Validation of reference genes for gene expression analysis in chicory (Cichorium intybus) using quantitative real-time PCR

Asad Maroufi1,2, Erik Van Bockstaele1,3, Marc De Loose2,4*

Abstract

Background: Quantitative real-time reverse transcriptase polymerase chain reaction (qRT-PCR) is a sensitive technique for quantifying gene expression levels. One or more appropriate reference genes must be selected to accurately compare mRNA transcripts across different samples and tissues. Thus far, only actin-2 has been used as a reference gene for qRT-PCR in chicory, and a full comparison of several candidate reference genes in chicory has not yet been reported.

Results: Seven candidate reference genes, including nicotinamide adenine dinucleotide dehydrogenase (NADHD), actin (ACT), β-tubulin (TUB), glyceraldehyde-3-phosphate-dehydrogenase (GADPH), histone H3 (H3), elongation factor 1-alpha (EF) and 18S rRNA (rRNA) were selected to study the expression stability for normalisation of gene expression in chicory. Primer specificity and amplification efficiency were verified for each gene. The expression stability of these genes was analysed across chicory root and leaf tissues using geNorm, NormFinder and BestKeeper software. ACT, EF, and rRNA were the most stable genes as identified by the three different analysis methods. In addition, the use of ACT, EF and GAPDH as reference genes was illustrated by analysing 1-FEHII (FEHII) expression in chicory root and leaf tissues. These analyses revealed the biological variation in FEHII transcript expression among the tissues studied, and between individual plants.

Conclusions: geNorm, NormFinder, and BestKeeper analyses indicated that ACT, EF and rRNA had the highest expression stability across leaf and root tissues, while GAPDH and NADHD showed relatively low expression stability. The results of this study emphasise the importance of validating reference genes for qRT-PCR analysis in chicory. The use of the most stable reference genes such as ACT and EF allows accurate normalisation of gene expression in chicory leaf and root tissues.

Background

Quantitative real-time reverse transcriptase polymerase chain reaction (qRT-PCR) is an efficient, sensitive and reliable technique to quantify transcript expression levels. qRT-PCR is fast, easy to use and provides simultaneous measurement of gene expression in many different samples for a limited number of genes [1,2]. qRT-PCR has various applications, such as clinical diagnostics [3], analysis of tissue-specific gene expression in humans [4], and gene expression studies in plants [5]. Appropriate normalisation is very important for quantification of transcript expression levels. The most accepted approach to quantification is normalisation of the expression level of a gene of interest (target gene) to the expression level of an internal stably expressed gene (control gene) [6-10]. The control gene, often termed reference gene, is a stably expressed gene that is experimentally verified in given species and tissues under given experimental conditions [6]. By normalising the transcript expression level of a target gene to the expression level of a reference gene, differences in the quality or quantity of template RNA and differences in efficiencies of the reverse transcription reaction between samples are accounted for. This allows the direct comparison of normalised transcript expression levels between samples. However, this approach requires the

* Correspondence: marc.deloose@ilvo.vlaanderen.be
2 Institute for Agricultural and Fisheries Research (ILVO), Burg Van Gansberghelaan 115 bus 1, 9820 Merelbeke, Belgium

© 2010 Maroufi et al; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
selection of at least one reference gene for validation of a corresponding qRT-PCR method.

This case study illustrates the use of qRT-PCR for improving inulin production in chicory. Chicory (Cichorium intybus) is an important crop for inulin production. Inulin is a group of naturally occurring polysaccharides that are produced by many types of plants including Jerusalem artichoke (Helianthus tuberosus) and chicory [11]. Inulin is widely used as an ingredient in functional foods, and there is growing interest in food and non-food industries to make new compounds from inulin and its derivatives [12-15]. Currently, chicory is the only plant species used on an industrial scale for the commercial extraction of inulin [15]. Investigation of the expression level of genes encoding enzymes involved in the inulin biosynthesis and degradation pathway will enable breeders to improve inulin content. This case study illustrates the use of qRT-PCR in chicory using the fructan 1-exohydrolasell a, 1-FEHIIa (FEHIIa), and fructan 1-exohydrolasell b, 1-FEHIIb (FEHIIb) genes involved in the inulin degradation pathway. The encoded enzymes are believed to catalyse fructan depolymerisation at the end of the growing season, as well as during storage and forcing of tubers and tuberous roots [16-19].

Up to now, Northern blot analysis has been the favoured means of studying the expression of the genes involved in the inulin biosynthesis and degradation pathway [16]. Northern blotting requires a relatively high amount of RNA, and it is laborious and time-consuming [20,21]. Further, in certain cases, the expression of low-expressed genes may be below the detection limit of Northern blotting. A more sensitive and efficient method, such as qRT-PCR, is thus desirable [5]. This requires selection of appropriate reference genes for normalisation. Only actin-2 has been used as a reference gene for qRT-PCR in chicory [22], and a full comparison of several candidate reference genes in chicory has not yet been reported. Taken together, the aims of this study are i) to rank the candidate reference genes according to expression stability across chicory root and leaf tissues using three different methods for expression stability measurements, ii) to develop and evaluate qRT-PCR methods for these genes in chicory, iii) to select appropriate reference genes to use for normalisation of gene expression by qRT-PCR in chicory and iv) to demonstrate their usefulness in qRT-PCR by analysing the expression level of fructan 1-exohydrolasell, 1-FEHII (FEHII) in chicory root and leaf tissues, as an example for the genes involved in inulin accumulation.

Results
Selection of candidate reference genes and primer design
We selected seven candidate reference genes to validate and develop a qRT-PCR method in chicory, including nicotinamide adenine dinucleotide dehydrogenase (NADHD) [23-25], actin (ACT) [21,26-31], β-tubulin (TUB) [28-33], glyceraldehyde-3-phosphate-dehydrogenase (GADPH) [21,26-31,34], histone H3 (H3) [34], elongation factor 1-alpha (EF) [21,26,27,29,34] and 18S rRNA (rRNA) [21,26-30,34]. For all selected genes except GAPDH, chicory transcript sequences are available in GenBank to design qRT-PCR primers (Table 1). Based on the Arabidopsis thaliana GAPDH sequence (GenBank accession number: AK317337) a primer pair (5’-TGGAGCTGACTTTGTTGTTGA-3’; 5’-TCCACCTCTCCAGTCTCTTC-3’) was designed that amplified a 298 bp fragment from chicory genomic DNA. This fragment was cloned and sequenced. TBLASTx analysis revealed that this fragment contains an 87 bp region that has at least 89% identity at the amino acid level to GAPDH genes across the plant kingdom. We subsequently designed a qRT-PCR primer pair to amplify a 91 bp amplicon covering this region from chicory GAPDH.

