



Bioinformatic analysis of the phenylpropanoid biosynthesis pathway in *Gomphrena globosa* for the study of gene expression levels

Análisis bioinformático de la ruta de biosíntesis de fenilpropanoides en *Gomphrena globosa* para el estudio de los niveles de expresión génica

Estefania del Carmen Urdiana-Arteaga¹, Edgar García-López*^{1,2}

¹Instituto de Biotecnología, División de estudios de Posgrado, Universidad del Papaloapan campus Tuxtepec, Oaxaca, CP. 68301, México.

²Cátedras CONACYT-UNPA, Universidad del Papaloapan campus Tuxtepec, Oaxaca, CP. 68301, México.

*Corresponding author.

E-mail address: egarcia@unpa.edu.mx (E. García-López).

Article history:

Received: 15 December 2017 / Received in revised form: 21 April 2018 / Accepted: 4 May 2018 / Published online: 1 July 2018.

<https://doi.org/10.29267/mxjb.2018.3.3.68>

ABSTRACT

Gomphrena globosa is a plant that produces different secondary metabolites that have an interest due to the biological activity they present, for example: anti-inflammatory, cytotoxic, antioxidant and antimicrobial activity. The main metabolites that have been found in *in vitro* cultures are simple phenols and flavonoids. The biosynthetic pathway of these compounds is well known, which is as a tool for genetic manipulation when there is enough information about their structural genes. However, there is not enough information about its regulation, which opens the field to investigate the correlation between the corresponding gene expression levels and the regulation of the pathway in *G. globosa*, and the applications in metabolic engineering. Therefore, a bioinformatic analysis was carried out to identify possible candidate enzymes that exert regulation or that could be branch points in the pathway. Nucleotide sequences were obtained from highly related enzymes, as there are no reported sequences for *G. globosa* enzymes at the moment; however the available information for these related molecules allowed selecting a group of six enzymes as regulation nodes in the biosynthetic pathway of phenylpropanoids in *G. globosa*. The oligonucleotides designed by sequence homology gathered the design considerations to be used in further gene expression analysis and they are an important contribution to the genetic and metabolic regulation research in *G. globosa*.

Keywords: *in silico*, metabolic engineering, phenylpropanoids.

RESUMEN

Gomphrena globosa es una planta que produce diferentes metabolitos secundarios que son de interés debido a la actividad biológica que presentan, por ejemplo: actividad antiinflamatoria, citotóxica, antioxidante y antimicrobiana. Los principales metabolitos que se han encontrado en cultivos *in vitro* son fenoles simples y flavonoides. La ruta biosintética de estos compuestos es bien conocida, siendo útil como herramienta para manipulación genética al existir amplia información sobre sus genes estructurales. Sin embargo, no hay suficiente información sobre su regulación, lo que abre campo para investigar la correlación entre los niveles de expresión de los genes correspondientes y la regulación de la ruta en *G. globosa*, con aplicaciones en la ingeniería metabólica. Por lo tanto, se realizó un análisis bioinformático para identificar posibles enzimas candidatas que ejerzan regulación o que fueran puntos de ramificación en la ruta. Las secuencias de nucleótidos se obtuvieron de enzimas altamente relacionadas, debido a que no hay secuencias reportadas de enzimas de *G. globosa* al momento, sin embargo la información disponible para estas moléculas relacionadas permitió seleccionar un grupo de seis enzimas como nodos de regulación en la vía biosintética de fenilpropanoides en *G. globosa*. Los oligonucleótidos diseñados por homología de secuencia reunieron las consideraciones de diseño para ser utilizados en subsecuentes análisis de expresión génica y son un importante aporte a la investigación genética y de regulación metabólica en *G. globosa*.

Palabras clave: fenilpropanoides, Ingeniería metabólica, *in silico*.

