Gene expression studies in different genotypes of an ectomycorrhizal fungus require a high number of reliable reference genes

Non Peer-reviewed author version

DOI: 10.7287/peerj.preprints.2125v1
Handle: http://hdl.handle.net/1942/22507

Gene expression studies in different genotypes of an ectomycorrhizal fungus require a high number of reliable reference genes.

Joske Ruytinx*, Tony Remans and Jan V. Colpaert

Hasselt University, Centre for Environmental Sciences, Environmental Biology, Agoralaan building D, 3590 Diepenbeek, Belgium

*for correspondence:

joske.ruytinx@uhasselt.be; phone: +32 11 268377; fax: +32 11 268301

Running title: Gene expression normalization in mycorrhiza

Abstract

Quantitative reverse transcription PCR (qRT-PCR) has become the standard technique for the expression analysis of a set of chosen genes of interest. The accuracy and reliability of qRT-PCR measurements strongly depends on the normalization with appropriate endogenous reference genes. In this study a set of candidate reference genes for the use in gene expression studies of a basidiomycete fungus, *Suillus luteus*, exposed to toxic concentrations of zinc or cadmium was identified, evaluated and validated. Seven candidate genes were selected from cDNA-AFLP as stably expressed and the algorithms geNorm and Normfinder were used to evaluate these genes alongside the traditionally used housekeeping genes (actin, tubulin) in different *S. luteus* isolates. The use of several *S. luteus* isolates revealed that each isolate has its own most stably expressed set of reference genes, regardless of the metal treatments, *in casu* metal exposures. Metal treatments had only a minor impact on the expression of the candidate reference genes. The validated reference genes outperform the in fungal research commonly used single, arbitrary chosen ("housekeeping") genes in terms of reliability, and have the potential to be suitable reference genes when studying the effect of other environmental factors. A relatively high number of reference genes is required to correct for intraspecific variability when studying natural populations.

Keywords

qRT-PCR normalization, gene expression, heavy metal tolerance, intraspecific variation, *Suillus*

1. Introduction

Gene expression studies have become indispensable in elucidating fungal development or response to environmental changes (e.g. Plett *et al.*, 2012; Doré *et al.*, 2015; Henke *et al.*, 2016; Hlozkova *et al.*, 2016). Expression patterns are a reflection of immediate cellular responses and provide a stepping stone to the identification of the biochemical pathways involved. For expression profiling of a limited selection of genes, quantitative reverse transcription PCR (qRT-PCR), with its high sensitivity and specificity, is the technique of choice over more conventional methods (e.g. Northern blotting; Bustin, 2002). However, qRT-PCR expression profiles are only meaningful when they are corrected for technical variability and for differences in overall transcriptional activity between different developmental stages, treatments, individuals, ecotypes and species. Technical variability can be estimated by measurement of an external spike of commercially available mRNA. Differences in transcriptional activity can partly be estimated by total RNA level. Endogenous reference genes can correct for all possible variability at least when they show an equal expression level in all experimental conditions (Huggett *et al.*, 2005). The first commonly used reference genes were the so-called housekeeping genes. These genes are involved in basic cellular processes and were assumed to be constitutively and stably expressed. However, this assumption was contradicted in several studies (Thellin *et al.*, 1999; Glare *et al.*, 2002). Beside the traditional “housekeeping” genes, rRNA (e.g. 18S) expression level is often used to normalize qRT-PCR data. Nevertheless, most of the time it is impossible to measure both rRNA and a transcript of interest in the same cDNA dilution because of the high abundance of rRNA in total RNA samples. Moreover, rRNA and mRNA transcription depend on different types of polymerases of which the activity is not always affected similarly by development or external factors (Huggett *et al.*, 2005). In fact there is no single gene, whether it is a “housekeeping” gene or ribosomal, that is stably expressed in every experimental

condition and therefore normalization to multiple reference genes that are validated under the particular experimental condition, is needed (Guénin *et al.*, 2009).

In plant, animal and human research, there are many publications available that describe the identification and evaluation of multiple reference genes for normalization of qRT-PCR data (e.g. Remans *et al.*, 2008; Ledderose *et al.*, 2011; Vensentini *et al.*, 2012). In fungal research these kind of studies are not yet common practice and restricted to model species or species with a high medical or economical importance (e.g. Nailis *et al.*, 2006; Teste *et al.*, 2009; Hacquard *et al.*, 2011; Vieira *et al.*, 2011). Here, we report the identification and validation of a large set of reference genes for the normalization of qRT-PCR data in different *S. luteus* isolates upon Cd and Zn exposure by using two algorithms, geNorm (Vandesompele *et al.*, 2002) and Normfinder (Andersen *et al.*, 2004).

Suillus luteus is a typical ectomycorrhizal basidiomycete with a pioneer lifestyle. For many physiological and functional traits, phenotypic variation within mycorrhizal species (i.e. intraspecific variation) can be of the same magnitude as that among species (Colpaert *et al.*, 2005; Johnson *et al.*, 2012; Branco *et al.*, 2015). Adaptive tolerance of *S. luteus* to one or more heavy metals has developed in populations thriving on metal-contaminated soils. Metal-tolerant isolates protect their host plant from metal toxicity, but the molecular mechanisms of this protection remain unknown (Colpaert *et al.*, 2011). From an ecological point of view, *S. luteus* is an excellent species to explore evolutionary adaptation and population dynamics. The identification of reliable reference genes is crucial to enable accurate gene expression studies as an essential part of functional population genetics in relation to metal tolerance in this species. In addition, the identified set of reference genes could be useful for data normalization with respect to other techniques (e.g. RNAseq, micro-array, in situ hybridization) and could serve as a starting point to identify reliable reference genes for studying gene expression in other biological processes or in related taxa.

Objectives of the current study are dual. Firstly, the study was designed to identify reliable reference genes for the normalization of gene expression data of three *S. luteus* genotypes showing a contrasting metal sensitivity and to evaluate to which extent the results of this pilot experiment could be used in a broader experimental context including multiple *S. luteus* genotypes. Secondly, with this study we want to stress the importance of proper normalization of gene expression data derived from natural populations of fungi and provide an example or guide to set up a qRT-PCR experiment with reliable reference genes, including reference gene selection, with a minimal effort of time and resources.

2. Materials and Methods

2.1 Fungal material, growth conditions and metal exposure

A zinc tolerant (UH Slu Lm8), a zinc/cadmium tolerant (UH Slu Lm2) and a zinc/cadmium sensitive (UH Slu P4) isolate of *S. luteus* (L.:Fr.) were used to assess stability of selected candidate reference genes in a pilot experiment. Isolate reference labels are further abbreviated to respectively Lm8, Lm2 and P4. In a subsequent experiment seven zinc tolerant isolates (UH Slu Lm8, Ls1, Ls4, OF3, OF8, DS10 and Lc2) and seven zinc sensitive isolates (UH Slu P4, P8, P13, MM4, HH19, HR1, and MG4) were included. All isolates were individually cultured on cellophane-covered solid modified Fries medium (28 mM glucose, 5.4 mM ammonium tartrate, 1.5 mM KH_2PO_4, 0.4 mM $MgSO_4 \cdot 7H_2O$, 5 µM $CuSO_4 \cdot 5H_2O$, 20 µM $ZnSO_4 \cdot 7H_2O$, 0.1 µM biotin, 0.5 µM pyridoxine, 0.3 µM riboflavin, 0.8 µM nicotinamide, 0.7 µM p-aminobenzoic acid, 0.3 µM thiamine, 0.2 µM Ca-pantothenate and 0.8% agar; pH-adjusted to 4.8) as described by Colpaert *et al.* (2004). One-week old fungal colonies were mixed with a kitchen blender in 150 ml liquid modified Fries medium and incubated in Erlenmeyer flasks on a shaking incubator in a climate room at 23°C. Fries medium was changed every three days. After a week, 1 g of spherical mycelia was transferred to a petri dish containing 30 ml modified liquid Fries medium supplemented with 200 µM $ZnSO_4$, 1000 µM $ZnSO_4$, 4.5 µM $CdSO_4$, 9 µM $CdSO_4$ or control Fries medium (20 µM Zn, 0 µM Cd) and incubated while shaking for 48h at 23°C. Metal concentrations were chosen to cause a mild metal stress condition. Metal exposure was done in triplicate. Fungal mycelia were flash frozen in liquid nitrogen as aliquots of 200 mg. Aliquots were stored at -70°C.

2.2 RNA isolation, quality control and cDNA synthesis

Frozen spherical mycelia (200 mg) were thoroughly ground in liquid nitrogen using a mortar and pestle. Total RNA was extracted from the ground tissue using the RNeasy Plant mini kit (Qiagen). Three biological replicates were included in the pilot experiment, five in the

subsequent experiment. RNA quality was assessed with the Agilent-2100 Bioanalyzer and RNA 6000 NanoChips (Agilent Technologies). To assure absence of DNA, RNA samples were analysed on an agarose gel. The RNA concentration was determined on a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies), all RNA samples were adjusted to the same concentration and measured again. One µg total RNA was used in a Quantiscript Reverse Transcription reaction (Qiagen), according to the manufacturer's instructions which includes a genomic DNA elimination step and makes use of random hexamer priming. The cDNA was diluted 10x in a tenfold dilution of TE buffer (1mM Tris-HCl, 0.1mM EDTA, pH 8.0) and stored at -20°C.

