



**HAL**  
open science

## Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions

Hua Cassan-Wang, Marçal Soler, Hong Yu, Eduardo Leal O. Camargo, Victor Carocha, Nathalie Ladouce, Bruno Savelli, Jorge A. Pinto Paiva, Jean-Charles Leplé, Jacqueline Grima Pettenati

### ► To cite this version:

Hua Cassan-Wang, Marçal Soler, Hong Yu, Eduardo Leal O. Camargo, Victor Carocha, et al.. Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions. *Plant and Cell Physiology*, 2012, 53 (12), pp.2101-2116. 10.1093/pcp/pcs152 . hal-01267884

**HAL Id: hal-01267884**

**<https://hal.science/hal-01267884>**

Submitted on 28 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript  
published in : Plant & Cell Physiology. 2012, 53(12), 2101-2116, doi: [10.1093/pcp/pcs152](https://doi.org/10.1093/pcp/pcs152)

## Reference Genes for High-Throughput Quantitative RT-PCR Analysis of Gene Expression in Organs and Tissues of Eucalyptus Grown in Various Environmental Conditions.

Hua Cassan-Wang<sup>1\*</sup>, Marçal Soler<sup>1</sup>, Hong Yu<sup>1</sup>, Eduardo Leal O. Camargo<sup>1,4</sup>, Victor Carocha<sup>1,2,3</sup>, Nathalie Ladouce<sup>1</sup>, Bruno Savelli<sup>1</sup>, Jorge A. Pinto Paiva<sup>2,3</sup>, Jean-Charles Leplé<sup>5</sup>, Jacqueline Grima-Pettenati<sup>1\*</sup>

\*corresponding authors

1 LRSV, Laboratoire de Recherche en Sciences Végétales, Université Toulouse III, UPS, CNRS, BP 42617, Auzeville, 31326 Castanet Tolosan, France.

2 Instituto de Investigação Científica e Tropical (IICT/MNE); Palácio Burnay - Rua da Junqueira, 30, 1349-007 Lisboa; Portugal.

3 Instituto de Biologia Experimental e Tecnológica (IBET) Av. da República, Quinta do Marquês, 2781-901 Oeiras, Portugal.

4 Universidade Estadual de Campinas; UNICAMP ; Instituto de Biologia ; Departamento de Genética, Evolução e Bioagentes; Laboratório de Genômica e Expressão, SP, Brasil.

5 INRA, UR0588 Amélioration Génétique et Physiologie Forestières (AGPF), F-45075 Orléans, France.

### Abbreviation:

ACT, actin; AS, all samples; CT, cycle threshold; EF1a, elongation factor1a; GAPDH, glyceraldehyde3phosphate dehydrogenase; GM, geometric mean; IDH, NADP Isocitrate dehydrogenase; L, leaves; M value, gene expression stability value; OT, organs and tissues; PP2A, protein phosphatase 2A subunits; qRT-PCR, quantitative reverse transcription PCR; TUA, atubulin; TUB,  $\beta$ tubulin; UBIQ, polyubiquitin (UBI); XA, xylem all; XC, xylem-cambium; XES, xylem environmental stimuli.

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A. P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions. *Plant and Cell Physiology*, 53 (12).

## Abstract

Interest in the genomics of Eucalyptus has skyrocketed thanks to the recent sequencing of the genome of Eucalyptus grandis and to a growing number of largescale transcriptomic studies. Quantitative RTPCR is the method of choice for gene expression analysis and can also be used as a highthroughput method now. The selection of appropriate internal controls is becoming of utmost importance to ensure accurate expression results in Eucalyptus. To this end, we selected 21 candidate reference genes and used highthroughput microfluidic dynamic arrays, to assess their expression among a large panel of developmental and environmental conditions with a special focus on woodforming tissues. We analyzed the expression stability of these genes by using three distinct statistical algorithms (geNorm, NormFinder and Ct) and used principal component analysis to compare methods and rankings. We showed that the most stable genes identified depended not only on the panel of biological samples considered but also on the statistical method used. We then developed a comprehensive integration of the rankings generated by the three methods and identified the optimal reference genes for 17 distinct experimental sets covering 13 organs and tissues as well as various developmental and environmental conditions. The expression patterns of Eucalyptus master genes EgMYB1 and EgMYB2 experimentally validated our selection. Our findings provide an important resource for the selection of appropriate reference genes for accurate and reliable normalization of gene expression data in the organs and tissues of Eucalyptus trees grown in a range of conditions including abiotic stresses.

## Keywords:

**abiotic stress, Eucalyptus, gene expression, normalisation, reference genes, xylem**

## Introduction:

Wood is the main source of terrestrial biomass; as well as providing fibres and solid wood products, it is a significant renewable and environmentally costeffective alternative feedstock to biofuel and a major sink for excess atmospheric CO<sub>2</sub> (Boudet et al. 2003; Plomion et al. 2001). Wood formation or xylogenesis involves a transition from meristematic cambium cells to highly differentiated xylem cells, and is one of the most remarkable examples of plant cell differentiation. It involves sequential stages of cell division, cell elongation, formation of lignified secondary cell walls and finally programmed cell death to produce tracheary elements. This irreversible developmental process requires complex networks of spatial and temporal regulation events in order to coordinate gene expression networks. Exhaustive sequencing of developing xylem tissues in several tree species and in model systems have increased our understanding of wood formation and highlighted the pivotal role of transcriptional regulation (Demura and Fukuda 2007). In Eucalyptus, the most planted hardwood tree in the world mainly for pulp production and also among the most appealing

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions. *Plant and Cell Physiology*, 53 (12).

lignocellulosic feedstocks for bioenergy production (Myburg et al. 2007), largescale transcriptomic studies including microarrays have revealed thousands of genes expressed in differentiating xylem (Foucart et al. 2006; Paux et al. 2004; Rengel et al. 2009). The roles of most of these genes, however, are still to be discovered. To this end, gene expression profiling of various organs and/or under different conditions is a powerful tool to identify those genes relevant to new biological processes, to provide insights into regulatory networks, and to select members of multigene families prior to functional characterization. With the sequence of the genome of *E. grandis* publicly accessible since 2011 in the Phytozome database (<http://www.phytozome.net/>) (Goodstein et al. 2011), interest in the genomics of *Eucalyptus* has skyrocketed. There are growing numbers of functional genomics studies and whole transcriptome sequencing studies linked to digital transcript counting (RNAseq) of various *Eucalyptus* species, organs, tissues, developmental stages and environmental conditions (Camargo et al. 2012; Paiva et al. 2011; Villar et al. 2011) and, thus, a pressing need for accurate techniques to validate and mine these high-throughput expression data.

Real-time quantitative reverse transcription PCR (qRT-PCR) is considered as the most sensitive and specific technique to assess expression patterns of a moderate number of genes and has become the standard method for validating microarray data. One limitation of qRT-PCR was the relatively low number of genes that can be assessed. This limitation has been overcome with the microfluidic technology allowing high-throughput expression measurements using dynamic arrays (Spurgeon et al. 2008). This technology which allow 9 216 simultaneous real time PCR gene expression measurements in a single run has been proven to be as reliable as conventional real time PCR (Spurgeon et al. 2008) and was successfully used in several applications as for instance in animal and human for single-cell gene expression analysis (Guo et al. 2010; Pang et al. 2011) but to the best of our knowledge, so far no study on gene expression measurement in plants has been reported.

When using either low- or high- throughput qRT-PCR, appropriate normalization is essential for obtaining an accurate and reliable quantification of gene expression level. The purpose of normalization is to correct technical or experimental variations in order to reveal the true biological changes in expression. This is especially relevant when the samples come from different individuals, different tissues and/organs, different time courses, different environmental conditions, etc. The success of the normalization strategy in correcting variability between samples is highly dependent on the choice of the appropriate internal control or reference gene since its expression level should be constant among the tissues or cells, the experimental treatments and the biological conditions tested. No universal reference gene has been found in any plant, animal or medical system. In the pre-genomic era, genes believed to play housekeeping roles in basic cellular processes such as 18S rRNA, Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), Elongation factor-1 $\alpha$  (EF-1 $\alpha$ ), Polyubiquitin (UBQ), Actin (ACT),  $\alpha$ -Tubulin and  $\beta$ -Tubulin (TUA and TUB, respectively) were frequently used as reference genes. The growing use of the very sensitive qRT-PCR technique, however, has shown that the expression of these genes varies under different experimental or biological conditions and their

systematic use without previous validation can lead to the misinterpretation of results (Czechowski et al. 2005; Gutierrez et al. 2008a). It has become clear that it is necessary to validate the expression stability of a candidate control gene in each experimental system prior to its use for normalization (Bustin et al. 2009; Guenin et al. 2009; Gutierrez et al. 2008a; Udvardi et al. 2008).

In their pioneering work, Scheible and colleagues (Czechowski et al. 2005) analyzed a very large set of data from Affymetrix ATH1 (whole-genome GeneChip microarrays of RNA from the model plant *Arabidopsis*) and selected hundreds of candidate reference genes that were expressed at similar levels in a wide range of experimental conditions, outperforming traditional housekeeping genes (GAPDH, ACT2, UBQ10, UBC and EF1 $\alpha$ ) in terms of expression stability. About 20 of these new-generation reference genes were assessed by qRT-PCR throughout development and in a range of environmental conditions confirming their superior expression stability and lower absolute expression levels when compared to traditional housekeeping genes. These genes included those encoding members of the polyubiquitin family, proteins with potential regulatory functions such as F-box proteins, protein phosphatase 2A subunits (PP2A), and ‘expressed proteins’ of unknown function. These new reference genes were successfully employed to search for orthologs in unrelated species such as *Vitis vinifera* (Reid et al. 2006), and cotton (Artico et al. 2010).

Most of reference gene evaluation studies of plants have been performed on model species and crop species (Czechowski et al. 2005; Exposito-Rodriguez et al. 2008; Guenin et al. 2009; Gutierrez et al. 2008b) and fewer on woody plants. Strauss and colleagues (Brunner et al. 2004) used single-factor ANOVA and linear regression analysis to study the expression of ten potential reference genes in eight tissues from the poplar tree, and Huang and colleagues (Xu et al. 2010) tested three algorithms – geNorm, NormFinder and BestKeeper – to assess the stability of nine housekeeping genes during adventitious rooting of poplar. In *Eucalyptus*, traditional housekeeping genes such as 18S rRNA (Foucart et al. 2006; Navarro et al. 2011), GAPDH (Navarro et al. 2011), and NADP Isocitrate dehydrogenase (IDH) (Gallo de Carvalho et al. 2008; Goicoechea et al. 2005; Legay et al. 2010; Paux et al. 2004) have been used frequently as internal controls for gene expression analysis. Recent studies, however, have reported the selection of reference genes for normalization during cold acclimation (Fernández et al. 2010) and in vitro adventitious rooting in *E. globulus* (de Almeida et al. 2010) as well as in *Eucalyptus* leaves exposed to biotic (*Puccinia psidii*) and abiotic (acibenzolar-S-methyl) stresses (Boava et al. 2010). Very recently, Pasquali and colleagues (de Oliveira et al. 2012) selected housekeeping genes based on the 50 genes exhibiting the least variation in microarrays of *E. grandis* leaves and xylem and of *E. globulus* xylem.