1. INTRODUCCIÓN

Gomphrena globosa es una planta nativa de América tropical que ha adquirido interés debido a la actividad biológica de sus metabolitos secundarios; los principales son: betalainas con actividad antioxidante, anticancerígena y antiinflamatoria (Tanaka *et al.*, 2008); saponinas con actividad citotóxica, antiinflamatoria y antimicrobiana (Kim *et al.*, 2011), y los fenilpropanoides con actividad antioxidante y antiinflamatoria (Silva *et al.*, 2012). Particularmente, los fenilpropanoides contribuyen en prácticamente todos los aspectos de las respuestas de las plantas a estímulos bióticos y abióticos; son indicadores de la respuesta al estrés y son mediadores de la resistencia a patógenos (Vogt, 2010). La diversidad de fenilpropanoides se debe a la modificación y amplificación de un núcleo limitado de intermediarios derivados de la vía del shikimato, cuyas transformaciones son catalizadas por la familia de enzimas que participan en la ruta: oxigenasas, ligasas, oxidoreductasas y transferasas (Vogt, 2010). Antecedentes directos de este trabajo han mostrado que los compuestos con mayor acumulación utilizando cultivos *in vitro* de en *G. globosa* son los flavonoides y fenoles simples (Méndez, 2017).

La ruta de biosíntesis de los fenilpropanoides cuenta con extensa información disponible relacionada con los genes estructurales y regulatorios en una amplia variedad de plantas

(Dixon *et al.*, 2002), aunque no para *G. globosa*, por lo que es un excelente modelo para mejorar el entendimiento de la regulación metabólica y, eventualmente, la manipulación genética de la misma con el objetivo de mejorar la acumulación de estos compuestos; sin embargo la información que correlacione los niveles de expresión de los genes de enzimas de la ruta biosintética con el control y regulación metabólica son escasos. Estudios de expresión genética de enzimas involucradas en la síntesis de fenoles y flavonoides realizados han permitido comprender, aunque parcialmente, el metabolismo de estos compuestos y además señalar puntos de regulación en la ruta en sistemas como *Vitis vinifera* (Sparvoli *et al.*, 1994), *Medicago truncatula* (Suzuki *et al.*, 2005), *Solanum tuberosum* (André *et al.*, 2009) y *Mangifera indica L.* (Palafox *et al.*, 2014). Por tal motivo el objetivo del presente trabajo fue analizar *in silico* la ruta de biosíntesis de fenilpropanoides en *G. globosa* para encontrar posibles enzimas regulatorias y diseñar oligonucleótidos para la próxima investigación de sus niveles de expresión que ayude a comprender los posibles mecanismos regulatorios postranscripcionales y, eventualmente, proponer estrategias de manipulación de dichos mecanismos para mejorar o dirigir el flujo metabólico hacia intermediarios de interés.

2. MATERIALES Y MÉTODOS

Se realizó la búsqueda de información sobre las enzimas clave que componen la ruta de fenilpropanoides, y sus ramificaciones hacia los principales derivados fenólicos en la base de datos KEGG (<http://www.genome.jp/kegg/>), así como el nivel de elucidación de la totalidad de los pasos catalíticos y la disponibilidad de secuencias reportadas de *G. globosa*. A través de la aplicación de criterios de selección como información disponible sobre regulación y su posición como nodo se identificaron enzimas candidatas que fueron investigadas más detalladamente en la base de datos BRENDA (<http://www.brenda-enzymes.org/>), lo que proporcionó los números de acceso a la base de datos UNIPROT (<http://www.uniprot.org>) donde se obtuvieron las secuencias aminoacídicas de las enzimas y además de los números de acceso a las secuencias nucleotídicas en GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>) para su posterior análisis. Mediante el software MEGA7 (<http://www.megasoftware.net/home>) se realizaron los alineamientos múltiples para las secuencias aminoacídicas y nucleotídicas de cada enzima seleccionada, buscando las islas de homología en los alineamientos de ADN para el diseño de oligonucleótidos de amplificación, considerando parámetros de diseño: el tamaño del oligonucleótido entre 18-24 nucleótidos de longitud, y un amplicón de entre 150-300 pb para hacerlo compatible con los protocolos de PCR cuantitativa (qPCR). Para el caso particular de la enzima *p*-cumarato 3-hidroxilasa C3H se utilizaron oligonucleótidos reportados por Palafox y colaboradores en 2014 (Oligonucleótido directo 5'-GGGTTGAACTTGGAGCTTC-3' y oligonucleótido reverso 5'-GACGAAATGATGCTTGACACC-3') en el modelo vegetal de *Mangifera indica L.* (mango Ataulfo). Todos los juegos de oligonucleótidos resultantes se analizaron usando la herramienta OligoAnalyzer (<https://idtdna.com/calc/analyzer>) para validar sus características, principalmente; temperatura de alineamiento (T_m por sus siglas en inglés) que fuera cercana a 60 °C con una diferencia menor a los 5 °C entre el oligo directo y reverso, porcentaje de guanina/citosina alrededor de 50 y que tuvieran mínima formación de estructuras secundarias estables. Se seleccionaron los pares de oligos que cumplieran con los parámetros de diseño mencionados.

