HAL
open science

Critical selection of reliable reference genes for gene expression study in the HepaRG cell line

Liesbeth Ceelen, Ward de Spiegelaere, Michael David, Jurgen de Craene, Mathieu Vinken, Tamara Vanhaecke, Vera Rogiers

To cite this version:

Liesbeth Ceelen, Ward de Spiegelaere, Michael David, Jurgen de Craene, Mathieu Vinken, et al.. Critical selection of reliable reference genes for gene expression study in the HepaRG cell line. Biochemical Pharmacology, 2011, 10.1016/j.bcp.2011.03.004 . hal-00690184

HAL Id: hal-00690184

https://hal.science/hal-00690184v1

Submitted on 22 Apr 2012

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

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

Accepted Manuscript

Title: Critical selection of reliable reference genes for gene expression study in the HepaRG cell line

Authors: Liesbeth Ceelen, Ward De Spiegelaere, Michael David, Jurgen De Craene, Mathieu Vinken, Tamara Vanhaecke, Vera Rogiers

PII: S0006-2952(11)00149-3
DOI: doi:10.1016/j.bcp.2011.03.004
Reference: BCP 10844

To appear in: *BCP*

Received date: 4-2-2011
Revised date: 4-3-2011
Accepted date: 8-3-2011

Please cite this article as: Ceelen L, De Spiegelaere W, David M, De Craene J, Vinken M, Vanhaecke T, Rogiers V, Critical selection of reliable reference genes for gene expression study in the HepaRG cell line, *Biochemical Pharmacology* (2010), doi:10.1016/j.bcp.2011.03.004

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Critical selection of reliable reference genes for gene expression study in the HepaRG cell line

Liesbeth Ceelen[1], Ward De Spiegelaere[2], Michael David[1], Jurgen De Craene[2], Mathieu Vinken[1], Tamara Vanhaecke[1], Vera Rogiers[1]

Author affiliations:

[1]Department of Toxicology, Faculty of Medicine and Pharmacy, "Center for Pharmaceutical Research", Vrije Universiteit Brussel, Laarbeeklaan 103, B-1090 Brussels, Belgium.

[2]Department of Morphology, Faculty of Veterinary Medicine, Ghent University, Salisburylaan 133, B-9820 Merelbeke, Belgium.

L.C. is postdoctoral research fellow of the Industrial Research Fund (IOF) of Vrije Universiteit Brussel; M.V. is postdoctoral research fellow of the Fund for Scientific Research Flanders (FWO-Vlaanderen), Belgium.

Correspondence: Liesbeth Ceelen (Ph.D, D.V.M., MScLAS), Department of Toxicology, Centre for Pharmaceutical Research, Faculty of Medicine and Pharmacy, Vrije Universiteit Brussel, Laarbeeklaan 103, B-1090 Brussels, Belgium, Tel: +32.2.477.45.20, Fax: +32.2.477.45.82, e-mail: lceelen@vub.ac.be

Short title: reference genes in HepaRG.

Word count: - Abstract: 252

- Text (excluding References, Table and Figure legends): 3791

Accepted Manuscript

ACCEPTED MANUSCRIPT

Abstract

The human HepaRG cell line has shown to be a valuable *in vitro* tool for repeated exposure to chemical compounds and to evaluate their potential toxic outcome. Seen the importance given by the actual EU legislation of cosmetics and chemical substances to the use of *in vitro* methods in human safety evaluation, one can expect that HepaRG cells will gain importance as human-relevant cell source. At the transcriptional level, RT-qPCR assays are often used to obtain quantitative results. The choice of internal control is important since it may affect the study outcome. Indeed, it is well-known that expression levels of traditional reference genes can vary across tissue types and across experimental settings within one specific tissue type. From a review of the scientific literature, it appears that, for HepaRG cells, *S18* often is used as internal control, but without any evidence of its expression stability in this cell line. Therefore, we aimed to select the most optimal reference genes for gene expression studies in HepaRG cells and to check whether *S18* is a suitable reference gene. Twelve candidate genes' expression stability level was analyzed by three algorithms (geNorm, BestKeeper, Normfinder), which identified the optimal single reference gene (*TBP*) and the most suitable set of reference genes (*TBP*, *UBC*, *SDHA*, *RLP13*, *YHWAZ*, *HMBS*, *B2M* and *HPRT1*) for HepaRG transcriptional profiling. This study provides a new set of reference genes that is suitable for testing whenever RT-qPCR data for HepaRG cells are generated. The most stable ones can then be selected for further normalization.

Keywords: HepaRG, RT-qPCR, reference gene, normalization

Accepted Manuscript

ACCEPTED MANUSCRIPT

1. Introduction

Drug-induced liver toxicity is among the main reasons to withdraw drugs from the market [1]. This may be partially explained by the poor capability of the currently available animal-based *in vitro* systems to predict human toxicity in preclinical testing and hence implicates an urgent need for a reliable *in vitro* system that appropriately reflects the human *in vivo* situation for hepatotoxicity testing. Notwithstanding primary human hepatocytes still represent the gold standard in the field of liver-based *in vitro* models, many efforts have been done for extended periods of time, to develop a human hepatic cell line expressing various liver-specific functions at an *in vivo*-like level. The hepaRG cell line, derived from a human hepatocellular carcinoma, seems to meet the majority of these criteria and therefore represents for the time being a suitable alternative for freshly isolated human hepatocytes in drug screening [2-4]. Seen the importance given by the actual EU legislation of cosmetics (76/768/EEC and 2003/15/EC) and chemical substances to the use of *in vitro* methods in human safety evaluation, it could be expected that in the near future the HepaRG cells will gain importance as human-relevant cell source. Actually, the human HepaRG CYP induction test method is currently under validation at the European Centre for the Validation of Alternative Methods (ECVAM Technical report on the Status of Alternative Methods for Cosmetics Testing, 2008-2009; http://ecvam.jrc.ec.europa.eu/). The HepaRG cell line has indeed shown high metabolic capacity compared to primary human hepatocytes [5, 6]. These cells have been used for long-term repeated exposure to evaluate the potential toxic effects of different chemical entities [5, 7]. Once differentiated using 2% dimethylsulfoxide (DMSO), cells at confluency show indeed a relative preserved expression of transcripts encoding various phase I, phase II and antioxidant enzymes, membrane (apical, canalicular and basolateral) transporters, the nuclear receptors constitutive androstane receptor and pregnane X receptor, aldolase B and albumin, up to 6 weeks in culture [3-5, 7, 8]. Expression of a plethora of markers/functions has been

ACCEPTED MANUSCRIPT

evaluated both at the transcriptional and the translational level. With respect to the former, reverse transcription quantitative PCR (RT-qPCR) assays were generally adopted often using the *S18* rRNA gene as single reference gene [7-13].