In this study, we report the selection and evaluation of 21 candidate reference genes to identify the most suitable internal control gene or gene combinations for normalization of qRT-PCR data from various species of *Eucalyptus* organs/tissues (xylem, cambium, stem, root and leaves), at different developmental stages, and exposed to environmental stimuli such as drought, cold, bending or nitrogen fertilization. We analyze the entire gene expression dataset by three different statistical

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*, 53 (12).

algorithms to compare the ranking of the potential reference genes. In addition, to illustrate the utility of the new reference genes, we provide a detailed expression analysis of two MYB transcription factors in a set of Eucalyptus tissues and organs.

## Results and discussion

### Selection of candidate reference genes

To identify candidate reference genes in Eucalyptus over a broad range of developmental and environmental conditions, we first surveyed the *E. grandis* Phytozome database (<http://www.phytozome.net/eucalyptus.php>) for putative orthologs of the top 27 reference genes in Arabidopsis as defined by Scheible and colleagues (Czechowski et al. 2005). Using the best hits in Phytozome, we selected 16 *E. grandis* genes with high orthology probabilities (E-value < 1.00E-60) (Table 1). Despite their weak scores, we also included genes encoding Ubiquitin10 and F-box protein because they are optimal reference genes in Arabidopsis and in cotton (Artico et al. 2010; Czechowski et al. 2005). Also, we added three genes commonly considered to be housekeeping genes in Eucalyptus: GAPDH, IDH, and TUB. In total, 21 candidate reference genes were evaluated for their expression levels by making use of microfluidic dynamic array technology, which is appropriate and reliable to perform high-throughput gene expression measurements by qRT-PCR (Spurgeon et al. 2008). The cycle threshold (Ct) values (medians and ranges) for each of the 21 candidate genes in 90 samples (including biological triplicates, as described in detail in the Methods section) are shown in Figure 1. The transcripts encoding GAPDH and Actin2 (ACT2) were the most abundant whereas those encoding F-box were the least abundant. The variability of Ct values in the 90 samples examined was widest for the F-box transcript, whereas the ACT2 transcripts had the least variable Ct values (Figure 1).

### Expression profiling of the candidate reference genes

We performed a preliminary analysis of the expression stability of these 21 genes, and decided to discard five of them (F-box, UBQ10, UBQ14, TUB and YLS8) because their expression varied substantially in different tissues and conditions (data not shown). To obtain an accurate view of the expression profiles of the 16 remaining candidate genes, we organized the 90 samples into five subsets based on their origin: the ‘Organs and Tissues’ (OT) subset included thirteen organs and/or tissues (38 samples); the ‘Leaves’ (L) subset comprised young and mature leaves from control and cold-treated plants (nine samples); the ‘Xylem All’ (XA) subset included all xylem samples from trees grown in various environmental conditions and from several species or hybrids (33 samples); the ‘Xylem–Cambium’ (XC) subset comprised samples from juvenile and mature xylem and cambium-enriched tissues (12 samples), and the ‘Xylem Environmental stimuli’ (XES) subset comprised samples from xylem tissues of trees under drought, mechanical stress and/or with different nitrogen fertilization

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions. *Plant and Cell Physiology*, 53 (12).



conditions (27 samples). Table S2 reports the relative Ct values for ten genes (ACT, EF-1 $\alpha$ , GAPDH, IDH, PP2A-1, PP2A-3, PPR2, PTB, SAND, and UBC2) in the five subsets indicated above. The mean Ct values (average of three independent biological samples) were in the range of 5.9–19.0, typical of moderately to highly expressed genes as assayed by using dynamic array technology.

### Expression stability analyses

The most widely used statistical algorithm to analyze the stability of gene expression is geNorm (Vandesompele et al. 2002), which calculates the expression stability value (M) of a gene defined as the average pairwise variation of that gene relative to all other potential reference genes in a panel of cDNA samples. Genes with the lowest M value have the most stable expression and useful reference genes should have M values < 0.5 (Vandesompele et al. 2002). We used geNorm to determine the M values of our 16 candidate genes by stepwise exclusion of the least stable control gene in the five sample subset panels described above and in a sixth panel composed of all the samples (AS, 90 samples; Figure 2). The subset including various organs and tissues showed the greatest variation in expression for all the genes tested (with M values of 0.85–0.47), when compared to the subsets including samples from only one organ (e.g. subset L had M values of 0.57–0.15) or from a single tissue (e.g. subset XA had M values of 0.65–0.28). Surprisingly, in the panel containing all samples (AS), gene expression was more variable (M values of 0.77–0.34) than in the OT panel. This may be due to the relatively high proportion of xylem samples (33/90). By applying a cut-off M value of 0.5, we found that in the OT panel only two genes, EF1 $\alpha$  and GAPDH, qualified as stable and even those had relatively high M values (0.47; Table 2). By contrast, many more candidate genes qualified as stable in the other panels: five in AS, nine in XA, ten in XES, 13 in L, and 14 in XC. The identity of the most stably expressed genes in a panel varied depending on the panel considered (Table 2). For instance, EF-1 $\alpha$  was one of the most stably expressed genes in the OT and XC panels, whereas PP2A-1 and IDH were the most stable in the L panel; PP2A-3 and SAND the most stably expressed genes in the XA, XES and, remarkably, in the AS panel in which all the samples were grouped together. In all panels, PPR2 was among the least stably expressed genes (ranked consistently 14–16). *Expressed1* ranked poorly in most panels except in L, where it ranked 3 with an M value of 0.2. The commonly used reference gene GAPDH ranked well in the OT panel (ranked top but with a relatively high M value of 0.47) and in XC (ranked 5, M value of 0.2); it was also stably expressed in L (ranked 10, M value of 0.35) and in XA (ranked 10, M value of 0.48). IDH ranked top in L (M value of 0.15) but less well in OT (ranked 4, M value of 0.54) and even worse in XA (rank 12, M value of 0.52).

The geNorm program also determines the ideal and/or minimal number of reference genes that would be required to calculate an accurate normalization factor, as the geometric mean of their relative quantities. A pairwise variation  $V_n/V_{n+1}$  of 0.15 provides a cut-off value below which the inclusion of an additional control gene is not necessary for reliable normalization. This pairwise analysis revealed that the ideal number of reference genes varied among the different data subsets. Use of the

two best reference genes was sufficient ( $V2/V3 < 0.15$ ) when the subsets were composed of only one type of organ or tissue, and also for the panel containing all samples (AS; Figure 3). When the panel was composed of less homogeneous samples, such as the OT set, at least three reference genes ( $V2/V3 = 0.17$ ,  $V3/V4 = 0.13$ ) was necessary for accurate normalization (Figure 3).

We also analyzed our data with NormFinder, a program taking into account the intra- and inter- group variations for normalization factor calculation and whose results are not affected by occasional co-regulated genes ((Andersen et al. 2004); Table 3). Remarkably, in this analysis PP2A-3 was the first or the second the best-ranked gene in each panel, although over all 16 genes there were some differences in the rankings in the different panels. Similarly, PPR2 was consistently the worst ranked gene in all panels. NormFinder also defines the best combination of two reference genes for each experimental set, which provides a corresponding lower stability value M (Table 3). A comparison of the outcomes using geNorm or NormFinder revealed common features but also significant differences (Table 3 versus Table 2). The four most stably expressed genes in the AS panel as ranked by geNorm were PP2A-3, SAND, UPL7 and PTB whereas when the same dataset was ranked by NormFinder the four most stably expressed genes were PP2A-3, CACS, GAPDH and SAND. Consistently, PP2A-3 was ranked top and SAND was found among the top four reference genes by both methods. The other genes ranked in the top four by one method (i.e. UPL7 and PTB for geNorm and CACS and GAPDH for NormFinder) were not well ranked by the other method.

As there were significant differences in the results obtained with geNorm and NormFinder, we employed a third analytical method called  $\Delta Ct$  (Silver et al. 2006). This method compares the relative expression of 'pairs of genes' within each sample to confidently identify reference gene. If the  $\Delta Ct$  value between two genes remains constant when analyzed in different samples, it means either both genes are stably expressed or are co-regulated. On this basis, the stability of expression of the candidate genes was ranked according to the reproducibility of the gene expression difference among all tested samples (Silver et al. 2006). Using this method, again, the ranking results differed among the different data panels, but PP2A-3 and SAND were consistently well ranked in most, whereas PPR2, Actin2 and Expressed1 were the least stably expressed genes (Table 4). Generally, these results were closer to those obtained with NormFinder than with geNorm. This was unexpected because the  $\Delta Ct$  method is similar to geNorm (Vandesompele et al. 2002), in that it compares pairs of genes (Silver et al. 2006).

Because these three independent methods produced similar results in some cases but discrepancies in others, we decided to make a comprehensive ranking by integrating the ranking obtained with the three methods. A simple algorithm was designed based on the geometric mean (GM) of each gene weight across the three methods (as detailed in the Methods section). The gene with the lowest GM was considered to be the most stable gene. Table 5 displays the results obtained for the six datasets and Table S3 contains the results for eleven additional datasets including stems, roots and cold stress.

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).



## Comparison of the four statistical methods by Principal Component Analysis

To compare visually the four methods, we used principal component analysis (PCA), which takes data points in a high dimensional space and defines new axes (components) that cut across that space such that the first component captures as much as the variance in the data as possible, the second component (orthogonal to the first) captures as much of the remaining variance as possible, and so on. In this study, the data points are the stability values of the sixteen genes in a 4-D space, with the coordinate in each dimension being the stability value of each gene calculated by each of the four methods. Importantly, because the components cut across this 4D space, each component has contributions from all sixteen genes, which allows the original variables (here the four methods) to be represented in new axes defined by the new components. This leads to a graphical representation of the correlations between variables (the four algorithms used to analyze gene expression stability), and also of the similarity between data points (the gene expression stability values). Applied to the expression stability values of the AS, OT, XA and L subsets, the first principal component (PC1) explained 88.0%, 82.1%, 91.0% and 90.7% of the observed variance, respectively, whereas the second principal component (PC2) explained only 8.4%, 15.9%, 5.6% and 7.1%, respectively. Thus, in all cases, more than 95% of the total information was explained by the first plane, defined by the first two components. The projection of the four 'gene stability' variables onto this plane was performed for each of four sample panels (Figure 4A). In each case, the four methods were grouped together on the extreme right of the PC1 axis (Figure 4A), which explains the major part of the variance. The distance between them along the PC2 axis was variable depending on the panel considered. In the most heterogeneous panel, OT, geNorm was somewhat distant (on both PC1 and PC2 axes) from NormFinder and Ct, which were close to each other (Figure 4A). For the XA panel and, to a lesser extent, for the L panel, which are the most homogeneous subsets of samples, all four methods grouped spatially (Figure 4A). As expected, the position of the values obtained by the comprehensive ranking method reflected the fact that they were the geometrical means (GM) of the three other methods. Because PCA components consist of contributions of all sixteen genes, we plotted the score of the stability of the expression of the 16 genes for the two PCA components and this, for each panel of samples (Figure 4B). Nearly all the genes projected on the left part of the central vertical axis were within the top six genes as ranked by at least one of the four methods. PP2A3 was always well ranked if not the top-ranked gene by the four methods, whatever the subset considered. Thus, it appears to be the best reference gene in these samples. By contrast, PPR2 was located in all cases on the extreme right of the plots; it is clearly unsuitable as a reference gene. In the AS subset, PP2A-3 (which ranked top by the four statistical methods) appeared as a singleton on the extreme left of the x axis, ahead of the second most stable gene, SAND (which was also well-ranked by the four methods). To the right of these two genes on the x axis were several genes with a very similar score for the first PCA component but with a diverse score for the second component on the y axis. Among them, only

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*, 53 (12).

