Evaluation of candidate reference genes for expression study in Saccharum spp. hybrids under heavy metal stress

Hui Ling, Ning Huang, Liping Xu, Long Huang, Qibin Wu, Jinlong Guo, Yachun Su, Youxiong Que

Heavy metal contamination has been a significant problem limiting agricultural development, but sugarcane has recently emerged as a valuable phytoremediator. To better understand the molecular mechanism behind sugarcane's metal tolerance, it is necessary to analyze the expression of a novel gene(s) by qRT-PCR. Importantly, introducing internal reference gene(s) should be selected based upon gene stable expression, the inclusion of which could enhance both the accuracy and reliability of this method. In this study, 13 candidate genes were selected and evaluated stability of each genes. The results derived by statistical algorithms were then validated by normalizing the expression of metal related gene ScMTP (GenBank Accession No. KP864146), ScMT2-1-3 (GenBank Accession No. JQ627644), ScMPP (GenBank Accession No. CA267392.1) and ScHMA1 (GenBank Accession No. CA156665.1). Collectively, our qRT-PCR results indicated that in heavy metal-exposed sugarcane, APRT was the better single internal control in expression quantification. Moreover, the combination of CAC + CUL provide for a more accurate normalization for gene transcript profiles under these same conditions. The gene expression quantification that included APRT and CAC + CUL suggested that ScMTP had a differential expression pattern, ScMT2-1-3 and ScMPP were slightly inhibited, and ScHMA1 had minimal induction of expression in response to Cd2+ and Cu2+ stresses in sugarcane. Taken together, the suitable reference genes identified in this study will benefit future work aimed at the sugarcane gene functional characterization.

Evaluation of candidate reference genes for expression study in *Saccharum* spp. hybrids under heavy metal stress

Hui Ling[1], Ning Huang[1], Liping Xu[1, 2], Long Huang[1], Qibin Wu[1], Jinlong Guo[1], Yachun Su[1], Youxiong Que[1, 2]

1 Key Laboratory of Sugarcane Biology and Genetic Breeding, Ministry of Agriculture, Fujian Agriculture and Forestry University, Fuzhou 350002, Fujian, China

2 Corresponding authors:

Xiadian Road Num. 15, Cangshan District, Fuzhou City, Fujian, 350002, China

E-mail: xlpmail@126.com;

Youxiong Que

Xiadian Road Num. 15, Cangshan District, Fuzhou City, Fujian, 350002, China

E-mail: queyouxiong@126.com.

This manuscript has been thoroughly edited by a native English speaker from Boston Professional Group (*BPG*) Editing. Editing Certificate will be provided upon request.

Abstract

Heavy metal contamination has been a significant problem limiting agricultural development, but sugarcane has recently emerged as a valuable phytoremediator. To better understand the molecular mechanism behind sugarcane's metal tolerance, it is necessary to analyze the expression of a novel gene(s) by qRT-PCR. Importantly, introducing internal reference gene(s) should be selected based upon gene stable expression, the inclusion of which could enhance both the accuracy and reliability of this method. In this study, 13 candidate genes were selected and evaluated stability of each genes. The results derived by statistical algorithms were then validated by normalizing the expression of metal related gene *ScMTP* (GenBank Accession No. KP864146), *ScMT2-1-3*

(GenBank Accession No. JQ627644), *ScMPP* (GenBank Accession No. CA267392.1) and *ScHMA1* (GenBank Accession No. CA156665.1). Collectively, our qRT-PCR results indicated that in heavy metal-exposed sugarcane, *APRT* was the better single internal control in expression quantification. Moreover, the combination of *CAC* + *CUL* provide for a more accurate normalization for gene transcript profiles under these same conditions. The gene expression quantification that included *APRT* and *CAC* + *CUL* suggested that *ScMTP* had a differential expression pattern, *ScMT2-1-3* and *ScMPP* were slightly inhibited, and *ScHMA1* had minimal induction of expression in response to Cd^{2+} and Cu^{2+} stresses in sugarcane. Taken together, the suitable reference genes identified in this study will benefit future work aimed at the sugarcane gene functional characterization.

Keywords

Sugarcane · Quantitative real-time PCR · Reference gene · heavy metal stress

Abbreviations

25 *rRNA*, 25S ribosomal RNA; *GAPDH*, glyceraldehyde-3-phosphate dehydrogenase; *eEF-1a / EF1-α*, Eukaryotic elongation factor 1-alpha; *eIF-4α*, Eukaryotic initiation factor 4-alpha; *CAC*, Clathrin adaptor complex; *CUL*, Cullin; *ACT* (2 / 7), β-actin (2 / 7); *TUB*, β-tubulin; *UBQ* (2 / 9), Ubiquitin (2 / 9); 18S *rRNA*, 18S ribosomal RNA; *TIPS-41*, Tonoplastic intrinsic proteins; *APRT*, Anthranilate phosphoribosyl transferase; *PRR*, Pseudo response regulator; *TIP41*, TIP41-like protein; *MTP*, metal tolerance protein gene; *MT*, metallothionein; *ScMTP*, sugarcane metal tolerance protein gene; *ScMT2-1-3*, sugarcane metallothionein type2-1-3; *ScMPP*, metallophosphoesterase gene; *ScHMA1*, heavy metal transporting ATPase gene; EST, Expressed sequence tags; $V_{n/n+1}$, pair-wise variation; SV(s), stability value(s); CSV, comprehensive stability value; Ct, Threshold cycle.

Introduction

Heavy metal contamination has steadily but significantly emerged over the years as one of the most serious environmental problem threatening our global ecosystem (Li et al. 2007). During the

process of economic development, such metal pollution has increased and limited the utilization of agricultural lands (He et al. 2013). Metal pollution derives from extensive sources, including mineral mining, industrial discharge, coal combustion, and increasing electronic-waste, and is becoming the major threat to both the environment and to local populations (Bindu ; Chen et al. 2015; Chirila & Draghici 2008; He et al. 2013; Kollikkathara et al. 2009; Li et al. 2007; Mukherjee et al. 2009; Pan & Wang 2012; Wong et al. 2006). Over the past few decades, the heavy metal pollution has become a serious concern to people and governments, leading to demand for a feasible approach to handle such pollutants (He et al., 2013).

What's the worse, Lead (Pb), copper (Cu), zinc (Zn), manganese (Mn), mercury (Hg), cadmium (Cd), chromium (Cr) and arsenic (As) are all commonly found in polluted regions, as these heavy metals are not only non-biodegradable but have high toxicity and devastate to local organismal populations (Ali et al. 2013; Chen et al. 2015; He et al. 2013; Hu et al. 2014a; Memon & Schröder 2009). Phytoremediation centers on using plants to remedy previously contaminated areas and combines biomass production, recovery of polluted-land, expansion of planting area, and the generation of extra-economic benefit for nearby farmers (Ali et al. 2013; Arunakumara et al. 2013; Chen et al. 2015; Hu et al. 2014a; Memon & Schröder 2009; Puschenreiter et al. 2005). It has showed large potential to improve the environment and has been regarded as promising bio-method to enrich and clean up metal contaminants (Ali et al. 2013; Arunakumara et al. 2013; Chen et al. 2015; Hu et al. 2014a; Memon & Schröder 2009; Puschenreiter et al. 2005). If developed, it would provide the most robust method for environmental clean-up, as it is economic-cost, efficient, novel and eco-friendly, as well as highly acceptable to locally peoples (Lee 2013; Puschenreiter et al. 2005).

Generally, heavy metal iron is harmful to plant protein and cellular physiology, but several plants can not only survive the damage caused by metal toxicity, but also thrive (Hong-Bo et al. 2010; Lee 2013). Plant species that can overcome such highly toxic levels of heavy metals are called metallophytes or phytoremediator (Hong-Bo et al. 2010). Owing to its high production of biomass, high water use efficiency, broad environmental adaptability, small energy input and good

77 ratooning ability, sugarcane (*Saccharum* spp. hybrids) is a potential phytoremediator that could
78 both accumulate metal iron and survive under high copper or cadmium concentration, thus
79 supporting the potential use for improving the polluted areas (Sereno et al. 2007; Zhang et al.
80 2014). Moreover, other two, the increasing demand for bioenergy production and not consumed
81 immediately by human, further facilitates its use as a remedy to metal pollution remedy (Gentile
82 et al. 2013; Puschenreiter et al. 2005; Zhang et al. 2014). Thus, using sugarcane as a bioenergy
83 crop and a phytoremediator in heavy metal contaminated subtropical and tropical areas is likely a
84 feasible strategy.

85 In plants, understanding the molecular basis of heavy metal tolerance would facilitate the
86 development of new strategies to create metal-tolerant crops, bio-fortified foods and suitable plants
87 for the phytoremediation of contaminated lands (Hong-Bo et al. 2010; Lee 2013; Memon &
88 Schröder 2009; Yang et al. 2005). The identification of genes related to plant heavy metal
89 phytoremediation, elucidating the molecular mechanism(s) of phytoremediation, and developing
90 transgenic or mutagenized plants to improve the hyper-accumulation individually or collectively
91 be beneficial for the improvement of heavy metal pollution (Hong-Bo et al. 2010; Memon &
92 Schröder 2009). To identify the function of these genes, such as metallothioneins and metal
93 tolerance protein genes, it is necessary to evaluate and analyze the gene expression pattern in such
94 plants under metal stress conditions. To this end, quantitative real-time polymerase chain reaction
95 (qRT-PCR) has become a widely used technique to quantity gene expression level in different
96 experimental samples (Fang et al. 2004; Guénin et al. 2009; Wang et al. 2014). However, due to
97 its high sensitivity, the accuracy of qRT-PCR is easily influenced by several factors, including the
98 inputted sample amount, RNA quantity, RNA integrity and purity, efficiency of cDNA synthesis,
99 and even by differences in materials activities (Andersen et al. 2004; Zhu et al. 2013). In order to
100 eliminate these negative effects during qRT-PCR gene normalization, it has been determined that
101 either one gene or the gene group that exhibit the most stable expression under a given set of
102 experimental conditions or across various developmental and growth periods should be used as an
103 internal reference control during data analysis (De Santis et al. 2011; Die et al. 2010; Guénin et al.