3. RESULTADOS

La construcción integral de la ruta de biosíntesis de fenilpropanoides fue resultado de la revisión bibliográfica, Dixon y colaboradores en 2002, muestran una ruta de biosíntesis condensada, además de describir los posibles mecanismos de regulación de la ruta y de enlistar las secuencias disponibles de las enzimas que intervienen. Dewick y colaboradores, también, en 2002, puntualizan cada reacción en la ruta de biosíntesis de fenilpropanoides. La base de datos KEGG (Kanehisa Laboratories, 2017) presenta una ruta de biosíntesis completa y actualizada (Fig. 1), esto permitió seleccionar las posibles enzimas reguladoras de esta ruta, debido a que modificaron sus niveles de expresión al exponerse a diferentes estímulos y diferentes especies vegetales (Sparvoli *et al.*, 1994; Suzuki *et al.*, 2005; André *et al.*, 2009; Palafox *et al.*, 2014). De esta manera se seleccionó un grupo de seis enzimas: fenilalanina amonio-liasa (PAL), p-cumarato 3-hidroxilasa (C3H), cinamato 4-hidroxilasa (C4H), 4-cumaroato CoA ligasa (4CL), hidroxicinamoil CoA transferasa HQT y cinamoil CoA reductasa (CCR). Estas enzimas y su efecto en la regulación metabólica en *G. globosa* se evaluarán utilizando los resultados del escrutinio bioinformático de este trabajo. Se buscaron y reunieron secuencias nucleotídicas de enzimas homólogas y/o con altos niveles de identidad respecto a *G. globosa*; dado que no se encontraron secuencias reportadas para esta planta. Segmentos de los alineamientos múltiples resultantes se muestran en la figura 2, resaltando en rojo las regiones que fueron consideradas de homología en las que se puede notar un moderado grado de identidad, aunque en los casos en los que las posiciones nucleotídicas no tuvieran coincidencia se seleccionó el nucleótido con mayor frecuencia de aparición como consenso. Los alineamientos mostrados en los incisos **a)**, **c)** y **d)** de la Figura 2, fueron los que en conjunto mostraron menos identidad general (75-87%) y, por ende, menos regiones de homología. Por otro lado, los alineamientos **b)** y **e)** tuvieron mayores valores de identidad (91-97%) y mayor número de regiones de homología. Las diferencias en identidad de las secuencias alineadas estuvieron asociadas probablemente al tipo de secuencia utilizada, ya que algunas de ellas eran transcritos completos y otras eran secuencias parciales de regiones codificantes. No obstante, se utilizaron las secuencias más completas disponibles. La selección de las secuencias consenso en las regiones de homología de los alineamientos y la consideración de los tamaños de amplicón permitieron obtener los juegos de oligonucleótidos para cada una de las enzimas analizadas (Tabla 1).