PP2A1, located on the horizontal axis, was well ranked by all methods. The main difference between the remaining genes was the method used for ranking. For instance CACS and UPL7, which were located at the two extremes of the y axis, were well ranked by only one method (NormFinder and geNorm, respectively).

In the very heterogeneous OT subset, PP2A1, EF1 $\alpha$  and PP2A3 were the three best genes. PP2A1, EF1 $\alpha$  were well ranked by all methods whereas PP2A3 was ranked first by NormFinder and Ct and only eighth by geNorm (Figure 4B). The spatial distribution of the values for the genes in the L subset was very different from the other three panels with a clear distinction between the group on the left containing the best reference genes versus the group on the right, which contained the worst ones (Figure 4B). There were equal numbers of genes in the two groups and they were distributed evenly along both the x and y axes, in contrast to those in the XA subset, which were distributed more closely on both axes (Figure 4B). This is consistent with the fact that the three statistical methods were closely grouped for the XA subset (Figure 4A), and the relative density along the y axis (second PCA component) reflects the fact that there are fewer differences between the stability rankings of the L subset genes by the three methods when compared to the other subsets.

One conclusion from this PCA is that the ranking result depends not only on the nature of the samples and the degree of heterogeneity of the gene panel considered but also on the analysis method applied to a particular subset of samples: the different rankings obtained by using different methods depend on which sample panel is being considered (Figure 4A). Interestingly, there is a very good agreement between the ranking of the three best genes obtained by the comprehensive ranking method (reflecting the geometric mean of the ranking of the three other methods; Table 5) and those identified by PCA analysis (Figure 4B). We used this comprehensive ranking method, therefore, for 11 further subsets of samples from other tissues, developmental stages and environmental conditions (Table S3). The expression of PP2A-3 was consistently stable in all eleven data sets. In roots and stems, UBC2 and CACS were stably expressed whereas PPR2, Actin2 and Helicase were, as already noticed in the other sample panels, the least stably expressed genes. Taking together the two comprehensive ranking tables (Table 5 and Table S3), PP2A-3, PP2A-1, EF-1a, SAND, IDH, CACS, UBC2, and PTB exhibited relatively stable expression, whereas PPR2, Actin2, Expressed1 and Helicase were the least stable.

### **Expression profiling of EgMYB1 and EgMYB2**

To evaluate the effect of the choice of reference gene on the expression profiling of other genes, we analysed the expression of two genes, EgMYB1 and EgMYB2, (which are master regulators of xylem formation in Eucalyptus; (Goicoechea et al. 2005; Legay et al. 2010)), in several tissues and organs by using the various combinations of top reference genes indicated by each of the four methods: EF-1 $\alpha$ , GAPDH and PP2A-1 as indicated by geNorm; PP2A-3 and PP2A-1 as indicated by NormFinder; PP2A-3, PP2A-1 and SAND as indicated by Ct, and PP2A-3, PP2A-1 and EF-1 $\alpha$  as indicated by the

comprehensive ranking method. We also used the overall best (PP2A3) and the worst (PPR2) ranked genes (Figure 5). relative expression data we obtained by using the various reference gene combinations were consistent for both EgMYB1 (Figure 5A) and EgMYB2 (Figure 5B) expression in all tissues, with the exception of the data obtained by using PPR2 as a reference gene. This was particularly obvious for the expression of EgMYB1 and EgMYB2 in secondary stems, which were largely overestimated by normalizing to PPR2 expression when comparing with the data obtained with the other reference genes. Using PPR2 also led to underestimation of EgMYB1 expression in the cambium samples as well as to an overestimation of EgMYB2 expression in primary stems. PPR2 was consistently poorly ranked by the four statistical methods independently of the experimental data set considered, thus it should not be used as a reference gene for expression studies in Eucalyptus. Conversely, the use of PP2A3 (best ranked by at least three algorithms) produced relative expression data very similar to those obtained by using combinations of two or three reference genes from all sample subsets. Speleman and colleagues (Vandesompele et al. 2002) recommended the use of three reference genes for pairwise analysis, however, we found that this did not improve normalization, when compared to the use of one or two carefully selected reference genes even in the case of a heterogeneous samples set such as the OT subset in this study (Figure 3).

By making use of a high-throughput gene expression platform based on microfluidic dynamic arrays, this in-depth study has identified and validated the optimal reference genes for gene expression analysis of a large panel of 90 samples corresponding to 30 distinct organs, tissues or developmental stages and/or environmental conditions (nitrogen fertilization, mechanical, cold and drought stress) in five different species of Eucalyptus. The stability of expression of sixteen Eucalyptus genes was evaluated in seventeen experimental datasets and the data were analyzed by using three statistical algorithms: geNorm (Vandesompele et al. 2002), NormFinder (Andersen et al. 2004) and [Ct (Silver et al. 2006). As illustrated by principal component analysis, the rankings of gene expression stability depended on the statistical method, on the nature of the biological samples and the degree of heterogeneity of the panel considered; also, the differences between the methods varied with the sample dataset considered. There was very good agreement between the ranking obtained by comprehensive integration ranking, which reflects the geometric mean of the ranking outcomes of the three statistical algorithms, and those identified by PCA analysis. This integrative ranking applied to the seventeen experimental panels enabled us to identify among the 16 candidate genes, the most stable genes: PP2A-3, PP2A-1, EF-1a, SAND, IDH, CACS, UBC2, and PTB as well as the least stable genes: PPR2, Actin2, Expressed1 and Helicase. The use of PPR2 as an internal reference gene was shown to be unsatisfactory because it led to erroneous patterns of expression of two Eucalyptus master genes EgMYB1 and EgMYB2. Interestingly, EF-1a and IDH have been used traditionally as internal controls for studies of Eucalyptus gene expression (Legay et al. 2010; Paux et al. 2005; Paux et al. 2004). Previously, Fett-Netto and colleagues using geNorm identified IDH and SAND as the most stable genes during in vitro adventitious rooting in Eucalyptus (de Almeida et al. 2010). Among the

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions. *Plant and Cell Physiology*, 53 (12).

stable genes, PP2A-3 was the most stable being among the four best-ranked genes in all the 17 sample panels. We therefore tested PP2A-3 alone as a reference gene in the most heterogeneous sample panel “organs/tissues” for which geNorm recommended at least three reference genes. It is worth noting that for this panel of samples, PP2A-3 was not in the first three top-ranked genes for geNorm whereas it was ranked first by NormFinder, Ct and by the comprehensive ranking. Interestingly, the use of PP2A-3 as the only reference gene gave consistent pattern of expression for both EgMYB1 and EgMYB2. Thus in our hands, a wellchosen single reference gene is able to provide an accurate and reliable expression pattern even in a complex and heterogeneous samples panel. This finding is interesting since the systematic use of multiple references genes is costly and laborious and sometimes even impossible as in the case of limited available amount of RNA such as samples harvested by laser micro-dissection. In conclusion, this study provided useful clues for accurate and reliable normalization of gene expression in a wide panel of organs, tissues and conditions for Eucalyptus, a tree of great economic importance.

## Material and methods

### Plant material

Shoot tips and vascular tissue samples (cambium-enriched fraction, secondary phloem and differentiating xylem) were harvested from seven-year-old Eucalyptus Gundal hybrids (gunnii x darympleana, genotype 850645) grown in south-west France (Longages) by the Institut Technologique FCBA (Forêt Cellulose Bois-construction Ameublement). Cambiumenriched fractions were also collected from for 25-year-old Gundal hybrids (genotype 821290). Vascular tissue sampling was performed as previously described (Paux et al. 2004). Juvenile and mature xylem samples (kindly provided by the RAIZ Institute of Forest and Paper Research, Portugal) were harvested in Herdade do Zambujal from, respectively, fourand 10-year-old E. globulus trees (genotype VC9). Tension and opposite xylem samples were collected at Quinta do Furadouro (Portugal) from two-year-old trees of three distinct genotypes of E. globulus (GM52, BB3 and MB43 kindly provided by Altri-Florestal, Portugal) after three weeks of bending (at 45°). Fruit capsules were harvested from E. globulus, genotype C33 (provided by Altri-Florestal, Portugal). Drought-stressed xylem samples were collected from non-irrigated 16-month-old trees (a clonal plantation of E.urophylla x E.grandis, hybrid 1850) after four months without rainfall during the dry season in Yanika, Republic of Congo (Villar et al. 2011). Control xylem samples were taken from irrigated trees. These samples were kindly provided by the CIRAD Forêt (France). For plants grown under conditions of nitrogen fertilization, three-month-old rooted cuttings of a Eucalyptus hybrid (E. urophylla x E.grandis, clone IPB2-H15, kindly provided by International Paper do Brasil) grown in greenhouse conditions were submitted for 30 days to three different fertilization treatments: 7.5 mM NH<sub>4</sub>NO<sub>3</sub> for limiting N (N-), 15mM NH<sub>4</sub>NO<sub>3</sub> for optimal

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions. *Plant and Cell Physiology*, 53 (12).

N (CT) and 30 mM NH<sub>4</sub>NO<sub>3</sub> for luxuriant N (N+) (Camargo et al. 2012). Each xylem sample consisted of a pool of 20 debarked stems. For cold experiments, one-year-old *E. globulus* trees (genotype GM258 provided by Altri-Florestal, Portugal) were submitted to cold (7°C) for 16h in the dark. In parallel, control plants were maintained for 16h in the dark in greenhouse conditions. Expanding leaves, fully expanded leaves, primary stems, secondary stems, and roots were harvested for each condition. Each sample consisted of pooled tissues from two trees. *E. grandis* calli were obtained from in vitro cultures. For each tissue or organ, three independent biological repetitions were collected, except the fruit capsules samples for which we had only two biological repetitions. Following harvesting all samples were immediately frozen in liquid nitrogen and stored at -80°C until extraction.

