

Selección de genes de normalización para
RT-PCR cuantitativa en *Petunia hybrida*

Izaskun Mallona González
(directores: Julia Weiss y Marcos Egea-Cortines)

Proyecto fin de máster en técnicas avanzadas
en investigación y desarrollo agrario y alimentario

30 de septiembre de 2008

Índice

1. Introducción	5
1.1. Cuantificación de la expresión génica	6
2. Materiales y métodos	7
2.1. Material vegetal	7
2.2. Extracción de ARN y retrotranscripción	7
2.3. Optimización de la PCR	7
2.4. PCR en tiempo real	8
2.5. Captura de datos	10
2.6. Algoritmo geNORM	10
2.6.1. Valores M de estabilidad de expresión	10
2.6.2. Determinación del número óptimo de genes normalizadores	11
2.7. Cuantificación relativa mediante qBasePlus	11
2.8. Análisis estadístico	12
3. Resultados y discusión	12
3.1. Confirmación de la especificidad de los cebadores	12
3.2. Valores M	12
3.3. Valores CT	13
3.4. Determinación del número óptimo de genes normalizadores . . .	15
3.5. Cuantificación relativa	18
4. Anexo	20

Lista de figuras

1. Diagrama de cajas y bigotes de los valores M del conjunto de las muestras (arriba, izquierda), de los valores del conjunto de piezas florales (arriba, centro) y foliares (arriba, derecha). Diagrama de los valores CT del conjunto de las muestras (abajo, izquierda), de los valores del conjunto de piezas florales (abajo, centro) y foliares (abajo, derecha) Las abreviaturas significan: AC, ACT; CY, CYP; EF, EF1 α ; GA, GAPDH; RA, RAN3, RP, RPS13; SA, SAND; TU, TUB; UB, UBQ. 14
2. Valores de CT para cada uno de los estadios de desarrollo. Las abreviaturas significan: AC, ACT; CY, CYP; EF, EF1 α ; GA, GAPDH; RA, RAN3, RP, RPS13; SA, SAND; TU, TUB; UB, UBQ. 16
3. Análisis mediante el algoritmo geNORM para los distintos tipos de material vegetal de los valores: M, que son inversamente proporcionales a la estabilidad de expresión (izquierda); y PV, que indican la necesidad de añadir nuevos genes de referencia (derecha). 17
4. Análisis de la expresión relativa de los distintos genes mediante el software qBasePlus. En el eje de ordenadas se indica, en unidades arbitrarias, el nivel de expresión; el el de abscisas se listan las flores A, B, C, D (tres réplicas cada uno), los controles sin molde (seis réplicas, una para cada tipo de muestra) y las hojas A y C (tres réplicas cada uno). Las abreviaturas significan: A, ACT; B, CYP; C, EF1 α ; D, GAPDH; E, RAN3, F, RPS13; G, SAND; H, TUB; I, UBQ. 19

Lista de tablas

1.	Características de los genes, cebadores y amplicones.	9
2.	Valores M según el algoritmo geNORM. Estos valores guardan relación inversamente proporcional con la estabilidad de la expresión.	13
3.	Valores PV según el algoritmo geNORM. Este valor, que representa la variación entre dos factores de normalización secuenciales con diferente número de genes normalizadores, indica si $PV > 0,15$ la necesidad de añadir nuevos genes internos (véase la sección 3).	17
4.	Valores M y factores de normalización según el algoritmo geNORM para todas las muestras analizadas.	21
5.	Valores M y factores de normalización según el algoritmo geNORM para todas las flores de tipo A.	21
6.	Valores M y factores de normalización según el algoritmo geNORM para todas las flores de tipo B.	22
7.	Valores M y factores de normalización según el algoritmo geNORM para todas las flores de tipo C.	22
8.	Valores M y factores de normalización según el algoritmo geNORM para todas las flores de tipo D.	22
9.	Valores M y factores de normalización según el algoritmo geNORM para todas las hojas de tipo A.	23
10.	Valores M y factores de normalización según el algoritmo geNORM para todas las hojas de tipo C.	23
11.	Valores M y factores de normalización según el algoritmo geNORM para el módulo vegetativo.	23
12.	Valores M y factores de normalización según el algoritmo geNORM para el módulo reproductivo.	24

Resumen

La cuantificación de la expresión génica de forma precisa mediante RT-PCR cuantitativa requiere utilizar como genes normalizadores aquéllos cuyo nivel de expresión sea constante, al menos bajo unas determinadas condiciones experimentales. Por ello, debe analizarse de forma absoluta la abundancia de los genes normalizadores putativos para cada especie y bajo distintas condiciones, a fin de escoger los más apropiados para cada caso. Existen trabajos realizados en los que se han determinado genes de referencia para *Arabidopsis thaliana*, vid, arroz o patata. Sin embargo, no se han identificado genes con expresión estable para estudios de desarrollo meristemático y floral. En este estudio analizamos la idoneidad de nueve genes de *Petunia hybrida*, como sistema modelo para Solanáceas, bajo distintos estados de desarrollo vegetativo y reproductivo.

Nuestros resultados sugieren que ninguno de los genes analizados se mantiene estable durante los cuatro estadios descritos de desarrollo floral, aunque sí durante el foliar (RPS13, UBQ). Del mismo modo, ninguno de ellos es válido para normalizar órganos foliares y florales simultáneamente. No obstante, describimos el número y las características de los genes de normalización precisos para cada estadio en particular y establecemos un ranking de estabilidad global: RPS13=UBQ>RAN3>ACT>TUB>CYP>>EF1 α >SAND>GAPDH.

1. Introducción

La reacción en cadena de la polimerasa en tiempo real precedida por la transcripción reversa (RT-qPCR) es una herramienta de detección y cuantificación de ARN mensajero (ARNm). Esta cuantificación de ARN, ya sea absoluta o relativa (es decir, en la cual se obtiene un nivel de expresión del gen de interés respecto de otro que, se asume, lo hace de forma constante), puede realizarse también mediante otras técnicas, como los chips de ADN y los Northern blots si bien la RT-qPCR posee como ventajas una gran sensibilidad, buena reproducibilidad y un amplio rango dinámico de cuantificación [Brunner et al., 2004].

La PCR en tiempo real difiere de la PCR clásica en que permite el análisis en continuo de la amplificación del ADN: es decir, es una técnica cinética, y como tal fue puesta a punto en 1993 [Higuchi et al., 1993]. Para ello emplea los productos convencionales de PCR, es decir: un par de oligonucleótidos, los cebadores, específicos; dNTPs; un tampón de reacción adecuado; una ADN polimerasa termoestable; y algún tipo de compuesto marcado con un fluorocromo que permita, en un termociclador con capacidad de excitarlo convenientemente, la emisión de fluorescencia detectable. Este último compuesto fluorescente puede emplearse fundamentalmente de dos formas: bien unido a un oligonucleótido específico que hibride con el producto de PCR a monitorizar, bien un compuesto que se una inespecíficamente y de forma dependiente del número de copias a todo ácido nucleico de doble cadena de la mezcla. En el presente trabajo se empleó

SYBR-Green, perteneciente al segundo grupo [T.B. Morrison & Wittwer, 1998].

La ventaja de la qRT-PCR es la alta sensibilidad en la detección de transcritos poco abundantes, lo cual redundaría en una detección precisa en diversas situaciones: en tejidos con concentraciones de ARN bajas, cuando se dispone de poca muestra o cuando se pretende discriminar pequeños cambios en la expresión génica [Schmittgen et al., 2000].