2009; Kundu et al. 2013; Vandesompele et al. 2002).

Till now, there currently exist five widely used statistical algorithms, geNorm (Vandesompele et al. 2002), NormFinder (Andersen et al. 2004), BestKeeper (Pfaffl et al. 2004), deltaCt method (Silver et al. 2006) and RefFinder (A WEB-based software) (Xie et al. 2012), have been commonly recognized for reference genes evaluation (Guénin et al. 2009). Work by Iskandar et al. (2004) validated *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase) in two sugarcane cultivars and three *Saccharum* species (Iskandar et al. 2004), and Que et al. (2009) validated 25 *rRNA* (25S ribosomal RNA) as a suitable reference gene for gene expression analysis by qRT-PCR method under the stress of sugarcane smut pathogen (Que et al. 2009). Furthermore, Ling et al. (2014) identified *GAPDH*, *eEF-1a* (Eukaryotic elongation factor 1-alpha) and *eIF-4α* (Eukaryotic initiation factor 4-alpha) as stable and suitable reference genes across various abiotic stresses and hormone treatments (Ling et al. 2014). Guo et al. (2014) also identified *GAPDH* and *eEF-1a* as suitable reference genes, but for gene expression under NaCl and PEG stresses (Guo et al. 2014). The aforementioned studies also recommended *CAC* (Clathrin adaptor complex) and *CUL* (Cullin) as the best gene combination (Guo et al. 2014; Ling et al. 2014). However, the report from Guo et al. (2014) also indicated that the reference genes recommended specifically for NaCl and PEG stresses were different from the reference genes for the abiotic stress conditions used in Ling et al. (2014) (Guo et al. 2014), which recommended by geNorm (*CAC* and *CUL*) and NormFinder (*eIF-4α*) (Ling et al. 2014).

Some studies have suggested that reference genes should be validated before being used for normalization under certain experimental conditions (Kozera & Rapacz 2013; Lilly et al. 2011; Nicot et al. 2005; Zhu et al. 2013). Though the report from Ling et al. (2014) had collectively taken into account NaCl, H_2O_2, PEG, $CuCl_2$ and $CdCl_2$ as sugarcane abiotic stresses (Ling et al. 2014), these results would likely confuse the future researchers who sought only gene expression detection under heavy metal stress. In fact, several systematic studies had validated that metal-related reference genes possess more stable performance than either *EF1-α* (named *eEF-1a* in the present study) or *GAPDH* (Borowski et al. 2014; Hu et al. 2014b; Sang et al. 2013; Wang et al.

2014). For instance, Wang et al. (2014) found that *ACT7* (β-actin7) and *TIP41* (TIP41-like protein) could serve as the best reference genes in *Brassica napus* under Cr^{6+} stress (Wang et al. 2014). All three programs (geNorm, NormFinder and BestKeeper) found that under Cd, Pb, Zn and Cu stresses, *UBC9* (Ubiquitin9) and *TUB* were the least variable reference for gene expression in *Sedum alfredii* (Sang et al. 2013). To NormFinder, *ACT1* (β-actin1) was the best choice in *Lactuca sativa* L. under 0.7 $g \cdot mL^{-1}$ sodium metathioarsenate stress (Hu et al. 2014b). In *Fortunella crassifolia* Swingle under Pb^{2+} and Zn^{2+} treatment, *ACT7* performed more stably when evaluated by geNorm evaluation, while *UBQ2* (Ubiquitin2) was better according to NormFinder evaluation (Borowski et al. 2014). These results support the idea the suitable reference gene(s) for normalization in qRT-PCR is various in different depending on plant species, suggesting that there is the specificity in which genes are suitable for normalization. Until now, no suitable reference gene has been reported for use in normalizing gene transcript profile under heavy metal stress in sugarcane.

In this present study, we evaluated the stability of six candidate reference genes (*GAPDH*, *25S rRNA*, *eEF-1a*, *eIF-4α*, *CAC* and *CUL*—all of which were chosen as suitable reference gene in previously reported studies (Guo et al. 2014; Iskandar et al. 2004; Ling et al. 2014; Que et al. 2009). Stability was determined using five statistical algorithms: geNorm (Vandesompele et al. 2002), NormFinder (Andersen et al. 2004), BestKeeper (Pfaffl et al. 2004), deltaCt method (Silver et al. 2006) and RefFinder (A WEB-based software) (Xie et al. 2012) in four sugarcane cultivars under copper chloride ($CuCl_2$) and cadmium chloride ($CdCl_2$) stresses. However, genes (β-actin, *ACT*; β-tubulin, *TUB*; Ubiquitin, *UBQ*) that had been evaluated and performed well in other plant species (Borowski et al. 2014; Hu et al. 2014b; Sang et al. 2013) and four genes (18S ribosomal RNA, 18S *rRNA*; Tonoplastic intrinsic proteins, *TIPS-41*; Anthranilate phosphoribosyl transferase, *APRT*; and Pseudo response regulator, *PRR*) that had been evaluated and performed less stably in previously evaluations (Guo et al. 2014; Ling et al. 2014) were also included in the present study. In addition to these seven, a sugarcane metallothionein gene (*ScMT2-1-3*, GenBank Accession No. JQ627644) reported in Guo et al. (2013) (Guo et al. 2013), a sugarcane metal tolerance protein

gene (*ScMTP*, GenBank Accession No. KP864146), a metallophosphoesterase gene (*ScMPP*, GenBank Accession No. CA267392.1) and a heavy metal transporting ATPase gene (*ScHMA1*, GenBank Accession No. CA156665.1), were used to further validate the availability and feasibility of those selected reference genes in the present study. Taken together, this study sought to determine the reference gene that was suited specifically for gene normalization in sugarcane under heavy metal stress and to facilitate the study of the molecular mechanism behind sugarcane metal tolerance.

Materials and Methods

Plant materials growth and treatment

Following the methods of Ling et al. (2014) and Guo et al. (2014), disease-free plantlets of four sugarcane cultivars ROC20, FN40, Liucheng03-182 and YC05-179 were generated and maintained in water solutions containing $CuCl_2$ (100 mM) and $CdCl_2$ (500 mM). Whole plantlets were sampled at 24 h after treatment along with untreated control samples (0 h sample). Simultaneously, ROC20 seedlings were treated with water solutions containing $CuCl_2$ (100 mM) and $CdCl_2$ (500 mM) for 0 h, 12 h, 24 h, 48 h and 96 h. Samples were then collected as described in Ling et al. (2014) and Guo et al. (2014). Each sample contained three biological replicates (three plantlets per replicate), was snap–frozen immediately and kept at 80°C until used for RNA extraction.

RNA isolation, DNase treatment and cDNA synthesis

RNAprep Pure Plant Kit (polysaccharides & polyphenolics-rich) (TIANGEN, Beijing, China) was employed to isolate total RNA from the above collected sugarcane plantlet-samples. Following that, the integrity of RNA samples were analyzed by agarose gel electrophoresis and the quality were quantified by a synergy H1 Microplate Reader Multi-Mode (Bio-Tek, Vermont, USA). Finally, 500ng total RNA samples were selected with good integrity and quality, of which the electrophoretic bands (28S *rRNA*, 18S *rRNA* and 5S *rRNA*) were clear, and 260/280 ratio was from

1.9 to 2.1 and 260/230 ratio from 2.0 to 2.5, were used for DNaseI treatment with Promega RQ1 RNase-Free DNase kit (Promega, Madison WI, USA). The first-strand cDNA synthesis was conducted with TAKARA PrimeScrit RT reagent Kit (Perfect for Real Time) (TAKARA Biotechnology, Dalian, China) according to the manufacturer's instructions.

qRT-PCR and data analyses

Using the previously published methods and primer pairs reported by Ling et al. (2014) and Guo et al. (2014), this study used qRT-PCR to evaluate stabilities of candidate reference genes. Threshold cycle (Ct) values, standard deviations, and covariation coefficient were calculated in Microsoft Excel 2013. The Ct value mean were calculated from three biological replicates, which were transformed and separately inputted into geNorm (trial version; Biogazelle, Zwijnaarde, Belgium) (Vandesompele et al. 2002) and NormFinder (ver. 0.953) (Andersen et al. 2004) according to their respective instruction manuals. Likewise, Ct value means were also directly inputted into RefFinder (A WEB-based software) (Xie et al. 2012).

The stability values (SV) of candidate genes obtained from geNorm (Vandesompele et al. 2002) and NormFinder (Andersen et al. 2004) together with those from BestKeeper (Pfaffl et al. 2004) and deltaCt (Silver et al. 2006) (calculated on RefFinder) were used to calculate the Pearson correlation values (r value) in IBM SPSS Statistics Version22.0. The candidate gene stability values obtained from each statistical algorithm that had significant correlation in this our correlation analysis between two of the statistical algorithms (geNorm, NormFinder, BestKeeper and deltaCt) were chosen to calculate the relative SVs. These values were obtained by transforming the SV with the following formula: Relative SV=SV of Rank N/ SV of Rank 1; N=1~13. Then, the comprehensive stability value (CSV), geometrical mean which is the —geometrical mean (GM) of the relative SVs of each candidate reference gene was further calculated and re-ranked according to a previously described method (Chen et al. 2011; Zhang et al. 2012b; Zhu et al. 2012). Based on the re-ranked list, the top two genes plus *GAPDH* and 25S *rRNA* were selected for further normalization to the expression of *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1* (Table S2). The top two candidate genes ranked in geNorm were also chosen as a combined gene set and used to further

normalize the expression of *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1*. Following the procedure reported in Guo et al. (Guo et al. 2014), this study identified suitable candidates, which would result in the expression characteristics and their maintenance of *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1* under heavy metal stress. Finally, the statistical package software DPS7.05 was used to analyze the differential expression of *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1*.