2.3 Real-time PCR and data-analysis

Candidate reference genes whose expression remained apparently invariable were selected from cDNA-AFLP expression data generated by Muller *et al.* (2007) and Ruytinx *et al.* (2011) and evaluated together with two general reference genes (*ACT1* and *TUB1*). Primer sequences (Table 1) with a melting temperature of 60 +/-1 °C were designed (using Primer 3; Rozen & Skaletsky, 2000) to yield PCR-amplicons of approximately 100 bp. *Suillus luteus* *ACT1* and *TUB1* primer sequences were based on gene sequences (AF155930 and AY112730) of the closely related *S. bovinus*. Real-time PCR was performed in an optical 96-well plate with an ABI PRISM 7500 sequence detection system (Applied Biosystems) and fast cycling conditions (20 s at 95°C, 40 cycles of 3 s at 95°C and 30 s at 60°C). Each 10 µl reaction contained fast 2x FAST SYBR Green Master Mix (Applied Biosystems), 300 nM gene-specific forward and reverse primer and 2.5 µl cDNA. "No template controls" contained 2.5 µl RNase free water instead. A melting curve was generated to assure specificity of amplification. Primer efficiencies were calculated for each *S. luteus* isolate, on a standard curve (Cq vs. Log(dilution)) that was generated through a fourfold dilution series of a pooled sample over at least four dilution points, measured in triplicate. Expression levels were

calculated using the standard curve which takes into account primer efficiency (E). If E approximates two (100%) and is comparable for different isolates the formula $2^{-\Delta Cq}$ may be used to calculate expression values. Before input into geNorm (Vandesompele *et al.*, 2002) or NormFinder (Andersen *et al.*, 2004) data were expressed relative to the sample with the highest expression level.

Expression of two genes of interest, a heat shock protein (*HSP70*, GR976103) and a Mn-superoxide dismutase (Mn-*SOD*, AM085202), was measured by qRT-PCR as described above in a sub-set of samples. The *HSP70* was studied after Cd exposure in a Cd-tolerant (Lm2) and sensitive (P4) isolate; the Mn-*SOD* after Zn exposure in a Zn-tolerant (Lm8) and sensitive (P4) isolate. The products of these genes are involved in general stress response and their gene expression is expected to change upon Cd (*HSP70*) or Zn (*Mn-SOD*) treatment, at least for some *S. luteus* isolates. Primer sequences were taken from Ruytinx *et al.* (*HSP70*, 2011) and Muller *et al.* (Mn-*SOD*, 2007). Calculation of expression levels was done using the formula $2^{-\Delta Cq}$. The data were normalized in five different ways; using (a) geNorm output, (b) Normfinder output, (c) actin, (d) tubulin and (e) the most stably expressed reference gene only. A normalization factor (NF) for each sample was calculated as the geometric mean of the expression value of the respective reference genes. Expression levels were rescaled relatively to the control (no metal treatment, sensitive *S. luteus* isolate). Statistical analysis was done using the non-parametric Kruskall-Wallis test.

Finally, the stability of the reference genes defined as most stable by the pilot experiment was assessed in an expanded group of *S. luteus* isolates to assess the value of the obtained results in a broader experimental context. Gene expression was measured in seven Zn-tolerant and seven Zn-sensitive isolates after 48h exposure to 200 or 1000 µM Zn and in control condition. Expresssion levels were expressed relatively to the sample with the highest expression using the formula $2^{-(\Delta Cq)}$ with $\Delta Cq = Cq_{sample} - Cq_{minimal}$. Five biological replicates were included for

each isolate and condition. Reference gene stability was assessed using geNorm. Initially, the three genes identified as most stable by the pilot experiment assessing stability in the same conditions for two contrasting isolates Lm8 and P4 (i.e. *TUB1*, *AM085168* and *AM085296*) were analysed. Additional reference genes were added one by one, in order of their ranking in the pilot experiment, until stability criterion was reached.

3. Results

3.1 Specificity, efficiency and expression levels

The specificity of each primer pair was assessed by the melting curve assay following the PCR. The presence of a single peak in the melting curve analysis was confirmed for all primer pairs with the exception of GR975706. This primer pair generated two peaks in the melting curve analysis of some samples and one primer dimer peak in the "no template control". Because of primer dimers GR975706 was excluded from further analysis.

A fourfold dilution series of a pooled sample of each isolate was used to prepare a standard curve from which primer efficiency was calculated using the formula $E = 10^{-1/SLOPE}$. All primer pairs except one resulted in a sufficient amplification (Table 2). The expression of GR975713 was too low to generate a standard curve for all *S. luteus* isolates. Amplification products in the undiluted samples only exceeded the threshold value starting from cycle 37 and this gene therefore was eliminated from further analysis.

Expression levels of the remaining candidate reference genes were measured as Cq values from the three different *S. luteus* isolates after Cd and Zn exposure. For each isolate three biological replicates were included. Figure 1 shows the distribution of the Cq values for all genes. When considering all samples (Fig. 1a) Cq's are distributed from 14.31 (TDFC) to 38.31 (AM085168) and median varies from 20.73 (TDFC) to 25.14 (GR975621). The range of distribution of Cq's for most of the candidate reference genes is comparable. However, when analysing sample sub-sets (Cd exposure for Cd-tolerant and Cd-sensitive isolate, Fig. 1c; Zn exposure for Zn-tolerant and Zn-sensitive isolate, Fig. 1b) a considerable variability in range of Cq value was observed for the 7 genes. The smallest range was measured for GR975621 in the Cd sub-set (Fig. 1c) and for AM085296 in the Zn sub-set (Fig. 1b), indicating a more stable expression of these genes in the respective sub-sets.

3.2 Expression stability analysis

In order to find the most suitable reference genes for normalization of gene expression data in different *S. luteus* isolates after Cd and/or Zn exposure two different algorithms, geNorm and Normfinder were used.

geNorm ranks candidate reference genes based on their average expression stability value M. The M value is the average pairwise variation of a particular gene with all other candidate reference genes, and is calculated for all genes in a first step. In subsequent steps the least stable genes (highest M values) are excluded one by one and new M values are calculated. When considering the whole data-set, no combination of candidate reference genes had an M value below 1.5 (Fig. 2a; Cd+Zn) and therefore none of them fulfils the criterion for high expression stability (M<1.5) proposed by Vandesompele *et al.* (2002). Exclusion of one of the two metal treatments does not affect M values and the ranking is only subtly changed (Table 3). In contrast, exclusion of isolates from the data-set causes a decrease of M values (Fig. 2b+c) and the stability criterion is now reached for some reference gene combinations. However, a comparative analysis of data-sets including only control conditions for different combinations of isolates shows that M values are not depending on the number of isolates included but rather on the combination of isolates (Fig. 2d). When considering only Lm2 and Lm8, M-value based on the 7 candidate reference genes even drops below 1.5 and therefore all candidate reference genes seem to be expressed in a relatively stable manner. For all other tested groups of isolates, whether they consist of three or two members, similar M-values were found when comparing the same amount of remaining candidate reference genes. Furthermore, ranking of the genes differed significantly depending on the isolate or isolate combinations (Table 3). GR975621 is consistently top-ranked in the isolates P4 and Lm2 whereas it is only ranked in fifth (Zn, Cd and Zn) or sixth (Zn) position for Lm8. On the other hand *ACT1* is the least stable gene in Lm8 and Lm2 whereas it is relatively stable in P4.

A normalization factor NF_n for each sample can be calculated as the geometric mean of the expression values of the n reference genes with the lowest M value. The optimal number of reference genes required for accurate normalization was determined by calculating the pairwise variation $V_{n/n+1}$ between the two sequential normalization factors NF_n and NF_{n+1} for all samples (Fig.3); and measures the effect of adding further reference genes to the normalization factor (Vandesompele *et al.*, 2002). Lowest $V_{n/n+1}$ values are measured for $V_{5/6}$ in the Cd sample sub-set (P4 + Lm2; Fig. 3, $V_{5/6}$ = 0.30) as well as in the Zn sample sub-set (P4 + Lm8; Fig. 3, $V_{5/6}$ = 0.27). For the Zn sub-set there is only a small difference (0.037) between the $V_{4/5}$ and $V_{5/6}$ value. Therefore NF calculation should include at least the 4 (Zn sub-set) or 5 (Cd sub-set) candidate reference genes with the lowest M values. Inclusion of an extra reference gene will not significantly alter the normalization factor. However, since $V_{5/6}$ values never reach 0.15 threshold (cut-off value suggested by Vandesompele *et al.*, 2002) and M values for 4, 5 or 6 remaining genes are all +/- 1.8 (Fig. 2) it is advisable to include the 6 most stable candidate reference genes to maximize robustness of the normalization. Inclusion of a 7^{th} gene is not advisable because of the instability of this gene (V-value increases again after reaching a minimum, and M value = +/- 2.2). In accordance with the effect on the M values, excluding isolates from the sample set reduces $V_{n/n+1}$ values (lowest value for $V_{3/4}$ or $V_{4/5}$ when considering single isolates) and the number of reference genes needed for accurate normalization (Fig 3a +b). Nevertheless as for M-values, the number of reference genes needed for accurate normalization is not only depending on the number of isolates included but also on the specific combination. Lowest $V_{n/n+1}$ value for the combination Lm8 and Lm2 is $V_{6/7,}$ whereas for the combination P4 and Lm8 $V_{5/6}$ is the lowest (Fig. 3c).