1.1. Cuantificación de la expresión génica

La cuantificación puede realizarse en términos absolutos o relativos. En el primer caso, la estrategia es relacionar la señal de amplificación obtenida con el contenido en ADN empleando una curva de calibrado; para este enfoque es vital que la PCR de la muestra y de los elementos de la recta de calibrado posean una misma eficiencia de amplificación. En el segundo caso, se expresa el cambio en los niveles de expresión de ARNm; esta cuantificación relativa es más fácil de realizar, puesto que no requiere curva de calibrado, y se sustenta en la comparación entre el nivel de expresión del gen a estudiar *versus* un gen control (también llamado de referencia, interno o normalizador o, en inglés, *housekeeping gene*).

Por tanto, en la cuantificación relativa es irrelevante en qué unidades se expresa la cuantificación, y sus resultados son comparables entre múltiples experimentos de RT-qPCR. De hecho, el propósito de emplear uno o más genes de normalización es corregir la variación no específica, como las diferencias en la cantidad y calidad del ARN empleado, que pueden afectar a las eficiencias de retrotranscripción y de PCR. No obstante, un aspecto crucial es que la estabilidad del gen de referencia sea una realidad [Brunner et al., 2004].

La selección de los genes internos se ha realizado clásicamente en Biología Molecular analizando la estabilidad de la expresión en estudios cualitativos o de baja sensibilidad, como el examen visual de geles de ARN, densitometría de Northern blots o qPCR semicuantitativa. En plena era de la genómica, es posible realizar una aproximación a gran escala empleando los chips de ADN para muchos organismos. [Czechowski et al., 2005] No obstante, se ha descrito que la mayoría de los genes empleados como normalizadores en la cuantificación de la expresión de ARN mensajero varían según las condiciones experimentales [Thellin et al., 1999] [Stürzenbaum & Kille, 2001] [Radonić et al., 2004]. Por ello, es preciso estudiar la estabilidad de los genes normalizadores bajo las condiciones experimentales a emplear a fin de seleccionar estadísticamente los más apropiados.

Se han desarrollado varios algoritmos estadísticos que detectan qué gen o genes son los más apropiados para la normalización de un conjunto de tejidos bajo unas condiciones dadas: algunos, como geNORM o BestKeeper, realizan, sobre una matriz de expresión de genes de referencia para diversos teji-

dos, comparaciones por pares y medias geométricas. [Vandesompele et al., 2002] [Pfaffl et al., 2004]

2. Materiales y métodos

2.1. Material vegetal

Se empleó la variedad “V30” de *Petunia × hybrida* Hort. ex Vilm. [Koes et al., a] germinada sobre vermiculita y crecida en bandejas con una mezcla de vermiculita, perlita, turba y fibra de coco en proporción 2:1:2:2 en una cámara de cultivo bajo condiciones controladas de 16L:8D y 25 °C de temperatura en L y 18 °C en D. A fin de analizar los cambios a nivel vegetativo se seleccionaron dos estadios de desarrollo foliar, uno adulto (hoja C, correspondiente al primer o segundo par, desechando las excesivamente cloróticas y/o dañadas) y otro temprano (hoja A, primer o segundo par anteriores a la flor); y para observar la evolución en la flor, se estudiaron cuatro fases de desarrollo: primordio temprano (flor A, de 1 a 1,5 cm), primordio elongado (flor B, en torno a 2,5 cmB), flor cerrada (flor C, de 3,5 a 4,5 cm). y flor abierta (flor D, plena antesis). Se realizaron tres réplicas por tratamiento, contando cada réplica con nueve flores u hojas.

2.2. Extracción de ARN y retrotranscripción

A fin de extraer el ARN total, cada réplica fue homogenada en mortero y procesada mediante el kit NucleoSpin® RNA Plant (Macherey-Nagel, Düren, Germany), que incluía un tratamiento con ADNasa. La calidad del ARN obtenido fue analizada mediante electroforesis desnaturizante en agarosa y su cantidad cuantificada mediante espectrofotometría a 260 nm, para lo cual se emplearon diluciones realizadas en agua DEPC. La retrotranscripción se realizó mediante la enzima SuperScript III (Invitrogen Corp., Carlsbad, CA) empleando oligodT₂₀ y escalando la reacción sobre una receta base de: 10 µL de 2x RT reaction mix, 2 µL de RT enzyme mix y 8 µL de ARN. El protocolo de retrotranscripción (consistente en 10 minutos a 25 °C, 30 minutos a 50 °C, 5 minutos a 85 °C) fue finalizado mediante la adición de 1 u de ARNasa H de *Escherichia coli* (Fermentas, Hunover, MD, USA) y su incubación durante dos horas a 15 °C.

2.3. Optimización de la PCR

Se diseñaron 10 pares de cebadores específicos para los genes: cadena 4 de beta tubulina (TUB), peptidilprolil cis-trans isomerasa (CYP), actina (ACT), factor de elongación 1-alfa (EF1α), proteína ribosomal citoplasmática S13 (RPS13), ubiquitina (UBQ), Ran3 (RAN3), gliceraldehído-3-fosfato deshidrogenasa (GAPDH) y proteína SAND [Nicot et al., 2005]. Véase la tabla 1 para consultar las características de los cebadores y amplicones empleados. Las condiciones de PCR

fueron optimizadas en un termociclador creador de gradientes (Robocycler gradient 96, Stratagene, La Jolla, CA) mediante GoTaq® Flexi DNA polymerase (Promega, Madison, WI, USA) en una reacción 2 mM de $MgCl_2$, 0,2 mM de cada dNTP, 0,4 μ M de cada cebador y 1,25 u de enzima. Los protocolos optimizados fueron transferidos a la qPCR.

2.4. PCR en tiempo real

La PCR en tiempo real se efectuó empleando SYBR Premix Ex Taq® (TaKaRa Biotechnology, Dalian, Jiangsu, China) en un termociclador Rotor-Gene 2000 (Corbett Research, Australia), realizándose cada reacción por triplicado, con el contenido que se detalla a continuación: 2 μ L de cDNA, 7,5 μ L de SYBR *Premix Ex Taq* 2x, 0,36 μ L de cada cebador 10 μ M y 4,78 μ L de agua. El protocolo empleado fue el siguiente: 30 s a 95 °C, 30 s a 55 ó 60 °C, 30 s a 72 °C y 15 s a 80 °C; repitiéndose 40 ciclos. Finalmente, se añadió un programa de curva de fusión de 50 a 95 °C con un incremento de 1 °C y lapsos entre pasos de 5 s, más 15 s al inicio del primero. A fin de maximizar la especificidad de la reacción, se efectuaron dos tipos de reacciones de qPCR: las relativas a TUB, CYP, ACT, EF1 α , GAPDH y SAND poseyeron una temperatura de alineamiento de los cebadores de 55 °C, mientras que para RPS13, UBQ y RAN3 fue de 60 °C. Se realizó un control negativo sin ADNc para cada par de cebadores.

Tabla 1: Características de los genes, cebadores y amplicones empleados.