Because the nucleotide sequences of chicory FEHIIa and FEHIIb transcripts are 94% identical [35], it was not possible to design qRT-PCR primers that could differentiate between the two transcripts. We thus designed a primer pair to amplify a specific region of both FEHIIa and FEHIIb (Table 1 and Additional file 1). The primer pair was designed to cross the exon5/6 junction of FEHIIa (Additional file 1), excluding the possibility that amplification occurs from any genomic DNA contamination. In this case, we studied FEHII expression as the combined expression of FEHIIa and FEHIIb transcripts in qRT-PCR.

Verification of amplicons, primer specificity, Cp data collection and gene-specific PCR amplification efficiency
Agarose gel electrophoresis (Figure 1a) and melting curve analysis (Figure 1b and Table 1) revealed that all primer pairs amplified a single PCR product with the expected size. Furthermore, sequence analysis of cloned amplicons revealed that all sequenced amplified fragments were identical or nearly identical (1 bp different for ACT and EF) to the sequences used for primer design from GenBank, except H3, which showed an 18 bp difference in the middle of the amplicon. Real-time RT-PCR was conducted on the 25 samples with eight primer pairs. To reveal the differences in transcript expression levels between studied genes, the average crossing point cycle number (Cp) value was calculated across all leaf and root samples of five individual plants (Table 1). Use of equal amounts of cDNA raised different values from real-time PCR amplification. As expected, the average Cp value varies between the different genes. rRNA was the most abundant (mean Cp = 7.11) reference transcript while GAPDH was the least...
abundant (mean Cp = 29.58). A standard curve using a dilution series of the cloned amplicons (spanning five orders of magnitude) was made to calculate the gene-specific PCR efficiency. The correlation coefficient (R²) of the slope of the standard curve used to calculate gene-specific PCR amplification efficiency (E), and PCR efficiency including standard deviation (S.D), of all selected genes are listed in Table 1. The PCR amplification efficiency for the various PCR methods varied between 82.24% and 98.17% (Table 1).

Expression stability of candidate reference genes

Three different software programmes were used to calculate the expression stability of the candidate reference genes: geNorm [36], NormFinder [37] and BestKeeper [38]. To find stably expressed genes, we first assayed gene expression stability across leaf and root tissues (L1-1/1, R1-1/1, R1-1/2, R2-1/1 and R2-1/2 samples; Figure 2). Cp data were collected for all selected tissues of each plant. These data were either used directly for stability calculations (BestKeeper analysis) or were first transformed to relative quantities using the delta-Ct method or were first transformed to relative quantities using the delta-Ct method and the gene-specific amplification efficiency implemented in qBase (geNorm and NormFinder analysis).

a) geNorm analysis

Average expression stability (M value) of all genes was calculated by geNorm (version 3.5). The M values of the candidate reference genes across chionic leaf and root tissues are shown in Table 2. geNorm recommends using an M value below the threshold of 1.5 to identify (sets of) reference genes with stable expression. The three genes ACT, EF, and rRNA had the highest expression stability in leaf and root tissues (the lowest M values). H3 and TUB had intermediary M values that are still below the threshold of 1.5. NADHD and GAPDH had M values higher then 1.5, indicating less stable expression across leaf and root tissues (Table 2). To determine the optimal number of reference genes, geNorm calculates the pairwise variation Vₙ/Vₙ₊₁ between two sequential normalisation factors NFₙ and NFₙ₊₁ that contain an increasing number of reference genes. A large variation means that the added gene has a significant effect on the normalisation and should preferably be included for calculation of a reliable normalisation factor. Ideally, extra reference genes are included until the variation Vₙ/Vₙ₊₁ drops below a given threshold. Vandesompele and colleagues recommended a threshold of 0.15, although this threshold should not be viewed as too strict of a cut-off [36]. In our data sets, the calculated NFₙ at increasing numbers of reference genes shows that the inclusion of the third, fourth and fifth genes (rRNA, H3 and TUB; M values < 1.5) still contribute significantly to the variation of the normalisation factor (V₂/₃, V₃/₄ and V₄/₅ > 0.15). NADHD and GAPDH had M values higher then 1.5, suggesting that they should not be included in the normalisation factor (Figure 3).

b) NormFinder analysis

The stability value of each gene was calculated by NormFinder (Table 2). Genes that are more stably expressed are indicated by lower average expression stability values. The analysis ranks ACT, EF, rRNA and H3 as the four most stable genes (Table 2). Thus, both geNorm and NormFinder rank the same four genes as the most stable and the entire order is identical. Both analyses rank GAPDH as the least stable gene.

---

Table 1. Selected candidate reference genes, primers and different parameters derived from qRT-PCR analysis

| Gene name | GenBank accession number | Primer sequences (forward/reverse) | Tm (°C) | Amplicon length (bp) | Amplification efficiency (%) | *S.D of efficiency | Average Cp of cDNA | **R² |
|-----------|--------------------------|------------------------------------|--------|---------------------|-----------------------------|-------------------|-------------------|------|
| ACT       | EF528575                 | CCAATCTGGCTTTGCAGTAATGGCATG        | 80.26  | 74                  | 94.31                       | 0.0123            | 24.29             | 0.9992 |
| TUB       | AF101419                 | GCACGCTATGATGTGACCAGGCTAGG         | 82.56  | 101                 | 95.77                       | 0.0094            | 20.65             | 0.9995 |
| NADHD     | L39390                   | TCGAAGCAAAATCGTCACTATTAATAAAT      | 75.84  | 102                 | 84.89                       | 0.0104            | 27.12             | 0.9989 |
| EF        | AY378166                 | CATGCGTCAGAGGCTGTCGTCGTCG          | 82.17  | 100                 | 98.17                       | 0.0055            | 19.11             | 0.9999 |
| H3        | AY378165                 | AGAGCTCTGGAATCAAGACGTCGTCG         | 83.79  | 100                 | 94.49                       | 0.0059            | 26.46             | 0.9998 |
| rRNA      | U42501                   | GCAGCGTATTAAAGAATGCACTGATTAAT      | 80.61  | 102                 | 91.27                       | 0.0107            | 7.11              | 0.9992 |
| GAPDH     | AY378166                 | AGGCGCGGTTCAAGAACAGCTCCTACCTATT   | 81.55  | 91                  | 91.04                       | 0.0037            | 29.58             | 0.9999 |
| FELHII/a  | AY378166                 | TGAAGCTTCAAGAAGAACAGTCGAACCTG     | 78.98  | 135                 | 82.24                       | 0.0116            | 31.53             | 0.9981 |