Normfinder calculates a stability value (SV) for each candidate reference gene and ranks them according to this value (low SV = high stability). The calculation is based on a statistical model that also measures the variation between sample subgroups and provides an estimation

of the systemic variation across subgroups. Table 4 shows SV and ranking for the 7 candidate reference genes. geNorm and Normfinder provide almost the same ranking when analysing all samples (P4, Lm8, Lm2; Cd and Zn). For the Cd sub-set (P4, Lm2) ranking is slightly different but the top 4 of genes is the same. For the Zn sub-set (P4, Lm8), the two algorithms result in a completely different ranking. The two most stable genes according to geNorm, AM085168 and *TUB1*, are only placed in position 5 and 6 by Normfinder. For the control sub-sets ranking by both algorithms is different (Table 3 and supplementary table 1).

The use of minimum two reference genes is suggested by Normfinder. However, it is possible to calculate a SV based on any number of reference genes as an accumulated standard deviation (SD_{acc}; Andersen *et al.*, 2004). Indeed, the use of a gene overexpressed in one group and underexpressed in another combined with a gene showing the opposite bias might be the most stable option. Therefore, the combination of reference genes leading to the lowest SD_{acc} is recommended to calculate a normalization factor. For the Cd sub-set (P4-Lm2), lowest SD_{acc} was obtained when including the five most stable reference genes (Fig. 4a). For the Zn sub-set (P4-Lm8), the six most stable references are required to minimize SD_{acc} (Fig. 4b). geNorm suggested exactly the same set of six reference genes for this data sub-set. Also for the different control sub-sets the use of a high number of reference genes is recommended and depends mainly on the isolates included (supplementary Fig. 1).

When considering both algorithms together, the most stably expressed candidate reference gene for both Zn (P4, Lm8) and Cd (P4, Lm2) sub-sets is GR975621 (Table 3, 4). This gene is placed in the first position by geNorm and Normfinder for the Cd sub-set and in the first (Normfinder) and fourth (geNorm) position for the Zn sub-set.

3.3 Expression level of genes of interest

To show the effect of using NFs derived from different reference genes on expression data, and to illustrate the need for accurate normalization, the expression level of two genes of

interest was normalized in five different ways: according to (a) geNorm (6 top ranked genes), (b) Normfinder (5 or 6 top ranked genes for Cd or Zn sub-set respectively), (c) using the traditional "housekeeping genes" *ACT1* or (d) *TUB1* and (e) using the single most stably expressed candidate reference gene GR975621. A *HSP70* was analysed for the Cd sub-set (Fig. 5), a *Mn-SOD* for the Zn sub-set (Fig. 6). For both genes, gene expression patterns are similar to each other after normalization according to geNorm and Normfinder. Since both algorithms suggested the use of the same six reference genes for the Zn sub-set, exactly the same significant increase compared to the control is demonstrated twice for the *Mn-SOD* in the Zn sensitive *S. luteus* isolate after exposure to 200 µM Zn (Fig. 6a, b). After exposure to 9 µM Cd a significant increase of *HSP70* expression was shown for the Cd tolerant isolate (Figs 5a, b). This increase in *HSP70* expression was also detected after normalization to *ACT1* (Fig. 5c). In all other cases normalization to a single "housekeeping" gene shows a divergent pattern and may lead to misinterpretation of the data (Figs 5d, e and Figs 6c, d, e). Standard errors are smallest after normalization according to geNorm. However, this choice of reference genes results in significant differences in expression level between the two isolates after exposure to 4.5 µM Cd and 9 µM Cd (Fig. 5a) whereas normalization according to Normfinder (Fig. 5b) does not result in significant differences in expression level for both isolates.

3.4 Expanding the group of genotypes

Finally, the stability of the reference genes defined as most stable by the pilot experiment (Table 3, row 6) was assessed in an expanded group of *S. luteus* isolates exposed to Zn to assess the value of the obtained results in a broader experimental context. *TUB1,* AM085168 and AM085296 were ranked by geNorm based on their stability value M and pairwise variation $V_{n/n+1}$ was calculated. M values were all below 1.5. The least stable gene in this extended sample set was *TUB1* with a $V_{2/3}$ of 0.523, failing the $V_{n/n+1}$ <0.15 criterion

(Vandesompele et al. 2002). We therefore added a fourth gene to the analysis which was the gene determined next most stable gene in the pilot experiment (GR975621). Resulting M values and ranking are presented in Fig. 7a. Again all M values were below 1.5. The lowest $V_{n/n+1}$ value for the 14 isolates and three experimental conditions (20, 200, 1000 µM Zn) was $V_{2/3}$ which equaled 0.28 (Fig. 7b). Again, the stability criterion as set by Vandesompele *et al.* (2002) was not reached as $V_{n/n+1}$ was above 0.15. However, the pilot study showed that when studying different isolates, the variability may not allow the identification of genes with $V_{n/n+1}$ <0.15 even though low M values (<1.5) may be obtained. Here, we show that similar M and $V_{n/n+1}$ values as in the pilot study could be obtained in an experiment using an extended set of isolates, in which GR975621, AM085168 and AM085296 were selected. Therefore, the reference genes proposed in the pilot study could be useful in different experiments and genes for normalization selected using geNorm analyses after measurement of all of these genes, or starting with the three best and adding additional ones, until M values <1.5 are reached and $V_{n/n+1}$ is minimal.

4. Discussion

Whole genome expression data of *S. luteus* exposed to Zn (Muller *et al.*, 2007) or Cd (Ruytinx *et al.*, 2011) were searched for equally expressed genes. Seven gene fragments (TDFC, AM085296, AM085177, AM085168, GR975621, GR975706 and GR975713) out of 458 seemed to be steadily expressed and were selected for further analysis. Since the genome of the studied organism became only recently available (Kohler *et al.*, 2015), researchers used a genome-wide cDNA-AFLP expression analysis tool, which does not require prior sequence information, and isolated fragments were sequenced afterwards. The sequenced gene fragments are relatively short (100-500 bp) and therefore choice of optimal primer pairs for qRT-PCR is limited. Without evaluation of a second primer set, GR975706 was excluded from further analysis because of nonspecific amplification, and GR975713 because of inefficient amplification due to its low expression level. The efficiency of the primer pairs amplifying the five remaining genes and two "housekeeping" genes was considered sufficient. For one primer pair, targeting AM085177, higher than optimal PCR efficiencies (> 2.15 or 115%; table 2) were obtained for two out of the three isolates. Using SYBR technology, which is detecting all double stranded DNA, efficiencies exceeding 2.20 (120%) can be obtained due to nonspecific amplification or primer-dimers (Lutfalla & Uze, 2006). Nevertheless, melt curve analysis confirmed the absence of additional PCR products or primer-dimers for all isolates in our experiment. The difference in efficiency between the different isolates might be caused by sequence variation in the primer region; the efficiency of > 120% by allelic diversity. If the amplification efficiency of one allele equals 100% and the amplification efficiency of a variant is less, overall efficiency will exceed 100% in dikaryotic individuals which are heterozygous. High genetic diversity was reported previously for *S. luteus* (Muller *et al.*, 2007b; Ruytinx *et al.*, 2011). Therefore, we recommend to prefer most conservative candidate reference genes and to exclude high polymorphic regions if enough

candidate reference genes are present. In metabarcoding studies primers are optimized by *in silico* analysis to equally amplify the targeted sequence in all species present (Op De Beeck *et al.*, 2014; Waud *et al.*, 2014). Such an approach, redesigning primers with different amplification efficiencies in different isolates, could result in efficiencies approximating 2 for all isolates. However, a priori knowledge of the target sequence for all examined isolates is required to optimize primers by *in silico* analysis.