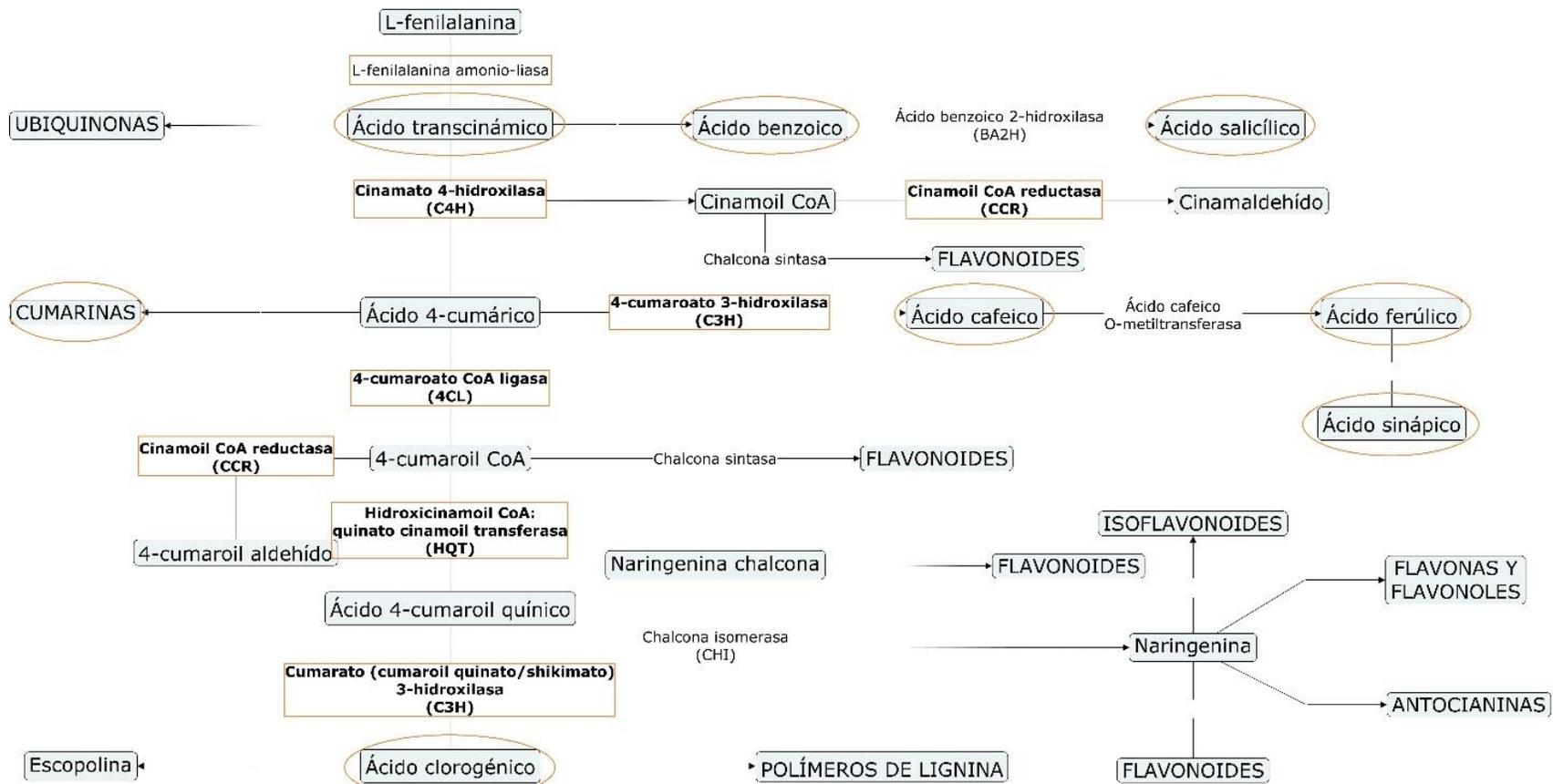


Fig. 1. Ruta de biosíntesis de fenilpropanoides. Las enzimas en negritas indican aquellas consideradas como reguladoras en la ruta de biosíntesis de fenilpropanoides. Los cuadros grises representan los principales intermediarios; los que se encuentran en letras mayúsculas involucran más de una reacción que no se especifica en la figura. Los intermediarios enmarcados en elipses son compuestos fenólicos simples. Compilada de Dewick 2002; Dixon *et al.*, 2002; Kanehisa Laboratories, 2017.

a)

```

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
          1310      1320      1330      1340      1350
AB300199.1 GGTGAGTGAC AATCCTATGA TCGATGTAGC CCGTGACAAG GCATTACACG
AY879309.1 TGTCAATGAC AACCCCTTGA TTGATGTGTC GAGGAACAAG GCCATTTCATG
AY986506.1 CGTCAACGAC AACCCCTTGA TTGATGTGTC AAGGAACAAG GCTTTACATG
EF031063.1 AGTCAATGAC AACCCCTTGA TTGATGTGTC AAGGAACAAG GCCTTGCATG
EF031064.1 AGTGAATGAT AACCCATTGA TTGATGTGTC AAGGAACAAG GCTTTACATG
FJ594466.2 GGTTAATGAT AACCCCTTGA TTGATGTGTC AAGGAACAAA GCATTACATG
FJ603650.2 GGTGAATGAT AATCCATTGA TTGATGTGTC GAGGAATAAA GCTCTACATG
GU456381.1 CGTGAACGAC AACCCATTGA TCGACGTGTC AAGGAACAAA GCCGTGCACG
HM114215.1 GGTCAATGAT AACCCCTTGA TTGATGTGTC AAGGAACAAG GCTATTTCACG
X75967.1_V TGTGAATGAT AACCCCTTGA TCGATGTGTC CAGGAACAAG GCTCTACATG