In the entire qPCR process, a number of variables, including RNA extraction (yield, integrity and DNA contamination), efficiency of reverse transcription and PCR cycling impose the necessity to include a number of controls to guarantee the integrity of every step and hence, to achieve reliable and accurate results. It is therefore important to use reliable reference genes as internal controls in order to normalize mRNA data [14-17]. Their efficacy, however, must be experimentally validated for a particular tissue type and/or experimental condition. Indeed, it is known that the expression profiles of commonly used control genes can vary across different cell types, but also within one cell type subjected to a different experimental treatment [15, 16, 18-20]. Consequently, the selection of inappropriate reference genes may significantly affect the study outcome by inaccurate data interpretation as highlighted in a number of previously reported studies involving a variety of cell types and experimental situations [17, 21, 22].

Therefore, the current study was set up to search for the best reference genes for RT-qPCR assays in the HepaRG cell line using the geNorm [17], BestKeeper [23] and Normfinder [24] software. The quality assessments described, are in line with the recently published MIQE guidelines (Minimum Information for Publications of Quantitative Real-Time PCR Experiments) that provide authors and reviewers with the minimum required information to ensure the quality of interpretation and repeatability of qPCR experiments [15].

2. Materials and methods

2.1.HepaRG Cell cultures

HepaRG cells, originally derived from a liver tumor of a female patient suffering from hepatocarcinoma [2], were purchased from Biopredic International (BPI, Rennes, France). For

ACCEPTED MANUSCRIPT

the present study, HepaRG cells were cultured as previously described. Thus, cells at three different passages were cultured at low density (*i.e.* 2.6×10^4 cells/cm^2) or high density (*i.e.* 4.5×10^5 cells/cm^2). Cells at low density were incubated in Williams' E medium (Gibco, Invitrogen, Merelbeke, Belgium) supplemented with 10% fetal calf serum (Gibco), 100 units/ml penicillin-100 µg/ml streptomycin (Gibco), 5 µg/ml insulin (Sigma-Aldrich, Bornem, Belgium), 2 mM glutamine (Sigma-Aldrich), and 5×10^{-5} M hydrocortisone hemisuccinate (Pfizer, Puurs, Belgium). After two weeks of culture, cells were shifted to the same culture medium supplemented with 2% DMSO (Sigma-Aldrich) for two more weeks in order to reach complete differentiation and were maintained for almost two additional weeks. Cells at high density were immediately incubated in the 2% DMSO (Sigma-Aldrich) containing culture medium. Culture media of both low and high density cells were renewed every two or three days [2, 9, 10].

2.2.Isolation of RNA

Total RNA was isolated with the GenElute Mammalian Total RNA Miniprep Kit (Sigma-Aldrich) which directly included a DNase treatment step from HepaRG cells sampled on days 7 and 15 of the proliferation period (n = 6; UNDIFF samples) and on days 1, 3, 10, 12, 15, 16, 18, 19 and 26 of the differentiation period (n = 17; DIFF samples which included cells treated with DMSO to reach complete differentiation and cells at fully differentiated stage, seeded at both densities).

RNA integrity was analyzed with the Agilent 2100 Bioanalyzer™ Automated Gel Electrophoresis System (Agilent Technologies, Diegem, Belgium) according to the manufacturer's protocol. Its algorithm gives an RNA Integrity Number (RIN) which scales from 1 to 10 with 1 being totally degraded RNA and 10 being totally intact RNA. Sample integrity was determined by the entire electrophoretic trace of the RNA sample [25]. This contains the presence or absence of degradation products. As the allocated RIN is independent

ACCEPTED MANUSCRIPT

of sample concentration, instrument and analyst, it is becoming a gold standard for RNA integrity [15]. The isolated RNA was reverse transcribed with the iScript™ cDNA Synthesis Kit (Bio-Rad, Nazareth, Belgium) followed by cDNA purification with the Genelute PCR clean up kit (Sigma-Aldrich) to remove potential PCR inhibitors [26]. cDNA was finally quantified using a Nanodrop® microvolume spectrophotometer (Thermo Scientific Nanodrop Products, Wilmington, Delaware).

2.3. *RT-qPCR analysis*

Based on literature search [17, 23, 27, 28], 12 putative reference genes in the present study were selected for evaluation of their expression profile (Table 1).

Quantative-PCR was performed using either TaqMan hydrolysis probes for *TBP*, *RLP13*, *ACTB*, *SDHA*, *UBC*, *B2M* (PrimeTime qPCR Assay 6-FAM/ZEN/IBFQ, synthesized by Integrated DNA Technology, IDT, Iowa, USA), *S18*, *GAPDH* and *HPRT1* (TaqMan Gene Expression Assays FAM™ including dye-labeled TaqMan® MGB probe, synthesized by Applied Biosystems, Halle, Belgium) or Perfecta™ SYBR Green FastMix (Quanta Biosciences Inc., Gaithersburg, USA) for *YHWAZ*, *TOP2B* and *HMBS* (synthesized by IDT). Sequences of primers and probes are outlined in Table 2. All samples were run in duplicate and each run included two negative controls and a serial dilution of a pooled cDNA mix from all samples to calculate the standard curve. The PCR reaction mix consisted of (i) 100 nM of each primer, 12.5 µL Perfecta™ SYBR Green FastMix (Quanta Biosciences Inc.) and 5 µl of DNA in a 25 µl volume adjusted with DNase/RNase-free HLPC water for *YHWAZ*, *TOP2B* and *HMBS*, (ii) 900 nM of each primer, 250nM of probe (both present in 1.25 µl Assay Demand, Applied Biosystems), 12.5 µl TaqMan Universal Master Mix (Applied Biosystems) and 1 or 2 µl of DNA (depending on the tested gene) in a 25 µl volume adjusted with DNase/RNase-free HLPC water for *S18*, *HPRT1* and *GAPDH* or (iii) 500 nM of each primer,

ACCEPTED MANUSCRIPT

250 nM of probe (both present in 1 µl 20X Prime Time Assay, IDT), 10 µl TaqMan Universal Master Mix (Applied Biosystems) and 1 to 3 µl of cDNA (depending on the tested gene) in a 20 µl volume adjusted with DNase/RNase-free HLPC water for *TBP*, *RLP13*, *UBC*, *B2M*, *SDHA* and *ACTB*. Reactions were performed on the iQ^{TM}5 Bio-rad system (Bio-rad), with cycling conditions as depicted in Table 2. To verify the accuracy of the *YHWAZ*, *TOP2B* and *HMBS* amplicons, a melting curve analysis was performed after amplification (for *YWHAZ* en *TOP2B* from 60° to 94° and for HMBS from 59°C to 94°C, with an interval of 0.5°C and a holding time of 30s).