Stability of the five remaining genes and two "housekeeping genes" was assessed by geNorm and Normfinder algorithms. geNorm and Normfinder did not always produce identical results for our data-set. Usually both algorithms rank the same genes at the highest and lowest position except for the Zn and control sub-set (Table 3, 4, supplementary table 1). Differences in geNorm and Normfinder results have been reported by several researchers (e.g. Petit *et al.*, 2012) and are not surprising since the algorithms rely on different methods. Normfinder uses a model-based algorithm that takes into account inter- and intragroup variation. geNorm is based on pairwise variation and ranking might be artefactual because of co-regulation. Except for *ACT1* and AM085177 (nucleus protein *Coprinopsis cinerea*, actin related), there is no reason to believe that the candidate reference genes we selected are co-regulated. These two genes together were never top ranked in our analysis. Moreover, geNorm suggest the use of more than 2 genes in our Zn and Cd data sub-set, which makes the normalization more robust. $V_{n/n+1}$ cut-off value (0.15) was never reached. Until now the $V_{n/n+1}$ and M cut-off values suggested by Vandesompele *et al.* (2002) appeared to be good threshold values for most data-sets (e.g. Guénin *et al.*, 2009; Remans *et al.*, 2008). However, they seem to be inadequate when studying natural populations with their high genetic variation. To ensure and maximize robustness of normalization in case of a short fall of the suggested cut-off values we recommend to strive for a minimal $V_{n/n+1}$ value and for a NF including the maximal number of reference genes without provoking a significant increase of M value. Also SD_{acc} calculated

by the Normfinder algorithm suggest the use of a high number of reference genes for most of our data sub-sets (Fig. 4).

In our experimental design, gene expression might be influenced by metal treatment and genotype. Elevated concentrations of Zn or Cd can act as a stress factor and induce a different systemic and cellular response depending on the metal and the concentration applied (Bellion *et al.*, 2006, Cuypers *et al.*, 2011). Yet, the expression of the analysed candidate reference genes seems to be little affected by the metal treatment in *S. luteus*. Excluding one of the metals from analysis did, according to geNorm, not result in an increased stability or an altered ranking of the genes (Fig. 2, Table 3).

Different *S. luteus* isolates (i.e. genotypes) can respond differentially to metal exposure. Some of them cannot tolerate elevated concentrations whereas others evolved an adaptive tolerance for one or more metals (Colpaert *et al.*, 2004). In general, *S. luteus* shows a high intraspecific variation (Ruytinx *et al.*, 2011). This intraspecific variation is retrieved in the expression of the candidate reference genes. Exclusion of isolates from the stability analysis results in decreased M-values and an altered ranking (Fig. 2, Table 3). Each isolate seems to have its own top ranked gene set regardless the metal treatment. This is the most pronounced for Lm2 with the genes GR975621 and AM085177 consistently in position 1 and 2. The top ranked genes might therefore also be suitable for the normalization of gene expression data in other stress situations.

Identification and validation of accurate reference genes is both, time consuming and expensive. Still, it is advisable to assess the usefulness of candidate reference genes in every new experimental design (Guénin *et al.*, 2009; Murphy & Bustin, 2009). We clearly demonstrate the benefit of this kind of reference gene validation by analysing the expression profile of 2 genes of interest. Both gene products are known to be involved in alleviating heavy metal stress (Hall, 2002) and were selected because of their responsiveness to Zn or Cd

in *S. luteus* (Muller *et al.*, 2007; Ruytinx *et al.*, 2011). The expression profile of the Mn-*SOD* and the *HSP70* altered as a consequence of the way of normalization (Figs. 5+6). Generally, normalization according to geNorm and Normfinder leads to comparable expression profiles whereas normalization to a single gene (*ACT1*, *TUB1* or GR975621) leads to different, unreliable expression profiles. For the Mn-*SOD*, the wrong expression profile (normalization to *ACT1*) even shows a statistically significant difference as a result of Zn exposure in the isolate P4. An influence of metal treatment on *ACT1* and *TUB1* expression could be expected since heavy metals were shown to act upon hyphal morphology and cytoskeletal components (Pawlowska & Charvat, 2004; Tuszynska, 2006). Also for other commonly used reference genes (e.g. glyceraldehyde-phosphate-dehydrogenase (GAPDH), elongation factors (EF), etc.) variations in expression upon heavy metal treatment were demonstrated previously (Muller *et al.*, 2007; Remans *et al.*, 2008; Ruytinx *et al.*, 2011). In fact there is no single gene, whether or not it is a housekeeping gene, which can correct for all possible variability. Even the most stably expressed candidate reference gene is subjected to minimal variations in expression level and can lead to erroneous conclusions when used for normalization of gene expression data. In all organisms, normalization to a single non-validated gene can result in unreliable conclusions (Garson *et al.*, 2009).

Significant differences in expression level among *S. luteus* isolates are found by normalization according to geNorm for the *HSP70* and Mn-*SOD* genes, differences that are absent for *HSP70* in Normfinder profiles. As long as the reference genes are not co-regulated and inter- and intragroup variation are comparable we recommend normalization according to geNorm. Normalization to six genes is more robust and reduces biological variability as shown by reduced standard errors in our experiments.

The genome sequence of *S. luteus* recently became available and future RNAseq or micro-array data may uncover new, more stable reference genes. These new candidate reference

genes may be necessary to obtain reliable gene expression results in larger population studies without the need to include an overwhelming amount of reference genes or to overcome a shortage of suitable reference genes. Yet, the number of reference genes needed for accurate and reliable normalization of qRT-PCR data is not directly depending on the amount of isolates included but rather on the combination of isolates (Fig. 3c, Fig. 7 and supplementary Fig. 1). However, the need for a large number of reference genes may be inherent to gene expression analysis in species showing high intraspecific variation in general, a condition which might be expected in natural populations of fungi (Ellison *et al.*, 2011; Johnson *et al.*, 2012).

Currently, normalization to external spikes is often used in gene expression studies of natural populations because of a lack of suitable reference genes (e.g. MacNeal Rehrig *et al.*, 2011). This kind of normalization can never correct for differences in overall transcriptional activity between genotypes or individuals (Huggett *et al.*, 2005), which might be considerable in natural populations of species showing a high intraspecific variability. Therefore, to overcome misconclusions because of unreliability of gene expression data, an assessment of candidate reference genes reaching beyond the traditional housekeeping genes is necessary and highly recommended even in populations of non-model species.

Acknowledgements

We thank the Research Foundation – Flanders for sponsoring the *Suillus* research (FWO-project G.0925.10N).

References

Andersen CL, Jensen JL, Orntoft TF (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 Research*, **64**, 5245-5250.

Bellion M, Courbot M, Jacob C, Blaudez D, Chalot M (2006) Extracellular and cellular mechanisms sustaining metal tolerance in ectomycorrhizal fungi. *FEMS Microbiology Letters*, **254**, 173-181.

Branco S, Gladieux P, Ellison CE, Kuo A, LaButti K, Lipzen A, Grigoriev IV, Liao HL, Vilgalys R, Peay KG, Taylor JW, Bruns TD (2015) Genetic isolation between two recently diverged populations of a symbiotic fungus. *Molecular Ecology*, **24**, 2747-2758.

Bustin SA (2002) Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *Journal of Molecular Endocrinology*, **29**, 23-39.

Colpaert JV, Muller LAH, Lambaerts M, Adriaensen K, Vangronsveld J (2004) Evolutionary adaptation to Zn toxicity in populations of Suilloid fungi. *New Phytologist*, **162**, 549-559.

Colpaert JV, Adriaensen K, Muller LAH, Lambaerts M, Faes C, Carleer R, Vangronsveld J (2005) Element profiles and growth in Zn-sensitive and Zn-resistant Suilloid fungi. *Mycorrhiza*, **15**, 628-634.

Colpaert JV, Wevers J, Krznaric E, Adriaensen K (2011) How metal-tolerant ecotypes of ectomycorrhizal fungi protect plants from heavy metal pollution. *Annals of Forest Science*, **68**, 17-24.

Cuypers A, Smeets K, Ruytinx J, Opdenakker K, Keunen E, Remans T, Horemans N, Vanhoudt N, Van Sanden S, Van Belleghem F, *et al.* (2011) The cellular redox state as

a modulator in cadmium and copper responses in *Arabidopsis thaliana* seedlings. *Journal of Plant Physiology*, **168**, 309-316.

Doré J, Perraud M, Dieryckx C, Kohler A, Morin E, Henrissat B, Lindquist E, Zimmermann SD, Girard V, Kuo A, *et al.* (2015) Comparative genomics, proteomics and transcriptomics give new insight into the exoproteome of the basidiomycete *Hebeloma cylindrosporum* and its involvement in ectomycorrhizal symbiosis. *New Phytologist*, **208**, 1169-87.

Ellison CE, Hall C, Kowbel D, Welch J, Brem RB, Glass NL, Taylor JW (2011) Population genomics and local adaptation in wild isolates of a model microbial eukaryote. *Proceedings of the National Academy of Science USA*, **108**, 2831-2836.

Garson JA, Huggett JF, Bustin SA, Pfaffl MW, Benes V, Vandesompele J, Shipley GL (2009) Unreliable real-time PCR analysis of human endogenous retrovirus-W (HERV-W) RNA expression and DNA copy number in multiple sclerosis. *Aids Research and Human Retroviruses*, **25**, 377-378.

Glare EM, Divjak M, Bailey MJ, Walters EH (2002) Beta-actin and GAPDH housekeeping gene expression is variable and not suitable for normalizing mRNA levels. *Thorax*, **57**, 765-770.

Guénin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C, Gutierrez L (2009) Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental condition-specific, validation of references. *Journal of Experimental Botany*, **60**, 487-493.