The relative expression level of *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1* were evaluated by qRT-PCR and normalized according to the $2^{-\Delta\Delta Ct}$ method (Livak & Schmittgen 2001). Primer pairs are shown in Table S2. All Ct values obtained from the three biological replicates were from the cultivar ROC20 samples that had been collected after $CdCl_2$ and $CuCl_2$ stress treatments for time points 0 h, 12 h, 24 h, 48 h and 96 h.

Results

Candidate reference genes expression in sugarcane

Our data showed that the mean Ct values of these 13 selected candidate genes ranged from 24.45 to 29.84, with the exception of 25S *rRNA* (Ct=14.20±0.893) and 18S *rRNA* (Ct=15.28±0.853) (Table 1). Moreover, these results also demonstrated that *PRR* was the least accumulated and the most variable gene among the13 candidate genes in the heavy metal treated sugarcane samples (Table 1). In comparison, *ACT* was found to be the least variable (Table 1). Ithe present study, the expression of *GAPDH*, *ACT*, and *eEF-1a* were all accumulated to approximately the same level, as were *TUB* and *UBQ* (26.47 and 26.65) (Table 1). Depending on the covariation coefficient, the least variable to the most variable genes were ranked in order as *ACT* > *TUB* > 18S *rRNA* > 25S *rRNA* > *TIPS-41* > *UBQ* > *APRT* > *eIF-4α* > *CAC* > *CUL* > *GAPDH* > *eEF-1a* > *PRR* (Table 1).

Analysis of stability of gene expression

Based on the four statistical algorithms evaluations of the expression stability of our 13 candidate reference genes from under heavy metal stress-exposed sugarcane, we found four types of stability values (Supplementary Table S1). The correlation level of the rank order for candidate reference

genes given by the geNorm, NormFinder, BestKeeper, and deltaCt method were calculated in IBM SPSS Statistics Version22.0 by directly inputting the stability values. The results suggested that the Pearson correlation coefficients were significantly higher between two of the rank-list candidates as evaluated by geNorm, NormFinder and deltaCt method (Table 2). Contrastingly, the rank results obtained from BestKeeper shared a lower correlation with the rank-lists obtained geNorm, NormFinder, and deltaCt method (Table 2).

The stability values (SVs) evaluated by geNorm, NormFinder and deltaCt were chosen and converted into relative SV (setting the minimum stability value as 1). After obtaining the relative SV, the geometrical mean of the relative SVs of each gene from three statistical algorithms were calculated and then evaluated by comprehensive rank (Table 3). The CSV indicated that *UBQ* was the most stable gene in sugarcane under heavy metal treatment, followed by *APRT*, *CUL*, *CAC* and *GAPDH* which all four had near-equal CSVs (Table 3). When sugarcane plantlets were exposed to heavy metal stress, the expression of *TIPS-41* varied more than for the remaining 12 candidate reference genes, revealing that it was the most unstable gene under these conditions (Table 3).

The comprehensive ranks of the remaining genes were *eEF-1a* > *eIF-4α* > 25S *rRNA* > 18S *rRNA* > *TUB* > *ACT* > *PRR* (Table 3). *UBQ* and *APRT*, which were ranked as the first two genes, were selected as the two most suitable reference genes and selected for subsequent study. *GAPDH* and 25S *rRNA*, which had been used as reference genes in Guo et al. (2013), were also selected for subsequent quantitative validation

Optimal combination of reference genes for gene expression normalization under heavy metal stress

Using geNorm, we analyzed the optimal number of reference genes and the optimal combination of reference genes for gene expression quantification under heavy metal stress (Fig. 1). A 0.15 cut-off level of the pair-wise variation $V_{n/n+1}$ was used, which was originally put forth by Vandesompele et al. (Vandesompele et al. 2002) and indicated the ineffectiveness of adding one

more reference gene to create gene combination (red line indicates cut-off value Fig. 1). As shown in Fig. 1, the $V_{2/3}$, $V_{3/4}$, $V_{4/5}$, and $V_{5/6}$ values for all samples were under the 0.15 cut-off level. However, and according to the suggestion put forth by Vandesompele et al. (Vandesompele et al. 2002). We found that using the geNorm calculated combination of the first two genes ($V_{2/3}$=0.130) was the best choice for the present study. Based on these data, *CAC + CUL* were selected and used in the following validation studies.

Expression analysis of *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1* genes based on selected reference gene(s) found under heavy metal stress conditions

To validate the application of the selected reference genes *UBQ*, *APRT*, *GAPDH*, 25S *rRNA* and *CAC + CUL*, we conducted expression normalization of *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1* in sugarcane cultivar ROC20 under heavy metal (Cd and Cu) stresses. Simultaneously, the relative fold of gene expression with the reference of *UBQ*, *APRT*, *GAPDH*, and 25S *rRNA* were compared with the reference combination of *CAC + CUL* (Fig. 2and Fig. 3). As shown in Fig. 2 and 3, the expression patterns of *ScMTP* (Fig. 2 and 3, Panel A), *ScMT2-1-3* (Fig. 2 and 3, Panel B), *ScMPP* (Fig. 2 and 3, Panel C) or *ScHMA1* (Fig. 2 and 3, Panel D) normalized to *APRT* and *CAC + CUL* were more similar than when compared with the reference genes *UBQ*, *GAPDH* and 25S *rRNA*. Statistical normalization to *UBQ*, *GAPDH*, or 25S *rRNA* led to more variable transcript trend of *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1* (Fig. 2 and 3). When compared with the expression level referenced by *CAC + CUL*, $CdCl_2$ treatment led to an initial decrease in the expression levels of *ScMTP*, *ScMT2-1-3*, *ScMPP*, and *ScHMA1* below the control level (normalized to 1, data not shown) (*ScMTP*, $p < 0.01$; *ScMT2-1-3*, $p < 0.01$; and *ScHMA1*, $p < 0.01$). $CdCl_2$ treatment also led to a more significant inhibition (*ScMPP*, $p < 0.01$) during the first 12 h of treatment when normalized to *UBQ* (Fig. 2). However, $CdCl_2$ treatment led to more significant induction of gene expression when normalized with *GAPDH* for these four genes (*ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1*; $p < 0.01$) at 12 h, 48 h, and 96 h, with the exception of *ScMTP* expression at 12 h ($p < 0.05$) (Fig. 2). When the expression of *ScMT2-1-3* (Fig. 2B) and *ScHMA1* (Fig. 2D) 48 h to 96 h were normalized with 25S *rRNA*, results showed the opposite expression trend when normalized

with *CAC* + *CUL*. This was also seen with the expression of *ScMPP* from 12 h to 48 h (Fig. 2C). When *APRT* and *CAC* + *CUL* were used as reference genes, the accumulation of *ScMTP* decreased gradually from a nearly 1.63-fold up-regulation at 12 h to 0.62-fold at 48 h Cd treatment. Expression increased to 3.78-fold at the last 96 h after treatment (Fig. 2A). However, these same samples showed a down-regulation of *ScMT2-1-3* along with ongoing treatment except at 12 h (Fig. 2B). Finally, our results also showed that with *CAC* + *CUL* normalization, there was both a continuing decrease of *ScMPP* (Fig. 2C) accumulation and a continuous small up-regulation of *ScHMA1* (Fig. 2D).

Under $CuCl_2$ treatment and using *APRT* and *CAC* + *CUL* as reference genes, the rise of *ScMTP* expression was found at 12 h and 48 h, although *ScMTP* would transcript levels decreased at following two time points (24 h and 96 h) (Fig. 3A). Though slight inhibition of *ScMT2-1-3* transcription was found when referenced by *APRT* and *CAC* + *CUL* during the first day of $CuCl_2$ treatment (12 h and 24 h), *ScMT2-1-3* recovered to control levels at later time points (48 h and 96 h, Fig. 3B). When normalized to the same genes (*APRT* and *CAC* + *CUL*), *ScMPP* and *ScHMA1* both had $CuCl_2$-mediated induction by at the firstly 12 h. However, as the time of exposure to heavy metal increasing, *ScMPP* had continuing reduction in its accumulation, ending at 0.77 / 0.79-fold of controls (Fig. 3C), Comparatively, *ScHMA1* maintained consistently higher expression levels when compared with control. As shown in Fig. 3, with exception of the 12 h treated sample, the relative transcript levels of four genes (*ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1*) relative to *GAPDH* showed significant difference levels when compared with levels referenced by *APRT* and *CAC* + *CUL* at later time points ($p < 0.01$). This same pattern was seen when referenced by *UBQ* at 12 h, 48 h, and 96 h ($p < 0.01$), although the opposite patterns emerged when referenced by *GAPDH* and 25S *rRNA* reference, when compared with *APRT* and *CAC* + *CUL* references (Fig. 3).