When possible, probes and primers were designed so that they spanned an exon-exon junction to avoid amplification of genomic DNA.

2.4.Software programs used for statistical analysis and qPCR data processing

Three different algorithms and software programs that calculate the most stable reference genes which are freely available on the Internet were used for statistical analysis and qPCR data processing.

geNorm is a Visual Basic Application for Microsoft Excel which defines the expression stability of a possible reference gene by allocating a so-called M value to each gene in a pool of candidate reference genes. Stepwise exclusion of the reference gene with the least stable expression finally assigns the two most stable genes [17].

Normfinder is also a Visual Basic applet which assigns a stability value to the candidate reference genes. This robust algorithm, unlike geNorm, adopts a model-based approach to give a score for the two most stable reference genes with the least intra-and inter-group variation. Stability is expressed as a stability value in arbitrary units. Additionally, Normfinder possesses the ability to discriminate between sample variability and bias between

Accepted Manuscript

ACCEPTED MANUSCRIPT

several groups. The application indeed estimates the intra- and intergroup variance which then describes the stability of the gene expression between the groups resulting in an optimum pair of reference genes [24].

BestKeeper, another Excel-based tool, determines the variability in expression of a set of reference genes by analyzing quantification cycle (Cq) values and classifying variability by the coefficient of variance (CV) and the standard deviation (SD). To define the most stable reference gene, the software generates a BestKeeper index which finally is compared to each candidate gene resulting in a value for the Pearson correlation coefficient (r) and probability (p) which are then allocated to each candidate reference gene [23].

3. Results

3.1. RNA and DNA quantity and quality assurance

The assessed RIN for all 23 samples containing 478 (± 355) ng/µl total RNA fell within a range from 8.6 to 10 indicating the presence of (almost) totally intact RNA, and was hence useful for further qPCR analysis.

Nanodrop®-quantified final cDNA concentrations in all 23 samples varied from 17.8 to 31.9 ng/µl (average = 22.9 ng/µl).

3.2. Statistical and RT-qPCR analysis

PCR-efficiency was calculated from the slope of the standard curves for each gene and fell within the required range of 90-110%. Accuracy of the *YHWAZ*, *TOP2B* and *HMBS* amplicons was proven by a melting curve analysis since one single amplicon was generated.

Figure 1 shows the cycle threshold value (Cq-value) of the 12 tested reference genes which display a substantial variation in expression levels. The Cq-value refers to the fractional PCR cycle at which the fluorescent signal considerably surmounts the background signal. Indeed,

ACCEPTED MANUSCRIPT

HMBS, *UBC*, *TBP*, *HPRT1* and *TOP2B* exhibit a rather moderate mRNA expression (Ct-value range: 25.45-26.45), while the Cq-value related to *S18* was 13.41.

Statistical analysis and determination of appropriate reference genes was done by geNorm, Normfinder and BestKeeper.

geNorm

Stepwise exclusion of the reference gene with the least stable expression assigned *TBP* and *SDHA* as most stably expressed reference genes in all analyzed samples (DIFF and UNDIFF) (Table 3) as well as in the group of only DIFF samples. In contrast, *UBC* and *YHWAZ* showed to be the most stably expressed reference genes in the UNDIFF samples. *TBP* and *SDHA* were in the latter group of samples clearly stably expressed as well (Figure 2a-b-c). Finally, by using the pairwise variation between two sequential normalization factors, the software indicated that reliable data for HepaRG can be obtained in RT-qPCR when normalization is performed using a set of the two most stably expressed genes (V2/3 < 0.15) (all samples: *TBP* and *SDHA*; UNDIFF samples: *UBC* and *YHWAZ*) or the three most stably expressed genes (V3/4 < 0.15) (DIFF samples: *TBP*, *SDHA* and *UBC*). Next to *TOP2B* and *ACTB*, the traditionally used reference gene *S18*, was unstably expressed in HepaRG cells (Table 3).

The solid position of the reference genes assessed by geNorm was appraised by comparison with two other Excel-based applications *i.e.* Normfinder [24] and BestKeeper [23].

Normfinder

Analysis of our data showed that *UBC* and *TBP* were the most stable reference gene in DIFF and UNDIFF HepaRG cells, respectively (Table 3). Distribution of the least stably expressed genes within the DIFF and UNDIFF group of HepaRG samples was slightly different as *RLP13*, *GAPDH* and *ACTB* were the least stably expressed genes in the UNDIFF samples,

ACCEPTED MANUSCRIPT

while *TOP2B*, *B2M* and *S18* displayed the least stable expression in the DIFF samples (Table 4). When taking into account all analyzed samples (DIFF and UNDIFF), *S18*, *TOP2B* and *ACTB* were again the most unstable genes (Table 3). Following estimation of the intra- and intergroup variance in all HepaRG samples in the present study, the most optimum pair of genes turned out to be *TBP* and *HPRT1*.

BestKeeper

Since all genes with a SD >1 are considered inappropriate as reference genes, we could exclude *ACTB* and *GAPDH* for analysis when DIFF and UNDIFF HepaRG cells were analyzed both jointly or separately. In addition, when all 23 samples were studied, an SD value of 1.05 was defined for the *TOP2B* gene and could therefore not be considered as a suitable internal control. Moreover, *TOP2B* and *S18* showed an SD value higher than 1 in the group containing the DIFF samples and were hence not fitting as reference genes in differentiating and differentiated HepaRG cells.
Again, *TBP*, *SDHA* and *UBC* were found to be the most stable reference genes in all 23 samples (Table 3). The BestKeeper analysis also assigned *TBP* as best reference gene in undifferentiated HepaRG cells, but ranked this gene as fourth stably expressed gene in differentiated cells. However, SD (0,89), r (0,961) and p (0.001) values of *TBP* were situated closely to the best ranked reference gene *RLP13* (SD = 0,76; r = 0.968; p = 0.001) considering *TBP* also as valuable reference gene when only the DIFF samples are evaluated in the BestKeeper software (Table 4).

When taking geNorm, BestKeeper and Normfinder-analyzed data of all HepaRG samples together, positions for top and bottom ranked genes were fairly constant between the geNorm and BestKeeper software programs. The top three ranked genes included *TBP*, *SDHA* and *UBC*, while *ACTB*, *S18* are *TOP2B* were listed (nearly) at the bottom of the ranking in both

Accepted Manuscript

ACCEPTED MANUSCRIPT

softwares. Rankings differed when using Normfinder, especially for the midfield positioned genes, but this applet also assigned *TBP*, *UBC* and *SDHA* as reliable genes when all 23 samples were analysed.