*SD, standard deviation; **R², correlation coefficient of the slope of the standard curve.
c) BestKeeper analysis

BestKeeper analysis determines the most stably expressed genes based on the coefficient of correlation ($r$) to the BestKeeper Index (BI), which is the geometric mean of Cp values of candidate reference genes. Variations of gene expression, displayed as the standard deviation of the Cp values, were determined. BestKeeper analysis revealed that $H3$ is the gene with the lowest overall variation, and GAPDH with the highest, from the list of selected genes with an S.D of 0.74 and 2.29, respectively (Table 3). The stability value of individual genes was calculated by BestKeeper based on the pairwise correlation between genes and BI (Table 2). The BestKeeper revealed that the best correlations were obtained for rRNA ($r = 0.948$), ACT ($r = 0.866$), EF ($r = 0.847$) and TUB ($r = 0.759$) with p value of 0.001 (Table 3). GAPDH is ranked as the fifth stable gene but it has the highest S.D. $H3$ and NADHD are ranked as the least stable genes.

In conclusion, the three most stable reference genes commonly identified by the three different analysis methods are ACT, EF, and rRNA. GAPDH ranks fifth according to BestKeeper, or seventh according to geNorm and NormFinder.

Evaluation of expression ratios of candidate reference genes

The EF and GAPDH were respectively selected as stable and relatively unstable reference genes to show their expression ratios when normalised by ACT in 25 samples including L1-1/1, R1-1/1, R1-1/2, R2-1/1 and R2-1/2 cDNAs (Figure 4a and Figure 4b). The expression pattern of EF normalised to ACT illustrates the relatively stable expression ratio of these genes across the leaf and

Figure 1 Confirmation of amplicon size and primer specificity of studied genes (a) Agarose gel electrophoresis showing specific reverse transcription PCR products of the expected size for each gene, (b) Melting curves generated for all genes. M represents DNA size marker.
root tissues (Figure 4a). The expression patterns of GAPDH normalised against ACT show that the expression levels are more variable (Figure 4b). In particular, these results reveal the relatively low level of GAPDH expression in the R1 samples in three out of five independent plants. This illustrates the reduced expression stability of the GAPDH gene in the sample set.

**Application of the qRT-PCR protocol to evaluate the expression of fructan exohydrolase II in chicory leaf and root tissues**

To demonstrate the usefulness of the validated candidate reference genes in qRT-PCR, we analysed the expression of FEHII across leaf and root tissues (25 samples of five individual plants from cultivar ‘Hera’ including L1-1/1, R1-1/1, R1-1/2, R2-1/1 and R2-1/2 cDNAs). The relative expression level of FEHII in leaf and root tissues was calculated in qBase using the two best reference genes (ACT and EF) for normalisation (Figure 4c). Expression analysis shows that, in general, FEHII transcript expression was detected in all tissues in the five individual plants investigated, except for four cases (Figure 4c). As expected, the highest levels of FEHII expression were observed in the mature root tissues (Wt1R2, Wt3R2 and Wt5R2). FEHII was generally expressed at a very low level in green leaves (L) and root tips (R1; Figure 4c). Among the leaf samples, the highest-fold change in expression is 3.52 between Wt5L1-1/1 and Wt4L1-1/1 and the lowest-fold change is 1.5 between Wt3L1-1/1 and Wt4L1-1/1. Among the R2 tissues, the highest-fold change is 12.33 between Wt5R2-1/2 and Wt2R2-1/2 and the lowest-fold change is 1.91 between Wt4R2-1/2 and Wt2R2-1/1. These analyses show the biological variation in FEHII transcript expression among the studied tissues and between individual plants of the same cultivar. These expression results are consistent with FEHII transcript analysis by Northern blot as reported by Van den Ende and

**Table 2 Ranking of the candidate reference genes according to their stability value using geNorm, NormFinder and BestKeeper analysis**

| Gene name | Stability value (geNorm) | Ranking order (geNorm) | Stability value (NormFinder) | Ranking order (NormFinder) | Stability value (BestKeeper) | Ranking order (BestKeeper) |
|-----------|--------------------------|------------------------|------------------------------|---------------------------|-----------------------------|-----------------------------|
| ACT       | 0.37                     | 1                      | 0.127                        | 1                         | 0.866                       | 2                           |
| EF        | 0.37                     | 1                      | 0.193                        | 2                         | 0.847                       | 3                           |
| rRNA      | 0.88                     | 2                      | 0.509                        | 3                         | 0.948                       | 1                           |
| H3        | 1.08                     | 3                      | 0.684                        | 4                         | 0.479                       | 6                           |
| NADHD     | 1.55                     | 5                      | 1.296                        | 6                         | -0.069                      | 7                           |
| TUB       | 1.30                     | 4                      | 1.243                        | 5                         | 0.759                       | 4                           |
| GAPDH     | 1.84                     | 6                      | 1.598                        | 7                         | 0.635                       | 5                           |
colleagues [16], who demonstrated that the expression of \textit{FEHII} was highly abundant in roots. The current study also suggests that qRT-PCR analysis makes it possible to detect and quantify low levels of \textit{FEHII} expression in leaf tissue.