Hacquard S, Venault-Fourrey C, Delaruelle C, Frey P, Martin F, Duplessis S (2011) Validation of *Melampsora larici-populina* reference genes for *in planta* RT-quantitative PCR expression profiling during time-course infection of poplar leaves. *Physiological and Molecular Plant Pathology*, **75**, 106-112.

Hall JL (2002) Cellular mechanisms for metal detoxification and tolerance. *Journal of Experimental Botany*, **53**, 1-11.

Henke C, Jung EM, Voit A, Kothe E, Krause K (2016) Dehydrogenase genes in the ectomycorrhizal fungus *Tricholoma vaccinum*: A role for Ald1 in mycorrhizal symbiosis. *Journal of Basic Microbiology*, **56**, 162-74.

Hlozkova K, Matenova M, Zackova P, Strnad H, Hrselova H, Hroudova M, Kotrba P (2016) Characterization of three distinct metallothionein genes of the AG-hyperaccumulating ectomycorrhizal fungus *Amanita strobiliformis*. *Fungal Biology*, **120**, 358-369.

Huggett J, Dheda K, Bustin SA, Zumla A (2005) Real-time RT-PCR normalization, strategies and considerations. *Genes Immunity*, **6**, 279-284.

Johnson D, Martin F, Cairney JWG, Anderson IC (2012) The importance of individuals: intraspecific diversity of mycorrhizal plants and fungi in ecosystems. *New Phytologist*, **194**, 614-628.

Kohler A, Kuo A, Nagy LG, Morin E, Barry KW, Buscot F, Canbäck B, Choi C, Cichocki N, Clum A, *et al.* 2015. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nature Genetics*, **47**, 410-415.

Ledderose C, Heyn J, Limbeck E, Kreth S (2011). Selection of reliable reference genes for quantitative real-time PCR in human T cells and neutrophils. *BMC Research Notes*, **4**, 427.

Lutfalla G, Uze G (2006). Performing quantitative reverse-transcribed polymerase chain reaction experiments. *Methods in Enzymology*, **410**, 386-400.

MacNeal Rehrig E, Appel HM, Schultz JC (2011) Measuring 'normalcy' in plant gene expression after herbivore attack. *Molecular Ecology Resources*, **11**, 294-304.

Muller LAH, Craciun AR, Ruytinx J, Lambaerts M, Verbruggen N, Vangronsveld J, Colpaert JV (2007) Gene expression profiling of a Zn-tolerant and a Zn-sensitive *Suillus luteus* isolate exposed to increased external zinc concentrations. *Mycorrhiza*, **17**, 571-580.

Muller LAH, Vangronsveld J, Colpaert JV (2007b) Genetic structure of *Suillus luteus* populations in heavy metal polluted and nonpolluted habitats. *Molecular Ecology*, **16**, 4728-4737.

Murphy J, Bustin SA (2009) Reliability of real-time reverse-transcription PCR in clinical diagnostics: gold standard or substandard? *Expert Review of Molecular Diagnostics*, **9**, 187-192.

Nailis H, Coenye T, Van Nieuwenborg F, Deforce D, Nelis HJ (2006) Development and evaluation of different normalization strategies for gene expression studies in *Candida albicans* biofilms by real-time PCR. *BMC Molecular Biology*, **7**, 25.

Op De Beeck M, Lievens B, Busschaert P, Declerck S, Vangronsveld J, Colpaert JV (2014) Comparison and validation of some ITS primer pairs useful for fungal metabarcoding studies. *PLOS one*, **9**, e97629.

Pawlowska TE, Charvat I (2004) Heavy-metal stress and developmental patterns of arbuscular mycorrhizal fungi. *Applied and Environmental Microbiology*, **70**, 6643-6649.

Petit C, Pernin F, Heydel JM, Délye C (2012) Validation of a set of reference genes to study response to herbicide stress in grasses. *BMC Research Notes*, **5**, 18.

Plett JM, Gibon J, Kohler A, Duffy K, Hoegger PJ, Velagapudi R, Han J, Kües U, Grigoriev IV, Martin F (2012) Phylogenetic, genomic organization and expression analysis of hydrophobin genes in the ectomycorrhizal basidiomycete *Laccaria bicolor*. *Fungal Genetics and Biology*, **49**, 199-209.

Remans T, Smeets K, Opdenakker K, Mathijsen D, Vangronsveld J, Cuypers A (2008) Normalisation of real-time RT-PCR gene expression measurements in *Arabidopsis thaliana* exposed to increased metal concentrations. *Planta*, **227**, 1343-1349.

Rozen S, Skaletsky HJ (2000) Primer 3 on the www for general users and for biologist programmers, in: Krawetz S., Misener S. (Eds.), Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp. 365-386.

Ruytinx J, Craciun AR, Verstraelen K, Vangronsveld J, Colpaert JV, Verbruggen N (2011) Transcriptome analysis by cDNA-AFLP of *Suillus luteus* Cd-tolerant and Cd-sensitive isolates. *Mycorrhiza*, **21**, 145-154.

Teste MA, Duquenne M, François JM, Parrou JL (2009) Validation of reference genes for quantitative expression analysis by real-time RT-PCR in *Saccharomyces cerevisiae*. *BMC Molecular Biology*, **10**, 99.

Thellin O, Zorzi W, Lakaye B, De Borman B, Coumans B, Hennen G, Grisar T, Igout A, Heinen E (1999) Housekeeping genes as internal standards: use and limits. *Journal of Biotechnology*, **75**, 291-295.

Tuszynska S (2006) Ni^{2+} induces changes in the morphology of vacuoles, mitochondria and microtubules in *Paxillus involutus* cells. *New Phytologist*, **169**, 819-828.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, **3**, research0034.1-0034.11

Vesentini N, Barsanti C, Martino A, Kusmic C, Ripoli A, Rossi A, L'abbate A (2012) Selection of reference genes in different myocardial regions of an in vivo ischemia/reperfusion rat model for normalization of antioxidant gene expression. *BMC Research Notes*, **5**, 124.

555 Vieira A, Talhinhas P, Loureiro A, Duplessis S, Fernandez D, Silva M, Paulo OS, Azinheira
556 HG (2011) Validation of RT-qPCR reference genes for in planta expression studies in
557 *Hemileia vastatrix*, the causal agent of coffee leaf rust. *Fungal Biology*, **115**, 891-901.
558 WaudM, Busschaert P, Ruyters S, Jacquemyn H, Lievens B (2014) Impact of primer choice
559 on characterization of orchid mycorrhizal communities using 454 pyrosequencing.
560 *Molecular Ecology Resources*, **14**, 679-699.

Figure 1. Distribution of Cq values for the candidate reference genes in different experimental sets. Boxes indicate the first and third percentile, black centre line indicates the median, whiskers show largest/smallest value that falls within a distance of 1.5 times the interquartile range; outliers are shown as black dots. (a) all isolates, Cd + Zn; (b) P4 and Lm8, Zn; (c) P4 and Lm2, Cd.

Figure 2. Average expression stability (M) as a function of the remaining candidate reference genes. geNorm analysis was run for (a) all isolates together after exposure to different metals, (b) after Zn exposure, (c) after Cd exposure and, (d) in control conditions for different combinations of isolates. The lower the M value, the higher the stability.

Figure 3. Pairwise variation (V) between two sequential normalization factors NF_n and NF_{n+1} to determine the optimal number of reference genes for different experimental sets. Different combinations of isolates after (a) Zn exposure, (b) Cd exposure and, (c) in control conditions were analysed.

Figure 4. Accumulated standard deviation (SDacc) as a function of the number of candidate reference genes. SDacc was analysed after (a) Cd exposure for the isolates Lm2 and P4 together and, (b) after Zn exposure for Lm8 and P4 together. The lower SDacc, the higher the overall stability of the reference gene set.

Figure 5. Relative *HSP70* expression obtained by different normalization strategies after Cd exposure. Data are normalized according to (a) geNorm, (b) Normfinder, by the "housekeeping" gene (c) *ACT1*, (d) *TUB1* or by (e) the most stably expressed candidate reference gene GR975621. (■) Cd sensitive *S. luteus* isolate P4 and (■) Cd tolerant *S. luteus*

isolate Lm2; significant differences ($p < 0.05$) as a result of Cd treatment (compared to control concentration) are indicated by *.

Figure 6. Relative Mn-*SOD* expression obtained by different normalization strategies after Zn exposure. Data are normalized according to (a) geNorm, (b) Normfinder, by the "housekeeping" gene (c) *ACT1*, (d) *TUB1* or by (e) the most stably expressed candidate reference gene GR975621. (■) Zn sensitive *S. luteus* isolate P4 and (■) Zn tolerant *S. luteus* isolate Lm8; significant differences ($p < 0.05$) as a result of Zn treatment (compared to control concentration) are indicated by *.

Figure 7. (a) Average expression stability (M) as a function of the remaining candidate reference genes and (b) pairwise variation (V) between two sequential normalization factors NF_n and NF_{n+1} to determine the optimal number of reference genes. geNorm analysis was run for seven Zn sensitive and seven Zn tolerant *S. luteus* isolates exposed to different concentrations of Zn (0, 200, 1000 µM).