4. Discussion

The establishment of a clear-cut set of reliable reference genes is an essential prerequisite for analyzing gene expression using RT-qPCR as shown by several studies [15, 17, 21, 22, 29]. Often 'traditional' reference genes, such as *S18*, are used as internal controls without extensively verifying their validity. As such, the *S18* rRNA gene has been used in most studies in which the HepaRG cell line is involved [7-13]. Next to human primary hepatocytes, this hepatocyte model seems to be a very suitable tool for *in vitro* toxicity testing of xenobiotics and even has some additional advantages [5, 6]. Indeed, the cell line can be used either undifferentiated (progenitor cells) or differentiated (hepatocyte-like and biliary-like cells) making it possible to distinguish between compounds that are directly toxic and those that need activation to be hazardous to either cell type. It generally acquires both the metabolic capacity of freshly isolated human hepatocytes as well as the potential to grow indefinitely like cancer cells [5, 6]. This commercially available cell line is indeed increasingly adopted at the academic as well as industrial level. Hence, when performing RT-qPCR experiments, it is crucial to include appropriate normalization controls due to their progressively apparent significance as the amount of biological studies applying this methodology has increased.

To diminish the menace of erroneous results because of the use of invalid reference genes, we validated 12 potential reference genes, using geNorm, BestKeeper and Normfinder algorithms, suitable for RT-qPCR profiling experiments in this promising *in vitro* liver-based model during proliferation (UNDIFF), during differentiation and at fully differentiated stage

ACCEPTED MANUSCRIPT

(DIFF). Cultivation was performed exactly as described in the literature [2, 9, 10].

As mentioned by Vandesompele et al. [17], mRNA expression data are much more reliable and accurate when they are normalized using the geometric mean of multiple reference genes. As such, geNorm analysis demonstrated that, for the current study design, two valuable reference genes (*TBP* and *SDHA*) are required for normalization when analyzing expression levels in HepaRG cells at proliferating and differentiating or differentiated stage (*i.e.* all samples and UNDIFF samples). However, when mRNA expression was analyzed in merely the group of differentiating and differentiated cells (DIFF samples), geNorm assessment showed that adding the 3rd candidate reference gene (*TBP* next to *UBC* and *YHWAZ*) to the normalization factor, had the largest influence on diminishing variability. This undoubtedly underscores the previously stated importance of selection of a sufficient set of reference genes because of the variability within a single cell type across different experimental treatments.
GeNorm additionally showed that *S18*, commonly used as single internal control in HepaRG studies, was nearly the least stably expressed reference gene in HepaRG cells at all stages. This finding was confirmed by Normfinder and BestKeeper analyses. Collectively, these data show that *S18* is not suitable for normalization in the HepaRG cell system - and definitely not as single reference gene - resulting in possible misinterpretation of qPCR data.

In addition, current analyses showed that geNorm and BestKeeper both ranked the examined candidate genes generally in comparable order, certainly those positioned in the top and bottom ranking. Normfinder, however, provided a slightly different list of ranked genes. It defined for example *HPRT1* as a reference gene of the most optimum pair next to *TBP*, whereas the *HPRT1* gene was only positioned as the 7th and 8th most stable reference gene in the geNorm and BestKeeper analyses, respectively. This may be explained by the fact that

ACCEPTED MANUSCRIPT

Normfinder evaluates the candidate as to whether it shows systemic variation across the sample subgroups and hence does not only take overall expression variation into account [24]. Nonetheless, *UBC* and *TBP* were also highly ranked as potential reference genes by this applet, while *TOP2B* and *S18* were considered less valuable or even useless as reference genes in the HepaRG cell line.

Discrepancy was also seen for some genes when DIFF and UNDIFF samples were separately analysed, for example for the *B2M* gene and the *ACTB* gene, which were differently ranked within both groups by the three software programs and Normfinder, respectively. This finding again emphasizes the need for testing several candidate reference genes in HepaRG cells since some reference genes are differently regulated across multiple experimental conditions. This may be, for example, the case for studies wherein HepaRG cells are treated with different inducers of biotransformation phase I and phase II enzymes.

Although as few as two or three reference genes can serve as a good standard [29, 30] which was in our study demonstrated by geNorm analysis too, we would suggest, when setting up RT-qPCR experiments in HepaRG, to use an initial set of minimum six to eight candidate reference genes which have been confirmed in the current study to be the most stably expressed. When analyses of the three software programs are combined, this set of genes contains *TBP*, *UBC*, *SDHA*, *RLP13*, *YHWAZ*, *HMBS*, *B2M* and *HPRT1*. From these candidates, minimum two reference genes can then be selected for normalization in a specific experimental design with HepaRG cells.

In conclusion, while *TBP* was proven to be overall the most stable reference gene, *TOP2B*, *ACTB* and *S18* were considered less valuable as internal control in HepaRG cells. In general, we propose to test a set of at least six to eight reference genes (*TBP*, *UBC*, *SDHA*, *RLP13*,

ACCEPTED MANUSCRIPT

YHWAZ, *HMBS*, *B2M* and *HPRT1*) of which then the most stable genes can be picked out for further normalization in future HepaRG expression studies. As such, reproducibility and validity of RT-qPCR experiments in this *in vitro* liver-based model may notably improve.

5. Acknowledgements

This work was supported by grants from the Industrial Research Funds, (IOF), Vrije Universiteit Brussel, Belgium. Special thanks go to the European Union (FP6 project carcinoGENOMICS). We also thank Prof. P. Simoens (Ghent University) for his critical comments.

6. References

[1] Lasser KE, Allen PD, Woolhandler SJ, Himmelstein DU, Wolfe SM, Bor DH. Timing of new black box warnings and withdrawals for prescription medications. JAMA 2002;287:2215-20.

[2] Gripon P, Rumin S, Urban S, Le Seyec J, Glaise D, Cannie I, et al. Infection of a human hepatoma cell line by hepatitis B virus. Proc Natl Acad Sci U S A 2002;99:15655-60.

[3] Guillouzo A, Corlu A, Aninat C, Glaise D, Morel F, Guguen-Guillouzo C. The human hepatoma HepaRG cells: a highly differentiated model for studies of liver metabolism and toxicity of xenobiotics. Chem Biol Interact 2007;168:66-73.

[4] Kanebratt KP, Andersson TB. Evaluation of HepaRG cells as an in vitro model for human drug metabolism studies. Drug Metab Dispos 2008;36:1444-52.

[5] Guguen-Guillouzo C, Corlu A, Guillouzo A. Stem cell-derived hepatocytes and their use in toxicology. Toxicology 2010;270:3-9.

