfoque resalta la importancia de integrar metabolómica y transcriptómica para descifrar los mecanismos que gobiernan la biosíntesis de estos compuestos.

Los análisis filogenéticos sugieren que la capacidad de producir withanolidos era ancestral en la subfamilia Solaneae y se perdió de manera independiente en ciertos linajes. La diversificación en la expresión de genes relacionados con withanolidos ilustra la evolución de estas vías en Solanaceae, proporcionando información sobrela evolución de defensas y compuestos terapéuticos en plantas.

La variabilidad genómica y la diversificación de los withanolidos se ven reflejadas en la estructura de la región genómica que contiene el clúster biosintético de withanolidos (WBC). Este patrón es un indicador clave de la aparición y desaparición de diferentes withanolidos en las especies. La mayoría de los genes duplicados o presentes exclusivamente en Physalideae codifican enzimas que podrían catalizar reacciones involucradas en la diversificación de withanolidos, incluyendo CYP450s y 2ODDs (Fan et al, 2020).

La variabilidad en la estructura de la región genómica que contiene el clúster biosintético de withanolidos (WBC) es un indicador clave de la aparición y desaparición de diferentes withanolidos en las especies. Los clústeres biosintéticos actúan como mediadores en la comunicación entre diferentes vías defensivas, dondelos "superclústeres" que contienen genes de múltiples vías son centros de control para la regulación y evolución de las defensas vegetales.

La formación y regulación de clústeres biosintéticos es un área de investigación activa, y nuestros resultados sugieren que la evolución de los withanolidos estuvo estrechamente ligada al surgimiento de un clúster biosintético en la región genómica que contiene el gen 24-ISO. La duplicación del gen DWF1 y la neofuncionalizacióndel 24-ISO probablemente fueron pasos clave en la evolución de estos compuestos,como lo sugieren Schwander et al. (2014) y eLife (2020).

La regulación transcripcional y enzimática juega un papel crucial en la diversificación de los withanolidos. La identificación de factores de transcripción como MYB y WRKY, y enzimas como las CYP450, resalta su papel en la regulación y evolución de los withanolidos. Estas enzimas, en particular las CYP450, son fundamentales en la evolución de metabolitos secundarios, ya que realizan funciones de oxidoreductasa diversificadas. Los genes del clúster biosintético de withanolidos, incluyendo varias CYP450s, están presentes y se expresan principalmente en especies productoras de withanolidos, lo que sugiere un papel en la diversificación de estos compuestos (Itkin et al., 2013; Fan et al, 2020).

4. Conclusiones

El estudio analiza los perfiles de withanolidos en las especies de la familia Solanaceae,observando una variación significativa en su presencia y expresión génica. Esto sugiere una red genética compleja subyacente. Se cree que la capacidad de sintetizar withanolidos era probablemente ancestral en la subfamilia Solaneae, pero se perdió en algunos miembros del género Solanum y la tribu Capsiceae. Además, el patrón de distribución de withanolidos y la variabilidad en la expresión de genes relacionados conellos indican un control genético complejo sobre su biosíntesis.

El estudio destaca la importancia de la vía MEP en la biosíntesis de withanolidos, mostrando una conexión con la biosíntesis de carotenoides y funciones del cloroplasto.Esta relación sugiere una interacción entre las vías MEP y MVA, esencial para la modulación de la biosíntesis de withanolidos. También se observa una estrecha relaciónevolutiva entre los withanolidos y los glucósidos esteroidales (SGAs), con una correlación negativa en la expresión de sus genes, reflejando un equilibrio bioquímico entre ambas vías metabólicas.

La investigación identificó genes reguladores y estructurales candidatos, con una prevalencia de genes de defensa en plantas y factores de transcripción como MYB y WRKY, asociados a la biosíntesis de withanolidos. La evolución de un clúster biosintético de withanolidos parece vinculada a la duplicación y neofuncionalización delgen DWF1 y su parálogo 24-ISO, esencial en la biosíntesis. La presencia y expresión de24-ISO respaldan la idea de que estos clústeres son cruciales en la interacción entre diferentes vías defensivas, formando 'superclusters' que actúan como centros de controlen la evolución de las defensas de las plantas.