Dicussion

Due to the serious global levels of heavy metal environmental pollution, plant ecologists,

physiologists, and biologists have paid increasingly large amounts of attention to the physiology and genetic mechanisms underlying natural heavy metal tolerance found in some plants (Arunakumara et al. 2013; Memon & Schröder 2009). Several metal-related genes found in these species, such as genes encoding heavy metal ATPase, metallothionein1/2 and metal tolerance protein, have been identified, including in *Saccharum* spp., *Thlaspi caerulescens*, *Nicotiana glauca*, *N. tabacum*, *Chloris virgata* Swartz, *Arabidopsis halleri*, *Arabidopsis thaliana*, *O. sativa*, *Cajanus cajan* L., *S. alfredii* Hance and poplar hidrid (*Populus trichocarpa* × *Populus deltoides*) (Arrivault et al. 2006; Blaudez et al. 2003; Courbot et al. 2007; Guo et al. 2013; Nishiuchi et al. 2007; Sekhar et al. 2011; Sereno et al. 2007; Shingu et al. 2005; Yuan et al. 2012; Zhang et al. 2011). In these studies, the investigation of gene expression was conducted with semi-quantitative RT-PCR, Northern blot, and/or quantitative real time PCR (qRT-PCR) (Arrivault et al. 2006; Blaudez et al. 2003; Courbot et al. 2007; Guo et al. 2013; Nishiuchi et al. 2007; Papoyan & Kochian 2004; Sekhar et al. 2011; Sereno et al. 2007; Shingu et al. 2005; Yuan et al. 2012; Zhang et al. 2011). In order to reduce the potentially confounding effects of difference in sample amount, RNA recovery, RNA quantity, RNA integrity, RNA purity, efficiency of cDNA synthesis and/or even differences in materials' activities during gene expression quantification, we used the following internal reference genes: *EF1α*, *rRNA*, *Actin7* and *GAPDH* (Arrivault et al. 2006; Blaudez et al. 2003; Courbot et al. 2007; Guo et al. 2013; Nishiuchi et al. 2007; Papoyan & Kochian 2004; Sekhar et al. 2011; Sereno et al. 2007; Shingu et al. 2005; Yuan et al. 2012; Zhang et al. 2011). However, these traditional reference genes have been used directly as the normalization reference factor without any previous experimental evaluation as to their respective stabilities (Borowski et al. 2014; Hu et al. 2014b; López-Landavery et al. 2014). Thus, to achieve an accurate quantification of gene expression, it is necessary to conduct an assessment of stability of internal reference genes under specific experimental conditions—including heavy metal stress—to allow for the accurate quantification of gene transcript profile (Lilly et al. 2011; Nicot et al. 2005; Wang et al. 2014). This, in turn, would facilitate the functional identification of genes(Lilly et al. 2011; Nicot et al. 2005; Wang et al. 2014).

Sugarcane is a putatively valuable phytoremediator species with the ability grow in metal polluted regions to improve the overall health of the local environment (Sereno et al. 2007; Xia et al. 2009). Though many studies have aimed at unraveling the biological characteristics of metal accumulation and tolerance in sugarcane, its molecular mechanism of action for metal—related transportation and enrichment have not been clearly elucidated (Guo et al. 2013; Sang et al. 2013; Sereno et al. 2007; Wang et al. 2014). Identifying the expression feature of responsive genes in sugarcane exposed to heavy metal stress would help to clarify the molecular mechanism behind the metal tolerance of sugarcane. However, until recently, only one kind of metal tolerance related genes—metallothionein—had been identified in sugarcane (Gentile et al. 2013; Guo et al. 2013; Sereno et al. 2007).

The relative quantitation method using qRT-PCR plus normalization by a stable internal reference gene is a reliable and simple method for identifying gene transcriptional characteristic. The stable reference gene(s) should be rigorously and rationally selected with a sound experimental method. To date, some reports have found suitable reference genes for the differential gene expression in sugarcane, normalizing under experimental abiotic, hormone, drought and osmotic stress treatments. However, no reference gene(s) in sugarcane have been specifically validated for approaches featuring heavy metal stress (Iskandar et al. 2004, Que et al. 2009, Ling et al. 2014, Silva et al. 2014 and Guo et al. 2014). Previous reports by both Ling et al. (2014) and Guo et al. (2014) have highlighted differences about reference genes related to the heavy metals which collectively indicated that the reference genes should be validated before using them for normalization purpose under certain heavy metal stress conditions in sugarcane. In the present study, suitable reference gene(s) were selected from 13 candidate genes under either experimental $CuCl_2$ or $CdCl_2$ water-solution treatments. These have been used as candidates in the previously published works by both Ling et al. (2014) and Guo et al. (2014). The Ct value and covariation coefficient analysis indicated that 25S *rRNA* and 18S *rRNA* had the most abundant expression in the heavy metal treated sugarcane plantlets, whereas *PRR* was the least abundant and most variable one. Wan et al. (2010) suggested that moderate expression of internal reference genes (i.e. Ct value

368 ranging from 15 to 30) could enhance the accurate quantification of target genes. Furthermore, that
369 near-equal expression levels between reference and target genes could decrease the inaccuracy of
370 functional identification (Wan et al. 2010). As shown in Table 1, with the exception of 25S *rRNA*,
371 18S *rRNA* and *PRR*, the remaining genes had moderate transcript levels, with the top five least
372 variable genes being *ACT*, *TUB*, *TIPS-41*, *UBQ*, and *APRT*. In fact, Sang et al. (2013) also found
373 moderate transcript levels and the least variable expression pattern of both *ACT* and *TUB* in *S.*
374 *alfredii*.

375 In this study, five different statistical algorithms were used and four final stability rank-lists
376 wereobtained. The Pearson correlation coefficients indicated that two of geNorm, NormFinder and
377 deltaCt method had relatively high correlations but that high discrepancy existed between
378 BestKeeper and three of the statistical algorithms (geNorm, NormFinder or deltaCt). Moreover,
379 these three algorithms—geNorm, NormFinder and deltaCt method offered nearly identical ranking
380 of the 13 candidate reference genes, especially when comparing the ranking provided by both the
381 geNorm and deltaCt method. Thus, the SVs that were obtained from geNorm, NormFinder and
382 deltaCt method were used to count GM of the relative SV and then all 13 of these reference genes
383 were re-ranked. As shown in the e-ranking list in Table 3, *UBQ* ranked first among all 13 candidate
384 reference genes, indicating that *UBQ* is likely the most stably expressed gene in sugarcane
385 experiencing heavy metal stress. Previous studies by both Sang et al. (2013) and Hu et al. (2013)
386 demonstrated that *UBQ9* and *UBQ2* varied less than the other candidate reference genes in *S.*
387 *alfredii* and kumquat respectively.

388 Similarly, *APRT*, *CUL*, *CAC* and *GAPDH* had nearly the same CSV in our study, but lagged
389 behind *UBQ* (Table 3). Though *GAPDH* has been shown to be stable in sugarcane under several
390 kinds of stress conditions (Guo et al. 2014; Iskandar et al. 2004; Ling et al. 2014; Que et al. 2009;
391 Silva et al. 2014), we found it to be less stable than *UBQ*, *APRT*, *CUL* or *CAC*. Up until this poin,
392 although *APRT* was found to exhibit a stable expression pattern across different developmental
393 periods for different tissues in *Solanum melongena* L. (Gantasala et al. 2013), as well as in different
394 tissues of sugarcane cultivars (Ling et al. 2014), it has never been used as a heavy metal related

reference gene in plant species. Interestingly, *UBQ* and *APRT* varied more substantially under collective abiotic stress, including NaCl, PEG, H_2O_2 $CuCl_2$, and $CdCl_2$, than separately under either heavy metal stressor ($CuCl_2$ and $CdCl_2$).

CAC and *CUL* have been previously shown to be a reference set for normalizing gene expression in sugarcane (Guo et al. 2014; Ling et al. 2014). As such, we utilized them in a similar manner for the present study. Interestingly, the $V_{2/3}$, $V_{3/4}$, $V_{4/5}$ and $V_{5/6}$ value of all sample sets were under the 0.15 cut-off level used in this study. However, according to Guo et al. (2014) suggestion, the analysis resulting from geNorm most likely indicates that *CAC* + *CUL* are capable of providing a more accurate quantification of gene transcript profile when the target samples are from a heavy metal-exposed sugarcane. Adding more genes beyond these two does not produce any more reliable quantification results. Thus the combination of *CAC* + *CUL* could serve as a reliable internal reference gene set in sugarcane samples, which have been exposed to different stress abiotic factors, such as NaCl and/or PEG (Guo et al. 2014; Ling et al. 2014). Finally, on the basis of the results obtained from the covariation coefficient analysis, comprehensive stability value re-ranking and geNorm recommendations, we selected *UBQ*, *APRT* and *CAC* + *CUL* for further experimental validation together with *GAPDH* and 25S *rRNA*.

In this study, the expression level of *ScMTP* in conjunction, with (i) the heavy metal-inducing gene *ScMT2-1-3* feature in Guo et al. (Guo et al. 2013) and (ii) two genes, *ScMPP* and *ScHMA1* dereived from RNA-seq data related to heavy metal-treated sugarcane (unpublished), were normalized in the experimental validation of candidate reference gene / gene set *viz*., *APRT*, *UBQ*, *GAPDH*, 25S *rRNA* and *CAC* + *CUL*. Our results indicated that if either *APRT* or *CAC* + *CUL* were used as the reference, the expression of *ScMTP* in $CdCl_2$ and $CuCl_2$ treated sugarcane samples had approximately no difference, with only one sample having significant difference ($P < 0.01$). Similar normalization results were found in most of samples used in the present study for *ScMT2-1-3*, *ScMPP* and *ScHMA1*.