Nombre	Descripción	Accession number	Secuencia (forward/reverse)	Long. (pb)	Eficiencia	T _m (°C)
TUB	Tubulina beta 4	UP Q9ARZ3 (Q9ARZ3)	TGGAAACTCAACCTCCATCCA / TTTGGTCCATTGCTTCAC- CTG	114	1,61	83
CYP	Ciclofilina	Homólogo a UP CYPH_LYCES (P21568)	AGGCTCATCATTCACCGGTGT / TCATCTGCGAACTTAG- CACCG	111	1,67	84
ACT	Actin 97	Homólogo a UP ACTB_SOLTU (P30171)	TGCACTCCCACATGCTATCCT / TCAGCCGAAGTGGTGAAA- GAG	110	1,75	84
EF1 α	Factor de elongación 1 α	UP P93769 (P93769)	CCTGGTCAAATTGGAACGG / CAGATCGCCTGTCAAT- CTTGG	103	1,62	83
RPS13	Proteína ribosomal S13	UP Q6L4B2 (Q6L4B2)	AAGCTCCCACCTGTCTGGAAA / AACAGATTGCCGGAAGCCA	103	1,70	85
UBQ	Ubiquitina	UP Q39257 (Q39257)	TGGAGGATGGAAGGACTTTGG / CAGGACGACAACAAGCAA- CAG	153	1,64	84
Ran3	RAN3 <i>Petunia</i>	SGN-E522578	CAGGCAGGTTAAGGCAAAGC / CTAGCAAGGTACA- GAAACGGC	114	1,72	81
GAPDH	GAPDH <i>Petunia</i>	SGN-E533874	AACAACCTCACTCCTACACCGG / GGTAGCACTAGAGACA- CAGCCTT	135	1,86	85
SAND	Proteína SAND	SGN-E527478	CTTAGCAGGAGTTCAGATGCC / TAAGTCCTCAACACG- CATGC	135	1,64	82

2.5. Captura de datos

Los datos de fluorescencia crudos fueron procesados mediante el software del termociclador, obteniéndose los parámetros de eficiencia (véase la tabla 1) y CT (*cycle threshold*, también denominado CP, del inglés *crossing point*), siendo este último parámetro el número del ciclo en el cual la fluorescencia sobrepasa un umbral dado. Este umbral fue definido manualmente como el 20 % de la fluorescencia total normalizada, que, bajo nuestras condiciones, se encuentra en todas las muestras analizadas en la fase exponencial del experimento. El valor de CT es proporcional al logaritmo negativo de la cantidad inicial de ADNc [Gibson et al., 1996].

Se calculó la eficiencia de amplificación para cada reacción, es decir, para cada repetición de cada réplica, según la fórmula $eficiencia = [10^{(-1/g)}] - 1$ donde $g = 1/m$ y m es la pendiente de la recta obtenida al representar el logaritmo de la fluorescencia frente al número de ciclo. Los valores medios para cada amplicón se muestran en la tabla 1.

Los datos de CT de salida del termociclador fueron procesados mediante el software qBasePlus[Hellemans et al., 2007], analizando su calidad y transformándolos para su empleo mediante los algoritmos de normalización.

2.6. Algoritmo geNORM

El algoritmo geNORM realiza básicamente dos análisis: uno de evaluación de la estabilidad en la expresión de los genes seleccionados y otro que valora la conveniencia de emplear múltiples genes normalizadores [Vandesompele et al., 2002].

2.6.1. Valores M de estabilidad de expresión

Con el objeto de cuantificar la estabilidad de cada gen, el algoritmo empleado se fundamenta en que el ratio de expresión de dos genes internos debe ser por definición idéntico en todas las muestras, independientemente de las condiciones ambientales o del tipo celular. Por esta razón, una variación en los ratios de expresión refleja que al menos uno, o bien los dos, no está siendo expresado de forma constante, por lo que su estabilidad es menor. Siguiendo este planteamiento, el algoritmo determina la variación de cada gen control comparándolo a pares con los valores de los otros genes; esta comparación se traduce en la desviación típica del ratio de expresión transformado logarítmicamente. Finalmente, se define el valor M de estabilidad del gen normalizador como la variación a pares media de un gen particular frente a todos los otros genes normalizadores. De este modo, existe una relación inversa entre la estabilidad y el valor M.

Además, el algoritmo cuantifica el error sistemático sufrido durante las repeticiones técnicas, es decir, aquéllas diferencias en las datos de expresión para

el mismo gen y bajo las mismas condiciones de qPCR; este error es debido a las variaciones inherentes a la máquina, a las enzimas empleadas o al pipeteo.

2.6.2. Determinación del número óptimo de genes normalizadores

Como se indicaba en la sección introductoria (1.1), se ha descrito que es preciso emplear múltiples genes normalizadores a fin de realizar una valoración reproducible de la expresión relativa de uno o más genes. Por ello, es preciso desarrollar un factor de normalizador que dependa de los niveles de expresión de los genes normalizadores con una mayor estabilidad; la aproximación de geNORM a este enfoque incluye el empleo de la media geométrica de los valores de estabilidad de los distintos genes control empleados. Un aspecto crucial es definir qué número de genes normalizadores es preciso para cada experimento; evidentemente, debe existir un equilibrio entre la exactitud y las consideraciones de tipo práctico. Por ello, se recomienda emplear un número mínimo de tres genes normalizadores (los más estables) y añadir otros nuevos hasta que la inclusión de éstos no contribuya significativamente a la modificación del factor de normalización. Esta inclusión se computa como la variación del factor de normalización con n genes (donde $n \geq 3$) respecto del equivalente para $n + 1$ genes (es decir, de nuevo el análisis de variación a pares) para todas las muestras en el mismo grupo de tejidos u órganos.

Una vez determinados los valores de variación a pares (PV, del inglés *pairwise variation*) obtenidos mediante la comparación antes mencionada, es preciso establecer un umbral por debajo del cual la adición de más genes normalizadores es no necesaria. Este valor depende de los datos en concreto, pero suele aceptarse 0,15 como umbral adecuado [Vandesompele et al., 2002] [Coulson et al., 2008] [Toegel et al., 2007] [Silvia et al., 2006].

2.7. Cuantificación relativa mediante qBasePlus

Pese a que el objetivo del presente trabajo requiere del empleo de estrategias de cuantificación absoluta, se empleó el algoritmo de qBasePlus para, una vez evaluada la normalización, analizar la pauta de expresión de cada gen empleando el factor normalizador seleccionado por el propio software. La cuantificación relativa se produce mediante el principio básico del delta-CT, que asume que la diferencia en el valor del CT entre dos muestras puede transformarse en cantidades relativas empleando la función exponencial en la que la eficiencia de la PCR actúa como base. El software, además, evalúa los genes normalizadores más estables mediante un método que, además de definir el valor M, equivalente al empleado por geNORM, utiliza un coeficiente de variación (CV) de las cantidades relativas de los genes de referencia normalizados. Es decir, el coeficiente de variación indica el grado de estabilidad de expresión del gen. [Hellemans et al., 2007]

2.8. Análisis estadístico

El análisis estadístico de los datos y su representación gráfica se realizaron mediante el paquete R: A Language and Environment for Statistical Computing [R Development Core Team, 2008]. Las pruebas empleadas fueron: estadístico de Shapiro-Wilk para evaluar la normalidad, test de la F para detectar la homocedasticidad y test de la t o test de suma de rangos de Wilcoxon a fin de realizar comparaciones.