[6] Guillouzo A, Guguen-Guillouzo C. Evolving concepts in liver tissue modeling and implications for in vitro toxicology. Expert Opin Drug Metab Toxicol 2008;4:1279-94.

[7] Josse R, Aninat C, Glaise D, Dumont J, Fessard V, Morel F, et al. Long-term functional stability of human HepaRG hepatocytes and use for chronic toxicity and genotoxicity studies. Drug Metab Dispos 2008;36:1111-8.

[8] Le Vee M, Jigorel E, Glaise D, Gripon P, Guguen-Guillouzo C, Fardel O. Functional expression of sinusoidal and canalicular hepatic drug transporters in the differentiated human hepatoma HepaRG cell line. Eur J Pharm Sci 2006;28:109-17.

[9] Aninat C, Piton A, Glaise D, Le Charpentier T, Langouet S, Morel F, et al. Expression of cytochromes P450, conjugating enzymes and nuclear receptors in human hepatoma HepaRG cells. Drug Metab Dispos 2006;34:75-83.

[10] Antherieu S, Chesne C, Li R, Camus S, Lahoz A, Picazo L, et al. Stable expression, activity, and inducibility of cytochromes P450 in differentiated HepaRG cells. Drug Metab Dispos 2010;38:516-25.

[11] Cerec V, Glaise D, Garnier D, Morosan S, Turlin B, Drenou B, et al. Transdifferentiation of hepatocyte-like cells from the human hepatoma HepaRG cell line through bipotent progenitor. Hepatology 2007;45:957-67.

ACCEPTED MANUSCRIPT

[12] Dumont J, Josse R, Lambert C, Antherieu S, Laurent V, Loyer P, et al. Preferential induction of the AhR gene battery in HepaRG cells after a single or repeated exposure to heterocyclic aromatic amines. Toxicol Appl Pharmacol 2010;249:91-100.

[13] Lambert CB, Spire C, Renaud MP, Claude N, Guillouzo A. Reproducible chemical-induced changes in gene expression profiles in human hepatoma HepaRG cells under various experimental conditions. Toxicol In Vitro 2009;23:466-75.

[14] Teste MA, Duquenne M, Francois JM, Parrou JL. Validation of reference genes for quantitative expression analysis by real-time RT-PCR in Saccharomyces cerevisiae. BMC Mol Biol 2009;10:99.

[15] Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 2009;55:611-22.

[16] Thellin O, Zorzi W, Lakaye B, De Borman B, Coumans B, Hennen G, et al. Housekeeping genes as internal standards: use and limits. J Biotechnol 1999;75:291-5.

[17] Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002;3:RESEARCH0034.

[18] Nolan T, Hands RE, Bustin SA. Quantification of mRNA using real-time RT-PCR. Nat Protoc 2006;1:1559-82.

[19] Gonzalez-Verdejo CI, Die JV, Nadal S, Jimenez-Marin A, Moreno MT, Roman B. Selection of housekeeping genes for normalization by real-time RT-PCR: analysis of Or-MYB1 gene expression in Orobanche ramosa development. Anal Biochem 2008;379:176-81.

[20] De Spiegelaere W, Cornillie P, Van Poucke M, Peelman L, Burvenich C, Van den Broeck W. Quantitative mRNA expression analysis in kidney glomeruli using microdissection techniques. Histol Histopathol 2011;26:267-75.

[21] Ferguson BS, Nam H, Hopkins RG, Morrison RF. Impact of Reference Gene Selection for Target Gene Normalization on Experimental Outcome Using Real-Time qRT-PCR in Adipocytes. PLoS One 2010;5:e15208.

[22] Nicot N, Hausman JF, Hoffmann L, Evers D. Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. J Exp Bot 2005;56:2907-14.

[23] Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. Biotechnol Lett 2004;26:509-15.

[24] Andersen CL, Jensen JL, Orntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res 2004;64:5245-50.

[25] Fleige S, Pfaffl MW. RNA integrity and the effect on the real-time qRT-PCR performance. Mol Aspects Med 2006;27:126-39.

[26] De Spiegelaere W, Erkens T, De Craene J, Burvenich C, Peelman L, Van den Broeck W. Elimination of amplification artifacts in real-time reverse transcription PCR using laser capture microdissected samples. Anal Biochem 2008;382:72-4.

[27] Goossens K, Van Poucke M, Van Soom A, Vandesompele J, Van Zeveren A, Peelman LJ. Selection of reference genes for quantitative real-time PCR in bovine preimplantation embryos. BMC Dev Biol 2005;5:27.

[28] Cicinnati VR, Shen Q, Sotiropoulos GC, Radtke A, Gerken G, Beckebaum S. Validation of putative reference genes for gene expression studies in human hepatocellular carcinoma using real-time quantitative RT-PCR. BMC Cancer 2008;8:350.

[29] Erickson HS, Albert PS, Gillespie JW, Wallis BS, Rodriguez-Canales J, Linehan WM, et al. Assessment of normalization strategies for quantitative RT-PCR using microdissected tissue samples. Lab Invest 2007;87:951-62.

ACCEPTED MANUSCRIPT

[30] Derveaux S, Vandesompele J, Hellemans J. How to do successful gene expression analysis using real-time PCR. Methods 2010;50:227-30.

[31] Field KG, Olsen GJ, Lane DJ, Giovannoni SJ, Ghiselin MT, Raff EC, et al. Molecular phylogeny of the animal kingdom. Science 1988;239:748-53.

[32] Dormoy-Raclet V, Menard I, Clair E, Kurban G, Mazroui R, Di Marco S, et al. The RNA-binding protein HuR promotes cell migration and cell invasion by stabilizing the beta-actin mRNA in a U-rich-element-dependent manner. Mol Cell Biol 2007;27:5365-80.

[33] Piek E, Van Dinther M, Parks WT, Sallee JM, Bottinger EP, Roberts AB, et al. RLP, a novel Ras-like protein, is an immediate-early transforming growth factor-beta (TGF-beta) target gene that negatively regulates transcriptional activity induced by TGF-beta. Biochem J 2004;383:187-99.

[34] Nerlov C, Ziff EB. CCAAT/enhancer binding protein-alpha amino acid motifs with dual TBP and TFIIB binding ability co-operate to activate transcription in both yeast and mammalian cells. EMBO J 1995;14:4318-28.

[35] Aitken A. 14-3-3 proteins: a historic overview. Semin Cancer Biol 2006;16:162-72.