Conversely, the expression of *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1* using *UBQ* and *GAPDH* as reference showed a significantly different pattern from those compared against

422 references *APRT* and *CAC* + *CUL* ($P < 0.01$). This was true in most of samples, but especially at
423 the 12 h (*UBQ*, $p < 0.01$), 48 h, and 96 h (*GAPDH*, $p < 0.01$) time point samples. It was previously
424 recognized that combining several genes as internal reference controls would allow for more
425 reliable gene expression quantification (Chen et al. 2009; Janska et al. 2013; Vandesompele et al.
426 2002). This suggested that the combination of *CAC* + *CUL* would provide more reliable
427 normalization for gene expression quantification in heavy metal-stressed sugarcane. In accordance
428 with *MTP* expression under Cd^{2+} stress in *O. sativa* (Yuan et al. 2012; Zhang et al. 2012a), *S.*
429 *alfredii* Hance shoot and Zn supplementation in *M. truncatula* (Dräger et al. 2004; Zhang et al.
430 2011), the expression of *ScMTP* was constitutively up-regulated. However, it remained at low
431 levels with increasing exposure to heavy metal treatment ($CdCl_2$ and $CuCl_2$) in sugarcane when
432 referenced with *APRT* and *CAC* + *CUL*. Marked inhibition of *ScMTP* expression with Cd
433 treatment and the high induction of *ScMTP* expression under Cu treatment were first observed at
434 12 h when normalized by *UBQ* (Fig. 3). Likewise, *ScMTP* exhibited notable up-regulation at 96 h
435 when referenced by *GAPDH*. Generally, *MTP* transcript level change a little, with several plant
436 species showing was slight induction under Cd^{2+} stress (Blaudez et al. 2003; Zhang et al. 2011).
437 In the present study, the up-regulation of *ScMT2-1-3* when referenced with *APRT* and *CAC* + *CUL*
438 under $CdCl_2$ treatment exhibited a similar pattern as reported by Guo et al. (2013). This result
439 indicated that, after slight induction at 12 h, *ScMT2-1-3* in sugarcane was ultimately inhibited by
440 $CdCl_2$ stress continuously when referenced with *APRT* and *CAC* + *CUL*. However, the expression
441 of *ScMT2-1-3* under $CuCl_2$ stress were first inhibited at both 12 h and 24 h, after which expression
442 recovered to control levels.

443 A previous report by Sereno et al. (2007) suggested that a higher copper concentration copper
444 would inhibit the expression of sugarcane type two metallothionein (*MT2*). A separate study in
445 *Salicornia brachiata* (halophyte) suggested that *MT2* would be induced concurrently under copper
446 stress, but exhibit no change under cadmium stress (Chaturvedi et al. 2012). Like sugarcane,
447 *Populus alba* L. is a good phytoremediator for its specific biological characteristics. Its *MT2*
448 exhibited a slight induction when the cellular suspension exposed to higher $CdSO_4$ concentration

(150 μM as compared to 75 μM). *MT2* also had differential induction when exposed to higher $CuCl_2$ concentration (100 μM as comparing to 50 μM) (Macovei et al. 2010). Unlike the whole sugarcane plantlets which was sampled in the present study, Yuan et al. (2008) suggested that *OsMT2b* had a less than 0.25-fold down-regulation in the root under $CuCl_2$ ($P < 0.01$). Contrastingly, it exhibited a nearly 1.5-fold up-regulation in the shoot under $CuCl_2$ stress ($p < 0.05$) (Yuan et al. 2008). The aforementioned study in *Salicornia brachiate* demonstrated that *SbMT-2* could be induced by copper stress, but remain unaffected by cadmium stress (Chaturvedi et al. 2012). The significant induction of *MT2* by copper treatment was found under either low copper concentrations (50μM, 250μM and 500μM) or in the early-stress state of high copper concentration (750 μM) in both *Avicennia marina* and *Bruguiera gymnorrhiza* (Huang & Wang 2009; Huang & Wang 2010). Collectively, these results suggest that $CdCl_2$ treatment first induced a small range change for *ScMT2-1-3* expression and then recovered from an initial inhibition seen at the later treatment. This change was more reasonable in sugarcane, in addition to the transcript profile of *ScMT2-1-3* which was affected by $CuCl_2$. Marked and continuous decrease in *ScMPP* levels was observed under both Cu and Cd stress when *APRT* and *CAC + CUL* were used as internal references and compared to the use of *UBQ*, *GAPDG* and 25S *rRNA*. Previous studies have suggested that HMA1 protein has a broad-range of activities, including transportation of different metal iron. Further, that *HMA1* may be predominantly influenced by Zn and Cu (Kim et al. 2009; Mikkelsen et al. 2012; Moreno et al. 2008; Seigneurin-Berny et al. 2006). In leaves, over-dose concentration of Zn (1000 μM, $ZnCl_2$), Cu (500 μM, $CuSO_4$) and Cd (20μM, $CdCl_2$) have been shown to inhibit *HvHMA1* expression (Mikkelsen et al. 2012). This means that high concentration of Cd (100 mM, $CdCl_2$) and Cu (500 mM, $CuCl_2$) likely resulted in a slight induction of the expression of *ScHMA1*, which varied slightly along with the exposure time to increase in sugarcane *viz.* The normalization of *APRT* and *CAC + CUL* gave a more accurate picture of *ScHMA1* under both Cd and Cu stress. Taken together, this work demonstrate that reference of *APRT* and *CAC + CUL* enhances the accurate quantification of gene expression when the sugarcane plantlets suffered the heavy metal stress, thus facilitating better understanding of the

476 molecular mechanism underpinning heavy metal tolerance in sugarcane.

477 **Conclusion**

478 In plants, the quantification of gene expression using qRT-PCR is a popular method to identify
479 the function of novel gene. The internal reference gene(s), which have been obtained from suitable
480 experimental selection and evaluation of their stability, have been shown to enhance the accuracy
481 and reliability of qRT-PCR analysis. In the present study, 13 candidate genes were selected and
482 evaluated in four sugarcane cultivars exposed to heavy metal stress conditions. Results indicated
483 that *APRT* was the most suitable reference gene for qRT-PCR gene expression quantification in
484 heavy metal-exposed sugarcane. Moreover, our results also indicated that the combination of *CUL*
485 and *CAC* provided more accurate quantification of the gene transcript profile under the same heavy
486 metal experimental conditions. The gene expression quantification with *APRT* and *CAC* + *CUL*
487 suggested that, in accordance with the innate function of these four genes, *ScMTP* had different
488 expression patterns in response to Cd^{2+} and Cu^{2+} stresses. Moreover, under the same havy metal
489 stress, *ScMT2-1-3* and *ScMPP* was slightly inhibited whereas *ScHMA1* was minimally induced.
490 Collectively, the work presented here identified a suitable reference gene in sugarcane
491 experiencing heavy metal stress. Ultimately, this will benefit future research aimed at charactering
492 sugarcane gene functionality, which is crucial for unraveling the molecular mechanisms of
493 sugarcane heavy metal tolerance.

494 **Acknowledgements**

495 This work was funded by the earmarked fund for the Modern Agriculture Technology of China
496 (CARS-20). The funders had no role in study design, data collection and analysis, decision to
497 publish, or manuscript preparation.

498 **Author Contributions**

499 The experiments included in this study were conceived and designed by Liping Xu and Youxiong
500 Que. All experiments were performed by Hui Ling, Long Huang, Qibin Wu, Junlong Guo and
501 Yachun Su. The results obtained from the present study were analyzed by Hui Ling, Liping Xu,
502 Junlong Guo and YouXion Que. Together, Hui Ling, Liping Xu, Junlong Guo and Youxiong Que
503 prepared all experimental reagents, materials, and analytical tools. Hui Ling, Liping Xu and
504 Youxiong Que wrote the manuscript together. Liping Xu and Youxiong Que approved of the
505 submitted of the manuscript.

506 **Supplemental Information**

507 Supplementary information: Table. S1; Table. S2

508 **Financial Disclosure statement:** The authors declare that they have no competing financial
509 interests.

510 **References**

511 Ali H, Khan E, and Sajad MA. 2013. Phytoremediation of heavy metals—concepts and
512 applications. *Chemosphere* 91:869-881.
513 Andersen CL, Jensen JL, and Ørntoft TF. 2004. Normalization of real-time quantitative reverse
514 transcription-PCR data: a model-based variance estimation approach to identify genes
515 suited for normalization, applied to bladder and colon cancer data sets. *Cancer research*
516 64:5245-5250.
517 Arrivault S, Senger T, and Krämer U. 2006. The Arabidopsis metal tolerance protein AtMTP3
518 maintains metal homeostasis by mediating Zn exclusion from the shoot under Fe deficiency
519 and Zn oversupply. *The Plant Journal* 46:861-879.
520 Arunakumara K, Walpola BC, and Yoon M-H. 2013. Current status of heavy metal contamination
521 in Asia's rice lands. *Reviews in Environmental Science and Bio/Technology* 12:355-377.
522 Bindu GH. IMPACT OF TOXIC METALS, MINERALS, SOLVENTS, E-WASTE AND
523 PLASTICS LEADING TO ENVIRONMENTAL POLLUTION. *Journal of Chemical and*
524 *Pharmaceutical Sciences ISSN* 974:2115.
525 Blaudez D, Kohler A, Martin F, Sanders D, and Chalot M. 2003. Poplar metal tolerance protein 1
526 confers zinc tolerance and is an oligomeric vacuolar zinc transporter with an essential
527 leucine zipper motif. *The Plant Cell Online* 15:2911-2928.
528 Borowski JM, Galli V, da Silva Messias R, Perin EC, Buss JH, e Silva SDdA, and Rombaldi CV.
529 2014. Selection of candidate reference genes for real-time PCR studies in lettuce under
530 abiotic stresses. *Planta* 239:1187-1200.

Chaturvedi AK, Mishra A, Tiwari V, and Jha B. 2012. Cloning and transcript analysis of type 2 metallothionein gene (SbMT-2) from extreme halophyte Salicornia brachiata and its heterologous expression in E. coli. *Gene* 499:280-287.

Chen D, Pan X, Xiao P, Farwell MA, and Zhang B. 2011. Evaluation and identification of reliable reference genes for pharmacogenomics, toxicogenomics, and small RNA expression analysis. *Journal of cellular physiology* 226:2469-2477.

Chen H, Teng Y, Lu S, Wang Y, and Wang J. 2015. Contamination features and health risk of soil heavy metals in China. *Science of The Total Environment* 512:143-153.

Chen M, Shen X, Li D, Ma L, Dong J, and Wang T. 2009. Identification and characterization of MtMTP1, a Zn transporter of CDF family, in the Medicago truncatula. *Plant Physiology and Biochemistry* 47:1089-1094.