A través del análisis comparativo y la secuenciación de genomas de Solanaceae, la investigación proporciona un marco para generar hipótesis sobre la evolución del clústerbiosintético de withanolidos (WBC) y sus posibles beneficios evolutivos. La manipulación del WBC mediante bioingeniería y biología sintética, particularmente en el contexto de Solanaceae y su metabolismo especializado, permitirá probar estas hipótesis.

La tesis sugiere que los clústeres biosintéticos no solo facilitan la producción de compuestos defensivos, sino que también median la interacción entre diferentes vías metabólicas defensivas. El descubrimiento de que el WBC contiene ortólogos de genesbiosintéticos de SGA y la agrupación de algunos genes de withanolidos en clústeres deSGA subraya un alto nivel de agrupación genética. Además, la duplicación y neofuncionalización de genes enzimáticos y reguladores son claves para la expansión de clústeres biosintéticos y la evolución de la diversidad bioquímica, aunque losmecanismos de formación de estos clústeres aún no se comprenden completamente. Sepropone que las reorganizaciones genómicas han sido fundamentales en la especiación de plantas, formando 'supergenes de especiación'.

5. Recomendaciones

Realizar avances de los estudios en la ingeniería de vías metabólicas y la biología sintética, impulsados por un conocimiento más profundo de las rutas biosintéticas de withanolidos, abre un horizonte prometedor. La posibilidad de manipular genéticamentey diseñar vías metabólicas para optimizar o incluso incrementar la producción de withanolidos es una meta alcanzable. Esto no solo mejoraría el rendimiento de los compuestos bioactivos en las plantas de Solanaceae, sino que también podría conducir a la creación de nuevos compuestos con propiedades farmacológicas mejoradas. Tales avances no solo tienen implicaciones en el ámbito de la fitofarmacología, sino que también podrían tener aplicaciones significativas en la agricultura y la biotecnología.

Además, un enfoque comparativo más amplio que implique análisis genómico y transcriptómico en diversas especies de Solanaceae podría revelar los mecanismos subyacentes de la evolución y regulación genética de la biosíntesis de withanolidos. Dicha investigación podría arrojar luz sobre cómo estas vías metabólicas se handiversificado y adaptado en diferentes contextos ambientales y evolutivos. El entendimiento detallado de estas diferencias y similitudes a nivel genético y bioquímico podría proporcionar claves cruciales para manipular la producción de estos compuestosen diferentes especies.

Finalmente, la realización de estudios comparativos entre especies de Solanaceae y otras familias relacionadas podría proporcionar una visión integral sobre la biosíntesis de withanolidos. Tales estudios no solo ayudarían a comprender las diferencias en las víasbiosintéticas entre diferentes familias de plantas, sino que también podrían identificar características únicas de la familia Solanaceae. Esta investigación podría revelar cómo la evolución ha moldeado la diversidad química

de estas plantas, lo que sería fundamental para desarrollar estrategias de conservación, mejoramiento y utilización sostenible de estas especies.

6. Bibliografia

1. Agarwal, A. V. et al. Comprehensive assessment of the genes involved in withanolide biosynthesis from Withania somnifera: chemotype-specific and elicitor- responsive expression. Funct. Integr. Genomics 17, 477–490 (2017).

2. Alfonso, D., & Kapetanidis, I. (1994). Withanolidos from Iochroma gesnerioides. Phytochemistry, 36(1), 179-183.

3. Alharbi, A., & Stevenson, M. (2020). Refining Boolean queries to identify relevantstudies for systematic review updates. Journal of the American Medical InformaticsAssociation, 27(11), 1658-1666.

4. Altenhoff, A. M., Levy, J., Zarowiecki, M., Tomiczek, B., Vesztrocy, A. W., Dalquen, D. A., ... & Dessimoz, C. (2019). OMA standalone: orthology inference among public and custom genomes and transcriptomes. Genome research, 29(7), 1152-1163.

5. Bharadwaj, R., Kumar, S. R., Sharma, A. & Sathishkumar, R. Plant metabolic geneclusters: evolution, organization, and their applications in synthetic biology. Front. Plant Sci. 12, 697318 (2021).

