



Sociedad  
Española  
de **Ciencias**  
**Hortícolas**

**90**

**Septiembre 2022**

# **ACTA DE HORTICULTURA**

**Comunicaciones Técnicas  
Sociedad Española de  
Ciencias Hortícolas**

**X Congreso Nacional de  
Mejora Genética de Plantas**

**Editores:  
Rosa Ana Malvar  
Pedro Fiz Rocha**

**Pontevedra, 19-22 de septiembre 2022**

## 57. Validación de nuevos genes de referencia para estudios de expresión diferencial de genes involucrados en la síntesis de antocianinas en lechuga y especies silvestres relacionadas

I. Medina-Lozano<sup>1,2</sup>, M.S. Arnedo<sup>3</sup>, J. Grimplet<sup>1,2</sup>, A. Díaz<sup>1,2,\*</sup>

<sup>1</sup>Departamento de Ciencia Vegetal, Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Avd. Montañana 930, 50059, Zaragoza, España

<sup>2</sup>Instituto Agroalimentario de Aragón – IA2 (CITA-Universidad de Zaragoza), 50013, Zaragoza, España

<sup>3</sup>Ramiro Arnedo S.A. Paraje La Molina 54, Las Norias de Daza, 04716 Almería, España

**Palabras clave:** antioxidantes, estrés hídrico, *Lactuca sativa*, PCR cuantitativa a tiempo real, RNA-seq

### RESUMEN

La lechuga (*Lactuca sativa* L.) es la hortaliza de hoja más popular a nivel mundial y su consumo sigue aumentando. Es además fuente de compuestos con efectos beneficiosos para la salud, como las antocianinas, potentes antioxidantes presentes en variedades de hoja roja o semirroja. En estudios de expresión génica mediante PCR cuantitativa (qPCR) es fundamental la selección de genes de referencia con expresión estable para la correcta normalización de los datos. En este trabajo se han seleccionado seis genes putativos de referencia para el análisis de la expresión diferencial de genes relacionados con la síntesis de antocianinas en tres estudios distintos: color de la hoja en variedades comerciales (verde vs. roja); tejidos en una especie silvestre relacionada (hoja vs. tallo); y condiciones de estrés hídrico en una variedad comercial, una variedad tradicional y una especie silvestre emparentada (control vs. DH3 y DH4). Con los datos de RNA-seq de dichas muestras se filtraron los genes con expresión estable y alta cobertura de secuencia y se realizó un ANOVA entre los grupos dentro de cada ensayo, seleccionándose seis sin diferencias significativas: *ADF2*, *CYB5*, *iPGAM*, *SCL13*, *TRXL3-3* y *VHA-H*. Su expresión se validó mediante qPCR, ofreciéndose rankings de estabilidad calculada con los algoritmos geNorm, NormFinder y BestKeeper, así como con el método Delta Ct. Englobando todos los resultados, *CYB5* y *TRXL3-3* fueron los genes más estables en los ensayos de expresión génica según el color de la hoja y el tejido (aunque en orden inverso), y *TRXL3-3*, seguido de *ADF2*, en los estudios de estrés hídrico. Los datos de RNA-seq son una nueva fuente para seleccionar genes de referencia, aunque la validación mediante qPCR sigue siendo aconsejable. Este ranking propuesto de nuevos genes de referencia podría resultar útil para futuros estudios de expresión génica en *Lactuca* spp. en las condiciones aquí descritas.

### INTRODUCCIÓN

En las últimas décadas los programas de mejora de cultivos hortícolas se han centrado en el aumento de la producción y en la incorporación de resistencias a enfermedades, a veces en detrimento de caracteres relacionados con la calidad nutricional. En este sentido, resulta interesante abordar el enriquecimiento en compuestos beneficiosos para la salud humana de la lechuga, una de las hortalizas más consumidas a nivel mundial.

Este es el caso de las antocianinas, que son compuestos polifenólicos con potente actividad antioxidante, presentes exclusivamente en variedades de hoja roja y semirroja, y con una distribución desigual dependiendo del tejido. Aunque se desconoce cuál es su papel, se ha comprobado que en algunos cultivos como la vid, su contenido aumenta con el estrés hídrico (Ju et al., 2019). Sin embargo, hasta la fecha, esto no se ha estudiado en lechuga.

Por otro lado, la selección y validación de genes de referencia con expresión estable es esencial para realizar una correcta normalización de los datos en estudios de expresión génica diferencial mediante qPCR a tiempo real. Así, el objetivo de este estudio ha sido la validación de genes de referencia preseleccionados a partir de datos de RNA-seq para el análisis de la expresión diferencial de genes relacionados con la síntesis de antocianinas en lechuga en tres ensayos distintos: *i*) comparación según el color de la hoja en dos variedades comerciales (verde vs. roja); *ii*) comparación de tejidos en una especie silvestre relacionada (hoja vs. tallo); *iii*) condiciones de estrés hídrico en una variedad comercial, una variedad tradicional y una especie silvestre de *Lactuca* (control vs. DH3 y DH4).