Next, we examined the effect of choosing a reference gene with relatively low expression stability across the target tissues. To this end, the expression of \textit{FEHII} using \textit{GAPDH} as a reference gene was calculated. In comparison with normalisation of \textit{FEHII} data against \textit{ACT}/\textit{EF} the overall expression profile appears similar (Figure 4c and 4d), but marked changes are introduced in individual samples. For instance, relatively low expression of \textit{GAPDH} in R1 tissues of Wt2, Wt4 and Wt5 (Figure 4d) leads to overestimation of \textit{FEHII} expression in the respective samples. In Wt4, the R1 sample has the lowest \textit{FEHII} expression compared to L and R2, if expression is normalised against \textit{GAPDH}. A similarly strong apparent increase is observed for the R1 samples in Wt3 and Wt5 when expression is normalised against \textit{GAPDH}.

Reproducibility of the cDNA synthesis (R1 and R2 samples) was assessed using the best normalisation factor (combination of \textit{ACT} and \textit{EF}; Figure 4c). The similarity between the two observations of each root sample per individual plant reveals high reproducibility between independent cDNA syntheses for the root tissues. In contrast, reproducibility appears to be much lower, especially in the R1 samples, when \textit{FEHII} expression is normalised against \textit{GAPDH}, as compared to normalisation against \textit{ACT} and \textit{EF}. This effect is clearly due to low stability of \textit{GAPDH} expression across samples. This illustrates the adverse effect of an unsuitable reference gene.

\textbf{Discussion}

qRT-PCR has become a valuable tool for accurate gene expression profiling in addition to Northern blotting [5,39,40]. qRT-PCR is a rapid, accurate and sensitive technique for relative quantification of transcript expression levels and requires a relatively low amount of RNA. Quantification of gene expression is affected by several factors, including experimental sources of variation and the normalisation method. Various experimental sources of variation exist in qRT-PCR, such as sample-to-sample variation in RNA integrity, differences in reverse transcriptase reaction efficiency and the amount of cDNA template used in each PCR reaction. Normalisation of the expression level of a target gene against a stably expressed internal gene can compensate for all these kinds of variations and results in the relative quantification of gene expression levels across samples [41]. Moreover, correct and accurate sample normalisation is required to reveal small but significant differences in expression when comparing samples from different organs or tissues. The accuracy of the results obtained by qRT-PCR therefore strongly depends on the choice of one, or preferably more, reference genes that are stably expressed across all tissues or organ samples [7,8].

Some genes, such as \textit{ACT}, ubiquitin 10 (\textit{UBQ10}), glucose-6-phosphate dehydrogenase (\textit{G6PD}), \textit{GAPDH}, ribosomal genes, cyclophilin, \textit{EF} and alpha-tubulin (\textit{TUA}) are commonly used housekeeping genes for gene expression studies in many plant species [21,26,27,29]. However, recent studies indicate that the traditional housekeeping genes are not always stably expressed when tested in

\begin{table}[h]
\centering
\caption{Statistics results by BestKeeper software for seven selected genes based on Cp values} \label{tab:stats}
\begin{tabular}{llllllll}
\hline
 & \textit{H3} & \textit{EF} & \textit{ACT} & \textit{NADH} & \textit{rRNA} & \textit{TUB} & \textit{GAPDH} & BI \\
\hline
\(n\) & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 \\
GM (Cp) & 26.59 & 19.29 & 24.39 & 26.65 & 7.09 & 21.12 & 29.24 & 20.38 \\
S.D (± Cp) & 0.74 & 1.10 & 1.11 & 1.22 & 1.24 & 2.14 & 2.29 & 1.16 \\
CV (\%Cp) & 2.79 & 5.67 & 4.53 & 4.59 & 17.10 & 10.09 & 7.81 & 5.66 \\
\hline
\end{tabular}
\begin{tablenotes}
\item \(n\), number of samples; Cp, Crossing point cycle number equivalent terminology for Ct; GM(Cp), the geometric mean of Cp; S.D (± Cp), Cp standard deviation; CV (\%Cp), variance coefficient expressed as percentage of Cp level; BI, BestKeeper Index.
\end{tablenotes}
\end{table}
Figure 4 The relative expression level of EF, GAPDH and FEHII in chicory leaf and root. (a) EF normalised by ACT, (b) GAPDH normalised by ACT, (c) FEHII normalised by combined normalisation factor using ACT and EF and (d) FEHII normalised by GAPDH in young green leaf (L) and root tissues (R) of five wild type (Wt) chicory plants. Data are obtained from two independent RNA extractions for leaf tissue and two independent cDNA syntheses for R1 and R2 tissues. Expression profiles are calculated in qBase using selected reference genes and the gene specific amplification efficiency. Wt4-R1-1 as calibrator, ND: FEHII transcript was not detected in respective samples. Error bars represent standard error of the mean (white bars L samples, gray bars R1 samples and black R2 samples).
other species or in a wider range of experimental treatments [42-44]. For example, Nicot and colleagues demonstrated that ACT did not appear to be the best gene to use as reference gene during the different treatments [31]. In addition, Gutierrez and colleagues have found high variability in the relative expression of common reference genes, including ACT, TUB, UBQ and EF, during various developmental stages in Arabidopsis [10].

This means that the most stable reference gene(s) should be identified for a specific species under study or in a new experimental set-up. Accordingly, for gene expression studies in chicory, the stability of reference genes needs to be verified prior to use in qRT-PCR. Actin-2 is the only reported gene used in qRT-PCR in chicory [22]. The direct transfer of traditional and recently proposed novel candidate reference genes by Czechowski and colleagues [34] to non-model plants such as chicory is hampered by the limited availability of genomic sequences. We thus selected a series of candidate reference genes for which such sequence information was available. We developed a qRT-PCR method for ACT, EF, rRNA, GAPDH, H3, TUB, NADHD and FEHII as the target gene. The specificity of the qRT-PCR primer pairs was confirmed by agarose gel electrophoresis, Tm analysis and sequencing of the amplicons. The PCR amplification efficiency was estimated, and the reference genes were ranked according to their expression level stability across chicory leaf and root tissues using three different methods.