3. Resultados y discusión

3.1. Confirmación de la especificidad de los cebadores

Para cada repetición se analizó la curva de fusión generada a fin de confirmar la consistencia en temperatura del pico principal generado, así como la ausencia de otros artefactos. Además, se efectuaron electroforesis en gel de agarosa al 1 % de los productos de la qPCR a fin de confirmar la longitud del amplicón y, de nuevo, la especificidad de la reacción.

3.2. Valores M

Una vez aglutinados los datos de CT mediante el software qBasePlus, se analizaron sus valores M mediante el algoritmo geNORM tanto para el conjunto de ellos como para desgloses en función del tipo de muestra (flor A, B, C o D y hojas A o C) y conjuntos de estos últimos, como las hojas (hojas A y C) y flores (flores A, B, C y D). Para consultar los resultados, véase la tabla 2. Estos valores se representan mediante diagramas de cajas y bigotes (en los que se representa la mediana y los percentiles 25° y 75°, así como los valores disonantes externos a ellos) en la figura 1.

Se observa, por tanto, que los valores para cuando se analiza un conjunto heterogéneo de datos son altos, de acuerdo a la literatura [Pfaffl et al., 2004]. No obstante, la comparación de los resultados de diferentes estadios concretos muestra una gran disparidad en la respuesta para un mismo gen. El ejemplo más dramático es TUB que, para hojas de tipo A ó C es una buena elección (con $M = 0,09$) pero que, para flores B es demasiado inestable ($M = 1,45$).

De emplear muestras complejas, como flores u hojas de diferentes edades, la dificultad de encontrar un buen normalizador se acrecienta: en el caso de las flores, ninguno de los genes en solitario ofrece un valor $M \leq 0,5$ (los valores mínimos, para ACT y RAN3, son de 0,79); en el de las hojas, sí que existen tres adecuados, TUB con $M = 0,5$ y RPS13 y UBQ con $M = 0,26$. Estas diferencias pueden ser debidas a el mayor número de estadios discretos analizados en el caso

Tabla 2: Valores M según el algoritmo geNORM. Estos valores guardan relación inversamente proporcional con la estabilidad de la expresión.

	GAPDH	SAND	EF1 α	CYP	TUB	ACT	RAN3	RPS13	UBQ
Flor A	0,56	0,15	0,15	1,27	0,15	0,15	0,15	0,84	0,46
Flor B	1,03	0,67	0,59	0,23	1,45	0,9	0,23	1,27	1,18
Flor C	1,01	0,56	0,84	0,44	1,25	0,74	0,63	0,28	0,28
Flor D	0,85	0,29	0,09	0,66	0,19	0,45	0,25	0,09	0,15
Hoja A	0,42	0,56	0,14	0,83	0,09	0,18	0,69	0,09	0,21
Hoja C	0,06	0,06	0,64	1,23	0,09	0,4	0,2	0,1	0,15
Flores	4	3,54	2,94	1,67	1,87	0,79	0,79	1,55	1,43
Hojas	0,73	0,59	0,81	1,23	0,5	0,68	0,9	0,26	0,26
Todos	3,8	3,5	2,95	1,95	1,86	1,74	1,5	0,97	0,97

de las hojas o a la mayor complejidad estructural y funcional de las flores, como órganos modulares que son [Endress, 2001] [Schlosser & Wagner, 2004].

De acuerdo a los resultados experimentales, se observa que seis genes cumplen el criterio de estabilidad seleccionado para las flores de estadio A (SAND, EF1 α , TUB, ACT, RAN3 y UBQ); sólo uno para las flores de tipo B, CYP; tres para las flores de tipo C (CYP, RPS13, UBQ); y seis para las flores del estadio D (SAND, EF1 α , TUB, RAN3, RPS13 y UBQ). En cuanto a las hojas: seis genes son adecuados para normalizar las de tipo A (GADP, EF1 α , TUB, ACT, RPS13 y UBQ); y siete para las de tipo C (GAPDH, SAND, TUB, ACT, RAN3, RPS13 y UBQ). Ninguno de los genes analizados se mantiene estable durante los cuatro estadios descritos de desarrollo floral, aunque sí durante el foliar (RPS13, UBQ). Del mismo modo, ninguno de ellos es válido para normalizar órganos foliares y florales simultáneamente. La diferente pauta de expresión se representa en la figura 3.

De este modo, el ranking de estabilidad para el conjunto de tejidos analizados es: RPS13=UBQ>RAN3>ACT>TUB>CYP>EF1 α >SAND>GAPDH.

3.3. Valores CT

En los diagramas de cajas y bigotes siguientes se representan los valores CT para los distintos genes normalizadores en el contexto del análisis de todas las muestras y de los estadios de desarrollo florales y foliares (figura 1).

Cabe destacar la amplitud de los rangos de CT cuando se evalúa la expresión de los diversos genes en muestras mixtas de diversos estadios de desarrollo. No obstante, esta amplitud es mucho menor cuando se desglosa la expresión por estadios por estadio de desarrollo (véase la figura 2), puesto que en la mayoría