Supplementary fig. 1. Accumulated standard deviation (SDacc) as a function of the number of candidate reference genes. SDacc was analysed in control conditions for different combinations of isolates. The lower SDacc, the higher the overall stability of the reference gene set.

606 **Table 1.** Description of the candidate reference genes and their primer set

Gene symbol	Genbank accession number	Annotation or blast hit	Forward primer	Reverse primer
ACT1	-	actin 1	GGCCACACGAAGCTCATTATAGAAT	CTTGACCCTAAAGTACCCTATCGAG
TUB1	-	β-tubulin	GTTCTGGGAGGTGGTTTCTGATG	ACGTATTTGTTTGACCCGATTTCATTG
TDFC	-	putative L-amino acid oxidase, *Serpula lacrymans*	TGAAGGCAAAGTCACGAATG	TTCCCGAGCATTGATGTCTAC
AM085296	[Genbank:AM085296]	no significant similarity	AAGTTGGTAATATGCGTGAACAAAGC	CGTTGAGTGATTTTCTGTCTTTGTTGC
AM085177	[Genbank:AM085177]	nucleus protein *Coprinopsis cinerea*, actin related	GGGGCGATGATAAAATCCGCAA	ATGCGCCATACCAGGTATCAAG
AM085168	[Genbank:AM085168]	predicted protein *Laccaria bicolor*, PCI domain	GGGCGTTTGACGAAGCTCATC	CGATCTCGAGGAGCTGTGTTCCA
GR975706	[Genbank:GR975706]	no significant similarity	GTCATCTCACCCTCACAACACTC	GCAGGAATGAGAGTGGGTGTG
GR975713	[Genbank:GR975713]	UBQ superfamily, conserved domain	GCAGTCTATGATGAGTCCTGAGTAAGAC	ACGTATTTGTTTGACCCGATTTCATTG
GR975621	[Genbank:GR975621]	HGTP-anticodon superfamily, conserved domain	CCGGTTTCAGTAATACAGAGTCCT	GGTCTGTTTACCTTACTTTATTATGTCC/

607

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2125v1 | CC BY 4.0 Open Access | rec: 15 Jun 2016, publ: 15 Jun 2016

608 **Table 2.** Amplification efficiencies of the candidate reference genes in the different *S. luteus* isolates measured by a fourfold dilution series

	ACT1	TUB1	TDFC	AM085296	AM085177	AM085168	GR75621
S. luteus P4	1.998	1.978	1.954	2.148	2.144	1.948	1.954
S. luteus Lm8	2.046	2.001	1.969	2.100	2.692	2.030	2.079
S. luteus Lm2	2.029	1.781	2.026	2.064	2.331	1.923	1.964

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2125v1 | CC BY 4.0 Open Access | rec: 15 Jun 2016, publ: 15 Jun 2016

609 **Table 3.** Ranking of the candidate reference genes according to their expression stability as calculated by geNorm for different experimental

610 designs

611

Isolate	Metal	least stable gene		---------->			most stable gene
P4, Lm8 and Lm2	Cd + Zn	AM085177	AM085168	ACT1	TDFC	TUB1	GR975621/AM085296
P4, Lm8 and Lm2	Cd	AM085177	ACT1	TDFC	GR975621	AM085296	AM085168/TUB1
P4, Lm8 and Lm2	Zn	AM085168	AM085177	ACT1	TDFC	GR975621	AM085296/TUB1
P4, Lm8 and Lm2	Control	AM085177	ACT1	GR975621	TDFC	AM085296	AM085168/TUB1
P4 and Lm2	Cd	AM085177	TDFC	AM085168	ACT1	TUB1	GR975621/AM085296
P4 and Lm8	Zn	AM085177	ACT1	TDFC	GR975621	AM085296	AM085168/TUB1
P4	Cd + Zn	TDFC	AM085296	AM085168	TUB1	ACT1	GR975621/AM085177
P4	Cd	TDFC	AM085296	TUB1	AM085168	AM085177	ACT1/GR975621
P4	Zn	AM085168	TDFC	AM085296	ACT1	TUB1	GR975621/AM085177
Lm8	Cd + Zn	ACT1	TDFC	GR975621	AM085177	AM085296	AM085168/TUB1
Lm8	Cd	ACT1	GR975621	TDFC	TUB1	AM085168	AM085296/AM085177
Lm8	Zn	ACT1	TUB1	GR975621	AM085168	AM085296	TDFC/AM085177
Lm2	Cd + Zn	ACT1	AM085168	TDFC	TUB1	AM085296	GR975621/AM085177
Lm2	Cd	ACT1	TUB1	AM085168	AM085296	TDFC	GR975621/AM085177
Lm2	Zn	AM085168	ACT1	TDFC	AM085296	TUB1	GR975621/AM085177

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2125v1 | CC BY 4.0 Open Access | rec: 15 Jun 2016, publ: 15 Jun 2016

612 **Table 4.** Candidate reference genes ranked by Normfinder according to their stability value

613 (SV) for different experimental designs

	P4,Lm8, Lm2		P4, Lm8,Lm2		P4, Lm8		P4, Lm2	
	Control		Cd + Zn		Zn		Cd	
	SV	ranking	SV	ranking	SV	ranking	SV	ranking
ACT1	0.941	3	1.329	5	1.072	5	0.949	3
TUB1	0.848	2	0.814	2	0.943	4	0.881	2
TDFC	0.949	4	1.053	4	0.860	2	1.106	5
AM085296	1.021	5	0.915	3	0.921	3	0.994	4
AM085177	1.766	7	1.632	7	1.966	7	1.603	7
AM085168	1.311	6	1.546	6	1.128	6	1.376	6
GR975621	0.799	1	0.422	1	0.535	1	0.490	1

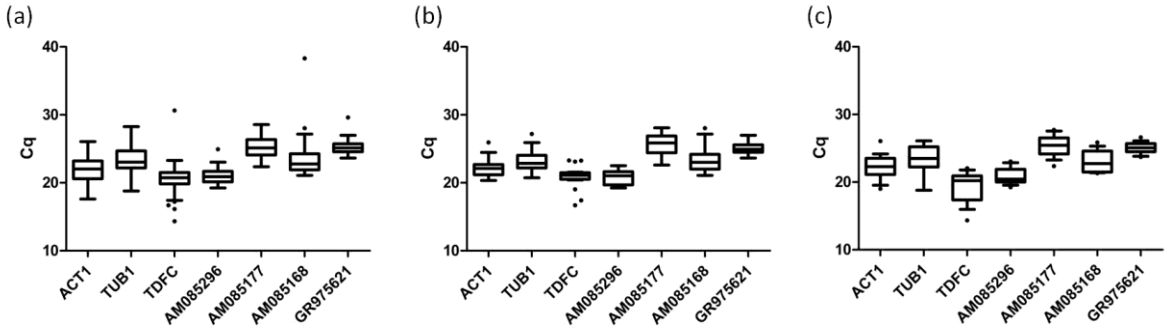
614

615 **Supplementary table 1.** Candidate reference genes ranked by Normfinder according to their
616 stability value (SV) for different experimental designs
617

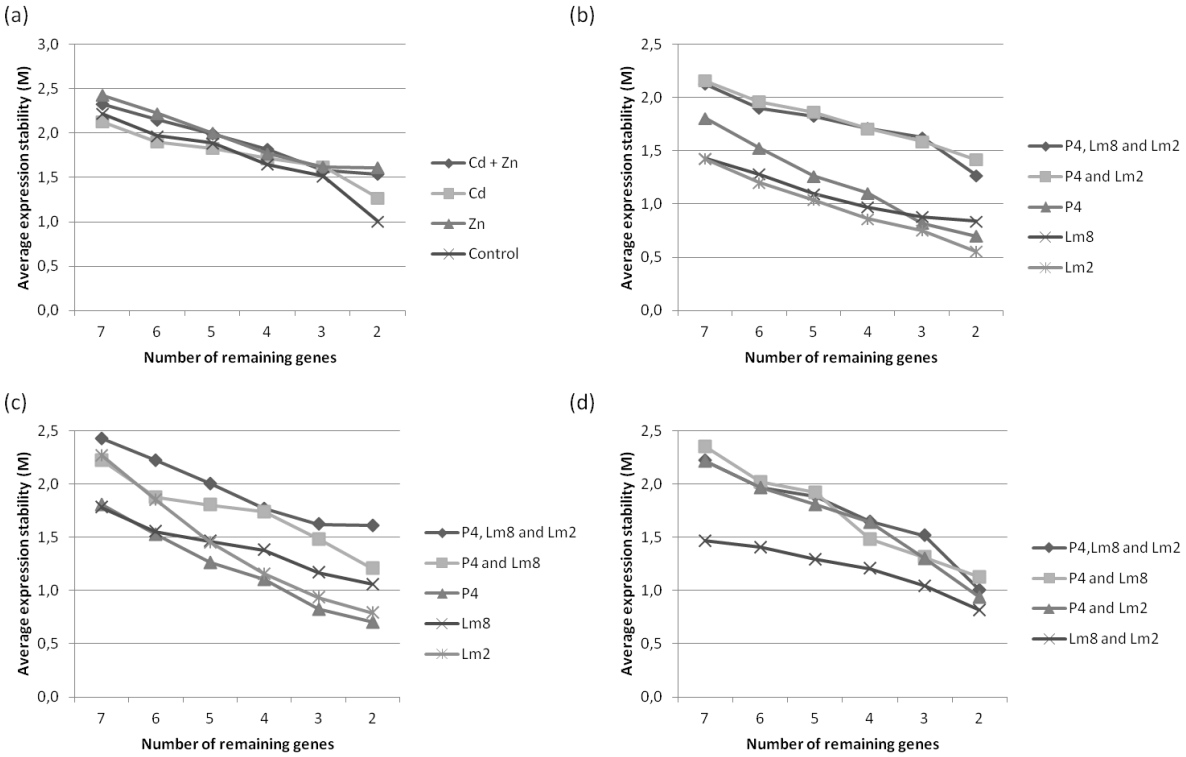
	P4,Lm8, Lm2		Lm8, Lm2		P4, Lm8		P4, Lm2	
	Control		Control		Control		Control	
	SV	ranking	SV	ranking	SV	ranking	SV	ranking
ACT1	0.941	3	0.840	5	0.900	2	1089	5
TUB1	0.848	2	0.922	7	0.904	3	0.956	3
TDFC	0.949	4	0.726	4	1.098	4	1.021	4
AM085296	1.021	5	0.892	6	1.234	5	0.677	2
AM085177	1.766	7	0.577	2	2.056	7	1.779	7
AM085168	1.311	6	0.551	1	1.255	6	1.593	6
GR975621	0.799	1	0.585	3	0.826	1	0.324	1