We used geNorm, which has been recently noted as one of the best methods to determine the most stably expressed genes for qRT-PCR analysis [10,42-44]. In addition, geNorm supplies more information about the optimal number of genes in a given experimental dataset. The analysis showed that the expression of ACT and EF are the most stable across tissues as compared with the other selected genes. Data analysis of expression stability (M value) and normalisation factor variation \( \frac{V_n}{V_{n+1}} \) determined that rRNA, H3 and TUB can be added in combination with ACT/EF to calculate a normalisation factor based on multiple reference genes [7]. The stability of a candidate gene is determined by pairwise comparison of variation of expression ratio in geNorm. In order to avoid co-regulation, following the lead of many other reports [44-46], we also determined the stability of the selected genes using Normfinder, which is less sensitive to co-regulation of the reference genes. NormFinder identified ACT, EF, rRNA and H3 as the four most stable genes, which supports the geneNorm analysis in this experiment. According to the results obtained from BestKeeper analysis ranking of the five most stable genes was as follows: rRNA, ACT, EF, TUB and GAPDH. Pfaffl and colleagues demonstrated that low-expressed genes (GAPDH in our dataset; Table 1), with Cp values around cycles 30-35, definitely show different variance compared to high-expressed genes (rRNA) with Cp values around 15 or even less. Such two genes can only be correlated on their ranking, not parametrically [41]. Thus, when comparing genes with very different expression levels, it is necessary to use a new model-based analysis that also employs non-parametric methods such as the Spearman and Kendall Tau correlation coefficient. Nevertheless, the latter algorithm is affected by a circular statistical problem; among these three algorithms Bestkeeper is inferior to the two other algorithms. Recently, other studies on validation of reference genes have shown that GAPDH is not stable in different tissues or environmental conditions [6,44,47]. Czechowski and colleagues [34] compared traditional and novel reference genes in Arabidopsis and found that GAPDH was ranked among the 100 most stably expressed genes only after omission of seed and pollen samples, while TUB6, EF-1α and ACT2 were never represented in the top 100. These results also indicate that there are no universal reference genes for all plant species. Validation is thus essential for any selected housekeeping gene used as reference gene in gene expression analysis. In conclusion, the three algorithms do not rank the candidate reference genes in the same order, but all indicated that rRNA, ACT and EF are the most stably expressed genes, given the experimental conditions applied in this study. H3, TUB, GAPDH and NADHD were ranked differently by different software programmes’ analysis.

rRNA is one of the highly stably expressed genes (Table 2). However, there are some drawbacks when using this gene as reference. One of the main problems of using total RNA for normalisation is the large quantity of 18S or 28S rRNA transcripts in comparison to the target mRNA transcripts [48], as revealed by a relatively low Cp value for rRNA in this experiment. Additionally, the mRNA fraction in total RNA may differ from sample to sample [48]. Due to the high abundance of 18S and 28S rRNA transcripts, when using rRNA as an internal control for quantification of genes with relatively low expression levels (such as FEHII), the cDNA templates may need additional dilution to improve comparison.

Increasing the number of reference genes for normalisation will improve the accuracy of the analysis, but it is expensive and time-consuming. Use of two stable reference genes is a valid normalisation strategy in most cases, and has already resulted in more accurate and reliable normalisation compared to the use of a single reference gene [36]. The current study suggests that ACT and EF would be two valid reference genes for gene expression study in chicory leaf and root tissues.
The qRT-PCR methods described here have important applications in quantifying gene expression levels in chicory as shown by the FEHII expression analysis. This analysis showed that the FEHII transcript in different plant and tissue samples varies widely. The relative expression analysis of FEHII also showed that the expression level is low, but still detectable, in leaf (L) and root tip tissues (R1) and is more highly expressed in mature root tissue (R2). We conclude that the qRT-PCR methods described here facilitate sensitive and accurate quantification of gene expression in chicory.

Conclusion
To the best of our knowledge, this article is the first attempt to validate a set of commonly used candidate reference genes in chicory for the normalisation of gene expression analysis using qRT-PCR. Analysis of stability using geNorm, NormFinder and BestKeeper revealed that the expression of ACT and EF is the most stable across root and leaf tissues. In addition, data analysis performed (RNA samples L1-1, R1-1 and R2-1; Figure 2) to analyse reproducibility of cDNA synthesis. For all samples (L1-1, R1-1 and R2-1), 423 ng total RNA was used to make cDNA samples. First strand cDNA synthesis was performed using the SuperScript® VILO™ cDNA Synthesis Kit (Invitrogen) following the manufacturer’s instructions in a final volume of 20 μl. The VILO™ reaction mix includes random primers to make cDNAs. The final cDNA products were diluted 10-fold prior to use in real-time PCR.

Primer design, verification of selected gene amplicons and gene-specific PCR amplification efficiency
For all genes, primer pairs (Table 1) were designed using Primer Express version 2.0.0 (Applied Biosystems). To check the specificity of all primers, PCR was performed on cDNA, and PCR products were analysed on a 2% agarose gel (Figure 1a). In order to confirm the sequences of the amplicons, PCR was performed on cDNA for all designed primer pairs. Each reaction contained 300 nM of each primer, 2 U of Pfu DNA polymerase (Promega), 400 μM dNTPs (Invitrogen), and 100 ng of cDNA in a total volume of 50 μl. Amplifications were performed with the following program: 95°C for 2 min and 35 cycles at 94°C for 30 sec, 58°C for 30 sec, and 72°C for 30 sec. PCR products were purified using the Qiaquick PCR purification kit (Qiagen) according to the manufacturer’s instructions. PCR products were cloned using the TOPO TA Cloning Kit (Invitrogen) according to the manufacturer’s instructions. Sequencing reactions were performed on 500 ng respective purified plasmids for each gene using 100 nM of either M13 forward or M13 reverse primer in separate reactions, 4 μl of Big Dye Mix (Big Dye Terminator v1.1 cycle sequencing kit; Applied Biosystems) and 2 μl of sequencing buffer in a total volume of 20 μl. The reactions were analysed on the ABI PRISM 3130 Genetic Analyzer (Applied Biosystems). Derived sequences for each gene were compared to the respective sequences designed primers with Vector NTI Advance™ 10 (Invitrogen). Purified plasmids containing the respective cloned amplicon for each gene were linearised using EcoRV restriction enzyme, separated on an agarose gel, excised and purified using Sigma Gel Extraction Kit following the manufacturer’s instructions (Sigma). Standard dilutions were constructed to determine the specific PCR amplification efficiency for each gene, using 10-fold of five dilution series of the purified fragments in 50 ng/μl yeast tRNA (Invitrogen). The PCR amplification efficiency of each primer pair is calculated from the slope of a standard curve as follows: for each gene, a standard curve is obtained using a 10-fold dilution series of the

Methods
Plant materials
We sowed seeds of the synthetic root chicory cultivar ‘Hera’, commonly used for commercial inulin production, then transferred seedlings to pots containing soil and grown under standard conditions in a greenhouse. Leaf (L) and root tissue samples were collected from five plants once the mature root developed (upper part of root (R2) and lower part of root (R1); Figure 2). The tissue samples were immediately frozen in liquid nitrogen. The R2 samples represent the part of the mature root that is normally used for inulin extraction.