```

b)

```

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
          560      570      580      590      600
HM062778.1 ATGACGGTCC CTTTCTTCAC CAACAAGTG GTGCAGCAGA ACCGCGCGGG
L38898.1_P ATGACAGTAC CTTTCTTTAC GAACAAGTT GTCCAGCAGT ATCGATTTCGG
DQ211885.1 ATGACGGTCC CTTTCTTCAC CAACAAGTG GTTCCAGCAGT ACCGCGCGGG
AY641731.2 ATGACGGTTC CTTTTTTTAC CAACAAGTG GTGCAGCAGT ACAGGTTTCGG
DQ485132.1 ATGACGGTTC CGTTTTTCAC CAACAAGTT GTGCAACGGA ACAGAGAAGG
DQ485130.1 ATGACGGTTC CGTTCTTCAC CAACAAGTG GTTCCAGCAGA ACAGAGAAGG
HM114216.1 ATGACAGTCC CTTTCTTTAC TAACAAGTT GTCCAGCAGT ATCGATTTGG
GU456382.1 ATGACTGTTC CTTTCTTCAC CAATAAGTA GTCAATCAGT ACAGGTATGG

```

c)

```

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
          810      820      830      840      850
EU616501.1 GGCTGGTGT CCAGCGTTGC CCAGCAGGTC GATGGTGAAA ATCCCAATCT
HM990148.1 AGCTTGATCA CCAGCATCGC GCAGCAAGTC GACGGCGAAA ACCCCAATTT
KC576841.1 AGCTTGATCA CCAGCATCGC GCAGCAAGTC GACGGTGAAA ACCCTAATTT
JX548316.1 AGTCTCGTAA CCAGCGTCGC TCAACATGTT GGTGGAGACA ATCCCAACAT
EF685345.1 AGTCTTGTGA CAAGCGTTGC TCAGCAGGTC GACGGGGAAA ATCCCAACTT
DQ194356.1 AGCTTGATAT CGAGCGTGGC TCAACAAGTC GATGGAGAGA ACCCAAATTT
AY587891.1 AGCTTGATCA CCAGCATCGC GCAGCAGGTC GACGGCGAGA ACCCTAATTT

```

d)

```

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
          360      370      380      390      400
AY149608.1 CTGGATTATC AGAGTTTGAG AGAAGCAATT TATGGCTGTG ACGGGGTTTT
AJ224986.1 CTTGATTATG AGTCTCTTAA AGAGGGTATT CAAGGGTGTG ACGGGTGTGTT
AF217958.1 CTTGATTATG AGTCTCTTAA AGAGGCTATT CAAGGGTGTG ACGGGTGTGTT
GQ450296.1 CTGGACTACG ACGCCATCTG CCGCGCCGTG GAGGGCTGCC ACGGCGTCTT
GQ450301.1 CTCGACCGGG CCAGCCTCCG CGCCGCCTTC GACGGCTGCG ACGGCGTCAT
DQ019126.1 CTAGACTATC AGAGTTTGAG AGAAGCAATT TATGGCTGTG ACGGAGTTTT
DQ019125.1 CTTGATTATC AGAGTTTGCG TGAAGCAATC AATGGCTGTG ACGGAGTTTT
AF278698.1 CTCGACTATG ATGCCATATG CGCCGCCGCC GAGGGCTGCC ACGGCGTGTT
GQ872418.1 CAGGACTACG AGGCTCTTAA AGCGGCGATC GATGGTTGCG ACGGCGTCTT
AY149607.1 CTCGACTACG ACGCCATCTG CGCCGCCGTC GAGGGCTGCC ACGGCGTGTT

```

e)