[36] Nebral K, Schmidt HH, Haas OA, Strehl S. NUP98 is fused to topoisomerase (DNA) IIbeta 180 kDa (TOP2B) in a patient with acute myeloid leukemia with a new t(3;11)(p24;p15). Clin Cancer Res 2005;11:6489-94.

http://ecvam.jrc.ec.europa.eu/. Zuang V, Barroso J, Bremer S, Casati S, Ceridono M, Coecke S, et al. ECVAM Technical Report on the Status of Alternative Methods for Cosmetics Testing (2008-2009). A report prepared in the framework of Directive 2003/15/EC (7th Amendment to the Cosmetics Directive)

Accepted Manuscript

ACCEPTED MANUSCRIPT

Brussels, 4th of March 2011

Dear reviewer,

The authors thank you for reviewing the manuscript entitled "Critical selection of reliable reference genes for gene expression study in the HepaRG cell line", by Ceelen L. *et al.*

We believe that it is an excellent proposal to check whether conventional inducers of phase I and II enzymes would influence the expression profile of the 12 candidate reference genes. However, the main message of this manuscript, as stated in the abstract and discussion, was to propose a set of approximately eight reference genes which should be validated in every new experimental design using the HepaRG cell line. We also wanted to investigate whether the commonly used *S18* gene was a valid reference gene, which turned out not to be. We believe that it is of utmost importance to analyze this set or reference genes when, as you correctly suggested, HepaRG cells are subjected to various experimental conditions. Following validation of this set of genes, the most stable ones can then be selected under these specific conditions. So, the take home message of the paper is that researchers should not use *S18*, or any other single reference gene for normalization in HepaRG gene profiling studies as such, but first to validate a set of genes and then pick out the most stable ones (a minimum of two reference genes is required as indicated by the geNorm applet (Vandesompele et al., 2002); MIQE guidlines).

In addition, in the current set, *TBP* was indeed the most stable reference gene, but the authors clearly declared not to use *TBP* as single reference gene (but as stated before to use a minimum of two genes). Depending on the experimental design (for example HepaRG cells treated with different inducers of Phase I and II enzymes), *TBP* may be again the most stable reference gene, but it should be validated against the other candidate reference genes in every new experimental setting, as you correctly observed.

The authors want to mention in addition, that at the current moment, experiments using HepaRG cells under different experimental conditions are running wherein the proposed set of candidate genes is being validated as well. These experiments have been set up in the framework of a large project and results will be published in the future.

The authors regret that the message of the paper was not entirely clear, but we hope by adapting the final conclusion in the abstract and by adding your proposal to the discussion, that the final point of the current paper is now clear.

Yours sincerely,

Liesbeth Ceelen (and co-authors)

[The authors prefer for the figures to appear in the printed version of the journal in grayscale.]

Accepted Manuscript

ACCEPTED MANUSCRIPT

Brussels, 4th of March 2011

Dear editor,

The authors thank you for taking the manuscript entitled "Critical selection of reliable reference genes for gene expression study in the HepaRG cell line", by Ceelen L. *et al.* into consideration for publication in "Biochemical Pharmacology".

We also want to thank the reviewer for his/her excellent proposal. To additionally underscore the take home message to the reader, we have changed the last sentence in the abstract (*i.e.* the main conclusion) and added a sentence to the discussion for further clarification. Moreover, due to impracticalities, it would not be possible to perform additional experiments as suggested. We also have added a letter to the reviewer in order to clarify the ultimate message of this paper.

The authors did a final review of the paper including a spelling and lay-out check (see below). We also want to confirm that all suppliers and their location are provided in the manuscript text for every chemical used in this study. Finally, we noticed that the melting curve analysis (for *YHWAZ*, *TOP2B* and *HMBS*) was not mentioned in the original text, although this analysis was performed. Therefore, to complete the RT-qPCR analysis, this information has been added to the revised text.

Textual changes include:

1. Page 3: These cells have been used for long-term repeated exposure to evaluate the potential toxic effects of different chemical entities [5, 7] and. -> "and" has been removed.
2. Page 5: 100 units/ml penicillin**,**100 µg/ml streptomycin (Gibco)-> 100 units/ml penicillin-100 µg/ml streptomycin (Gibco)
3. Page 6 and 8: Nanodrop has been replaced by Nanodrop®
4. Page 7: iQ5 Bio-rad system has been replaced by iQ™5 Bio-rad system

The authors prefer for the figures to appear in the printed version of the journal in grayscale.

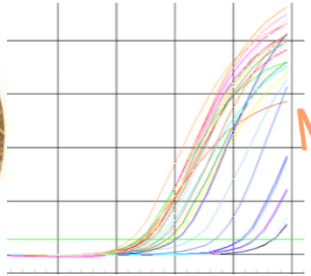
Hoping for a final affirmative answer, I remain,

Yours sincerely,

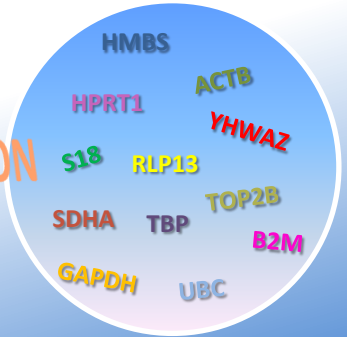
Liesbeth Ceelen (and co-authors)

Accepted Manuscript

*Graphical Abstract



NORMALIZATION



ACCEPTED MANUSCRIPT

Table 1. Twelve selected candidate genes, their encoded proteins and biological function.

Abbreviated gene name	Encoded protein	Biological function [reference]
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase	Oxidoreductase in glycolysis and gluconeogenesis [30]
HPRT1	Hypoxanthine phosphoribosyltransferase 1	Purine synthesis in salvage pathway [30]
18S rRNA, S18 or RPS18	Ribosomal protein S (Svedberg unit) 18	Constituent of 40S subunit of eukaryotic ribosomes. 18S rRNA is involved in the initiation of polypeptide synthesis in eukaryotes [31]
ACTB	β-actin	Maintenance of cell shape, growth and motility [32]
B2M	β2-microglobulin	Beta-chain of major histocompatibility complex class I molecules [30]
SDHA	succinate dehydrogenase, subunit A	Electron transporter in the TCA cycle and respiratory chain [30]
UBC	ubiquitin C	Playes a role in protein degradation [30]
RLP13	Ras-like Protein-13	Plays a role in TGF-β signal transduction [33]
TBP	TATA box-binding protein	Essential component of the RNA polymerase II basal transcriptional apparatus [34]
YHWAZ	14-3-3 protein zeta/delta	Protein involved in cell signaling, regulation of cell cycle progression, cytoskeletal structure, and transcription [35]
TOP2B	topoisomerase (DNA) II b	Enzyme controlling and altering the topologic states of DNA during transcription [36]
HMBS	Hydroxymethylbilane synthase	Heme synthesis, porphyrin metabolism [30]

Accepted Manuscript

ACCEPTED MANUSCRIPT

Table 2. Primers used for RT-qPCR and PCR cycling conditions.