Total RNA isolation and first strand cDNA synthesis
Total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen). For all tissues a single RNA extraction was performed (RNA samples L1-1, R1-1 and R2-1; Figure 2). Total RNA concentration and purity was determined using a Nanodrop ND1000 spectrophotometer (Thermo SCIENTIFIC). Total RNAs were treated with TURBO™ DNase (Ambion) according to the manufacturer’s instructions to remove all genomic DNA. A single cDNA synthesis was performed on the L1-1 sample to give the L1-1/1. Two independent cDNA syntheses were performed on the R1-1 and R2-1 RNAs (giving R1-1/1, R1-1/2 cDNA samples for R1-1 RNA and R2-1/1, R2-1/2 cDNA samples for R2-1 RNA; Figure 2) to analyse reproducibility of cDNA synthesis. For all samples (L1-1, R1-1 and R2-1), 423 ng total RNA was used to make cDNA samples. First strand cDNA synthesis was performed using the SuperScript® VILO™ cDNA Synthesis Kit (Invitrogen) following the manufacturer’s instructions in a final volume of 20 μl. The VILO™ reaction mix includes random primers to make cDNAs. The final cDNA products were diluted 10-fold prior to use in real-time PCR.
respective verified cloned amplicons, spanning five orders of magnitude. Based on the Cp values for all dilution points in a series, a standard curve was generated using linear regression and the slope. qBase calculates the gene specific PCR amplification efficiency using the following equation: Efficiency % = \((10^{-\left(\frac{1}{slope}\right)} - 1) \times 100\%\) [38].

Real-time PCR

Real-time PCR amplification reactions were performed in 384-well plates in a Lightcycler480 (Roche). Each reaction contained 2.5 μl 10-fold diluted cDNA template, 300 nM of each primer, and 1x LightCycler® 480 SYBR Green I Master (Roche), in a final volume of 10 μl. All reactions were carried out in duplicate for each cDNA sample. As a control for genomic DNA contamination, an equivalent amount of total RNA without reverse transcription was tested for each sample per gene. A no-template control (NTC) was also included in each run for each gene. This experiment was repeated two times in independent runs for all selected genes per plant per tissue. The thermal profile of the reaction was 95°C for 5 min activation and denaturation, followed by 45 cycles of 95°C for 10 sec, and 59°C for 10 sec. Finally, a dissociation curve was generated by increasing temperature starting from 56°C to 95°C to determine the specificity of the reactions. The crossing cycle number (Cp) was automatically determined for each reaction by the LightCycler480 SW 1.5 software with default parameters using the second derivative method.

Determination of reference gene expression stability using geNorm, NormFinder and BestKeeper

a) geNorm

Cp values of all samples were exported to Excel, ordered for use in qBase software and transformed to relative quantities using the gene-specific PCR amplification efficiency [49]. These relative quantities were then exported to geNorm (version 3.5) to analyse gene expression stability. The approach of reference gene selection implemented in geNorm relies on the principle that the expression ratio of two ideal reference genes should be identical in all samples, independent of the treatment, condition, or tissue type. Increasing variations in the expression ratio between two genes correspond to lower expression stability across samples. geNorm determines the level of pairwise variation for each reference gene with all other reference genes as the standard deviation of the logarithmically transformed expression ratios. In this way, the reference gene expression stability measure (M value) is calculated as the average pairwise variation of a particular gene with all other control genes included in the analysis [7]. Low M values characterise genes with the most stable expression. Sequential elimination of the least stable gene (highest M value) generates a ranking of genes according to their M values and results in the identification of the genes with the most stable expression in the samples under analysis. geNorm was also used to estimate the normalisation factor (NF_n) using n multiple reference genes, by calculating the geometric mean of the expression levels of the n best reference genes [7]. The optimisation of the number of reference genes starts with the inclusion of the two genes with the lowest M value, and continues by sequentially adding genes with increasing values of M. Thus, geNorm calculates the pairwise variation \(V_{i/n} / V_{n+1}\) between two sequential normalisation factors NF_i and NF_{n+1} containing an increasing number of reference genes [7,36]. A large variation means that the added gene has a significant effect on the normalisation and should preferably be included for calculation of a reliable normalisation factor. Ideally, extra reference genes are included until the variation \(V_i/V_{n+1}\) drops below a given threshold. Vandesompele and colleagues recommended a threshold of 0.15, although this threshold should not be taken as too strict of a cut-off [36].

b) NormFinder

For each gene, the average Cp value of each duplicate reaction was converted to relative quantity data as described for geNorm, to calculate the stability value with NormFinder program [50]. The NormFinder reference tool was applied to rank the candidate reference gene expression stability for all samples with no subgroup determination. According to the analysis, the lowest stability value will be top ranked.

c) BestKeeper

The average Cp value of each duplicate reaction is used (without conversion to relative quantity) to analyse the stability value of studied genes [41]. BestKeeper creates a pairwise correlation coefficient between each gene and the BestKeeper index. This index is the geometric mean of the Cp values of all candidate reference genes grouped together. BestKeeper also calculates standard deviation (S.D) of the Cp values between the whole data set. The gene with the highest coefficient of correlation with the BI indicates the highest stability.

qBase analysis

First the acquired Cps for each gene from LightCycler for all samples were exported to Excel, then ordered for use in qBase software [49] and imported to qBase. The relative gene expression analysis of the target gene was measured using gene-specific efficiency acquired from dilution series and selected reference genes for normalisation [49]. qBase performs relative quantification using a modified delta-Ct method with the possibility to adjust for PCR efficiency and to use multiple reference genes for normalisation. The algorithm of qBase for
calculation of relative quantities selecting different reference genes and specific efficiencies has four steps: 1) calculation of the average Cp value for all replicates of the same gene/sample combination within a given run, 2) transformation of mean Cp value into relative quantity using the gene specific PCR efficiency, 3) calculation of the normalisation factor and 4) calculation of the normalised relative quantity for gene of interest for each sample [49]. The relative expression of target genes for all samples can be collected from results in the qBase menu bar.