### **Total RNA extraction, cDNA synthesis and quality controls**

All the procedures used for the qRT-PCR, from the RNA extraction to the calculation of transcript abundance, were performed as described by Udvardi et al (Udvardi et al. 2008). and Derveaux et al.(Derveaux et al. 2010). Total RNAs were extracted from 1–5 g of frozen material as described by (Southerton et al. 1998). RNA concentration and purity were determined by using a NanoDrop spectrophotometer ND-1000 (Thermo Scientific, Wilmington, USA). RNA samples were then treated with Turbo DNA-free™ kit (Ambion, TX, USA). The absence of remaining genomic DNA was confirmed by PCR using ubiquitin primers (EgUbi1\_F: GCGGCTTTTAAGTCTCTTGC GAA, EgUbi1\_R: TTCGAAGCATAGCTTCGCCATATG). The integrity of RNAs was assessed by using the Agilent 2100 Bioanalyzer, and only samples with a RNA integrity number over 7 (RIN>7) were retained for reverse transcription performed by using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, CA, USA), according to the manufacturer's instructions and using up to 1µg of total RNA. The quality of each cDNA was assessed by using two pairs of primers located approximately 1.2 kb apart from each other in the 5' and in the 3' regions of the genes encoding IDH [5' end primers; (F\_AATCGACCTGCTTCGACCCTTC; R\_TCGACCTTGATCTTCTCGAAACCC)] [3' end primers; (F\_TGCTGTGGCAGCTGAACTCAAG; R\_ATGTTGTCCGCCAGTCACCTAC)] and PP2A3 [5' end primers; (F\_CGGAAGAAGTGGGTGTGTTT; R\_CACAGAGGGTCTCCAATGGT)] [3' end primers; (F\_CAGCGGCAAACA AACTTGAAGCG; R\_ATTATGTGCTGCATTGCCAGTC)]. The majority of our cDNA samples showed a Ct value of the 5'-end pair that did not exceed that of the 3'-end pair by more than one cycle number for both genes. These good quality cDNA samples were diluted five-fold and stored at -20°C until used for qPCR.

### **PCR primer design**

The sequences of the putative *E. grandis* orthologs of *A. thaliana* genes (best hits) were identified by BLASTP from the Phytozome database (<http://www.phytozome.net/eucalyptus.php>)

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*, 53 (12).



based on an E-value of less than 1.00E-60. Only F-box and Ubiquitin10 did not meet this criterion, as shown in Table 1. Primer pairs were designed using the software QuantPrime (qPCR primer design tool: <http://www.quantprime.de/> (Arvidsson et al. 2008) benefiting from exon–intron border, splice variant information of the *E. grandis* genome annotations (Phytozome version 6). Primers were preferentially selected to be as close to 3' end of the transcripts as possible and their sequences are shown in Table 1. Oligonucleotides were synthesized by Sigma Life Science (France).

### High-throughput quantitative qRT-PCR

High-throughput qRT-PCR was performed by the Genotoul service in Toulouse (<http://genomique.genotoul.fr/>) using the BioMark® 96:96 Dynamic Array integrated fluidic circuits (Fluidigm Corporation, San Francisco, CA, USA) according to the manufacturer's protocol. Briefly, each cDNA sample was pre-amplified with a pool of primers specific to the target genes by using the following program: 95°C for 10 minutes, then 14 cycles of 95°C for 15 s and 60°C for 4 min. The pre-amplified products were diluted 1:5 in 10mM Tris–HCl, pH 8; 0.1mM EDTA and analysed by qRT-PCR using the following conditions: 95°C for 10 min, then 35 cycles of 95°C for 15 s and 60°C for 30 s. The specificity of the PCR products was confirmed by analysing melting curves, following the final PCR cycle. Only primers that produced a linear amplification and qPCR products with a single-peak melting curves were used for further analysis. The efficiency of each pair of primers was determined by plotting the Ct values obtained for serial dilutions of a mixture of all cDNAs and the equation  $\text{Efficiency} = 10^{(-1/\text{slope})} - 1$ . Primer efficiencies for 16 candidate reference genes were higher than 95% and lower than 110% (Table S1) except for EF-1a (89%) and SAND (115%). For all primer pairs, amplicon sizes were around 70 bp, and annealing temperatures around 63 °C as indicated in Table S1. For all reference gene candidates, the amplification plots and melting curves are presented in Figure S1.

### Statistical analyses

To analyze the stability of the candidate reference genes, we used three different tools: geNorm, NormFinder and Ct. The geNorm software v3.4 is a Visual Basic Application tool for Microsoft Excel that relies on the principle that the expression ratio of two ideal reference genes should be constant in samples from different experimental conditions or cell types (Vandesompele et al. 2002). The NormFinder program uses a model-based approach for identifying the optimal normalization genes among a set of candidates (Andersen et al. 2004). It is rooted in a mathematical model of gene expression that enables estimation both the overall variation candidate genes and the variation between samples subgroups. For technical reasons, the sample set should contain a minimum of eight samples per group and at least three, but ideally five or more candidate genes. In this study, the subgroups were defined either by organ, tissue type or experimental condition. The Ct method uses a similar strategy to geNorm by comparing relative expression of 'pairs of genes' within each

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).



sample to identify reference genes (Silver et al. 2006). If the Ct value between two genes remains constant when analyzed in different samples, it means either both genes are stably expressed or they are co-regulated. Introduction of a third, fourth or more genes into comparisons will provide information on which pairs show less variability and hence which gene(s) are stably expressed among samples tested. Finally, we made our own comprehensive ranking of the best reference genes by integrating the ranking obtained by the three methods (geNorm, NormFinder and Ct). For this we developed a simple algorithm to present an overall ranking of the best reference genes. Briefly, we first made a linear transformation of the stability values by assigning a series of weight values from 1 to 10 to each reference gene according to the ranking obtained by each algorithm from the most stable gene to the least stable gene; then we calculated the geometric mean (GM) of each gene weight across the three methods and finally re-ranked these reference genes. The gene with the lowest GM was considered to be the most stably expressed reference gene.

### **Determination of EgMYB1 and EgMYB2 expression profiles**

We investigated the expression of Eucalyptus transcription factor EgMYB1 (Eucgr.G01774 ; F\_ACCATGACGAGCCCACCATTTTC, R\_ TCAGGTCAGGACACCTTTCTCG) and EgMYB2 (Eucgr.G03385 ; F\_AGGCATTGCACCGGTCAGTATG, R\_ TTCTCCTCGGTGGTGGTTGTG) in 90 samples as described under the subheading 'Plant materials' above. The primer design and the qRT-PCR conditions were carried out following the same parameters used for the analysis of reference genes. The relative expression profiles analysis were obtained through the E-[[CT method (Pfaffl 2001) using efficiency of each MYB primers and each reference genes. The 'leaves' sample was adopted as the calibrator reference tissue.

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).

## Funding

This work was supported by grants from the Agence Nationale pour la Recherche (ANR) (Project Tree For Joules ANR-2010-KBBE-007-01; Fundação para a Ciência e Tecnologia (FCT), P-KBBE/AGR\_GPL/0001/2010), the Centre National pour la Recherche Scientifique (CNRS), the University Paul Sabatier Toulouse III (UPS), the FCT project microEGo (PTDC/AGR-GPL/098179/2008) and (PEst-OE/EQB/LA0004/2011) and the INTEREG IVB SudoE project Interbio. This work is part of the Laboratoire d'Excellence (LABEX) project entitled TULIP (ANR - 10-LABX-41). HY, ELOC, VC were supported by PhD grants from the China Scholarship Council (to HY), the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, 2008/53520-3; to ELOC) and the FCT (SFRH/BD/72982/2010; to VC). MS had a postdoctoral fellowship 'Beatriu de Pinós' from the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya. JPP acknowledges a research contract from the Ciència 2008 program (FCT).

## Acknowledgements

We are grateful to J.M. Gion and E. Villar (CIRAD, FR), F. Melun and L. Harvengt (FCBA, France), C. Araujo and L. Neves (AltriFlorestal, Portugal) and C. Marques (RAIZ, Portugal) for kindly providing and/or allowing collection of Eucalyptus organ and tissue samples, C. Graça (IICT) and M.N. Saidi (LRSV) for help with sample collection and RNA extraction. We also warmly acknowledge the advice of C. Brière (LRSV) for the statistical analyses of our results, and S. Arvidsson for his kind help to link the *E. grandis* genome to the Quantprime software. Thanks to the Plateforme Génomique Génopole Toulouse/Midi-Pyrénées (Genotoul) for advice and technical assistance with high-throughput Biomark Fluidigm qRT-PCR amplifications. Finally, the authors acknowledge the Eucagene consortium led by A. Myburg and the Department of the Energy (USA) for making available the *E. grandis* genome.

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).

## References

Andersen, C.L., Jensen, J.L. and Orntoft, 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.

Artico, S., Nardeli, S.M., Brilhante, O., Grossi-de-Sa, M.F. and Alves-Ferreira, M. (2010) Identification and evaluation of new reference genes in *Gossypium hirsutum* for accurate normalization of real-time quantitative RT-PCR data. *BMC Plant Biol* 10: 49.

Arvidsson, S., Kwasniewski, M., Riano-Pachon, D.M. and Mueller-Roeber, B. (2008) QuantPrime-a flexible tool for reliable high-throughput primer design for quantitative PCR. *BMC Bioinformatics* 9: 465.

Boava, L.P., Laia, M.L., Jacob, T.R., Dabbas, K.M., Goncalves, J.F., Ferro, J.A., Ferro, M.I. and Furtado, E.L. (2010) Selection of endogenous genes for gene expression studies in *Eucalyptus* under biotic (*Puccinia psidii*) and abiotic (acibenzolar-S-methyl) stresses using RT-qPCR. *BMC Res Notes* 3: 43.

Boudet, A.M., Kajita, S., Grima-Pettenati, J. and Goffner, D. (2003) Lignins and lignocellulosics: a better control of synthesis for new and improved uses. *Trends Plant Sci* 8: 576-581.

Brunner, A.M., Yakovlev, I.A. and Strauss, S.H. (2004) Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biol* 4: 14.

Bustin, S.A., Benes, V., Garson, J.A., Hellems, J., Huggett, J., et al. (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55: 611-622.

Camargo, E., Costa, L., Soler, M., Salazar, M., Lepikson, J., Gonçalves, D., Marques, W., Carazzolle, M., Martinez, Y., Grima-Pettenati, J. and Pereira, G. (2012) Effects of nitrogen fertilization on global xylem transcript profiling of *Eucalyptus urophylla* x *grandis* evaluated by RNA-seq technology. *BMC Proceedings* 5(Suppl 7): P106.

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K. and Scheible, W.R. (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol* 139: 5-17.

de Almeida, M.R., Ruedell, C.M., Ricachenevsky, F.K., Sperotto, R.A., Pasquali, G. and Fett-Neto, A.G. (2010) Reference gene selection for quantitative reverse transcription-polymerase chain reaction normalization during in vitro adventitious rooting in *Eucalyptus globulus* Labill. *BMC Mol Biol* 11: 73.

de Oliveira, L.A., Breton, M.C., Bastolla, F.M., Camargo Sda, S., Margis, R., Frazzon, J. and Pasquali, G. (2012) Reference genes for the normalization of gene expression in eucalyptus species. *Plant Cell Physiol* 53: 405-422.