6. BLAST: Basic Local Alignment Search Tool. (n.d.). https://blast.ncbi.nlm.nih.gov/

7. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer forIllumina sequence data. Bioinformatics 30, 2114–2120 (2014).

8. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. (n.d.). http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

9. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using

DIAMOND. Nat. Methods 12, 59–60 (2015).

10. Chaurasiya, N. D., Sangwan, N. S., Sabir, F., Misra, L., & Sangwan, R. S. (2012). Withanolide biosynthesis recruits both mevalonate and DOXP pathways of isoprenogenesis in Ashwagandha Withania somnifera L.(Dunal). Plant cell reports, 31, 1889-1897.

11. Conesa, A., Madrigal, P., Tarazona, S. et al. A survey of best practices for RNA-seqdata analysis. Genome Biol 17, 13 (2016). https://doi.org/10.1186/s13059-016-0881-8.

12. Chen, Y., & Guo, D. (2016). Molecular mechanisms of coronavirus RNA capping and methylation. Virologica Sinica, 31, 3-11.

13. Chow, K. S., Ghazali, A. K., Hoh, C. C., & Mohd-Zainuddin, Z. (2014). RNA sequencing read depth requirement for optimal transcriptome coverage in Hevea brasiliensis. BMC research notes, 7, 1-13.

14. Das, S., & Bansal, M. (2019). Variation of gene expression in plants is influenced by gene architecture and structural properties of promoters. PLoS One, 14(3), e0212678.

15. Dhar, N. et al. A decade of molecular understanding of withanolide biosynthesis andin vitro studies in Withania somnifera (L.) Dunal: prospects and perspectives for pathway engineering. Front. Plant Sci. 6, 1031 (2015).

16. Deanna, R., Larter, M. D., Barboza, G. E. & Smith, S. D. Repeated evolution of a morphological novelty: a phylogenetic analysis of the inflated fruiting calyx in the Physalideae tribe (Solanaceae). Am. J. Bot. 106, 270–279 (2019).

17. Dunn, W. B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., ... & Human Serum Metabolome (HUSERMET) Consortium. (2011).Procedures for large-scale metabolic profiling of serum and plasma using

gaschromatography and liquid chromatography coupled to mass spectrometry. Nature protocols, 6(7), 1060-1083.

18. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 20, 1–14 (2019).

19. Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. Nature Reviews Genetics, 17(7), 422-433.

20. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–7 (2004).

21. Erazo, S., Rocco, G., Zaldivar, M., Delporte, C., Backhouse, N., Castro, C., ... & Garcıía, R. (2008). Active metabolites from Dunalia spinosa resinous exudates. Zeitschrift für Naturforschung C, 63(7-8), 492-496.

22. Espinoza, J., Echeverría, J., Urzua, A., & Niemeyer, H. M. (2012). Withanolid amineand nicotine from Dunalia spinosa (Solanaceae). Boletín Latinoamericano y del Caribe de Plantas Medicinales y Aromáticas, 11(3), 278-284.

23. Fan, P. et al. Evolution of a plant gene cluster in Solanaceae and emergence of metabolic diversity. Elife 9, 1–26 (2020).

24. Fang, S. T., Li, B., & Liu, J. K. (2009). Two new withanolidos from Physalis peruviana. Helvetica Chimica Acta, 92(7), 1304-1308.

25. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28, 3150–3152 (2012).

26. Garzón-Martínez, G. A., Zhu, Z. I., Landsman, D., Barrero, L. S., & Mariño-Ramírez, L. (2012). The Physalis peruviana leaf transcriptome: assembly, annotationand gene model prediction. BMC genomics, 13, 1-12.

27. Gao, X., Dong, J., Rasouli, F., Pouya, A. K., Tahir, A. T., & Kang, J. (2022). Transcriptome analysis provides new insights into plants responses under phosphatestarvation in association with chilling stress. BMC plant biology, 22(1), 1-14.

28. Gupta, P. et al. Comparative transcriptome analysis of different chemotypes elucidates withanolide biosynthesis pathway from medicinal plant Withania somnifera. Sci. Rep. 5, 18611 (2015).

29. Gupta, P. et al. De Novo Assembly, Functional Annotation and Comparative Analysis of Withania somnifera Leaf and Root Transcriptomes to Identify Putative Genes Involved in the Withanolidos Biosynthesis. PLoS One 8, (2013).