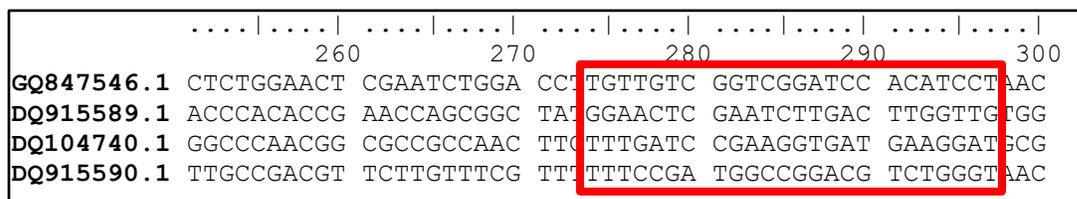


Fig. 2. Multi-alineamiento de secuencias nucleotídicas de las enzimas regulatorias seleccionadas. **a)** PAL, **b)** C4H, **c)** 4CL, **d)** CCR y **e)** HQT respectivamente. El código en negritas indica el número de acceso de GenBank de cada secuencia y en la parte superior la región del alineamiento parcial. Las cajas muestran las regiones de mayor homología utilizadas para el diseño de los iniciadores.

Tabla 1. Oligonucleótidos diseñados y sus características, para la amplificación de genes codificantes de las enzimas seleccionadas.

	Dirección	Secuencia (5'-3')	Longitud (pb)	Tm (°C)	% GC
PAL	D	TACATGGTGGTAACTTCCAAGGGA	24	57.7	45.8
	R	ATCTTGGTTGTGTTGCTTGAC	23	58.2	47.8
C4H	D	ATGACGGTCCCTTTCTTCAC	20	54.7	50.0
	R	GTTGTCCTCGTTGATCTCTC	20	52.6	50.0
4CL	D	CAAGTCGACGGTGAAAACCCCAAT	24	59.8	50.0
	R	CTCTGTCATACCGTAACCCTGT	22	56.0	50.0
HQT	D	TGGAACTCGAATCTTGACCTGG	22	56.7	50.0
	R	CAGAGAAACCCCGCCACAGTTGAA	24	61.5	54.2
CCR	D	GAATTGGTACTGCTACGGGAAG	22	55.4	50.0
	R	GGTAGGGATGGGGTACTC	18	53.7	51.1

D: oligo directo, R: oligo reverso. Ninguno de los oligonucleótidos mostrados en la tabla forma estructuras secundarias estables a las temperaturas de alineamiento convencionales entre 55-65°C.

4. DISCUSIONES

La determinación de la lista de posibles enzimas reguladoras se hizo en base a los resultados de estudios de expresión genética en especies vegetales, tal es el caso de *Vitis vinifera*, donde se encontró que enzimas que biosintetizan fenilpropanoides, PAL (L-fenilalanina amonio-liasa) y estilbeno sintasa se expresan de manera constitutiva, y que las enzimas CHS (chalcona sintasa), CHI (chalcona isomerasa), F3H (flavonona 3-hidroxilasa), DFR (dihidroflavonol 4-reductasa), LDOX (leucoantocianidina dioxigenasa) y UFGT (uridín difosfato glucosa: flavonoide 3-O-glucosil transferasa) se inducen al exponer las semillas a la luz para su germinación (Sparvoli *et al.*, 1994). En *Medicago truncatula* se incrementaron los transcritos que codifican para β -amirina sintasa (involucrada en la síntesis de saponinas), con metil jasmonato, y se incrementaron los transcritos para PAL y CHS exponiendo los cultivos de células a levadura, según lo reportado por Suzuki y