618

Figure 1

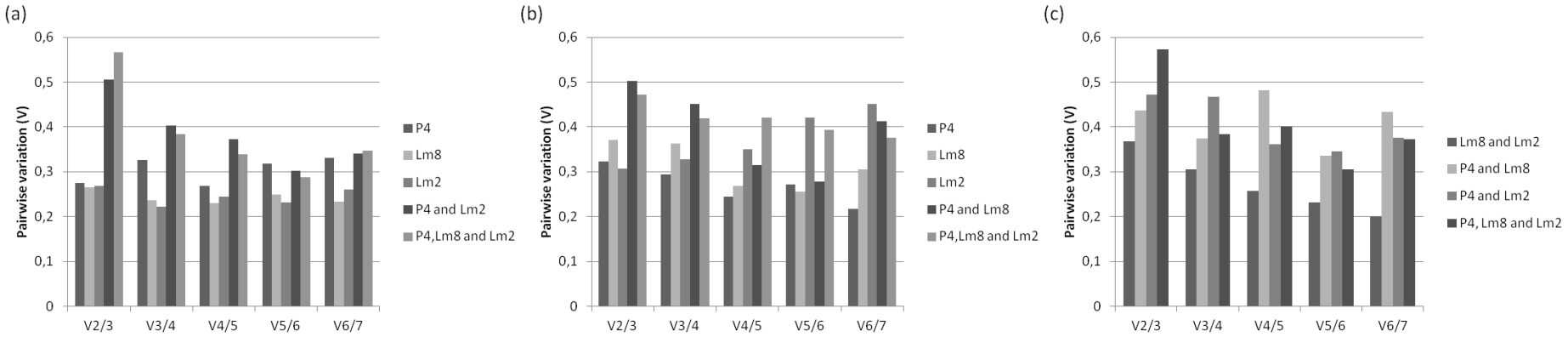


622 **Figure 2**



623

624 **Figure 3**

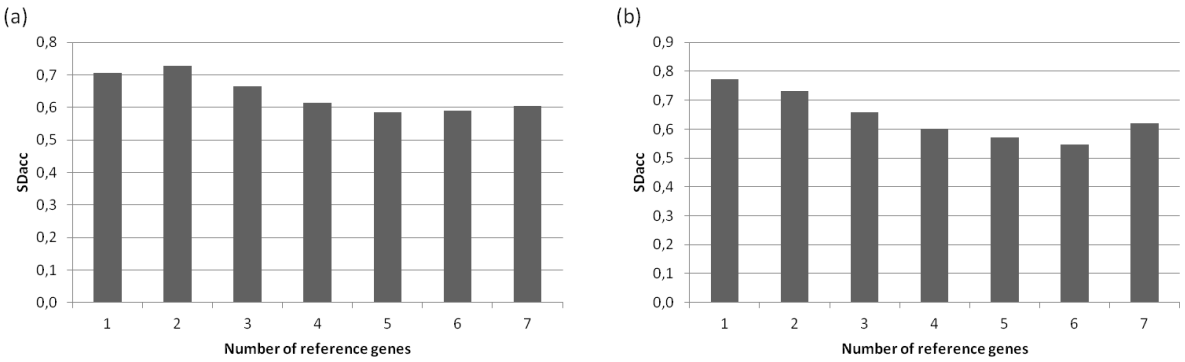


625

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2125v1 | CC BY 4.0 Open Access | rec: 15 Jun 2016, publ: 15 Jun 2016