30. Haas, B. J. et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8, 1494–1512(2013).

31. Hansen, C. C., Nelson, D. R., Møller, B. L. & Werck-Reichhart, D. Plant cytochromeP450 plasticity and evolution. Mol. Plant 14, 1244–1265 (2021).

32. He, Q. P., Ma, L., Luo, J. Y., He, F. Y., Lou, L. G., & Hu, L. H. (2007). Cytotoxic withanolidos from Physalis angulata L. Chemistry & Biodiversity, 4(3), 443-449.

33. Huang, M. et al. Withanolidos from the genus Physalis: a review on their phytochemical and pharmacological aspects. J. Pharm. Pharmacol. 72, 649–669 (2020).

34. Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrativeanalysis of large gene lists using DAVID bioinformatics resources. Nature protocols, 4(1), 44-57.

35. Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.

K., Cook, H., ... & Bork, P. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic acids research, 47(D1), D309-D314

36. Itkin, M. et al. Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. Science (80-.). 341, 175–179 (2013).

37. Knoch, E. et al. Third DWF1 paralog in Solanaceae, sterol Δ 24 -isomerase, branches withanolide biosynthesis from the general phytosterol pathway. Proc. Natl. Acad. Sci. 115, E8096–E8103 (2018).

38. Kriventseva, E. V., Kuznetsov, D., Tegenfeldt, F., Manni, M., Dias, R., Simão, F. A., & Zdobnov, E. M. (2019). OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. Nucleic acids research, 47(D1), D807-D811.

39. Lan, Y. H., Chang, F. R., Pan, M. J., Wu, C. C., Wu, S. J., Chen, S. L., ... & Wu, Y. C. (2009). New cytotoxic withanolidos from Physalis peruviana. Food Chemistry, 116(2), 462-469.

40. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9, 559 (2008).

41. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large setsof protein or nucleotide sequences. Bioinformatics 22, 1658–1659 (2006).

42. LibGuides: Literature Searching: Searching techniques: using Boolean. (n.d.). https://libguides.dundee.ac.uk/literaturesearching/searchtechniquesboolean

43. Lobatto VL, García ME, Nicotra VE, Orozco CI, Casero CN. Antibacterial activityof withanolidos and their structure-activity relationship. Steroids. 2023 Nov;199:109297. doi: 10.1016/j.steroids.2023.109297. Epub 2023 Aug 19. PMID: 37598738.

44. Makarov, A., Denisov, E., & Lange, O. (2009). Performance evaluation of a high-field Orbitrap mass analyzer. Journal of the American Society for Mass Spectrometry, 20(8), 1391-1396.

45. Mehta, V., Chander, H., & Munshi, A. (2021). Mechanisms of anti-tumor activity ofWithania somnifera (Ashwagandha). Nutrition and Cancer, 73(6), 914-926.

46. Misico, R. I. et al. Withanolidos and related Steroids. in Fortschritte der Chemie organischer Naturstoffe. Progress in the chemistry of organic natural products. Progrès dans la chimie des substances organiques naturelles 94, 127–229 (2011).

47. Misra, L. et al. Withanolidos from Withania somnifera roots. Phytochemistry 69, 1000–1004 (2008).

48. Munro, C., Zapata, F., Howison, M., Siebert, S. & Dunn, C. W. Evolution of Gene Expression across Species and Specialized Zooids in Siphonophora. Mol. Biol. Evol.39, 1–16 (2022).

49. Nützmann, H.-W., Scazzocchio, C. & Osbourn, A. Metabolic Gene Clusters in Eukaryotes. Annu. Rev. Genet. 52, annurev-genet-120417-031237 (2018).

50. Pabón-Mora, N., Suárez-Baron, H., Ambrose, B. A., & González, F. (2015). Flowerdevelopment and perianth identity candidate genes in the basal angiosperm Aristolochia fimbriata (Piperales: Aristolochiaceae). Frontiers in Plant Science, 6, 1095.

51. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 14, 417–419 (2017).

52. Pandey, S. S., Singh, S., Pandey, H., Srivastava, M., Ray, T., Soni, S., ... & Kalra, A. (2018). Endophytes of Withania somnifera modulate in planta content and the

site of withanolide biosynthesis. Scientific reports, 8(1), 5450.