colaboradores en 2005. André y colaboradores en 2009 reportaron que en *Solanum tuberosum* L., al evaluar la expresión de 13 genes involucrados en la biosíntesis de polifenoles (antocianinas) después de exponer a sequía, esto indujo cambios en la expresión de estos genes y originando diferentes perfiles metabólicos de acuerdo con esta variación, por lo que su propuesta es que la regulación en la expresión de estos genes es trascendental en la producción de polifenoles. Al investigar la expresión de los genes de las enzimas PAL y C3H (4-cumaroato 3-hidroxilasa; cumarato [cumaroil quinato/shikimato] 3-hidroxilasa) en diferentes estados de madurez, se observó que los niveles de expresión y la actividad de PAL fueron similares al inicio y al final del proceso de madurez, por lo que su regulación es a nivel postranscripcional; mientras que para C3H se observó una regulación a nivel transcripcional al mostrar cambios en los niveles de expresión y actividad dependiendo de la madurez del fruto (Palafox *et al.*, 2014), esto sugiere que la producción de compuestos fenólicos varía de acuerdo al estado de madurez del fruto, y esto se relaciona con la variación en la expresión de las enzimas dependiendo del estado de madurez en el que se encuentra el fruto. Por lo que se puede concluir que con base al análisis bioinformático realizado se encontró que las enzimas PAL, C3H, C4H y 4CL como puntos relevantes de regulación en la síntesis de compuestos fenólicos simples. Por otro lado, las enzimas HQT y CCR podrían ser también importantes puntos que dirigen el flujo metabólico hacia la biosíntesis de compuestos fenilpropanoides más complejos, debido a que dichas enzimas modificaron sus niveles de expresión al ser evaluadas bajo diferentes estímulos. Adicional a esto se diseñaron oligonucleótidos por homología de secuencias para todas las secuencias de enzimas seleccionadas con características adecuadas para la futura amplificación por qPCR e investigación de la relación entre los niveles de expresión y la regulación de la ruta de biosíntesis de fenilpropanoides en *G. globosa*

AGRADECIMIENTOS

Al proyecto 3212 Cátedra-CONACYT (No. 183958) y al CONACYT por la beca otorgada al primer autor.

CONFLICTO DE INTERESES

Los autores declaran que no existe conflicto de intereses.

REFERENCIAS

André C., Schafleitner R., Legay S., Lefèvre I., Alvarado-Aliaga C., Nomberto G., Evers D. 2009. Gene expression changes related to the production of phenolic compounds in potato tubers grown under drought stress. *Phytochemistry*. 70: 1107-1116.

Dewick P. 2002. The shikimate pathway: aromatic amino acids and phenylpropanoids. En: *Medicinal natural products: a biosynthetic approach*. Chichester: John Wiley & Sons. 121–166.

Dixon R., Achnine L., Kota P., Liu C., Reddy M., & Wang, L. 2002. The phenylpropanoid pathway and plant defence a genomics perspective. *Molecular plant pathology*. 3(5): 371-390.

Palafox C., Contreras-Vergara C., Muhlia-Almazán A., Islas-Osuna M., & González-Aguilar G. 2014. Expression and enzymatic activity of phenylalanine ammonia-lyase and p-coumarate 3-hydroxylase in mango (*Mangifera indica* 'Ataulfo') during ripening. *Genetics and Molecular Research*. 13(2): 3850-3858.

Kanehisa Laboratories. Phenylpropanoid biosynthesis, Japón, http://www.genome.jp/kegg-bin/show_pathway?map00940, (consultado 6 de julio 2017).

Kim Y., Cho J. H., Park S., Han J. Y., Back K., & Choi Y. E. 2011. Gene regulation patterns in triterpene biosynthetic pathway driven by overexpression of squalene synthase and methyl jasmonate elicitation in *Bupleurum falcatum*. *Planta*. 233: 343-355.

Méndez-Sánchez L. I. 2017. Producción de betalaínas y compuestos fenólicos en cultivos celulares de *Gomphrena globosa* L. San Juan Bautista Tuxtepec, México: Tesis de Maestría. Universidad del Papaloapan.

Silva L. R., Valentao P., Faria J., Ferreres F., Sousa C., Gil-Izquierdo A., Pinho B. R., & Andrade P. B. 2012. Phytochemical investigations and biological potential screening with cellular and non-cellular models of globe amaranth (*Gomphrena globosa* L.) inflorescences. *Food Chemistry*. 135: 756-763.

Sparvoli F., Martin C., Scienza A., Gavazzi G., & Tonelli C. 1994. Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). *Plant Molecular Biology*. 24: 743-755.

Suzuki H., Reddy M., Naoumkina M., Aziz N., May G. D., Huhman D. V., & Dixon R. A. 2005. Methyl jasmonate and yeast elicitor induce differential transcriptional and metabolic re-programming in cell suspension cultures of the model legume *Medicago truncatula*. *Planta*. 220: 696-707.

Tanaka Y., Sasaki N., & Ohmiya, A. 2008. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *The Plant Journal*. 54: 733-749.

Vogt T. 2010. Phenylpropanoid biosynthesis. *Mol Plant*. 3(1):2-20.