Gene		Primer sequence (5'-3')	Amplicon length (bp)	Cycling conditions	Ta (°C)	Genbank accession nr or reference	Supplier
GAPDH	F	AATCCCATCACCATCTTCCAG	122	10’15’’30’’	60	NM_002046.3°	AB
	Probe	CCAGCATCGCCCCACTTGATTTT					
	R	AAATGAGCCCCAGCCTTC					
HPRT1	F	-*	100	3’10’’30’’	60	NM_000194.2°	AB
	Probe	-					
	R	-					
S18	F	-	187	3’10’’30’’	60	Hs99999901_s1/X03205.1°	AB
	Probe	-					
	R	-					
ACTB	F	ACCTTCTACAATGAGCTGCG	148	10’15’’45’’	60	NM_001101°°	IDT
	Probe	ATCTGGGTCATCTTCTCGCGGTTG					
	R	CCTGGATAGCAACGTACATGG					
B2M	F	GGCATTCCTGAAGCTGACAG	135	10’15’’45’’	60	NM_004048°°	IDT
	Probe	CTAAGGCCACGGAGCGAGACATC					
	R	TGGATGACGTGAGTAAACCTG					
SDHA	F	TGGTTGTCTTTGGTCGGG	85	10’15’’45’’	60	NM_004168°°	IDT
	Probe	ATGACTCTTCGATGCTCAGGGCAC					
	R	GCGTTTGGTTTAATTGGAGGG					
UBC	F	GCCTTAGAACCCCAGTATCAG	74	10’15’’45’’	60	NM_021009°°	IDT
	Probe	CCCAAGTCCCGTCCTAAAATGTCCT					
	R	AAGAAAACCAGTGCCCTAGAG					
RLP13	F	CAAACTCATCCTCTTCCCCAG	127	10’15’’45’’	60	NM_000977°°	IDT
	Probe	TTCAGCAGAACTGTCTCCCTTCTTGG					
	R	CTCCTTCTTATAGACGTTCCGG					
TBP	F	GAGAGTTCTGGGATTGTACCG	143	10’15’’45’’	60	NM_003194°°	IDT
	Probe	TGGGATTATATTCGGCGTTTCGGGC					
	R	ATCCTCATGATTACCGCAGC					
YHWAZ	F	ATGCAACCAACACATCCTATC	178	3’10’’30’’	60	[26]	[26]
	R	GCATTATTAGCGTGCTGTCTT					
TOP2B	F	AACTGGATGATGCTAATGATGCT	137	3’10’’30’’	60	[26]	[26]
	R	TGGAAAAACTCCGTATCTGTCTC					
HMBS	F	CTGTTTACCAAGGAGCTGGAAC	100	3’10’’30’’	59	[26]	[26]
	R	TGAAGCCAGGAGGAAGCA					

*, Primer sequence is not published by Applied Biosystems; BA, Applied Biosystems; IDT, Integrated DNA Technologies; F, forward; R, reverse; °, The Applied Biosystems custom probe design service was used to assist in the design of primers and patented minor-groove-binding, non-fluorescent quencher (MGB-NFQ) TaqMan probes; °°, The IDT Real Time PCR SciTool was used to assist in the design of primers and Prime Time 6-FAM/ZEN/IBFQ quenched probes.

Accepted Manuscript

ACCEPTED MANUSCRIPT

Table 3. Ranking of candidate reference genes in all 23 analyzed samples based on geNorm, Normfinder and BestKeeper analyses.

Gene	**geNorm**		**BestKeeper**			**Normfinder**	
	M value	Ranking	r	p	Ranking	Stability value	Ranking
GADPH	0.896	8	Nd	Nd	Nd	0,888	9
HPRT1	0.854	7	0.864	0.001	8	**0,300**	**2**
S18	1.381	11	0.694	0.001	9	0,915	10
ACTB	1.049	10	Nd	Nd	Nd	1,213	12
B2M	0,903	9	0.906	0.001	6	0,499	4
SDHA	**0,730**	**2**	**0.961**	**0.001**	**2**	0,534	5
UBC	**0,745**	**3**	**0.953**	**0.001**	**3**	**0,397**	**3**
RLP13	0,746	4	0.936	0.001	4	0,876	8
TBP	**0,712**	**1**	**0.986**	**0.001**	**1**	**0,159**	**1**
YHWAZ	0,756	5	0.900	0.001	7	0,551	6
TOP2B	1,459	12	Nd	Nd	Nd	1,134	11
HMBS	0,763	6	0.910	0.001	5	0,717	7

M value, expression stability measured in geNorm (should be < 1.5); Nd, Not determined (excluded for final BestKeeper analysis because of SD >1); r, correlation coefficient (BestKeeper) (should be close to 1); p, probability value (BestKeeper); Stability value (Normfinder) should be as low as possible.

Accepted Manuscript

ACCEPTED MANUSCRIPT

Table 4. Ranking of candidate reference genes in the group of HepaRG cells at differentiating and differentiated (DIFF) and proliferating (UNDIFF) stage based on Normfinder and BestKeeper analyses.

	DIFF				**UNDIFF**			
Gene	**BestKeeper**		**Normfinder**		**BestKeeper**		**Normfinder**	
	R	Ranking	Stability value	Ranking	r	Ranking	Stability value	Ranking
GADPH	Nd	Nd	0.404	5	Nd	Nd	1.656	11
HPRT1	0.907	8	0.414	6	0.983	3	0.644	5
S18	Nd	Nd	1.231	12	0.406	10	1.190	8
ACTB	Nd	Nd	**0.354**	**3**	Nd	Nd	1.873	12
B2M	0.909	7	0.710	11	0.995	2	0.623	2
SDHA	**0.965**	**2**	0.557	9	0.969	5	1.114	7
UBC	0.958	5	**0.268**	**1**	0.974	4	0.623	3
RLP13	**0.968**	**1**	0.429	7	0.956	8	1.409	10
TBP	**0.961**	**4**	**0.280**	**2**	0.995	1	0.090	1
YHWAZ	0.936	6	0.403	4	0.966	6	0.645	4
TOP2B	Nd	Nd	0.611	10	0.965	7	1.251	9
HMBS	**0.963**	**3**	0.490	8	0.927	9	0.858	6

Nd, Not determined (excluded for final BestKeeper analysis because of SD >1); r, correlation coefficient (BestKeeper) (should be close to 1); Stability value (Normfinder) should be as low as possible.