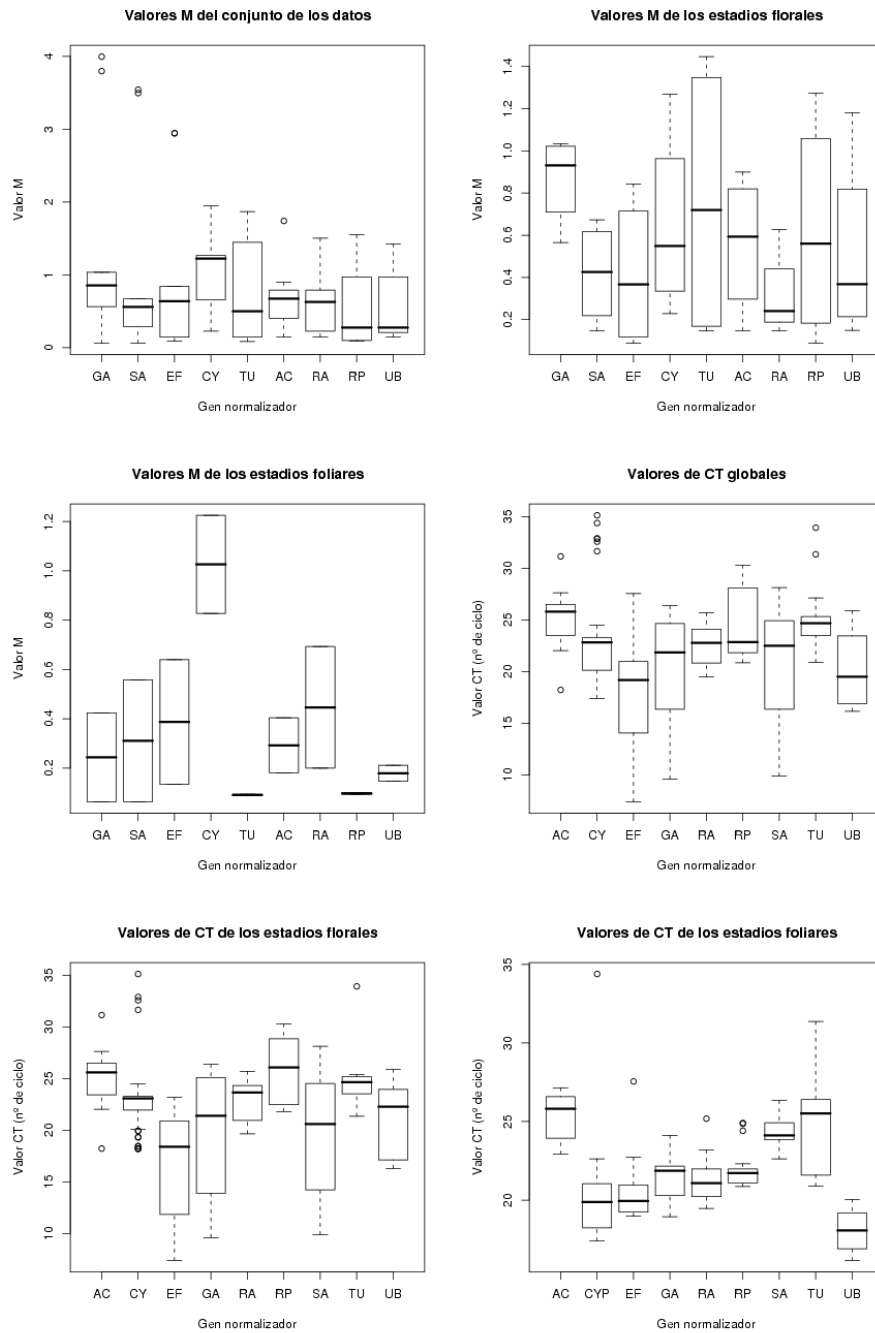


Figura 1: Diagrama de cajas y bigotes de los valores M del conjunto de las muestras (arriba, izquierda), de los valores del conjunto de piezas florales (arriba, centro) y foliares (arriba, derecha). Diagrama de los valores CT del conjunto de las muestras (abajo, izquierda), de los valores del conjunto de piezas florales (abajo, centro) y foliares (abajo, derecha) Las abreviaturas significan: AC, ACT; CY, CYP; EF, EF1 α ; GA, GAPDH; RA, RAN3, RP, RPS13; SA, SAND; TU, TUB; UB, UBQ.

de los casos existe una gran homogeneidad de valores CT dentro de cada estadio y para cada gen normalizador. Los análisis de comparación entre los bloques foliar y floral arrojó diferencias significativas ($\alpha < 0,05$) en los valores CT de CYP, SAND, EF1 α , RAN3, RPS13 y UBQ. Estas diferencias en los valores CT están relacionadas de forma inversa con la expresión génica, luego el presente trabajo reitera que los *housekeeping genes* no se expresan de forma constante en todo tejido: detectamos diferencias en seis de los nueve genes estudiados. No obstante, resulta llamativa la ausencia de diferencias significativas en la expresión de TUB, gen del que se conoce una menor expresión en las hojas conforme éstas maduran, si bien las diferencias más dramáticas se han encontrado en monocotiledóneas [Hellmann et al., 1995].

Asimismo, se representan los valores de CT para cada estadio de desarrollo concreto (figura 2). En este caso, destaca la escasa dispersión de los datos en los estadios florales A y D, así como en la hoja C. En cuanto al estadio de desarrollo B, esta variabilidad podría achacarse a su menor entidad morfológica: es decir, debido a su definición como paso intermedio entre dos estados claramente definidos (el primordio floral temprano y el tubo previo a la antesis), es posible que el material vegetal original sea más diverso. De hecho, los genes EF1 α , GAPDH y SAND, los de rango más amplio, están acotados en el estadio B mediante el máximo del estadio previo y el mínimo del posterior.

3.4. Determinación del número óptimo de genes normalizadores

Los datos que muestran el número mínimo necesario de genes normalizadores para el material vegetal analizado se muestran en la tabla 3 y se representan en la figura 3. Con este análisis se definen cuatro tipos de materiales vegetales, empleando un umbral de decisión en el cual el PV al añadir un nuevo gen de normalización sea menor de 0,15:

- Estadios cuya normalización requiere de tan sólo dos genes: flores A y D y hojas A y C. No obstante, la bibliografía aconseja emplear un mínimo de tres [Vandesompele et al., 2002].
- Material vegetal que requiere de tres genes de normalización: hojas en general.
- Estadios que precisan de cuatro genes: flores de tipo C.
- Estadios y material vegetal cuya evaluación no mejora al añadir más genes normalizadores: flores de tipo B, material floral mixto y mezclas de órganos foliares y florales.

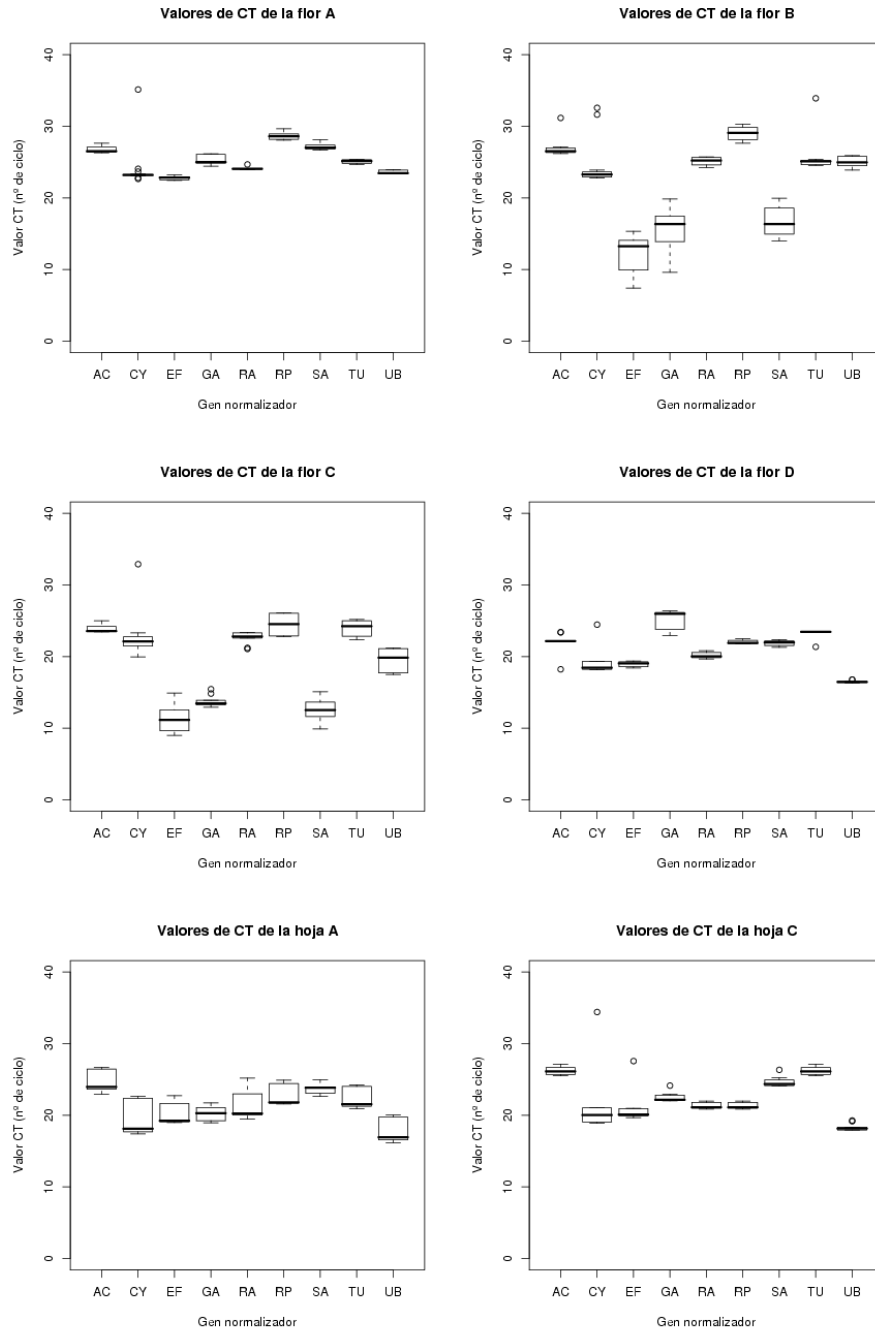


Figura 2: Valores de CT para cada uno de los estadios de desarrollo. Las abreviaturas significan: AC, ACT; CY, CYP; EF, EF1 α ; GA, GAPDH; RA, RAN3; RP, RPS13; SA, SAND; TU, TUB; UB, UBQ.