Demura, T. and Fukuda, H. (2007) Transcriptional regulation in wood formation. *Trends Plant Sci* 12: 64-70.

Derveaux, S., Vandesompele, J. and Hellemans, J. (2010) How to do successful gene expression analysis using real-time PCR. *Methods* 50: 227-230.

Exposito-Rodriguez, M., Borges, A.A., Borges-Perez, A. and Perez, J.A. (2008) Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. *BMC Plant Biol* 8: 131.

Fernández, M., Villarroel, C., Balbontín, C. and S., V. (2010) Validation of reference genes for real-time qRT-PCR normalization during cold acclimation in *Eucalyptus globulus*. *Trees - Structure and Function* 24: 1109-1116.

Foucart, C., Paux, E., Ladouce, N., San-Clemente, H., Grima-Pettenati, J. and Sivadon, P. (2006) Transcript profiling of a xylem vs phloem cDNA subtractive library identifies new genes expressed during xylogenesis in *Eucalyptus*. *New Phytol* 170: 739-752.

Gallo de Carvalho, M.C., Caldas, D.G., Carneiro, R.T., Moon, D.H., Salvatierra, G.R., Franceschini, L.M., de Andrade, A., Celedon, P.A., Oda, S. and Labate, C.A. (2008) SAGE transcript profiling of the juvenile cambial region of *Eucalyptus grandis*. *Tree Physiol* 28: 905-919.

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).

Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A., Lapierre, C., Pollet, B., Verhaegen, D., Chaubet-Gigot, N. and Grima-Pettenati, J. (2005) EgMYB2, a new transcriptional activator from Eucalyptus xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J* 43: 553-567.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N. and Rokhsar, D.S. (2011) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res* 40: D1178-1186.

Guenin, S., Mauriat, M., Pelloux, J., Van Wuytswinkel, O., Bellini, C. and Gutierrez, L. (2009) Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *J Exp Bot* 60: 487-493.

Guo, G., Huss, M., Tong, G.Q., Wang, C., Li Sun, L., Clarke, N.D. and Robson, P. (2010) Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Dev Cell* 18: 675-685.

Gutierrez, L., Mauriat, M., Guenin, S., Pelloux, J., Lefebvre, J.F., Louvet, R., Rusterucci, C., Moritz, T., Guerineau, F., Bellini, C. and Van Wuytswinkel, O. (2008a) The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. *Plant Biotechnol J* 6: 609-618.

Gutierrez, L., Mauriat, M., Pelloux, J., Bellini, C. and Van Wuytswinkel, O. (2008b) Towards a systematic validation of references in real-time rt-PCR. *Plant Cell* 20: 1734-1735.

Legay, S., Sivadon, P., Blervacq, A.S., Pavy, N., Baghdady, A., et al. (2010) EgMYB1, an R2R3 MYB transcription factor from eucalyptus negatively regulates secondary cell wall formation in *Arabidopsis* and poplar. *New Phytologist* 188: 774-786.

Myburg, A.A., Potts, B.M., Marques, C.M.P., Kirst, M., Gion, J.M., Grattapaglia, D. and Grima-Pettenati, J. (2007) *Eucalyptus*. In *Genome Mapping & Molecular Breeding in Plants*. Edited by Kole, C.R. Springer, Heidelberg.

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).

Navarro, M., Ayax, C., Martinez, Y., Laur, J., El Kayal, W., Marque, C. and Teulieres, C. (2011) Two EguCBF1 genes overexpressed in Eucalyptus display a different impact on stress tolerance and plant development. *Plant Biotechnology Journal* 9: 50-63.

Paiva, J.A.P., Prat, E., Vautrin, S., Santos, M.D., San-Clemente, H., et al. (2011) Advancing Eucalyptus genomics: identification and sequencing of lignin biosynthesis genes from deep-coverage BAC libraries. *Bmc Genomics* 12: -.

Pang, Z.P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D.R., Yang, T.Q., Citri, A., Sebastiano, V., Marro, S., Sudhof, T.C. and Wernig, M. (2011) Induction of human neuronal cells by defined transcription factors. *Nature* 476: 220-223.

Paux, E., Carocha, V., Marques, C., Mendes de Sousa, A., Borralho, N., Sivadon, P. and Grima-Pettenati, J. (2005) Transcript profiling of Eucalyptus xylem genes during tension wood formation. *New Phytol* 167: 89-100.

Paux, E., Tamasloukht, M., Ladouce, N., Sivadon, P. and Grima-Pettenati, J. (2004) Identification of genes preferentially expressed during wood formation in Eucalyptus. *Plant Mol Biol* 55: 263-280.

Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29: e45.

Plomion, C., Leprovost, G. and Stokes, A. (2001) Wood formation in trees. *Plant Physiol* 127: 1513-1523.

Reid, K.E., Olsson, N., Schlosser, J., Peng, F. and Lund, S.T. (2006) An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development. *BMC Plant Biol* 6: 27.

Rengel, D., San Clemente, H., Servant, F., Ladouce, N., Paux, E., Wincker, P., Couloux, A., Sivadon, P. and Grima-Pettenati, J. (2009) A new genomic resource dedicated to wood formation in Eucalyptus. *BMC Plant Biol* 9: 36.

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).



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.

Southerton, S.G., Marshall, H., Mouradov, A. and Teasdale, R.D. (1998) Eucalypt MADS-box genes expressed in developing flowers. *Plant Physiol* 118: 365-372.

Spurgeon, S.L., Jones, R.C. and Ramakrishnan, R. (2008) High throughput gene expression measurement with real time PCR in a microfluidic dynamic array. *PLoS One* 3: e1662.

Udvardi, M.K., Czechowski, T. and Scheible, W.R. (2008) Eleven golden rules of quantitative RT-PCR. *Plant Cell* 20: 1736-1737.

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: RESEARCH0034.

Villar, E., Klopp, C., Noirot, C., Novaes, E., Kirst, M., Plomion, C. and Gion, J.M. (2011) RNA-Seq reveals genotype-specific molecular responses to water deficit in eucalyptus. *Bmc Genomics* 12: 538.

Xu, M., Zhang, B., Su, X., Zhang, S. and Huang, M. (2010) Reference gene selection for quantitative real-time polymerase chain reaction in *Populus*. *Anal Biochem* 408: 337-339.

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).

Table 1. Twenty one potential reference genes and their primer sequences

*Gene abbreviation	Gene model	A. thaliana ortholog	A. thaliana annotation	Blast E-value	Identity (%)	**Protein size	***Blast alignment	Primer fw_start	Primer sequence (Forward/reverse 5'-3')
EF-1 $\alpha$	Eucgr.B02473	AT5G60390	Elongation factor 1-alpha	0	95.9%	449	434/449	1530	ATCGCTCAGACTGTGGCTGTG/ ATCGCTCAGACTGTGGCTGTG
****GAPDH	Eucgr.H04673	AT1G13440	Glyceraldehyde-3-phosphate dehydrogenase C2	6.50E-165	92.8%	433	320/338	688	TTGTGGGTGTGAACGAGAAGGAG/ TTGGTAGTGCAACTGGCGTTGG
UBC9	Eucgr.D01776	AT4G27960	Ubiquitin conjugating enzyme 9	1.90E-82	96.6%	148	148/178	36	ATCTGCGAGGTCTCGTTTC/ GGACCCGCTTCTCAATATCAACG
Actin2	Eucgr.I00241	AT3G18780	Actin 2	0	92.8%	377	377/377	1492	AGTCTTGCAGCCATAGTCAGG/ AGAAGCACCACAATCCCAATCCC
CACS	Eucgr.E00287	AT5G46630	Clathrin adaptor complex subunit,	0	94.2%	438	413/441	1384	TGGACAAGGCCACCAATTCAGATG/ AGCGGACTGTAATCCAGATGC
YLS8	Eucgr.G02580	AT5G08290	mRNA splicing factor, thioredoxin-like U5 snRNP	2.80E-82	98.6%	317	142/142	650	TGAAGTGTGGCATCAGTTGCG/ AATCGGGCACCTCAGTTATGTCC
Expressed1	Eucgr.F00240	AT4G33380	Expressed protein	1.70E-101	62.2%	348	315/328	737	TGAGCAGCAGCGAGATGCTATTG/ TTCCAAGGGTCACTCGTTTGGC
UBC2	Eucgr.E03515	AT5G25760	Peroxin-4, ubiquitin conjugating enzyme 2	2.90E-86	93.0%	157	157/157	379	ACTGCTCTATCAAGGGACCATCG/ TGCTCAGGTATAGCAAAGGCAAGC
PP2A-3	Eucgr.B03031	AT1G13320	Protein phosphatase 2A subunit A3	0	90.3%	587	587/587	1720	CAGCGGCAACAACCTGAAGCG/ ATTATGTGCTGCATTGCCAGTC
SAND	Eucgr.B02502	AT2G28390	SAND family, trafficking protein Mon1	0	66.7%	612	580/607	1863	TTGATCCACTTGCAGACAAGGC/ TCACCCATTGACATACAGATTGC
Expressed3	Eucgr.J01745	AT4G26410	Expressed protein	1.00E-74	59.3%	242	214/263	647	TCGCAAGCAGGAGGTCTAATGG/ ACCTCAGCCCGAAGTTTCAGTG
PTB	Eucgr.J01358	AT3G01150	Polypyrimidine tract-binding protein 1	0	82.4%	459	364/399	1277	TGTGGGAAGAGACGTTTGGAAGG/ AGACAACTGAGCCACTGAAGC
Helicase	Eucgr.G00345	AT1G58050	RNA helicase family protein	0	65.1%	852	854/1417	2775	AGGAGCAAACGCTGTCAACTG/ ACATGCCCAATTGTGAACAGGAG
UPL7	Eucgr.A01586	AT3G53090	Ubiquitin-protein ligase 7	0	59.7%	1161	1142/1142	3450	TTGGAGGTCAGATGTCGAGAGAC/ TGCACGCTTGATGTGGGAAGC
F-box	Eucgr.E04314	AT5G15710	Galactose oxidase/kelch repeat superfamily protein	3.20E-27	27.8%	445	371/448	688	TTGATGCTGGCGGTTCTTCTC/ TCCTGATTCGAGGCTGCAAAGC
PPR2	Eucgr.A00809	AT1G62930	Pentatricopeptide repeat (PPR) superfamily protein	2.40E-121	39.1%	589	580/629	1775	CCATTGTGTGAGATCTGTCCCTC/ CTGCTGTTAATCTGGCGTGCTC
PP2A-1	Eucgr.B03386	AT1G59830	Protein phosphatase 2A-2	1.50E-175	98.3%	306	305/306	720	TCGAGCTTTGGACCGCATAAAG/ ACCACAAGAGGTCACACATTGGC
UBQ10	Eucgr.H03021	AT4G05320	Polyubiquitin 10	5.60E-46	37.8%	321	296/457	612	TCGGCAGAGGTGATGAGCTTAC/ AGGCTTTCGATCCTCAAGTTGC
UBQ14	Eucgr.F04448	AT4G02890	Polyubiquitin family gene	1.90E-65	80.9%	154	152/305	436	ATTCCACCTGTCCAGCAAAGGC/ TCTCAGCGCAAGCACAAGATG
****Tubulin	Eucgr.K00264	AT5G23860	Tubulin beta 8	0	99.8%	447	430/449	978	CACTCAGCAAATGTGGGATGCG/ TAACGGCCATGACGTGGATCTG
****IDH	Eucgr.F02901	AT1G65930	NADP-isocitrate dehydrogenase	0	87.6%	416	410/410	1356	TGCTGTGGCAGCTGAACTCAAG/ ATGTTGTCCGCCAGTCACCTAC

\* Genes were named according to the most similar gene from Arabidopsis.