List of abbreviations

qRT-PCR: quantitative real-time reverse transcriptase polymerase chain reaction; ACT: actin; TUBB: β-tubulin; GADPH: glyceraldehyde-3-phosphate dehydrogenase; H3: histone H3; EF: elongation factor 1-α; rRNA: 18S rRNA; NADHD: nicotinamide adenine dinucleotide dehydrogenase; 1-FEHII: fructan 1-exohydrolaseII; NTC: no-template control; Cp: crossing point cycle number; Tm: melting temperature; R2: correlation coefficient; S: D: standard deviation; BI: BestKeeper Index; TIIA: alpha-tubulin; UBQ: ubiquitin; G6PD: glucose-6-phosphate dehydrogenase.

Additional file 1: A partial segment alignment of FEHiloand FEHilb amplicon used for qRT-PCR and position of exon and intron in FEHilogenes. 

Click here for file [http://www.biomedcentral.com/content/supplementary/1471-2199-11-S1.PDF].

Acknowledgements

This study was financially supported by the Institute for Agricultural and Fisheries Research (ILVO) in Ghent, Belgium, and the Ministry of Science, Research and Technology of Iran. Special thanks are due to Tom Ruttink for revising the manuscript and to Miriam Levenson for correcting the English.

Author details

1Department of Plant Production, Faculty of Bioscience Engineering, Couperus Links 653, Gent University, 9000 Gent, Belgium. 2Institute for Agricultural and Fisheries Research (ILVO), Burg Van Gansberghelaan 115 bus 1, 9830 Merelbeke, Belgium. 3Institute for Agricultural and Fisheries Research (ILVO), Burg Van Gansberghelaan 96 bus 1, 9820 Merelbeke, Belgium. 4Department of Plant Biotechnology and Genetics, Faculty of Sciences, Gent University, KL Ledeganckstraat 35, 9000 Ghent, Belgium.

Authors’ contributions

AM performed all the experimental procedures, data analysis, and drafted the manuscript. EVB and MDL discussed the paper and assisted with manuscript revision. All authors read and approved the final manuscript.

Received: 24 July 2009
Accepted: 15 February 2010 Published: 15 February 2010

References

1. Bustin SA. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. J Mol Endocriol 2002, 29:23-39.
2. Ginzinger DG. Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. Exp Hematol 2002, 30:503-512.
3. Bustin SA, Dorudi S. Molecular assessment of tumour stage and disease recurrence using PCR-based assays. Mol Med Today 1998, 4:389-396.
4. Bustin SA. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J Mol Endocriol 2000, 25:169-193.
5. Gachon C, Mingam A, Charrier B. Real-time PCR: what relevance to plant studies? J Exp Bot 2004, 55:1445-1454.
6. Lovdal T, Lillo C. Reference gene selection for quantitative real-time PCR normalisation in tomato subjected to nitrogen, cold, and light stress. Anal Biochem 2009, 387:238-242.
7. Vandesompole J, De Preter K, Patryn F, Poppe B, Van Roy N, De Paepe A, Spelman F. Accurate normalisation of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002, 3:1-11.
8. Radonic A, Thulke S, Mackay IM, Landt O, Siergert W, Nitsche A. Guideline to reference gene selection for quantitative real-time PCR. Biotechnol Bioprocess Eng Commun 2004, 23:856-862.
9. Huggett J, Dheda K, Bustin S, Zumla A. Real-time RT-PCR normalisation: strategies and considerations. Genes Immun 2005, 6:279-284.
10. Gutierrez L, Mauriat M, Guinlin S, Pellow J, LeFevbre JF, Louvet R, Rusterucci C, Moritz T, Guerineau F, Bellini C, Van Wuytswinkel O. 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 2008, 6:609-618.
11. Roberfroid MB. Introducing inulin-type fructans. Br J Nutr 2005, 93:513-525.
12. Madrigal L, Sangronis E. Inulin and derivates as key ingredients in functional foods. ALAN 2007, 5:37-39.
13. Heyer AG, Lloyd JR, Kossmann J. Production of modified polymeric carbohydrates. Curr Opin Biotechnol 1999, 10:169-174.
14. Morri A, Amaducci MT, Pritoni G, Venture G. Growth, fructan yield, and quality of chicory (Cichorium intybus L) as related to photosynthetic capacity, harvest time, and water regime. J Exp Bot 2005, 56:1389-1395.
15. Van Laere A, Ende Van Den W. Inulin metabolism in dicots: chicory as a model system. Plant Cell Environ 2002, 25:803-813.
16. Ende Van den W, Michels A, Van Wonterghem D, Clevers S, De Roover J, Van Laere A. Defoliation induces 1-FEH II (fructan 1-exohydrolase II) in wiltlocчив rooty clones. Cloning and purification of two isoforms (1-FEH Ia and 1-FEH Iib). Mass fingerprinting of the 1-FEH II enzymes. Plant Physiol 2001, 126:1186-1195.
17. De Roover J, De Winter M, Van Laere A, Timmermans JW, Ende Van Den W. Purification and properties of a second fructan exohydrolase from the roots of Cichorium intybus L. Physiol Plant 1999, 105:26-34.
18. Ende Van den W, Mintens A, Spekers H, Onuoha AA, Van Laere A. The metabolism of fructans in roots of Cichorium intybus L. during growth, storage and forcing. New PhytoL 1996, 132:555-563.
19. Ende Van den W, Michels A, De Roover J, Vertaert P, Van Laere A. Cloning and functional analysis of chicory root fructan 1-exohydrolase I (1-FEH I): a vacuolar enzyme derived from a cell-wall invertase ancestor? Mass fingerprint of the 1-FEH I enzyme. Plant J 2000, 24:447-456.
20. Zhou JW, Sun MX, Yang HY. Single-embryo RT-PCR assay to study gene expression dynamics during embryogenesis in Arabidopsis thaliana. Plant Mol Biol Rep 2002, 20:19-26.
21. Dean JD, Goodwin PH, Hsiang T. Comparison of relative RT-PCR and northern blot analyses to measure expression of b-1, 3-glucanase in Nicotiana benthamiana infected with Colletotrichum destructivum. Plant Mol Biol Rep 2002, 20:347-356.
22. Legrand S, Hendriks T, Hillert J, Quillet MC. Characterization of expressed sequence tags obtained by SSH during somatic embryogenesis in Cichorium intybus L. BMC Plant Biol 2007, 7:27.
23. Beikler T, Peters U, Prior K, Eisenacher M, Flemmig TF. Gene expression in periodontal tissues following treatment. BMC Med Genomics 2008, 1:30.
24. Ramírez-Gómez F, Ortiz-Pineda PA, Rivera-Cardona G, García-Arárras JE. LPS-Induced Genes in Intestinal Tissue of the Sea Cucumber Holothuria glaberrima. PLoS One 2009, 4(7):e6178.
25. Zhou X, Qi FM, Scharf ME. Social exploitation of hexamerin: RNAi reveals a major caste-regulatory factor in termites. Proc Natl Acad Sci USA 2006, 103(12):4499-504.
26. Strukenbaumer SR, Kille P. Control genes in quantitative molecular biological techniques: the variability of invariance. Comp Biochem Physiol B Biochem Mol Biol 2001, 130:281-289.
27. Bezier A, Lambert B, Bailliel F. Study of defense-related gene expression in grapevine leaves and berries infected with Botrytis cinerea. Eur J Plant Pathol 2002, 108:111-120.