Chirila E, and Draghici C. 2008. Contamination of Soils by Waste Deposits. In: Simeonov L, and Sargsyan V, eds. *Soil Chemical Pollution, Risk Assessment, Remediation and Security*: Springer Netherlands, 13-25.

Courbot M, Willems G, Motte P, Arvidsson S, Roosens N, Saumitou-Laprade P, and Verbruggen N. 2007. A major quantitative trait locus for cadmium tolerance in Arabidopsis halleri colocalizes with HMA4, a gene encoding a heavy metal ATPase. *Plant Physiology* 144:1052-1065.

De Santis C, Smith-Keune C, and Jerry DR. 2011. Normalizing RT-qPCR data: are we getting the right answers? An appraisal of normalization approaches and internal reference genes from a case study in the finfish Lates calcarifer. *Marine Biotechnology* 13:170-180.

Die JV, Roman B, Nadal S, and Gonzalez-Verdejo CI. 2010. Evaluation of candidate reference genes for expression studies in Pisum sativum under different experimental conditions. *Planta* 232:145-153. 10.1007/s00425-010-1158-1

Dräger DB, Desbrosses‐Fonrouge AG, Krach C, Chardonnens AN, Meyer RC, Saumitou‐Laprade P, and Krämer U. 2004. Two genes encoding Arabidopsis halleri MTP1 metal transport proteins co‐segregate with zinc tolerance and account for high MTP1 transcript levels. *The Plant Journal* 39:425-439.

Fang X, Willis RC, Hoang Q, Kelnar K, and Xu W. 2004. High-throughput sample preparation for gene expression profiling and in vitro target validation. *Journal of the Association for Laboratory Automation* 9:140-145.

Gantasala NP, Papolu PK, Thakur PK, Kamaraju D, Sreevathsa R, and Rao U. 2013. Selection and validation of reference genes for quantitative gene expression studies by real-time PCR in eggplant (Solanum melongena L). *BMC Res Notes* 6:312.

Gentile A, Ferreira TH, Mattos RS, Dias LI, Hoshino AA, Carneiro MS, Souza GM, Calsa T, Jr., Nogueira RM, Endres L, and Menossi M. 2013. Effects of drought on the microtranscriptome of field-grown sugarcane plants. *Planta* 237:783-798. 10.1007/s00425-012-1795-7

569 Guénin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C, and Gutierrez L. 2009.
570 Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental
571 conditions-specific, validation of references. *Journal of Experimental Botany* 60:487-493.
572
573 Guo J, Ling H, Wu Q, Xu L, and Que Y. 2014. The choice of reference genes for assessing gene
574 expression in sugarcane under salinity and drought stresses. *Scientific Reports* 4.
575 Guo J, Xu L, Su Y, Wang H, Gao S, Xu J, and Que Y. 2013. ScMT2-1-3, a metallothionein gene
576 of sugarcane, plays an important role in the regulation of heavy metal
577 tolerance/accumulation. *BioMed research international* 2013.
578 He B, Yun Z, Shi J, and Jiang G. 2013. Research progress of heavy metal pollution in China:
579 sources, analytical methods, status, and toxicity. *Chinese Science Bulletin* 58:134-140.
580 Hong-Bo S, Li-Ye C, Cheng-Jiang R, Hua L, Dong-Gang G, and Wei-Xiang L. 2010.
581 Understanding molecular mechanisms for improving phytoremediation of heavy metal-
582 contaminated soils. *Critical reviews in biotechnology* 30:23-30.
583 Hu X-F, Jiang Y, Shu Y, Hu X, Liu L, and Luo F. 2014a. Effects of mining wastewater discharges
584 on heavy metal pollution and soil enzyme activity of the paddy fields. *Journal of*
585 *Geochemical Exploration* 147:139-150.
586 Hu Y, Chen H, Luo C, Dong L, Zhang S, He X, and Huang G. 2014b. Selection of reference genes
587 for real-time quantitative PCR studies of kumquat in various tissues and under abiotic
588 stress. *Scientia Horticulturae* 174:207-216.
589 Huang G-Y, and Wang Y-S. 2009. Expression analysis of type 2 metallothionein gene in mangrove
590 species (Bruguiera gymnorrhiza) under heavy metal stress. *Chemosphere* 77:1026-1029.
591 Huang G-Y, and Wang Y-S. 2010. Expression and characterization analysis of type 2
592 metallothionein from grey mangrove species (Avicennia marina) in response to metal
593 stress. *Aquatic toxicology* 99:86-92.
594 Iskandar HM, Simpson RS, Casu RE, Bonnett GD, Maclean DJ, and Manners JM. 2004.
595 Comparison of reference genes for quantitative real-time polymerase chain reaction
596 analysis of gene expression in sugarcane. *Plant Molecular Biology Reporter* 22:325-337.
597 Janska A, Hodek J, Svoboda P, Zamecnik J, Prasil IT, Vlasakova E, Milella L, and Ovesna J. 2013.
598 The choice of reference gene set for assessing gene expression in barley (Hordeum vulgare
599 L.) under low temperature and drought stress. *Molecular Genetics and Genomics* 288:639-
600 649.
601 Kim YY, Choi H, Segami S, Cho HT, Martinoia E, Maeshima M, and Lee Y. 2009. AtHMA1
602 contributes to the detoxification of excess Zn (II) in Arabidopsis. *The Plant Journal*
603 58:737-753.
604 Kollikkathara N, Feng H, and Stern E. 2009. A purview of waste management evolution: Special
605 emphasis on USA. *Waste Management* 29:974-985.
606 http://dx.doi.org/10.1016/j.wasman.2008.06.032
607 Kozera B, and Rapacz M. 2013. Reference genes in real-time PCR. *Journal of applied genetics*
608 54:391-406.
609 Kundu A, Patel A, and Pal A. 2013. Defining reference genes for qPCR normalization to study

610 biotic and abiotic stress responses in Vigna mungo. *Plant Cell Reports* 32:1647-1658.
611 López-Landavery EA, Portillo-López A, Gallardo-Escárate C, and Del Río-Portilla MA. 2014.
612 Selection of reference genes as internal controls for gene expression in tissues of red
613 abalone Haliotis rufescens (Mollusca, Vetigastropoda; Swainson, 1822). *Gene* 549:258-
614 265.
615 Lee JH. 2013. An overview of phytoremediation as a potentially promising technology for
616 environmental pollution control. *Biotechnology and bioprocess engineering* 18:431-439.
617 Li M, Luo Y, and Su Z. 2007. Heavy metal concentrations in soils and plant accumulation in a
618 restored manganese mineland in Guangxi, South China. *Environmental pollution* 147:168-
619 175.
620 Lilly S, Drummond R, Pearson M, and MacDiarmid R. 2011. Identification and validation of
621 reference genes for normalization of transcripts from virus-infected Arabidopsis thaliana.
622 *Molecular Plant-Microbe Interactions* 24:294-304.
623 Ling H, Wu Q, Guo J, Xu L, and Que Y. 2014. Comprehensive Selection of Reference Genes for
624 Gene Expression Normalization in Sugarcane by Real Time Quantitative RT-PCR. *Plos*
625 *One* 9. 10.1371/journal.pone.0097469
626 Livak KJ, and Schmittgen TD. 2001. Analysis of relative gene expression data using real-time
627 quantitative PCR and the 2− ΔΔCT method. *methods* 25:402-408.
628 Macovei A, Ventura L, Donà M, Faè M, Balestrazzi A, and Carbonera D. 2010. Effect of heavy
629 metal treatments on metallothionein expression profiles in white poplar (Populus alba L.)
630 cell suspension cultures. *Ann Univ Oradea-Fascic Biol* 2:274-279.
631 Memon AR, and Schröder P. 2009. Implications of metal accumulation mechanisms to
632 phytoremediation. *Environmental Science and Pollution Research* 16:162-175.
633 Mikkelsen MD, Pedas P, Schiller M, Vincze E, Mills RF, Borg S, Møller A, Schjoerring JK,
634 Williams LE, and Baekgaard L. 2012. Barley HvHMA1 is a heavy metal pump involved
635 in mobilizing organellar Zn and Cu and plays a role in metal loading into grains. *Plos One*
636 7. doi:10.1371/journal.pone.0049027
637 Moreno I, Norambuena L, Maturana D, Toro M, Vergara C, Orellana A, Zurita-Silva A, and
638 Ordenes VR. 2008. AtHMA1 is a thapsigargin-sensitive Ca2+/heavy metal pump. *Journal*
639 *of Biological Chemistry* 283:9633-9641.
640 Mukherjee A, Bhattacharya P, Sarkar A, and Zevenhoven R. 2009. Mercury emissions from
641 industrial sources in India and its effects in the environment. In: Mason R, and Pirrone N,
642 eds. *Mercury Fate and Transport in the Global Atmosphere*: Springer US, 81-112.
643 Nicot N, Hausman J-F, Hoffmann L, and Evers D. 2005. Housekeeping gene selection for real-
644 time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of*
645 *Experimental Botany* 56:2907-2914.
646 Nishiuchi S, Liu S, and Takano T. 2007. Isolation and characterization of a metallothionein-1
647 protein in Chloris virgata Swartz that enhances stress tolerances to oxidative, salinity and
648 carbonate stress in Saccharomyces cerevisiae. *Biotechnology letters* 29:1301-1305.
649 Pan K, and Wang WX. 2012. Trace metal contamination in estuarine and coastal environments in
650 China. *Sci Total Environ* 421-422:3-16. 10.1016/j.scitotenv.2011.03.013