\*\*Eucalyptus protein size in amino acid

\*\*\*Size of the Blastp alignment in amino acid (Eucalyptus vs. Arabidopsis)

\*\*\*\*Eucalyptus genes used as reference genes in previous studies.

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A. P., Lep le, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).

**Table 2.** *Eucalyptus* candidate reference genes ranked according to their expression stability evaluated using geNorm algorithm

Rank	Organs/Tissues (OT)		Leaves (L)		Xylem (XA) (All samples)		Xylem&Cambium (XC) (Development)		Xylem (XES) (Environmental stimuli)		All samples (AS)	
	Gene	M	Gene	M	Gene	M	Gene	M	Gene	M	Gene	M
1 / 1	<b>EF-1<math>\alpha</math>/ GAPDH</b>	0.47	<b>PP2A-1/IDH</b>	0.15	<b>PP2A-3/SAND</b>	0.28	<b>EF-1<math>\alpha</math>/UBC9</b>	0.13	<b>PP2A-3/SAND</b>	0.24	<b>PP2A-3/SAND</b>	0.34
3	PP2A-1	0.53	<b>Expressed1</b>	0.20	<b>UPL7</b>	0.32	PP2A-1	0.19	UPL7	0.31	UPL7	0.40
4	UBC9	0.56	<b>CACS</b>	0.23	<b>UBC2</b>	0.37	<b>CACS</b>	0.21	<b>UBC2</b>	0.37	<b>PTB</b>	0.44
5	IDH	0.59	<b>Expressed3</b>	0.24	<b>PTB</b>	0.39	<b>GAPDH</b>	0.24	<b>PTB</b>	0.40	<b>UBC2</b>	0.49
6	CACS	0.62	<b>UBC2</b>	0.26	<b>EF-1<math>\alpha</math></b>	0.41	<b>PTB</b>	0.27	<b>EF-1<math>\alpha</math></b>	0.43	PP2A-1	0.53
7	Expressed1	0.63	<b>PP2A-3</b>	0.28	<b>UBC9</b>	0.43	<b>PP2A-3</b>	0.29	<b>UBC9</b>	0.45	EF-1 $\alpha$	0.55
8	PP2A-3	0.65	<b>EF-1<math>\alpha</math></b>	0.30	<b>Expressed3</b>	0.45	<b>UBC2</b>	0.31	<b>Expressed3</b>	0.48	GAPDH	0.58
9	Expressed3	0.67	<b>UBC9</b>	0.31	<b>PP2A-1</b>	0.47	<b>Actin-2</b>	0.32	<b>Helicase</b>	0.49	CACS	0.60
10	Actin-2	0.68	<b>GAPDH</b>	0.35	<b>GAPDH</b>	0.48	<b>Helicase</b>	0.34	PP2A-1	0.51	IDH	0.62
11	UBC2	0.71	<b>Actin-2</b>	0.40	CACS	0.50	<b>IDH</b>	0.38	GAPDH	0.52	UBC9	0.64
12	SAND	0.74	<b>PTB</b>	0.45	IDH	0.52	<b>UPL7</b>	0.41	CACS	0.54	Expressed3	0.66
13	PTB	0.75	<b>UPL7</b>	0.49	Actin-2	0.55	<b>SAND</b>	0.43	IDH	0.56	Helicase	0.68
14	Helicase	0.77	PPR2	0.52	Helicase	0.57	<b>PPR2</b>	0.48	Actin-2	0.58	Actin-2	0.71
15	UPL7	0.78	SAND	0.55	Expressed1	0.59	Expressed1	0.52	Expressed1	0.61	Expressed1	0.74
16	PPR2	0.85	Helicase	0.57	PPR2	0.65	Expressed3	0.59	PPR2	0.67	PPR2	0.77

Genes in bold : M value &lt; 0.5 (cut-off value)

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A. P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).

**Table 3.** *Eucalyptus* reference genes ranked according to their expression stability evaluated by NormFinder method

Rank	Organs/Tissues (OT)		Leaves (L)		Xylem (XA) (All samples)		Xylem&Cambium (XC) (Development)		Xylem (XES) (Environmental stimuli)		All samples (AS)	
	Gene	SV*	Gene	SV	Gene	SV	Gene	SV	Gene	SV	Gene	SV
1	PP2A-3	0.15	PP2A-3	0.04	SAND	0.08	PP2A-3	0.08	SAND	0.09	PP2A-3	0.08
2	PP2A-1	0.23	CACS	0.11	PP2A-3	0.13	GAPDH	0.10	PP2A-3	0.12	CACS	0.11
3	PTB	0.30	EF-1 $\alpha$	0.14	EF-1 $\alpha$	0.13	UBC2	0.13	EF-1 $\alpha$	0.14	GAPDH	0.12
4	SAND	0.30	IDH	0.15	UBC9	0.16	PTB	0.15	UBC9	0.18	SAND	0.13
5	Expressed1	0.30	PP2A-1	0.15	UPL7	0.17	CACS	0.17	UPL7	0.19	UBC9	0.14
6	EF-1 $\alpha$	0.31	UBC2	0.16	Expressed3	0.17	Actin2	0.17	Expressed3	0.19	PP2A-1	0.14
7	IDH	0.31	Expressed3	0.20	UBC2	0.18	PP2A-1	0.20	UBC2	0.21	PTB	0.14
8	CACS	0.32	Expressed1	0.20	GAPDH	0.19	EF-1 $\alpha$	0.21	GAPDH	0.22	IDH	0.14
9	UBC2	0.33	SAND	0.27	PTB	0.20	UBC9	0.23	PTB	0.23	UBC2	0.14
10	Expressed3	0.33	PTB	0.27	PP2A-1	0.26	IDH	0.24	Helicase	0.25	EF-1 $\alpha$	0.14
11	GAPDH	0.36	Actin2	0.32	CACS	0.27	Helicase	0.32	Actin2	0.28	UPL7	0.14
12	Helicase	0.39	Helicase	0.34	Actin2	0.28	UPL7	0.33	PP2A-1	0.28	Helicase	0.15
13	Actin2	0.41	UBC9	0.34	IDH	0.30	SAND	0.33	CACS	0.29	Expressed1	0.15
14	UPL7	0.42	GAPDH	0.35	Helicase	0.30	PPR2	0.52	IDH	0.32	Actin2	0.18
15	UBC9	0.44	UPL7	0.37	Expressed1	0.30	Expressed1	0.54	Expressed1	0.32	Expressed3	0.19
16	PPR2	0.63	PPR2	0.38	PPR2	0.39	Expressed3	0.74	PPR2	0.42	PPR2	0.22
<b>Best combination**</b>												
	PP2A-3		PP2A-3		PP2A-3		PP2A-3		PP2A-3		SAND	
	PP2A-1	0.09	CACS	0.06	Expressed3	0.06	GAPDH	0.05	Expressed3	0.07	PP2A-3	0.05

\*SV: Stability value

\*\*The best combination of two reference genes in each experimental set

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A. P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).

**Table 4.** *Eucalyptus* reference genes ranked according to their expression stability evaluated by  $\Delta Ct$  method

Rank	Organs/Tissues (OT)		Leaves (L)		Xylem (XA) (All samples)		Xylem&Cambium (XC) (Development)		Xylem (XES) (Environmental stimuli)		All samples (AS)	
	Ranking	$\overline{SD}$	Ranking	$\overline{SD}$	Ranking	$\overline{SD}$	Ranking	$\overline{SD}$	Ranking	$\overline{SD}$	Ranking	$\overline{SD}$
1	PP2A-3	0.67	PP2A-3	0.43	PP2A-3	0.52	PP2A-3	0.44	PP2A-3	0.53	PP2A-3	0.62
2	PP2A-1	0.72	EF-1 $\alpha$	0.47	SAND	0.52	EF-1 $\alpha$	0.44	SAND	0.54	SAND	0.66
3	SAND	0.73	Expressed1	0.47	UBC9	0.54	GAPDH	0.45	UBC9	0.57	PP2A-1	0.68
4	EF-1 $\alpha$	0.75	CACS	0.49	Expressed3	0.57	Actin2	0.45	Expressed3	0.6	EF-1 $\alpha$	0.71
5	GAPDH	0.78	UBC2	0.5	PP2A-1	0.57	UBC2	0.46	EF-1 $\alpha$	0.6	GAPDH	0.71
6	PTB	0.79	PP2A-1	0.5	UBC2	0.59	PTB	0.46	PP2A-1	0.61	PTB	0.73
7	UBC2	0.81	IDH	0.5	EF-1 $\alpha$	0.59	CACS	0.47	GAPDH	0.61	CACS	0.73
8	IDH	0.82	Expressed3	0.51	GAPDH	0.6	SAND	0.5	UBC2	0.62	UBC2	0.73
9	CACS	0.82	SAND	0.59	UPL7	0.6	PP2A-1	0.5	UPL7	0.62	UBC9	0.76
10	Expressed1	0.83	UBC9	0.59	PTB	0.6	UBC9	0.51	PTB	0.63	IDH	0.8
11	Expressed3	0.84	GAPDH	0.6	CACS	0.65	UPL7	0.57	Helicase	0.68	Helicase	0.82
12	UPL7	0.86	PTB	0.61	IDH	0.67	IDH	0.58	CACS	0.68	UPL7	0.83
13	Helicase	0.87	UPL7	0.64	Actin2	0.71	Helicase	0.59	IDH	0.71	Expressed3	0.83
14	UBC9	0.88	Actin2	0.65	Helicase	0.74	Expressed1	0.79	Actin2	0.74	Expressed1	0.93
15	Actin2	0.89	PPR2	0.69	Expressed1	0.76	PPR2	0.85	Expressed1	0.78	Actin2	0.97
16	PPR2	1.32	Helicase	0.76	PPR2	1.08	Expressed3	1.11	PPR2	1.15	PPR2	1.06

$\overline{SD}$ : average of standard deviation

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A. P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).