## MATERIALES Y MÉTODOS

En la comparación de lechugas de hoja verde y roja se usaron las variedades comerciales ‘Begoña’ y ‘Romired’, respectivamente; en el estudio de distintos tejidos (hoja y tallo) se empleó la especie silvestre de crecimiento arbustivo *Lactuca squarrosa* (Thunb.) Miq.; y en el ensayo de estrés hídrico se utilizaron la variedad comercial roja ‘Romired’, la variedad tradicional semirroja ‘Morada de Belchite’ y la especie silvestre *Lactuca homblei* de Wild. En este último, los tratamientos consistieron en tres regímenes de riego distintos aplicados las 3 semanas previas a la recolección: C (control o riego a demanda), DH3 (semana 1: 450 mL; semanas 2-3: 150 mL) y DH4 (semanas 1-3: 0 mL). En todos los casos se cultivaron tres plantas por accesión y, en el ensayo de estrés hídrico, también por tratamiento (3 réplicas biológicas), siguiendo un diseño de bloques aleatorizados. Tras aproximadamente 3 meses, se recogieron dos hojas por planta (exterior e interior), además de tallo en el caso de *L. squarrosa*.

Posteriormente se realizó la extracción de ARN total de las muestras liofilizadas utilizando el kit NZY Total RNA Isolation (NZYtech Lda.-Genes and Enzymes), seguida de una digestión con ADNasa empleando el kit Turbo DNA-free<sup>TM</sup> (Invitrogen). Se construyeron un total de 39 librerías de ADN copia (ADNC) procedentes de los tres estudios: hoja verde vs. hoja roja (2 accesiones x 3 repeticiones), hoja vs. tallo (2 tejidos x 3 repeticiones) y C vs. DH3 y DH4 (3 accesiones x 3 repeticiones x 3 tratamientos). Estas se secuenciaron mediante el protocolo TruSeq Stranded mRNA de Illumina en el Centro Nacional de Análisis Genómicos (CNAG-CRG, Barcelona), utilizándose la herramienta Galaxy (Afgan et al., 2018) para analizar las secuencias obtenidas.

A partir de los datos de RNA-seq se filtraron los genes con expresión estable ( $|\log_2(\text{fold change})| < 1$  y  $p\text{-valor ajustado} > 0,05$ ) y con una cobertura de secuencia amplia. Por último, se realizó un ANOVA y una comparación de medias entre los grupos dentro de cada ensayo y se seleccionaron seis genes sin diferencias significativas: *ADF2* (actin-depolymerizing factor 2), *CYB5* (cytochrome B5), *iPGAM* (probable 2,3-bisphosphoglycerate-independent phosphoglycerate mutase), *SCL13* (scarecrow-like protein 13), *TRXL3-3* (thioredoxin-like 3-3) y *VHA-H* (V-type proton ATPase subunit H).

Se aisló el ARN mensajero (ARN<sub>m</sub>) del ARN total con el kit Dynabeads<sup>®</sup> mRNA DIRECT<sup>™</sup> (Invitrogen) y se sintetizó el ADNC con el kit NZY M-MuLV First-Strand cDNA Synthesis, separate oligos (NZYTech). Las reacciones de qPCR se llevaron a cabo en un volumen final de 12 µL (2 µL de una dilución 1:40 de ADNC, 0,40 mM de cada primer y 1x de la máster mix NZYSupreme qPCR Green (2x), ROX plus (NZYTech)), en el equipo StepOnePlus<sup>™</sup> (Applied Biosystems). Se realizaron dos réplicas técnicas.

Se obtuvieron rankings de los valores de estabilidad calculados con los algoritmos geNorm (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004), BestKeeper (Pfaffl et al., 2004) y con el método Delta Ct (Silver et al., 2006); todos ellos integrados en el software RefFinder (Xie et al., 2012). Para BestKeeper y el método Delta Ct se utilizaron los valores medios de los C<sub>q</sub> (quantification cycles), mientras que para geNorm y NormFinder se utilizaron valores de C<sub>q</sub> corregidos (C<sub>qE</sub>) con los datos de la eficiencia (E) según la fórmula  $C_{qE} = C_q(\log(E)/\log(2))$  (Garrido et al., 2020). En todos los casos se usaron las medias de las dos réplicas técnicas para cada muestra.