53. Peng, Y., Wang, Z., Li, M., Wang, T., & Su, Y. (2024). Characterization and analysisof multi-organ full-length transcriptomes in Sphaeropteris brunoniana and Alsophilalatebrosa highlight secondary metabolism and chloroplast RNA editing pattern of tree ferns. BMC Plant Biology, 24(1), 1-23.

54. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS One 5, e9490 (2010).

55. PubChem. (n.d.). Withanoside v. PubChem. https://pubchem.ncbi.nlm.nih.gov/compound/Withanoside-V

56. Rana, S. et al. NADPH-Cytochrome P450 Reductase: Molecular Cloning and Functional Characterization of Two Paralogs from Withania somnifera (L.) Dunal. PLoS One 8, (2013).

57. Razdan, S. et al. Molecular characterization of DWF1 from Withania somnifera (L.) Dunal: its implications in withanolide biosynthesis. J. Plant Biochem. Biotechnol. 26, 52–63 (2017).

58. Revell, L. J. & Harrison, A. S. PCCA: a program for phylogenetic canonical correlation analysis. Bioinformatics 24, 1018–1020 (2008).

59. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. bioinformatics 26, 139–140 (2010).

60. Samadi, A. K. (2015). Potential anticancer properties and mechanisms of action of withanolidos. The enzymes, 37, 73-94.

61. Sang-Ngern, M., Youn, U. J., Park, E. J., Kondratyuk, T. P., Simmons, C. J., Wall, M. M., ... & Chang, L. C. (2016). Withanolidos derived from Physalis peruviana

(Poha) with potential anti-inflammatory activity. Bioorganic & Medicinal Chemistry Letters, 26(12), 2755-2759.

62. Schirmer, M., Stražar, M., Ávila-Pacheco, J., Rojas-Tapias, D. F., Brown, E.,Temple, E. R., Deik, A., Bullock, K., Jeanfavre, S., Pierce, K. A., Jin, S., Invernizzi,R., Pust, M., Costliow, Z., Mack, D. R., Griffiths, A. M., Walters, T. D., Boyle, B. M., Kugathasan, S., . . . Xavier, R. J. (2024). Linking microbial genes to plasma andstool metabolites uncovers host-microbial interactions underlying ulcerative colitis disease course. Cell Host & Microbe. https://doi.org/10.1016/j.chom.2023.12.013

63. Schwander, T., Libbrecht, R. & Keller, L. Supergenes and complex phenotypes. Curr. Biol. 24, R288–R294 (2014).

64. Sharma, A. et al. Characterization and overexpression of sterol Δ22-desaturase, a key enzyme modulates the biosyntheses of stigmasterol and withanolidos in Withania somnifera (L.) Dunal. Plant Sci. 301, 110642 (2020).

65. Sharma, A., Rather, G. A., Misra, P., Dhar, M. K. & Lattoo, S. K. Jasmonate responsive transcription factor WsMYC2 regulates the biosynthesis of triterpenoid withanolidos and phytosterol via key pathway genes in Withania somnifera (L.) Dunal. Plant Mol. Biol. 100, 543–560 (2019).

66. Shilpashree, H. B., Sudharshan, S. J., Shasany, A. K. & Nagegowda, D. A. Molecular characterization of three CYP450 genes reveals their role in withanolidosformation and defense in Withania somnifera, the Indian Ginseng. Sci. Rep. 12, 1– 12 (2022).

67. Singh, A., Raza, A., Amin, S., Damodaran, C. & Sharma, A. K. Recent advances inthe chemistry and therapeutic evaluation of naturally occurring and synthetic withanolidos. Molecules 27, 886 (2022).

68. Smit, S. J. & Lichman, B. R. Plant biosynthetic gene clusters in the context of metabolic evolution. Nat. Prod. Rep. 39, 1465–1482 (2022).

69. Stein, A. Withanolidos: Elucidating steroidal lactone biosynthesis in Nightshades. (2022).

70. Sukanya, D. H., Lokesha, A. N., Datta, G., & Himabindu, K. (2010). Phytochemical diversity in ashwagandha (Withania somnifera). Open Access Journal of Medicinaland Aromatic Plants, 1(2).