651 Papoyan A, and Kochian LV. 2004. Identification of Thlaspi caerulescens genes that may be
652 involved in heavy metal hyperaccumulation and tolerance. Characterization of a novel
653 heavy metal transporting ATPase. *Plant Physiology* 136:3814-3823.
654 Pfaffl MW, Tichopad A, Prgomet C, and Neuvians TP. 2004. Determination of stable
655 housekeeping genes, differentially regulated target genes and sample integrity:
656 BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnology letters* 26:509-
657 515.
658 Puschenreiter M, Horak O, Friesl W, and Hartl W. 2005. Low-cost agricultural measures to reduce
659 heavy metal transfer into the food chain–a review. *Plant Soil Environ* 51:1-11.
660 Que Y, Xu L, Xu J, Zhang J, Zhang M, and Chen R. 2009. Selection of control genes in real-time
661 qPCR analysis of gene expression in sugarcane. *Chinese Journal of Tropical Crops* 30:274-
662 278.
663 Sang J, Han X, Liu M, Qiao G, Jiang J, and Zhuo R. 2013. Selection and Validation of Reference
664 Genes for Real-Time Quantitative PCR in Hyperaccumulating Ecotype ofSedum alfredii
665 under Different Heavy Metals Stresses. *Plos One* 8:e82927.
666 Seigneurin-Berny D, Gravot A, Auroy P, Mazard C, Kraut A, Finazzi G, Grunwald D, Rappaport
667 F, Vavasseur A, and Joyard J. 2006. HMA1, a New Cu-ATPase of the Chloro plast
668 Envelope, Is Essential for Growth under Adverse Light Conditions. *Journal of Biological*
669 *Chemistry* 281:2882-2892.
670 Sekhar K, Priyanka B, Reddy V, and Rao K. 2011. Metallothionein 1 (CcMT1) of pigeonpea
671 (Cajanus cajan, L.) confers enhanced tolerance to copper and cadmium in Escherichia coli
672 and Arabidopsis thaliana. *Environmental and Experimental Botany* 72:131-139.
673 Sereno ML, Almeida RS, Nishimura DS, and Figueira A. 2007. Response of sugarcane to
674 increasing concentrations of copper and cadmium and expression of metallothionein genes.
675 *Journal of Plant Physiology* 164:1499-1515.
676 Shingu Y, Kudo T, Ohsato S, Kimura M, Ono Y, Yamaguchi I, and Hamamoto H. 2005.
677 Characterization of genes encoding metal tolerance proteins isolated from Nicotiana glauca
678 and Nicotiana tabacum. *Biochemical and Biophysical Research Communications* 331:675-
679 680.
680 Silva RLdO, Silva MD, Ferreira Neto JRC, Nardi CHd, Chabregas SM, Burnquist WL, Kahl G,
681 Benko-Iseppon AM, and Kido EA. 2014. Validation of Novel Reference Genes for Reverse
682 Transcription Quantitative Real-Time PCR in Drought-Stressed Sugarcane. *The Scientific*
683 *World Journal* 2014:12. http://dx.doi.org/10.1155/2014/357052
684 Silver N, Best S, Jiang J, and Thein SL. 2006. Selection of housekeeping genes for gene expression
685 studies in human reticulocytes using real-time PCR. *BMC Mol Biol* 7:33.
686 Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, and Speleman F. 2002.
687 Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of
688 multiple internal control genes. *Genome Biology* 3:research0034.
689 Wan H, Zhao Z, Qian C, Sui Y, Malik AA, and Chen J. 2010. Selection of appropriate reference
690 genes for gene expression studies by quantitative real-time polymerase chain reaction in
691 cucumber. *Anal Biochem* 399:257-261. 10.1016/j.ab.2009.12.008

692 Wang Z, Chen Y, Fang H, Shi H, Chen K, Zhang Z, and Tan X. 2014. Selection of reference genes
693 for quantitative reverse-transcription polymerase chain reaction normalization in Brassica
694 napus under various stress conditions. *Molecular Genetics and Genomics* 289:1023-1035.
695 Wong CSC, Duzgoren-Aydin NS, Aydin A, and Wong MH. 2006. Sources and trends of
696 environmental mercury emissions in Asia. *Science of The Total Environment* 368:649-662.
697 http://dx.doi.org/10.1016/j.scitotenv.2005.11.024
698 Xia H, Yan Z, Chi X, and Cheng W. 2009. Evaluation of the phytoremediation potential of
699 Saccharum officinarum for Cd-contaminated soil. Energy and Environment Technology,
700 2009 ICEET'09 International Conference on: IEEE. p 314-318.
701 Xie F, Xiao P, Chen D, Xu L, and Zhang B. 2012. miRDeepFinder: a miRNA analysis tool for
702 deep sequencing of plant small RNAs. *Plant Molecular Biology* 80:75-84.
703 Yang X, Feng Y, He Z, and Stoffella PJ. 2005. Molecular mechanisms of heavy metal
704 hyperaccumulation and phytoremediation. *Journal of Trace Elements in Medicine and*
705 *Biology* 18:339-353. http://dx.doi.org/10.1016/j.jtemb.2005.02.007
706 Yuan J, Chen D, Ren Y, Zhang X, and Zhao J. 2008. Characteristic and expression analysis of a
707 metallothionein gene, OsMT2b, down-regulated by cytokinin suggests functions in root
708 development and seed embryo germination of rice. *Plant Physiology* 146:1637-1650.
709 Yuan L, Yang S, Liu B, Zhang M, and Wu K. 2012. Molecular characterization of a rice metal
710 tolerance protein, OsMTP1. *Plant Cell Reports* 31:67-79.
711 Zhang M, Liu X, Yuan L, Wu K, Duan J, Wang X, and Yang L. 2012a. Transcriptional profiling
712 in cadmium-treated rice seedling roots using suppressive subtractive hybridization. *Plant*
713 *Physiology and Biochemistry* 50:79-86.
714 Zhang M, Senoura T, Yang X, and Nishizawa NK. 2011. Functional analysis of metal tolerance
715 proteins isolated from Zn/Cd hyperaccumulating ecotype and non-hyperaccumulating
716 ecotype of Sedum alfredii Hance. *Febs Letters* 585:2604-2609.
717 Zhang X, Zhu Y, Zhang Y, Liu Y, Liu S, Guo J, Li R, Wu S, and Chen B. 2014. Growth and metal
718 uptake of energy sugarcane (Saccharum spp.) in different metal mine tailings with soil
719 amendments. *Journal of Environmental Sciences* 26:1080-1089.
720 Zhang Y, Chen D, Smith MA, Zhang B, and Pan X. 2012b. Selection of reliable reference genes
721 in Caenorhabditis elegans for analysis of nanotoxicity. *Plos One* 7:e31849.
722 Zhu J, Zhang L, Li W, Han S, Yang W, and Qi L. 2013. Reference gene selection for quantitative
723 real-time PCR normalization in Caragana intermedia under different abiotic stress
724 conditions. *Plos One* 8:e53196.
725 Zhu X, Li X, Chen W, Chen J, Lu W, Chen L, and Fu D. 2012. Evaluation of new reference genes
726 in papaya for accurate transcript normalization under different experimental conditions.
727 *Plos One* 7:e44405.
728
729
730

Table 1(on next page)

Table 1 The expression of candidate reference genes in four cultivated sugarcane.

PeerJ PrePrints | https://dx.doi.org/10.7287/peerj.preprints.1487v1 | CC-BY 4.0 Open Access | rec: 9 Nov 2015, publ: 9 Nov 2015

1 **Table 1 The expression of candidate reference genes in four cultivated sugarcane**

gene	Mean Ct	SD	CV (%)
25S *rRNA*	14.20	0.893	6.29%
GAPDH	24.67	2.222	9.01%
ACT	24.45	0.683	2.80%
TUB	26.47	1.207	4.56%
18S *rRNA*	15.28	0.853	5.58%
UBQ	26.65	2.063	7.74%
eEF-1a	24.62	2.403	9.76%
eIF-4α	28.22	2.246	7.96%
CUL	27.51	2.35	8.54%
CAC	28.07	2.254	8.03%
TIPS-41	27.22	2.017	7.41%
APRT	27.59	2.148	7.79%
PRR	29.84	3.543	11.87%

SD, Standard Deviation;
CV, Covariation coefficient.

2

3

Table 2 (on next page)

Table 2 Correlation coefficients based on the visualizing reference genes ranked by geNorm, NormFinder and deltaCt.

Table 2 Correlation coefficients based on the visualizing reference genes ranked by geNorm, NormFinder and deltaCt.

	Correlations
geNorm *VS* NormFinder	0.932**
geNorm *VS* deltaCt	0.921**
geNorm *VS* BestKeeper	0.180
NormFinder *VS* deltaCt	0.972**
NormFinder *VS* BestKeeper	0.059
deltaCt *VS* BestKeeper	0.056

** Correlation is significant at the 0.01 level (2-tailed);

Table 3(on next page)

Table 3 Stability evaluation of 13 candidate reference genes by three statistical algorithms in *Saccharum* spp.

Table 3 Stability evaluation of 13 candidate reference genes by three statistical algorithms in *Saccharum* spp.

	geNorm		NormFinder		deltaCt		Comprehensive Rank	
	gene	SV	gene	SV	gene	SV	gene	CSV
1	*CAC*	1.00	*UBQ*	1.00	*APRT*	1.00	*UBQ*	1.09
2	*CUL*	1.00	*GAPDH*	1.82	*CAC*	1.00	*APRT*	1.32
3	*APRT*	1.22	*APRT*	1.89	*UBQ*	1.01	*CUL*	1.34
4	*UBQ*	1.29	*eEF-1α*	2.33	*CUL*	1.02	*CAC*	1.36
5	*GAPDH*	1.36	*CUL*	2.33	*GAPDH*	1.05	*GAPDH*	1.38
6	*eEF-1α*	1.53	*eIF-4a*	2.47	*eIF-4a*	1.12	*eEF-1α*	1.59
7	*eIF-4a*	1.77	*CAC*	2.49	*eEF-1α*	1.13	*eIF-4a*	1.70
8	25S *rRNA*	2.44	25S *rRNA*	3.09	25S *rRNA*	1.19	25S *rRNA*	2.08
9	18S *rRNA*	2.95	18S *rRNA*	3.43	*TUB*	1.30	18S *rRNA*	2.39
10	*TUB*	3.41	*ACT*	4.59	18S *rRNA*	1.35	*TUB*	2.76
11	*ACT*	3.79	*TUB*	4.73	*ACT*	1.70	*ACT*	3.09
12	*PRR*	4.39	*PRR*	7.85	*PRR*	2.17	*PRR*	4.21
13	*TIPS-41*	5.24	*TIPS-41*	9.92	*TIPS-41*	2.82	*TIPS-41*	5.27

SV, stability value; CSV, comprehensive stability value.