## RESULTADOS Y DISCUSIÓN

La utilización de C<sub>qE</sub> con los algoritmos geNorm y NormFinder resulta conveniente, ya que la eficiencia de la qPCR varía e influye en gran medida en los resultados. Así, los rankings de estabilidad obtenidos son distintos según el análisis (Tabla 1), como está ampliamente descrito en la literatura. Por ello, se han utilizado los resultados de los cuatro métodos para la obtención de un ranking comprehensivo (Tabla 2) que permite determinar cuáles son los genes de referencia más estables en cada ensayo, con este material y bajo estas condiciones concretas. En los ensayos de expresión génica según el color de la hoja y el tejido, los genes más estables fueron *CYB5* y *TRXL3-3* (en orden inverso); y en el de estrés hídrico, *TRXL3-3* también fue el gen más estable, seguido de *ADF2* (Tabla 2).

En lechuga se han descrito genes nuevos de referencia para estudios de respuesta a estreses abióticos (sequía, salinidad, rayos UV-C y metales pesados) mediante qPCR (Borowski et al., 2014), si bien esta selección no está basada en datos de RNA-seq. Este trabajo ofrece un ranking de nuevos genes de referencia según sus valores de estabilidad en el material ensayado, seleccionados a partir de datos de RNA-seq y validados mediante qPCR. Algunos genes resultan adecuados para todos los ensayos (ej.: *TRXL3-3*); mientras que otros son claramente descartables (ej.: *VHA-H*) (Tablas 1 y 2). El uso de RNA-Seq y qPCR resulta una estrategia eficaz y fiable para la selección de genes de referencia distintos a los que se han venido utilizando hasta ahora y que no siempre resultan adecuados.

## AGRADECIMIENTOS

Agradecemos al Banco de Germoplasma de Especies Hortícolas de Zaragoza (BGHZ-CITA, España) y al Centre for Genetic Resources (CNG, Wageningen, Países Bajos) por suministrarnos las semillas. Las fuentes de financiación de este trabajo son: RTA2017-00093-00-00 (Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria), LMP164\_18, LMP148\_21 y Grupo Consolidado A12-17R (Gobierno de Aragón). IML disfruta un contrato predoctoral para la formación de doctores del Ministerio de Ciencia, Innovación y Universidades y la Agencia Estatal de Investigación.

## REFERENCIAS

Afgan, E., Baker, D., Batut, B., Van Den Beek, M., Bouvier, D., Ech, M., Chilton, J., et al. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46: W537–W544.

- Andersen, C.L., Jensen, J.L., and Ørntoft, T.F. 2004. Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 64: 5245–5250.
- Borowski, J.M., Galli, V., da Silva Messias, R., Perin, E.C., Buss, J.H., dos Anjos e Silva, S.D., and Rombaldi, C.V. 2014. Selection of candidate reference genes for real-time PCR studies in lettuce under abiotic stresses. *Planta* 239: 1187–1200.
- Garrido, J., Aguilar, M., and Prieto, P. 2020. Identification and validation of reference genes for RT-qPCR normalization in wheat meiosis. *Sci. Rep.* 10: 2726.
- Ju, Y., Yang, B., He, S., Tu, T., Min, Z., Fang, Y., and Sun, X. 2019. Anthocyanin accumulation and biosynthesis are modulated by regulated deficit irrigation in Cabernet Sauvignon (*Vitis Vinifera* L.) grapes and wines. *Plant Physiol. Biochem.* 135: 469–479.
- Pfaffl, M.W., Tichopad, A., Prgomet, C., and Neuvians, T. 2004. Determination of most stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper Excel-based tool using pair-wise correlations. *Biotechnol. Lett.* 26: 509-515.
- Silver, N., Best, S., Jiang, J., and Thein, S.L. 2006. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol. Biol.* 7: 33.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3.
- Xie, F., Xiao, P., Chen, D., Xu, L., and Zhang, B. 2012. miRDeepFinder: A miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Mol. Biol.* 80: 75–84.

## TABLAS

**Tabla 1.** Valores de estabilidad de los genes de referencia candidatos según los algoritmos geNorm (gN), NormFinder (NF) y BestKeeper (BK) y el método Delta Ct ( $\Delta$ Ct) en tres ensayos distintos: hoja verde vs. hoja roja, hoja vs. tallo y estrés hídrico.