Tabla 3: Valores PV según el algoritmo geNORM. Este valor, que representa la variación entre dos factores de normalización secuenciales con diferente número de genes normalizadores, indica si $PV > 0,15$ la necesidad de añadir nuevos genes internos (véase la sección 3).

	V2/3	V3/4	V4/5	V5/6	V6/7	V7/8	V8/9
Flor A	0,1	0,12	0,09	0,08	0,11	0,21	0,3
Flor B	0,25	0,19	0,24	0,21	0,2	0,18	0,22
Flor C	0,17	0,16	0,13	0,15	0,15	0,18	0,23
Flor D	0,06	0,06	0,07	0,06	0,12	0,16	0,16
Hoja A	0,05	0,05	0,05	0,14	0,12	0,13	0,14
Hoja C	0,03	0,03	0,04	0,05	0,13	0,17	0,36
Hojas	0,2	0,15	0,14	0,12	0,13	0,13	0,25
Flores	0,57	0,37	0,33	0,34	0,78	0,64	0,58
Todos	0,57	0,45	0,35	0,29	0,76	0,61	0,48

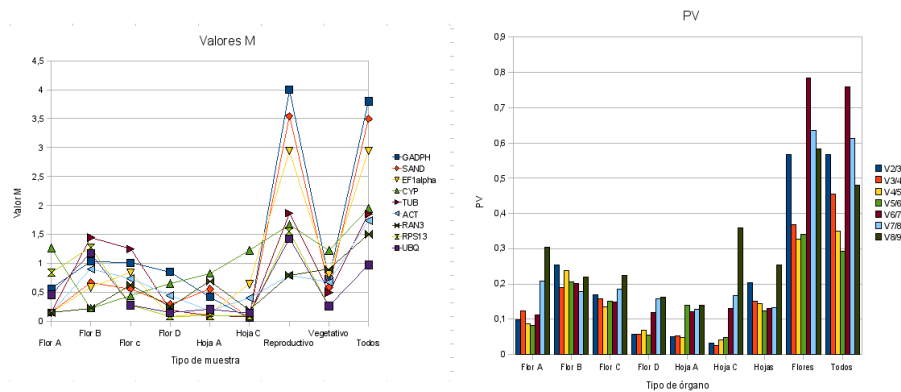


Figura 3: Análisis mediante el algoritmo geNORM para los distintos tipos de material vegetal de los valores: M, que son inversamente proporcionales a la estabilidad de expresión (izquierda); y PV, que indican la necesidad de añadir nuevos genes de referencia (derecha).

3.5. Cuantificación relativa

Los resultados del análisis de cuantificación relativa de cada gen de normalización, realizados mediante el software qBasePlus, se muestran en la figura 4. De forma consistente con el análisis de valores CT, se observa la pauta de expresión diferenciada entre el bloque vegetativo y el reproductivo: ACT posee un bajo nivel de expresión en flores de tipo B y una pauta similar en los demás estadios; CYP, bajo en flores de tipo B, disminuye de forma drástica su expresión en flores de tipo C; EF1 α y GAPDH se expresan en mayor medida en flores en estadios B y C, si bien en este último la expresión en flores de tipo D es muy inferior; RAN3 y RPS13 presentan una baja expresión en flores de tipo B y alta en hojas y flores de tipo D; SAND, de baja expresión en flores A y hojas, posee niveles altos en flores B y C; TUB se expresa de forma baja en flores de tipo B; y UBQ posee una baja expresión en flores B y C. En cuanto a la variabilidad detectada, el gen que la presenta en mayor medida es el de CYP, gen que, además, es el único capaz de normalizar ($M = 0, 23$) las flores de tipo B, el estadio más variable.

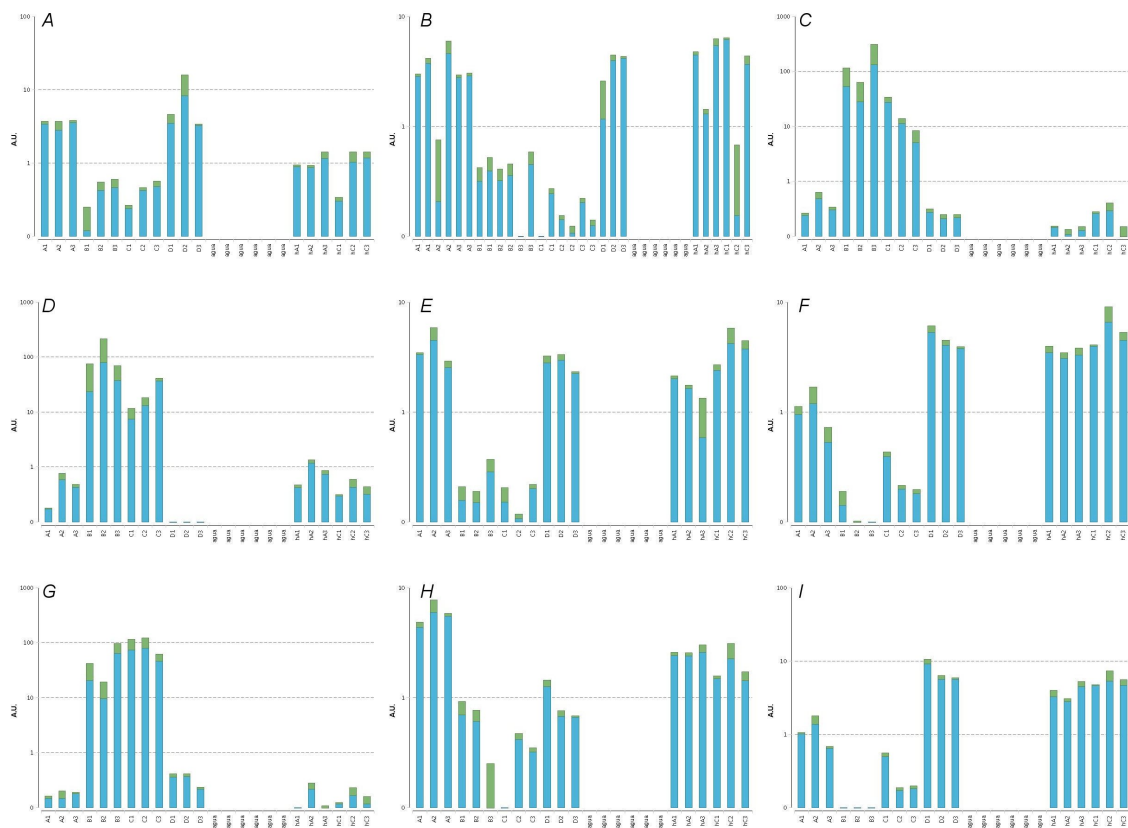


Figura 4: Análisis de la expresión relativa de los distintos genes mediante el software qBasePlus. En el eje de ordenadas se indica, en unidades arbitrarias, el nivel de expresión; el el de abscisas se listan las flores A, B, C, D (tres réplicas cada uno), los controles sin molde (seis réplicas, una para cada tipo de muestra) y las hojas A y C (tres réplicas cada uno). Las abreviaturas significan: A, ACT; B, CYP; C, EF1 α ; D, GAPDH; E, RAN3, F, RPS13; G, SAND; H, TUB; I, UBQ.