71. The cytochrome P450 enzyme WsCYP71B35 from Withania somnifera has a role in withanolidos biosynthesis and defense against bacteria. (2023).

72. Tong, X., Zhang, H., & Timmermann, B. N. (2011). Chlorinated withanolidos fromWithania somnifera. Phytochemistry Letters, 4(4), 411–414. https://doi.org/10.1016/j.phytol.2011.04.016

73. Villas-Bas, S. G., Roessner, U., Hansen, M. A. E., Smedsgaard, J., & Nielsen, J. (2007). Metabolome Analysis. doi:10.1002/0470105518

74. Xia, G., Cao, S., Chen, L. & Qiu, F. Natural withanolidos, an update. Nat. Prod. Rep.39, 784–813 (2022).

75. Xu, Z., Chang, L. (2017). Solanaceae. In: Identification and Control of Common Weeds: Volume 3. Springer, Singapore. https://doi.org/10.1007/978-981-10-5403- 7_9

76. Yang, Y., Xiang, K., Sun, D., Zheng, M., Song, Z., Li, M., ... & Chen, L. (2021). Withanolidos from dietary tomatillo suppress HT1080 cancer cell growth by targeting mutant IDH1. Bioorganic & Medicinal Chemistry, 36, 116095.

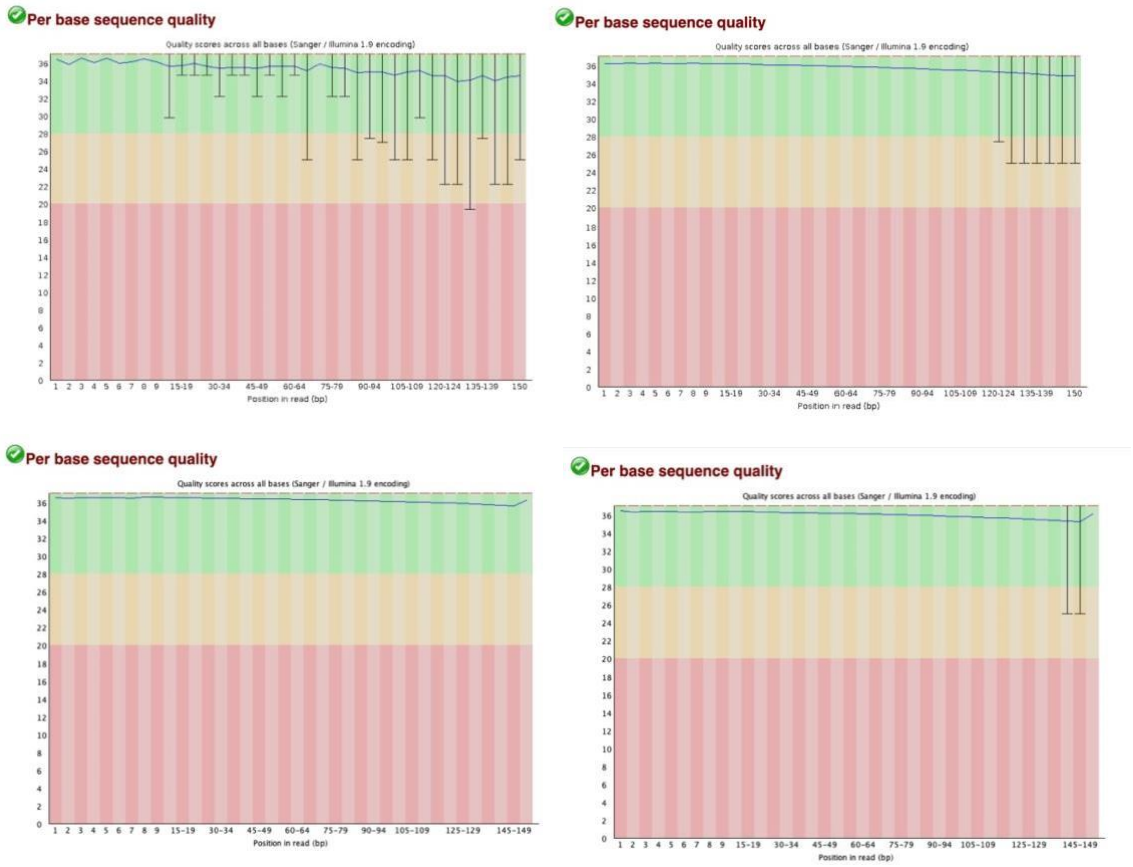
77. Zhang, C. & Mirarab, S. ASTRAL-Pro 2: ultrafast species tree reconstruction

frommulti-copy gene family trees. Bioinformatics 38, 4949–4950 (2022).