**Table 5.** Comprehensive ranking of Eucalyptus reference genes according to GM method\*

Rank	Organs/Tissues (OT)		Leaves (L)		Xylem (XA) (All samples)		Xylem&Cambium (XC) (Development)		Xylem (XES) (Environmental stimuli)		All samples (AS)	
	Ranking	GM**	Ranking	GM	Ranking	GM	Ranking	GM	Ranking	GM	Ranking	GM
1	PP2A-3	4.01	PP2A-3	1.42	SAND	2.04	EF-1 $\alpha$	1.52	SAND	2.05	PP2A-3	2.53
2	PP2A-1	4.33	IDH	1.74	PP2A-3	2.77	PP2A-3	1.71	PP2A-3	2.14	SAND	3.00
3	EF-1 $\alpha$	4.63	PP2A-1	1.74	UPL7	2.94	GAPDH	1.74	UPL7	3.04	PTB	3.57
4	GAPDH	4.95	CACS	1.90	UBC9	2.96	UBC9	1.84	EF-1 $\alpha$	3.10	PP2A-1	3.64
5	IDH	5.33	Expressed1	2.02	UBC2	3.11	CACS	1.98	UBC9	3.25	GAPDH	3.70
6	SAND	5.34	EF-1 $\alpha$	2.16	EF-1 $\alpha$	3.20	PTB	2.06	UBC2	3.35	UBC2	3.72
7	Expressed1	5.46	UBC2	2.26	PTB	3.29	UBC2	2.09	Expressed3	3.50	UPL7	3.75
8	CACS	5.46	Expressed3	2.36	GAPDH	3.46	PP2A-1	2.10	PTB	3.58	CACS	3.81
9	PTB	5.63	UBC9	3.49	Expressed3	3.56	Actin2	2.25	GAPDH	3.81	EF-1 $\alpha$	3.84
10	Expressed3	5.75	GAPDH	3.70	Actin2	3.60	IDH	3.32	PP2A-1	4.06	UBC9	4.16
11	UBC2	5.76	PTB	3.86	PP2A-1	3.81	SAND	3.39	Helicase	4.19	IDH	4.28
12	UBC9	6.04	Actin2	4.03	CACS	4.29	Helicase	3.54	CACS	4.52	Helicase	4.58
13	Actin2	6.37	SAND	4.03	Expressed1	4.44	UPL7	3.72	IDH	4.82	Expressed3	4.83
14	Helicase	6.46	UPL7	4.50	IDH	4.59	Expressed1	5.87	Actin2	4.84	Expressed1	5.05
15	UPL7	6.60	PPR2	4.87	Helicase	4.62	PPR2	5.88	Expressed1	5.27	Actin2	5.33
16	PPR2	9.49	Helicase	5.16	PPR2	6.79	Expressed3	8.13	PPR2	7.28	PPR2	6.04

\*Calculated by the integration of geNorm, NormFinder and  $\Delta$ Ct methods

\*\*geometric mean of ranking values

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A. P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).



## Legends to figures

### Figure 1. Range of Ct values of 21 candidate reference genes obtained for 90 Eucalyptus samples.

Comparisons of the Cycle Threshold (Ct) values variability of the 21 candidate reference genes among all 90 samples are shown as medians (line in box), 25th percentile to the 75<sup>th</sup> percentile (boxes) and ranges (whiskers).

### Figure 2. Average expression stability values (M) evaluated by geNorm.

Average expression stability values (M) evaluated by geNorm of remaining control genes during stepwise exclusion of the least stable control gene in the different tissue panels. See also Table 2 for the ranking of the genes according to their expression stability.

### Figure 3. Determination the optimal number of reference genes for normalization according to geNorm.

Pairwise variation ( $V_n/V_{n+1}$ ) was analysed between the normalization factors  $NF_n$  and  $NF_{n+1}$  by the geNorm software. The cut-off value of 0.15, below which the inclusion of an additional reference gene is not required, is indicated by a discontinuous line. \*Optimal number of reference genes for normalization.

### Figure 4. Principal component analysis of four different methods to measure gene stability values.

A). Factor loading of the 4 methods (geNorm, NormFinder, Ct and GM) used to measure gene stability for the two main PCA axes (PC1 and PC2) calculated for four different samples panels (AS, OT, XA and L). Each method is represented by the position of the black quarter Downloaded from <http://pcp.oxfordjournals.org/> at INRA Institut National de la Recherche Agronomique on January 30, 2013 in the disc: right upper quarter: geNorm; right lower quarter: NormFinder; left upper quarter: GM (comprehensive method) and left lower quarter: Ct.

B) PCA scores of the gene stability values of the 16 reference genes analyzed for the two main PCA axes extracted (PC1 and PC2). Each disc corresponds to a candidate reference gene. Following the same rules as in A), the position of the black quarter indicates the methods. A black right upper quarter in the disc means that the gene is ranked amongst the sixth best-classed genes by the geNorm method. Dots with one, two or three black quarter indicate that the gene is ranked amongst the best-classed sixth genes by one, two or three methods. Black disc (four black quarters) represent genes ranked amongst the first sixth best-classed genes by all four methods. All the genes well-ranked by at least one method are generally found on the left of the vertical axis.

### Figure 5. Expression profiles of EgMYB1 and EgMYB2 in different organs and tissues.

Expression ratios of EgMYB1 and EgMYB2 for the experimental set of different organs/tissues calculated using (i) the best-ranked reference gene PP2A-3, (ii) the worse-ranked reference gene PPR2, (iii) combinations of genes recommended by each algorithm i.e. PP2A-1, EF-1 $\alpha$  and GAPDH for geNorm; PP2A-3 & PP2A-1 for NormFinder; PP2A-3, PP2A-1 and SAND for Ct and PP2A-3, PP2A-1 and EF-1 $\alpha$  for the comprehensive ranking method (GM). Results are expressed in % of the value in xylem which was arbitrary set to 100%. The mature leaves sample was used as the calibrator.

Figure 1

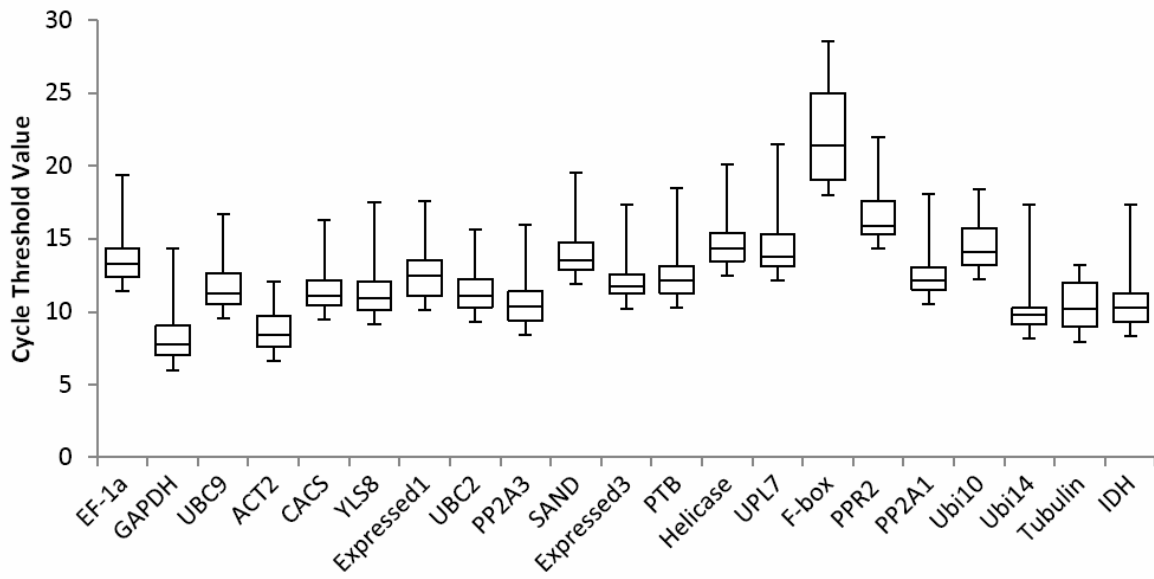
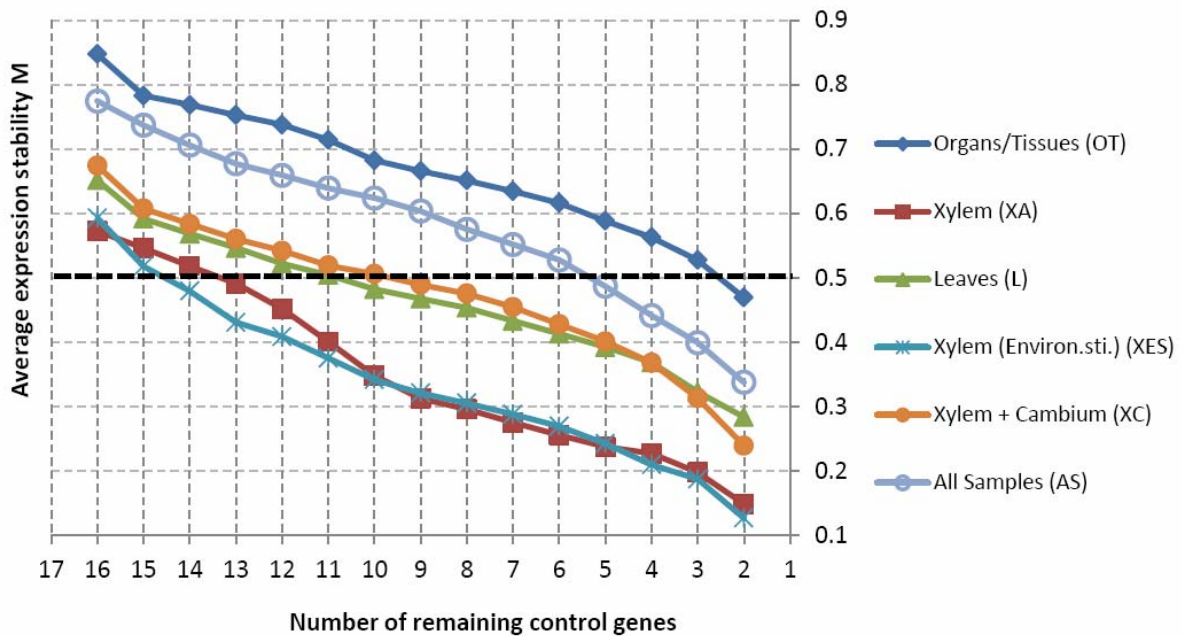


Figure 2



Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*, 53 (12).