4. Anexo

En las tablas siguientes se detallan los datos obtenidos mediante el algoritmo geNORM para todas las muestras realizadas (tabla 4) así como para los distintos grupos de tejidos: flores de tipo A (tabla 5), B (tabla 6), C (tabla 7) y D (tabla 8); hojas de tipo A (tabla 9), C (tabla 10); y el módulo reproductivo, es decir, todas las flores (tabla 12) y vegetativo, es decir, todas las hojas (tabla 11).

Tabla 4: Valores M y factores de normalización según el algoritmo geNORM para todas las muestras analizadas.

	ACT	CYP	EF1 α	GAPDH	RPS13	RAN3	SAND	TUB	UBQ	Factor de normalización
A1	0,3612	0,3939	0,0257	0,0183	0,0109	0,3553	0,0157	0,4680	0,1082	0,0839
A2	0,2032	0,0148	0,0351	0,0420	0,0985	0,3262	0,0107	0,4331	0,1003	0,0740
A3	0,4017	0,3090	0,0343	0,0481	0,0689	0,2879	0,0201	0,6283	0,0729	0,1158
B1	0,1189	0,3141	52,45	23,300	0,1624	0,1572	20,330	0,7003	0,0579	0,9954
B2	0,3058	0,2278	20,060	56,710	0,0657	0,1086	6,9280	0,4381	0,0362	0,7343
B3	0,4226	0,4068	119,10	33,960	0,0358	0,2607	58,180	0,0734	0,0191	0,9316
C1	2,8760	2,8860	326,20	88,990	5,3930	1,8160	874,60	0,5010	5,9460	12,2361
C2	3,0190	1,0060	80,050	93,850	1,6330	0,7687	566,40	2,9890	1,2270	7,3087
C3	1,3080	0,5545	13,940	101,30	0,5709	0,5563	126,40	0,8926	0,5030	2,9787
D1	4,3090	1,4070	0,3383	0,0188	7,4360	3,4280	0,4441	1,5430	11,280	1,2530
D2	18,600	8,8920	0,4729	0,0212	10,4900	6,6990	0,8269	1,5420	12,720	2,3121
D3	7,9490	10,110	0,5332	0,1320	10,5900	5,4920	0,5393	1,6350	13,880	2,5151
hA1	2,2460	11,020	0,3617	1,0460	9,9720	5,0650	0,2048	6,0890	8,1790	2,5661
hA2	0,3570	0,5282	0,0448	0,4848	1,4480	0,6754	0,0896	0,9859	1,1530	0,4201
hA3	3,7250	17,180	0,4145	2,3600	12,0800	1,8880	0,2551	8,2790	14,240	3,2717
hC1	0,2988	5,9210	0,2534	0,2910	4,4420	2,3570	0,1144	1,4600	4,5160	1,0033
hC2	0,7124	0,1065	0,2039	0,2964	5,2100	2,9500	0,1147	1,5720	3,7110	0,7122
hC3	0,4799	1,4870	0,0240	0,1311	2,1110	1,5300	0,0475	0,5875	1,9250	0,4180
Valor M	3,114	3,359	4,547	4,850	3,487	3,144	4,827	3,280	3,569	

Tabla 5: Valores M y factores de normalización según el algoritmo geNORM para todas las flores de tipo A.

	ACT	CYP	EF1 α	GAPDH	RPS13 RQ	RAN3	SAND RQ	TUB	UBQ	Factor de normalización
A1	0,3612	0,3939	0,0257	0,0183	0,0109	0,3553	0,0157	0,4680	0,1082	0,9366
A2	0,2032	0,0148	0,0351	0,0420	0,0985	0,3262	0,0107	0,4331	0,1003	0,8259
A3	0,4017	0,3090	0,0343	0,0481	0,0689	0,2879	0,0201	0,6283	0,0729	1,2927
Valor M	0,958	2,760	0,900	1,129	1,965	0,931	0,897	0,844	1,033	

Tabla 6: Valores M y factores de normalización según el algoritmo geNORM para todas las flores de tipo B.

	ACT	CYP	EF1 α	GAPDH	RPS13 RQ	RAN3	SAND RQ	TUB	UBQ	Factor de normali- zación
B1	0,1189	0,3141	52,4500	23,3000	0,1624	0,1572	20,3300	0,7003	0,0579	1,1315
B2	0,3058	0,2278	20,0600	56,7100	0,0657	0,1086	6,9280	0,4381	0,0362	0,8347
B3	0,4226	0,4068	119,1000	33,9600	0,0358	0,2607	58,1800	0,0734	0,0191	1,0589
Valor M	1,500	1,070	1,448	1,394	1,472	1,116	1,651	2,049	1,319	

Tabla 7: Valores M y factores de normalización según el algoritmo geNORM para todas las flores de tipo C.

	ACT	CYP	EF1 α	GAPDH	RPS13 RQ	RAN3	SAND RQ	TUB	UBQ	Factor de normali- zación
C1	2,8760	2,8860	326,2000	88,9900	5,3930	1,8160	874,6000	0,5010	5,9460	1,9017
C2	3,0190	1,0060	80,0500	93,8500	1,6330	0,7687	566,4000	2,9890	1,2270	1,1359
C3	1,3080	0,5545	13,9400	101,3000	0,5709	0,5563	126,4000	0,8926	0,5030	0,4629
Valor M	1,080	0,935	1,484	1,483	1,002	0,994	1,025	2,088	1,155	

Tabla 8: Valores M y factores de normalización según el algoritmo geNORM para todas las flores de tipo D.

	ACT	CYP	EF1 α	GAPDH	RPS13 RQ	RAN3	SAND RQ	TUB	UBQ	Factor de normali- zación
D1	4,3090	1,4070	0,3383	0,0188	7,4360	3,4280	0,4441	1,5430	11,2800	0,6463
D2	18,6000	8,8920	0,4729	0,0212	10,4900	6,6990	0,8269	1,5420	12,7200	1,1926
D3	7,9490	10,1100	0,5332	0,1320	10,5900	5,4920	0,5393	1,6350	13,8800	1,2973
Valores M	0,993	1,309	0,593	1,542	0,582	0,609	0,681	0,726	0,649	

Tabla 9: Valores M y factores de normalización según el algoritmo geNORM para todas las hojas de tipo A.