Figure 1(on next page)

Figure 1 The optimal combination of reference genes for gene expression normalization under heavy metal stresses in sugarcane.

The pairwise variation (V_n/V_{n+1}) was analyzed between normalization factors NF_n and NF_{n+1} by geNorm program to determine the optimal combination of reference genes for accurate normalization in samples from different sugarcane cultivar samples.

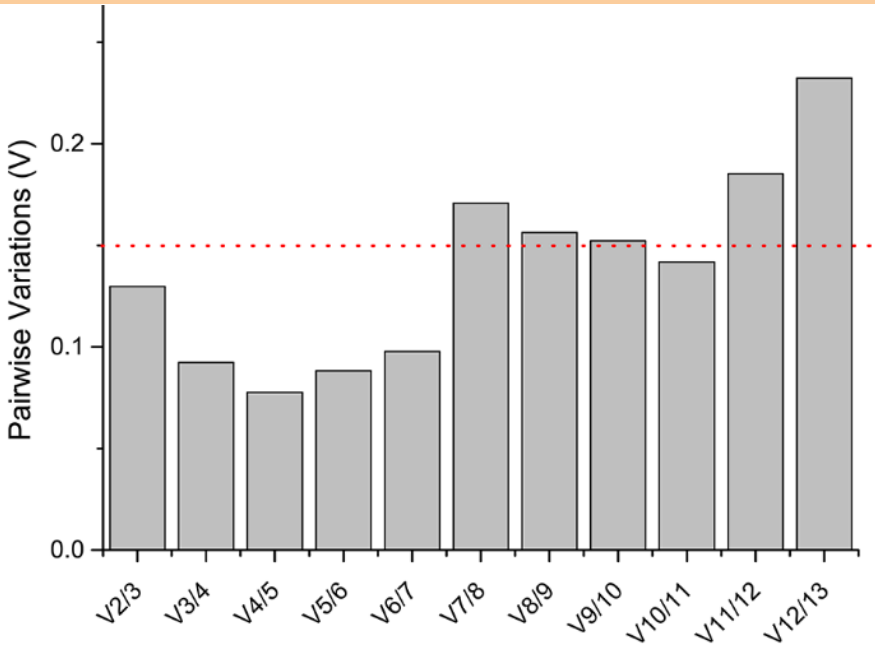
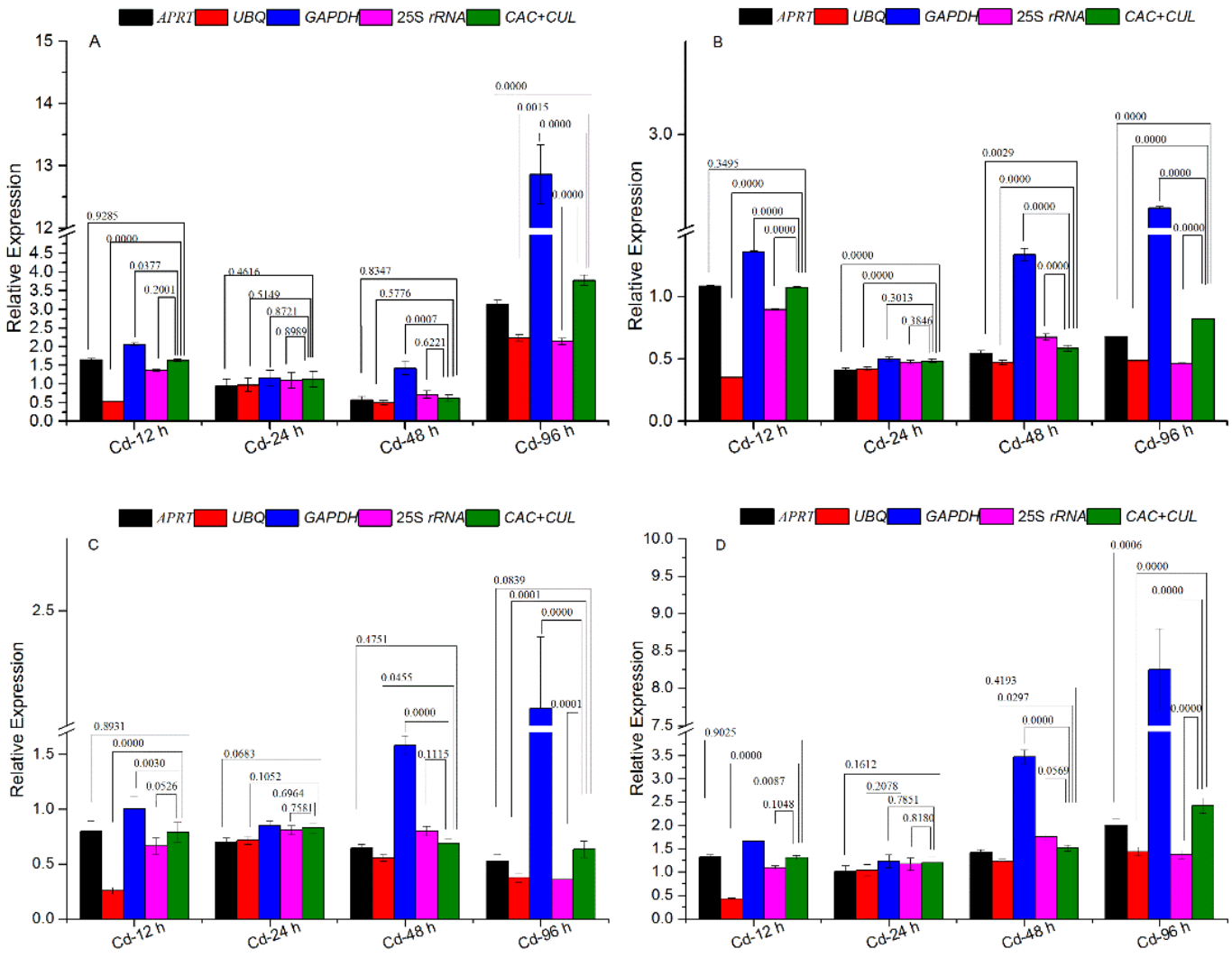


Figure 2(on next page)

Figure 2 Expression analysis of *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1* genes based on selected reference gene / genes under cadmium stress.

ScMTP (Sugarcane metal tolerance protein gene, GenBank Accession No. KP864146), *ScMT2-1-3* (Sugarcane metallothionein, GenBank Accession No. JQ627644), *ScMPP* (sugarcane metallophosphoesterase, GenBank Accession No. CA267392.1**)** and *ScHMA1* (sugarcane heavy metal transporting ATPase 1, GenBank Accession No. CA156665.1) were heavy metal stress response genes in sugarcane. In this study, the normalization of *ScMTP* (A), *ScMT2-1-3* (B), *ScMPP* (C) and *ScHMA1* (D) employed a single reference gene, *UBQ*, *APRT*, *GAPDH* or 25S *rRNA*, or the reference gene set, *CAC* + *CUL* as reference control under cadmium chloride ($CdCl_2$) treatment. Using $2^{-\Delta\Delta Ct}$ to normalize the *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1*, the control sample were converted into 1. Significant different expression of *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1* were marked with p value ($p < 0.01$ level (highly significant) and $p < 0.05$ level (significant)) on the line respectively when comparing the normalization by *UBQ*, *APRT*, *GAPDH* and 25S *rRNA* with the normalization by *CAC* + *CUL*.



PeerJ PrePrints | https://dx.doi.org/10.7287/peerj.preprints.1487v1 | CC-BY 4.0 Open Access | rec: 9 Nov 2015, publ: 9 Nov 2015

Figure 3(on next page)

Figure 3 Expression analysis of *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1* genes based on selected reference gene / genes under copper stress.

ScMTP (Sugarcane metal tolerance protein gene, GenBank Accession No. KP864146), *ScMT2-1-3* (Sugarcane metallothionein, GenBank Accession No. JQ627644), *ScMPP* (sugarcane metallophosphoesterase, GenBank Accession No. CA267392.1**)** and *ScHMA1* (sugarcane heavy metal transporting ATPase 1, GenBank Accession No. CA156665.1) were heavy metal stress response genes in sugarcane. In this study, the normalization of *ScMTP* (A), *ScMT2-1-3* (B), *ScMPP* (C) and *ScHMA1* (D) employed a single reference gene, *UBQ*, *APRT*, *GAPDH* or 25S *rRNA*, or the reference gene set, *CAC* + *CUL* as reference control under copper chloride ($CuCl_2$) treatment. Using $2^{-\Delta\Delta Ct}$ to normalize the *ScMTP*, *ScMT2-1-3*, *ScMPP*and *ScHMA1*, the control sample were converted into 1. Significant different expression of *ScMTP*, *ScMT2-1-3*, *ScMPP* and *ScHMA1* were marked with p value ($p < 0.01$ level (2-tailed) and the 0.05 level (1-tailed) on the line respectively when comparing the normalization by *UBQ*, *APRT*, *GAPDH* and 25S *rRNA* wtih the normalization by *CAC* + *CUL*.