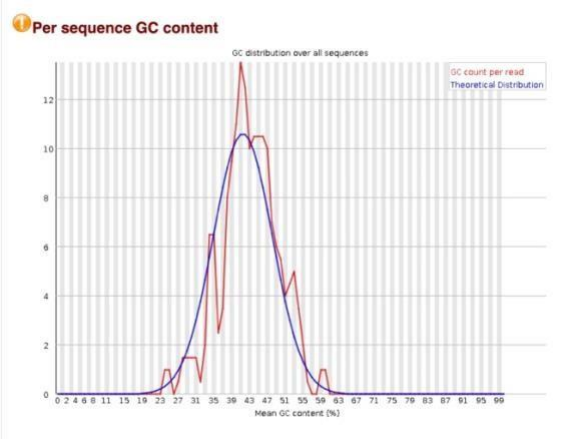
78. Zhan, C. et al. Plant metabolic gene clusters in the multi-omics era. Trends Plant Sci.(2022).

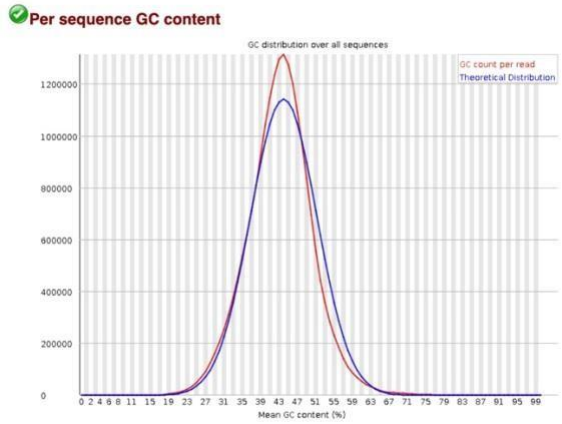
79. Zhao, Y., Li, M. C., Konaté, M. M., Chen, L., Das, B., Karlovich, C., ... & McShane, L. M. (2021). TPM, FPKM, or normalized counts? A comparative study of quantification measures for the analysis of RNA-seq data from the NCI patient-derived models repository. Journal of translational medicine, 19(1), 1-15.