	ACT	CYP	EF1 α	GAPDH	RPS13 RQ	RAN3	SAND RQ	TUB	UBQ	Factor de normali- zación
hA1	2,2460	11,0200	0,3617	1,0460	9,9720	5,0650	0,2048	6,0890	8,1790	1,6858
hA2	0,3570	0,5282	0,0448	0,4848	1,4480	0,6754	0,0896	0,9859	1,1530	0,2760
hA3	3,7250	17,1800	0,4145	2,3600	12,0800	1,8880	0,2551	8,2790	14,2400	2,1493
	0,592	1,296	0,597	0,976	0,562	1,183	1,037	0,557	0,645	

Tabla 10: Valores M y factores de normalización según el algoritmo geNORM para todas las hojas de tipo C.

	ACT	CYP	EF1 α	GAPDH	RPS13 RQ	RAN3	SAND RQ	TUB	UBQ	Factor de normali- zación
hC1	0,2988	5,9210	0,2534	0,2910	4,4420	2,3570	0,1144	1,4600	4,5160	1,5010
hC2	0,7124	0,1065	0,2039	0,2964	5,2100	2,9500	0,1147	1,5720	3,7110	1,0655
hC3	0,4799	1,4870	0,0240	0,1311	2,1110	1,5300	0,0475	0,5875	1,9250	0,6253
Valor M	1,389	3,273	1,635	0,751	0,774	0,858	0,758	0,783	0,806	

Tabla 11: Valores M y factores de normalización según el algoritmo geNORM para el módulo vegetativo.

	ACT	CYP	EF1 α	GAPDH	RPS13 RQ	RAN3	SAND RQ	TUB	UBQ	Factor de normali- zación
hA1	2,2460	11,0200	0,3617	1,0460	9,9720	5,0650	0,2048	6,0890	8,1790	2,5439
hA2	0,3570	0,5282	0,0448	0,4848	1,4480	0,6754	0,0896	0,9859	1,1530	0,4165
hA3	3,7250	17,1800	0,4145	2,3600	12,0800	1,8880	0,2551	8,2790	14,2400	3,2433
hC1	0,2988	5,9210	0,2534	0,2910	4,4420	2,3570	0,1144	1,4600	4,5160	0,9947
hC2	0,7124	0,1065	0,2039	0,2964	5,2100	2,9500	0,1147	1,5720	3,7110	0,7061
hC3	0,4799	1,4870	0,0240	0,1311	2,1110	1,5300	0,0475	0,5875	1,9250	0,4144
Valor M	1,144	2,367	1,229	1,172	0,903	1,354	1,013	0,932	0,925	

Tabla 12: Valores M y factores de normalización según el algoritmo geNORM para el módulo reproductivo.

	ACT	CYP	EF1 α	GAPDH	RPS13 RQ	RAN3	SAND RQ	TUB	UBQ	Factor de normali- zación
A1	0,361	0,394	0,026	0,018	0,011	0,355	0,016	0,468	0,108	0,0842
A2	0,203	0,015	0,035	0,042	0,099	0,326	0,011	0,433	0,100	0,0743
A3	0,402	0,309	0,034	0,048	0,069	0,288	0,020	0,628	0,073	0,1163
B1	0,119	0,314	52,450	23,300	0,162	0,157	20,330	0,700	0,058	0,9998
B2	0,306	0,228	20,060	56,710	0,066	0,109	6,928	0,438	0,036	0,7375
B3	0,423	0,407	119,100	33,960	0,036	0,261	58,180	0,073	0,019	0,9357
C1	2,876	2,886	326,200	88,990	5,393	1,816	874,600	0,501	5,946	12,2893
C2	3,019	1,006	80,050	93,850	1,633	0,769	566,400	2,989	1,227	7,3405
C3	1,308	0,555	13,940	101,300	0,571	0,556	126,400	0,893	0,503	2,9917
D1	4,309	1,407	0,338	0,019	7,436	3,428	0,444	1,543	11,280	1,2584
D2	18,600	8,892	0,473	0,021	10,490	6,699	0,827	1,542	12,720	2,3222
D3	7,949	10,110	0,533	0,132	10,590	5,492	0,539	1,635	13,880	2,5260
Valor M	3,165	3,337	4,691	5,582	3,556	3,284	5,006	3,602	3,751	

Referencias

- [Brunner et al., 2004] Brunner, A., Yakovlev, I. & Strauss, S. (2004). *BMC Plant Biology* 4, 14.
- [Coulson et al., 2008] Coulson, D., Brockbank, S., Quinn, J., Murphy, S., Ravid, R., Irvine, G. & Johnston, J. (2008). *BMC Molecular Biology* 9, 46.
- [Czechowski et al., 2005] Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. & Scheible, W. (2005).
- [Endress, 2001] Endress, P. (2001). *Journal of Experimental Zoology (Molecular Development and Evolution)* 291, 105–115.
- [Gibson et al., 1996] Gibson, U., Heid, C. & Williams, P. (1996). *Genome Research* 6, 995–1001.
- [Hellemans et al., 2007] Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. & Vandesompele, J. (2007). *Genome Biology* 8, R19.
- [Hellmann et al., 1995] Hellmann, A., Meyer, C. & Wernicke, W. (1995). *Cell Motility and the Cytoskeleton* 30, 67–72.
- [Higuchi et al., 1993] Higuchi, R., Fockler, C., Dollinger, G. & Watson, R. (1993). *Bio/Technology* 11, 1026–1030.
- [Koes et al., a] Koes, R., Spelt, C., Reif, H., van den Elzen, P., Veltkamp, E. & Mol, J. *Nucleic Acids Research* 14, 5229–5239.
- [Nicot et al., 2005] Nicot, N., Hausman, J.-F., Hoffmann, L. & Evers, D. (2005). *J. Exp. Bot.* 56, 2907–2914.
- [Pfaffl et al., 2004] Pfaffl, M., Tichopad, A., Prgomet, C. & Neuvians, T. (2004). *Biotechnology Letters* 26, 509–515.
- [R Development Core Team, 2008] R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria. ISBN 3-900051-07-0.
- [Radonić et al., 2004] Radonić, A., Thulke, S., Mackay, I., Landt, O., Siegert, W. & Nitsche, A. (2004). *Biochemical and Biophysical Research Communications* 313, 856–862.
- [Schlosser & Wagner, 2004] Schlosser, G. & Wagner, G. (2004). *Modularity in development and evolution*. University of Chicago Press Chicago.
- [Schmittgen et al., 2000] Schmittgen, T., Zakrajsek, B., Mills, A., Gorn, V., Singer, M. & Reed, M. (2000). *Analytical Biochemistry* 285, 194–204.
- [Silvia et al., 2006] Silvia, S., Francesca, C., Marco, L., Silvia, N. et al. (2006). *BMC Cancer* 6, 200.

- [Stürzenbaum & Kille, 2001] Stürzenbaum, S. & Kille, P. (2001). *Comparative Biochemistry and Physiology, Part B* 130, 281–289.
- [T.B. Morrison & Wittwer, 1998] T.B. Morrison, J. W. & Wittwer, C. (1998). *BioTechniques* 24, 954–962.
- [Thellin et al., 1999] Thellin, O., Zorzi, W., Lakaye, B., De Borman, B., Coumans, B., Hennen, G., Grisar, T., Igout, A. & Heinen, E. (1999). *Journal of Biotechnology* 75, 291–295.
- [Toegel et al., 2007] Toegel, S., Huang, W., Piana, C., Unger, F., Wirth, M., Goldring, M., Gabor, F. & Viernstein, H. (2007). *BMC Molecular Biology* 8, 13.
- [Vandesompele et al., 2002] Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. & Speleman, F. (2002). *Genome Biology* 3, research0034.1–0034.11.