626 **Figure 4**

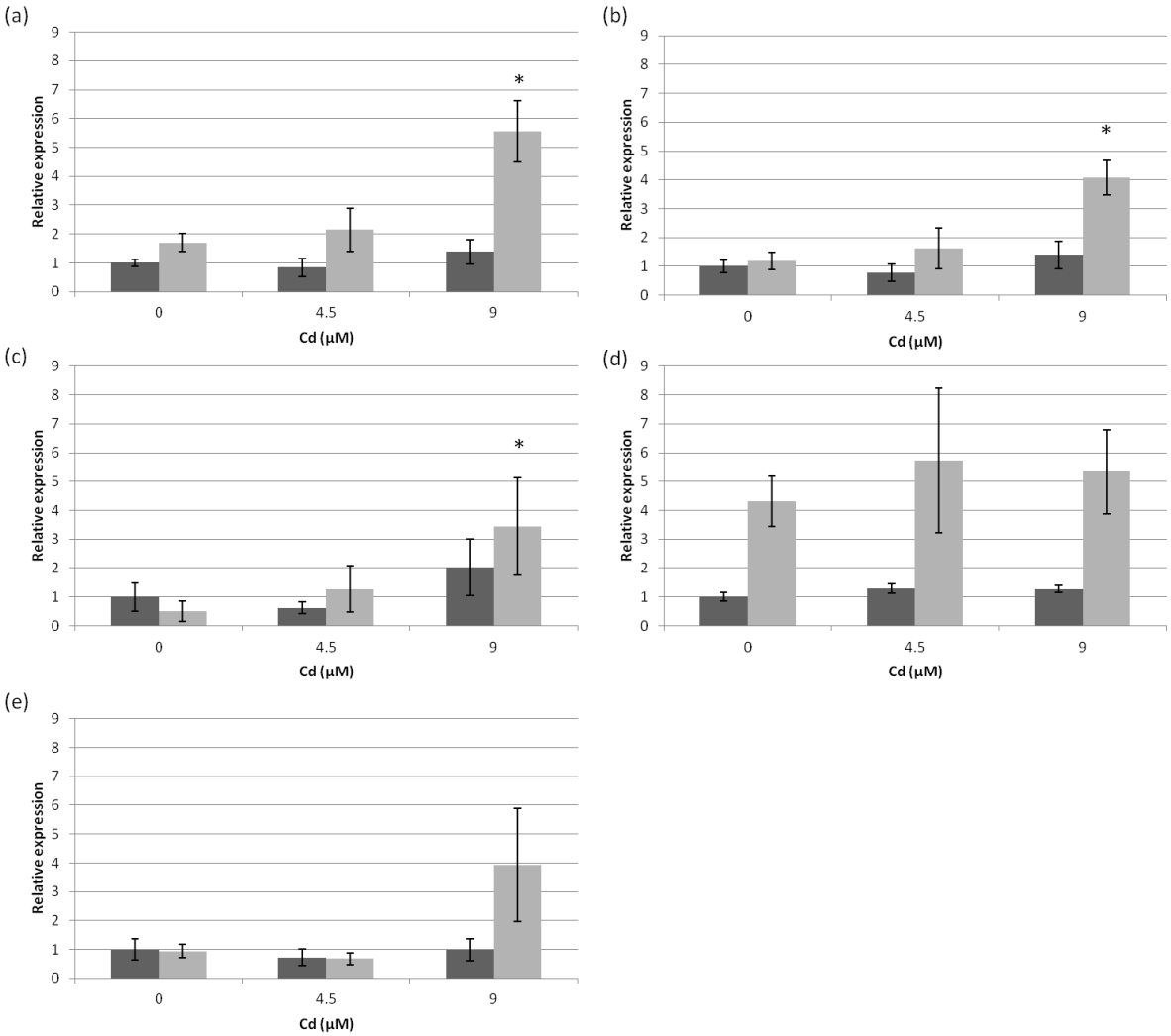
627



628

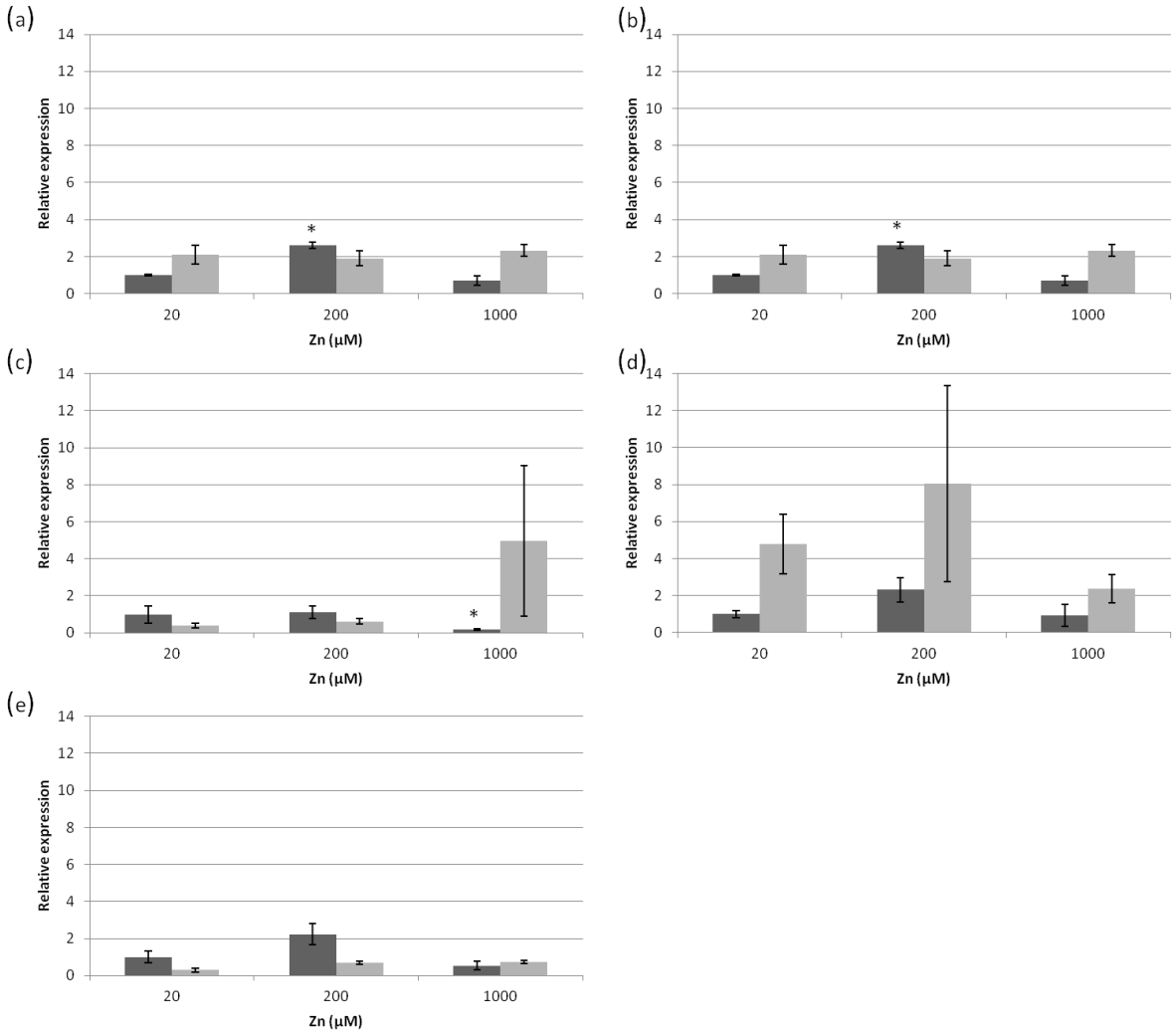
629 **Figure 5**

630



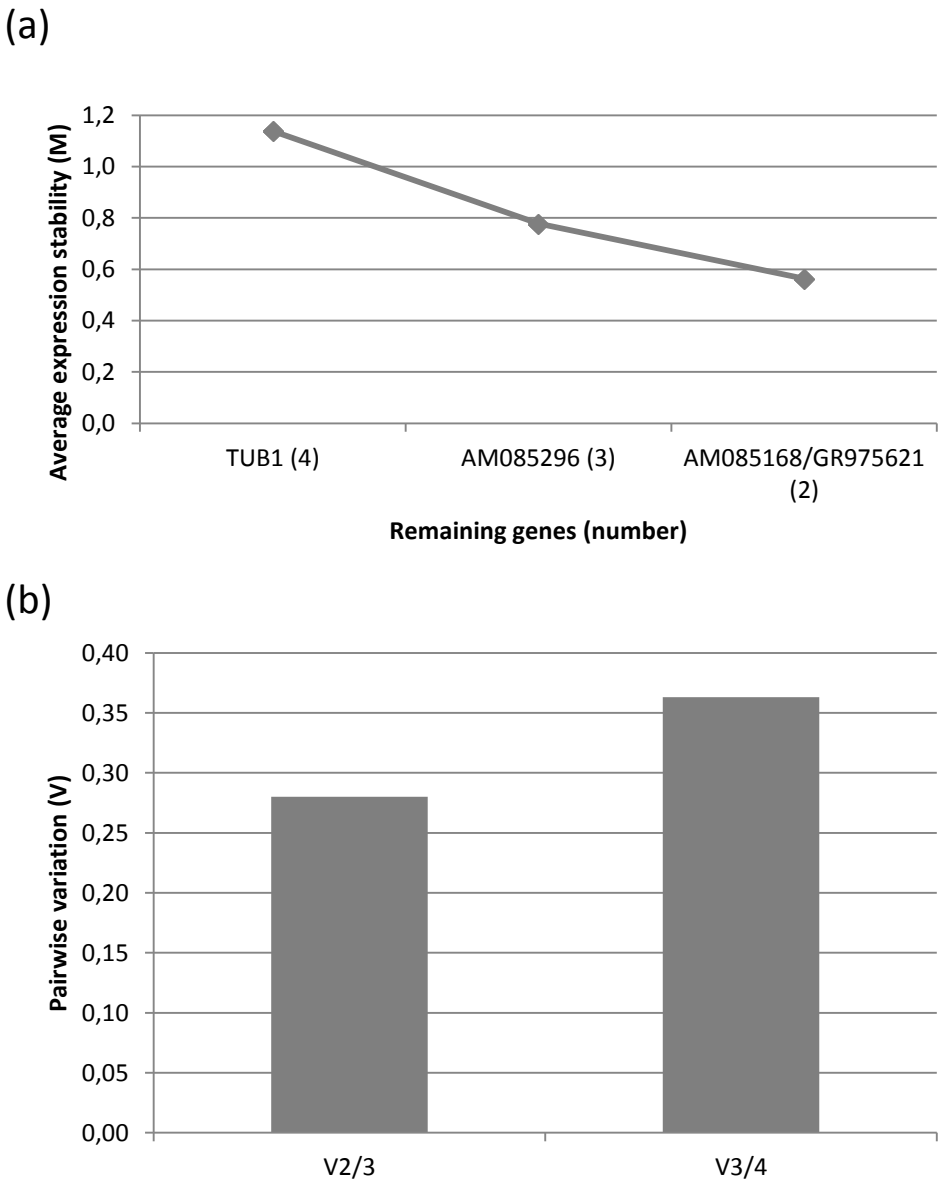
631

632 **Figure 6**



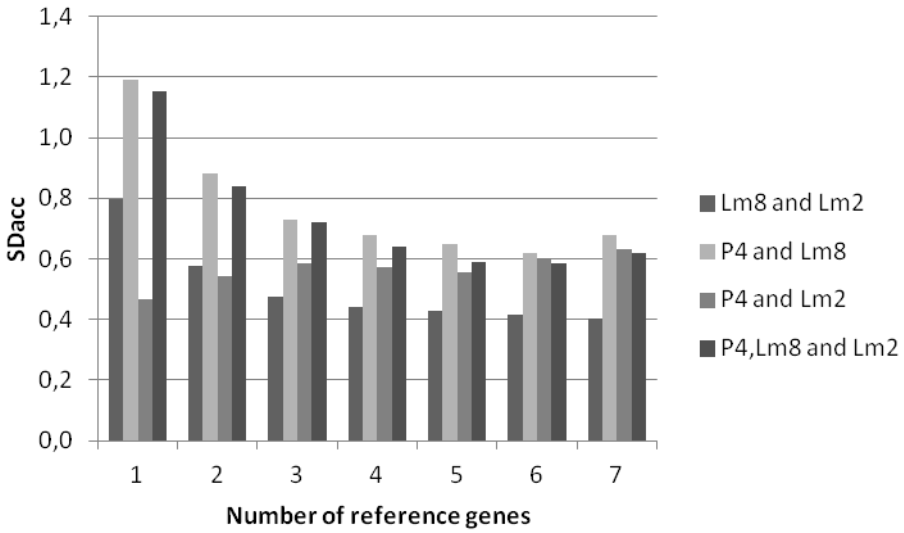
633

634 **Figure 7**



635

636 **Supplementary figure 1**



637