28. Kim BR, Nam HY, Kim SJ, Kim SI, Chang YJ. Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. Biotechnol Lett 2003, 25(21):1869-72.

29. Thomas C, Meyer D, Wolff M, Himber C, Alioua M, Steinmetz A: Molecular characterization and spatial expression of the sunflower ABP1 gene. Plant Mol Biol 2003, 52:1025-1036.

30. Brunner AM, Yakovlev IA, Strauss SH. Validating internal controls for quantitative plant gene expression studies. BMC Plant Biol 2004, 18:14-14.

31. Nicot N, Hausman JF, Hoffmann L, Evers D: Housekeeping gene selection for real-time RT-PCR normalisation in potato during biotic and abiotic stress. J Exp Bot 2005, 56:2907-2914.

32. Oztur ZN, Talamé V, Deyholos M, Michalovz CB, Galbraith DW, Gouzukirmizi N, Tuberosa R, Bohnert H: Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley. Plant Mol Biol 2002, 48:551-73.

33. Williams TD, Ginsberg K, Minchin SD, Chipman J: A DNA expression array to detect toxic stress response in European flounder (Platichthys flesus). Aquat Toxicol 2003, 65:141-157.

34. Czechowski T, Stitt M, Altman T, Udvardi MK, Scheible WR: Genome-wide identification and testing of superior reference genes for transcript normalisation in Arabidopsis. Plant physiol 2005, 139(1):5-17.

35. Michels A, Van Laere A, End Van den W, Tucker M: Expression analysis of a chicory fructan 1-exohydrolase gene reveals complex regulation by cold. J Exp Bot 2004, 55:1325-1333.

36. geNorm software manual update. http://medgen.ugent.be/~jvdesomp/genorm/.

37. NormFinder software manual update. http://www.mdl.dk/publications/normfinder.htm.

38. BestKeeper software manual update. http://gene-quantification.com/bestkeeper.html#download.

39. Wong ML, Medrano JF: Real-time PCR for mRNA quantitation. BioTechniques 2005, 39:75-85.

40. Nolan T, Hands RE, Bustin SA: Quantification of mRNA using real-time RT-PCR. Nat Protocols 2006, 1:1559-1582.

41. Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP: Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper-Excel-based tool using pairwise correlations. Biotechnol Lett 2004, 26:509-515.

42. Mukesh J, Nijhawan A, Tragi AK, Khurana JP: Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. Biochem Biophys Res Commun 2006, 345:646-651.

43. Jian B, Liu B, Bi Y, Hou W, Wu C, Han T: Validation of internal control for gene expression study in soybean by quantitative real-time PCR. BMC Mol Biol 2008, 9:59.

44. Marino ER, Borges AA, Perez AB, Perez JA: Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. BMC Plant Biol 2008, 8:131.

45. Hibberler S, Schanack JP, Becke S: Housekeeping genes for quantitative expression studies in the three-spined stickleback Gasterosteus aculeatus. BMC Mol Biol 2008, 9:18.

46. Cruz F, Kalaoun S, Nobile P, Colombo C, Almeida J, Barros LMG, Romano E, Grossi-de-Sa M, Vaslin M, Alves-Ferreira M: Evaluation of coffee reference genes for relative expression studies by quantitative real-time RT-PCR. Mol Breed 2009, 23:607-616.

47. Argyropoulos D, Psaila C, Spyropoulos CG: Generic normalisation method for real-time PCR: application for the analysis of the mannanase gene expressed in germinating tomato seeds. FEBS J 2006, 273:770-777.

48. Schmitting TD, Zakrajsek BA: Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. J Biochem Biophys Methods 2000, 46:69-81.

49. Hellermann JJ, Mertier G, De Paeppe A, Speleman F, Vandesompele J: qBase: relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol 2007, 8:R23.

50. Andersen CL, Jensen JL, Orntoft TF: Normalisation of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalisation, applied to bladder and colon cancer data sets. Cancer Res 2004, 64:5245-5250.

doi:10.1186/1471-2199-11-15

Cite this article as: Maroufi et al.: Validation of reference genes for gene expression analysis in chicory (Cichorium intybus) using quantitative real-time PCR. BMC Molecular Biology 2010 11:15.