Figure 3

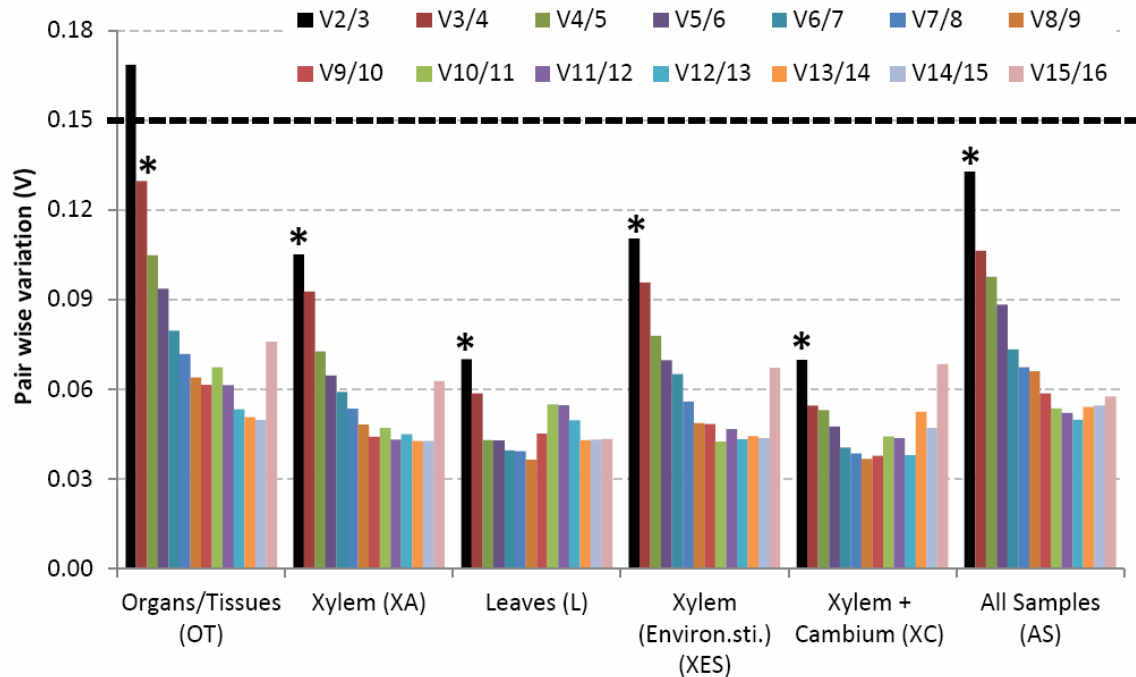
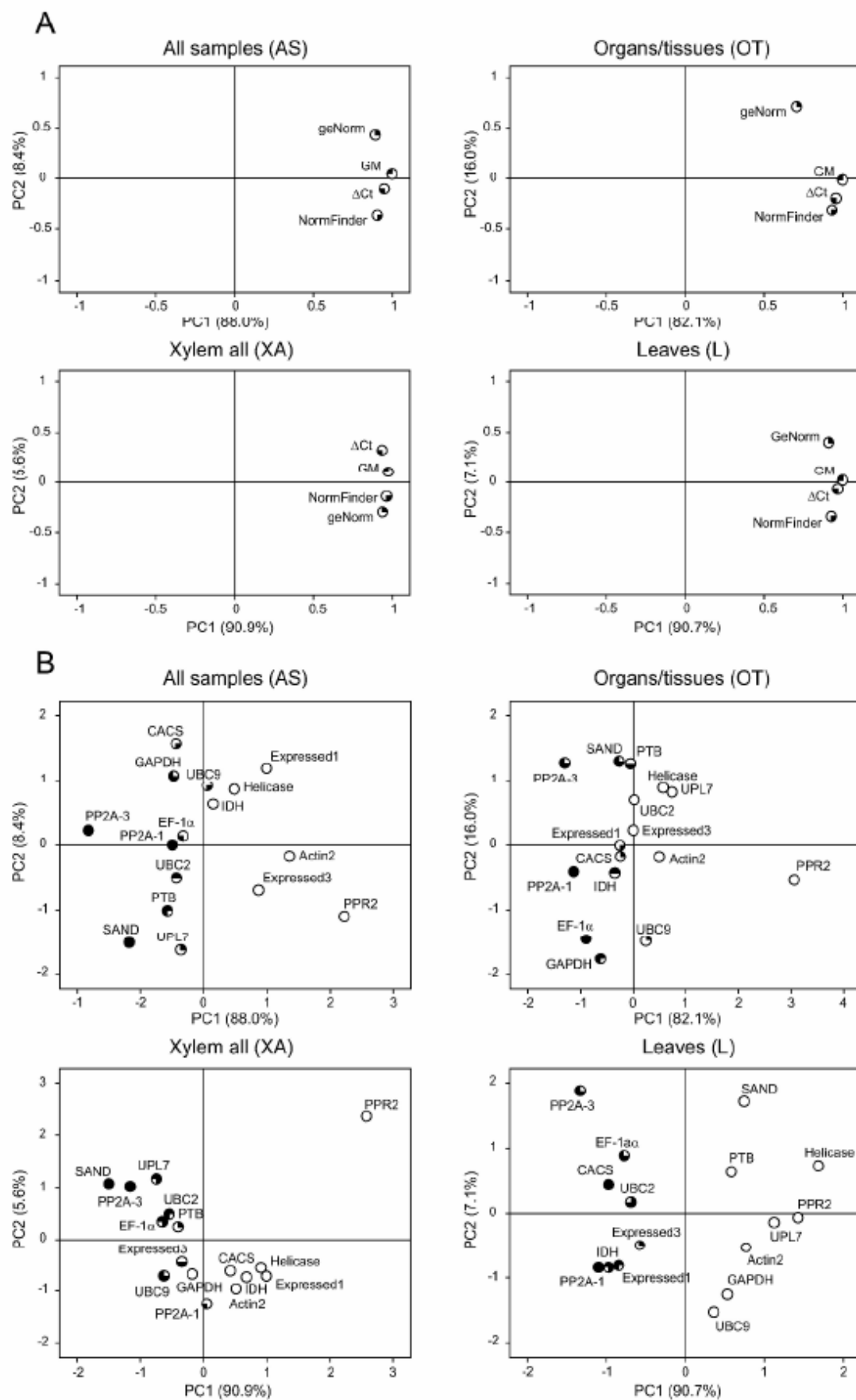


Figure 4

Figure 4. Principal component analysis of four different methods to measure gene stability values.

A). Factor loading of the 4 methods (geNorm, NormFinder, Ct and GM) used to measure gene stability for the two main PCA axes (PC1 and PC2) calculated for four different samples panels (AS, OT, XA and L). Each method is represented by the position of the black quarter in the disc: right upper quarter: geNorm; right lower quarter: NormFinder; left upper quarter: GM (comprehensive method) and left lower quarter: Ct.

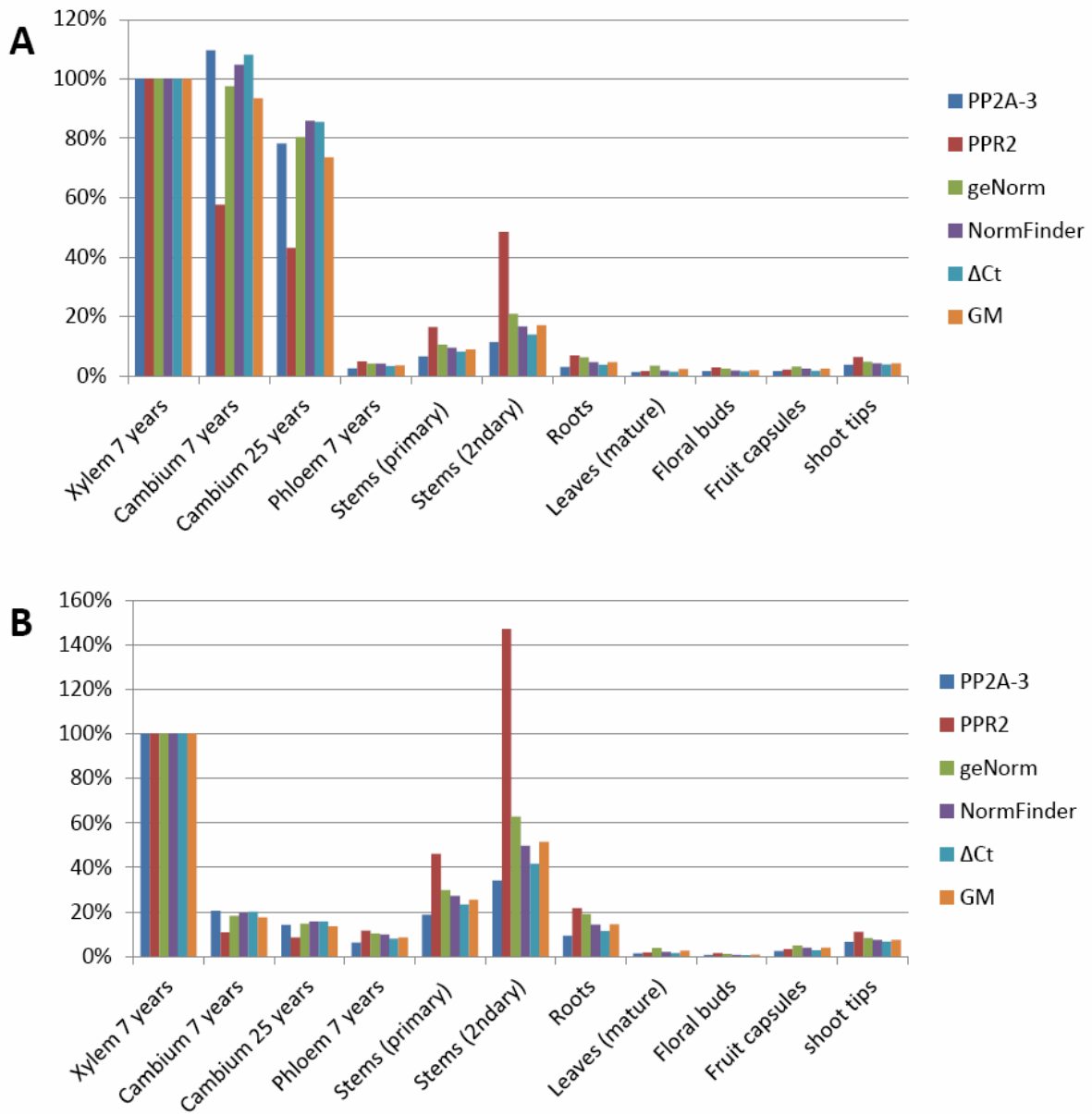
B) PCA scores of the gene stability values of the 16 reference genes analyzed for the two main PCA axes extracted (PC1 and PC2). Each disc corresponds to a candidate reference gene. Following the same rules as in A), the position of the black quarter indicates the methods. A black right upper quarter in the disc means that the gene is ranked amongst the sixth best-classed genes by the geNorm method. Dots with one, two or three black quarter indicate that the gene is ranked amongst the best-classed sixth genes by one, two or three methods. Black disc (four black quarters) represent genes ranked amongst the first sixth best-classed genes by all four methods. All the genes well-ranked by at least one method are generally found on the left of the vertical axis.



Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).

Figure 5



Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A.P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).

Comment citer ce document :

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L. O., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A. P., Leplé, J.-C., Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions. *Plant and Cell Physiology*. 53 (12).