Gen	Verde vs. roja				Hoja vs. tallo				Estrés hídrico			
	gN	NF	BK	$\Delta$ Ct	gN	NF	BK	$\Delta$ Ct	gN	NF	BK	$\Delta$ Ct
<i>ADF2</i>	0,3 3	1,09	0,57	0,8 3	0,77	1,18	0,89	1,1 0	1,5 4	0,77	1,23	4,04
<i>CYB5</i>	0,3 3	0,81	0,53	0,8 3	0,77	1,34	0,77	1,0 3	1,9 8	2,91	1,50	4,07
<i>iPGAM</i>	1,3 3	1,49	0,72	1,2 8	1,22	1,55	0,47	1,0 8	3,4 0	2,81	1,11	4,01
<i>SCL13</i>	1,9 0	2,88	0,53	0,9 2	2,41	4,73	0,92	1,5 4	4,4 1	6,62	0,92	3,89
<i>TRXL3-3</i>	0,5 4	0,71	0,40	1,1 0	1,01	0,30	0,44	0,9 3	1,5 4	0,77	0,85	3,90
<i>VHA-H</i>	1,1 1	0,77	1,12	1,4 1	1,07	0,30	1,23	1,3 8	6,8 2	11,27	10,3 5	14,0 8

**Tabla 2.** Rankings de estabilidad de los genes candidatos para tres ensayos distintos: hoja verde vs. hoja roja, hoja vs. tallo y estrés hídrico.

Carácter	Nº SNPs/QTL	SNP representativo QTL	valor <i>P</i>	Efecto <sup>1</sup>	No <sup>2</sup>	R <sup>2</sup> <sup>3</sup>
MO	1 SNP/QTL-1	S5_42298790	1.25E-05	0.45	176/18	0.10
MO	1 SNP/QTL-2	S2_27557208	2.91E-05	0.43	202/18	0.09
MO	4 SNPs/QTL-3	S2_176009554	4.10E-05	0.45	193/16	0.08
MO	2 SNPs/QTL-4	S3_35748221	4.91E-05	0.47	202/14	0.08
MO	1 SNP/QTL-5	S10_141002338	6.66E-05	0.47	209/13	0.07
MO	1 SNP/QTL-6	S5_10213553	7.02E-05	0.27	143/56	0.08
MO	1 SNP/QTL-7	S2_198391601	7.83E-05	0.28	157/44	0.07
MO	1 SNP/QTL-8	S3_20175107	8.16E-05	0.31	185/33	0.07
DMO	1 SNP/QTL-1	S9_137984657	1.20E-05	1.47	188/21	0.09
PB	1 SNP/QTL-1	S7_152965425	2.24E-05	0.34	121/84	0.09
PB	2 SNPs/QTL-2	S3_21856762	5.02E-05	0.33	164/68	0.07
PB	1 SNP/QTL-3	S2_111581192	9.90E-05	0.46	27/188	0.07
FAD	1 SNP/QTL-1	S1_216932338	1.29E-05	1.04	38/164	0.10
FAD	1 SNP/QTL-2	S1_179829720	1.30E-05	1.52	14/208	0.09
FAD	2 SNPs/QTL-3	S9_26818485	3.88E-05	0.78	123/79	0.09
FAD	1 SNP/QTL-4	S9_110319382	4.74E-05	1.11	27/199	0.08
FAD	1 SNP/QTL-5	S1_297910173	5.49E-05	1.42	14/191	0.08
FAD	1 SNP/QTL-5	S9_55272065	6.72E-05	1.17	20/209	0.07
FAD	1 SNP/QTL-7	S7_165744678	6.99E-05	0.71	92/118	0.08
FAD	1 SNP/QTL-8	S1_85018538	9.29E-05	0.74	154/64	0.07
FND	1 SNP/QTL-1	S1_216932338	1.24E-05	1.17	38/164	0.10
FND	1 SNP/QTL-2	S9_24359357	8.23E-05	1.46	17/197	0.07
CEL	1 SNP/QTL-1	S3_208629545	5.30E-05	0.47	48/171	0.08
CEL	1 SNP/QTL2	S1_294096858	6.57E-05	0.76	15/180	0.09
CEL	1 SNP/QTL-3	S7_989120	6.94E-05	0.41	77/122	0.09
CEL	1 SNP/QTL-4	S9_132814119	8.26E-05	0.41	134/77	0.07
Biomasa	1 SNP/QTL-1	S4_237680603	5.65E-05	0.57	62/148	0.08
Biomasa	1 SNP/QTL-2	S4_2693758	8.18E-05	0.50	81/135	0.08

<sup>1</sup>Efecto aditivo: calculado como la mitad de la diferencia entre la media de los homocigotos para el alelo de mayor valor y la media de los homocigotos para el alelo de menor valor

<sup>2</sup>No: el número antes de la barra representa el número de homocigotos con el mayor valor medio; y el número después de la diagonal el número de homocigotos con el menor valor medio.

<sup>3</sup>Varianza fenotípica explicada por cada SNP.