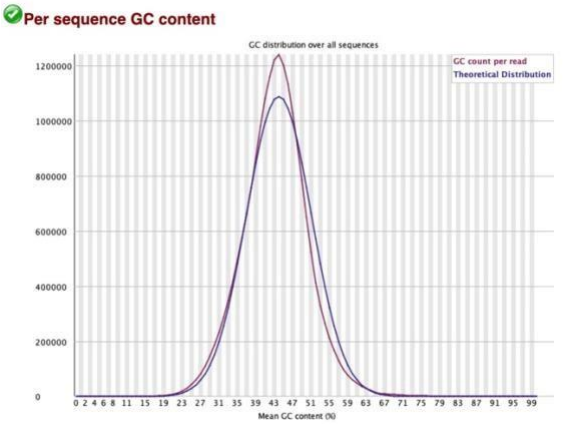
7. Anexos

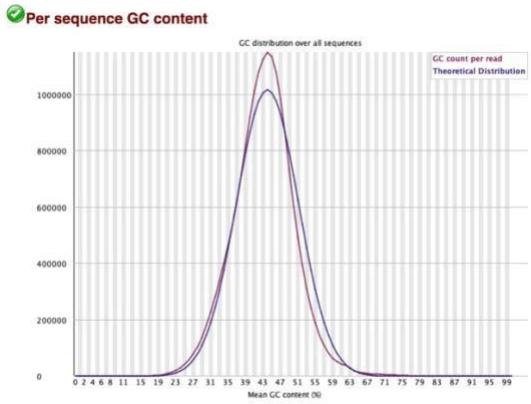


Anexo 1. Calidad de la secuencia antes y después de Trimmomatic. Las imágenes superiores muestran la calidad de las secuencias antes del procesamiento con Trimmomatic (izquierda: lecturas forward; derecha: lecturas reverse). Las imágenes inferiores muestran la calidad de las secuencias después del procesamiento con Trimmomatic (mismo orden: izquierda forward, derecha reverse). Cada gráfico muestra la calidad por base a lo largo de la secuencia, donde la línea verde indica alta calidad, la línea amarilla indica calidad media, y la línea roja indica baja calidad.


Per sequence GC content
GC distribution over all sequences
GC count per read
Theoretical Distribution
Mean GC content (%)


Per sequence GC content
GC distribution over all sequences
GC count per read
Theoretical Distribution
Mean GC content (%)


Per sequence GC content
GC distribution over all sequences
GC count per read
Theoretical Distribution
Mean GC content (%)



Anexo 2. Contenido de GC antes y después de Trimmomatic. Las imágenes superiores muestran la distribución del contenido de GC en las secuencias antes del procesamiento con Trimmomatic (izquierda: lecturas forward; derecha: lecturas reverse). Las imágenes inferiores muestran la distribución del contenido de GC después del procesamiento con Trimmomatic (mismo orden: izquierda forward, derecha reverse). Cada gráfico muestra la distribución del contenido de GC por secuencia, donde la curva azul representa la distribución teórica esperada y la curva roja representa la distribución observada en los datos.

Anexo 3: Análisis de correlaciones entre HOG_ID y producción de withanólidos (Adjunto al archivo).

Este anexo presenta una tabla detallada que resume las correlaciones entre diferentes genes (HOG_ID) y la producción de withanólidos en las especies de la tribu Physalideae y el Outgroup *Lycianthes*. La tabla incluye múltiples columnas que se describen a continuación:

- **HOG_ID**: Identificador de los genes homólogos analizados.
- **Pathway**: Ruta metabólica en la que participa el gen.
- **Short_name**: Nombre corto del gen o grupo de genes.
- **Cluster**: Agrupación de genes basada en sus características funcionales o expresión.
- **Keywords_DAVID_Annotation**: Palabras clave obtenidas a partir del análisis DAVID para la anotación funcional de los genes.
- **Function**: Función biológica del gen.
- **Entrez_Gene_ID**: Identificador del gen en la base de datos Entrez Gene.
- **Annotation_DAVID**: Anotación funcional detallada del gen, según el análisis DAVID.
- **moduleColor_A_TPM10K** y **moduleColor_Logcpm**: Colores de los módulos obtenidos del análisis de coexpresión para diferentes métodos de normalización de datos (TPM10K y Logcpm).
- **Traits_Signif**: Significancia de las correlaciones con los rasgos fenotípicos estudiados.
- **Correlations set**: Conjunto de correlaciones consideradas en el análisis.
- **Count_Neg_A**, **Count_Neg_Logcpm**, **Count_Pos_A**, **Count_Pos_Logcpm**: Conteo de correlaciones negativas y positivas para las normalizaciones TPM10K y Logcpm.
- **GS_PCP_05301_bicor_A**, **GS_PCP_05352_bicor_A**, etc.: Correlaciones bicor (biweight midcorrelation) específicas para diferentes genes y condiciones experimentales.
- **GS_PCP_05301_cor_Logcpm**, **GS_PCP_05352_cor_Logcpm**, etc.: Correlaciones estándar para diferentes genes y condiciones experimentales utilizando Logcpm.
- **p.GS_PCP_05301_bicor_A**, **p.GS_PCP_05352_bicor_A**, etc.: Valores de p asociados con las correlaciones bicor.
- **Sig_PCP_05301_bicor_A**, **Sig_PCP_05352_bicor_A**, etc.: Significancia de las correlaciones bicor.

Esta tabla proporciona una visión exhaustiva de cómo diferentes genes están correlacionados con la producción de withanólidos, permitiendo identificar genes clave y rutas metabólicas importantes para la biosíntesis de estos compuestos.