Accepted Manuscript

ACCEPTED MANUSCRIPT

FIGURE LEGENDS

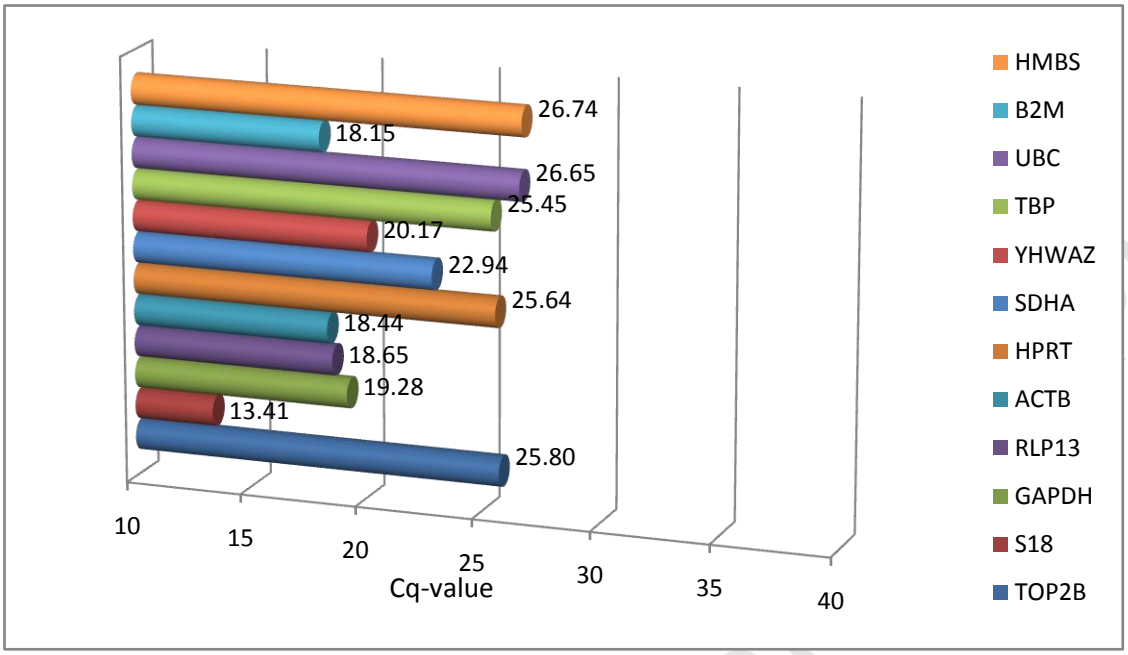
Figure 1. Cq-value range of the reference genes.

Figure 2. Reference gene mRNA expression stability according to geNorm (a) for all analyzed samples; (b) for DIFF samples; (c) for UNDIFF samples.

Accepted Manuscript

ACCEPTED MANUSCRIPT

Figure 1.



Accepted Manuscript

ACCEPTED MANUSCRIPT

Figure 2a.

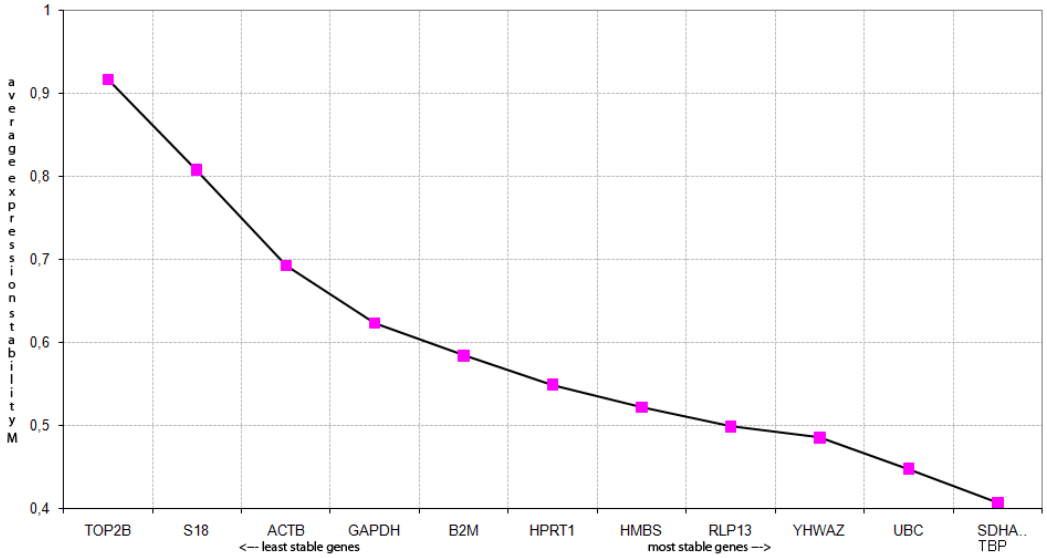
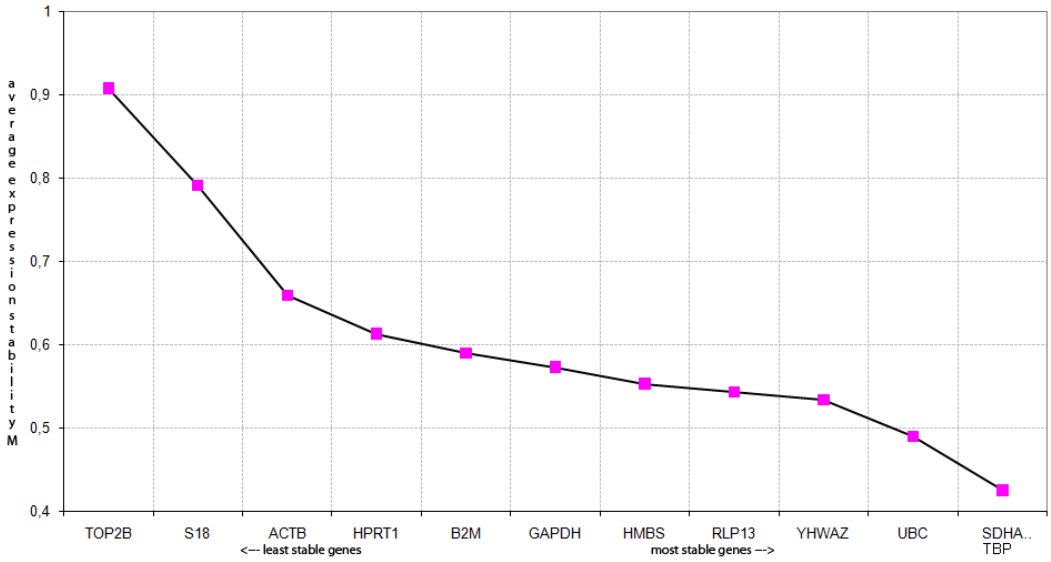
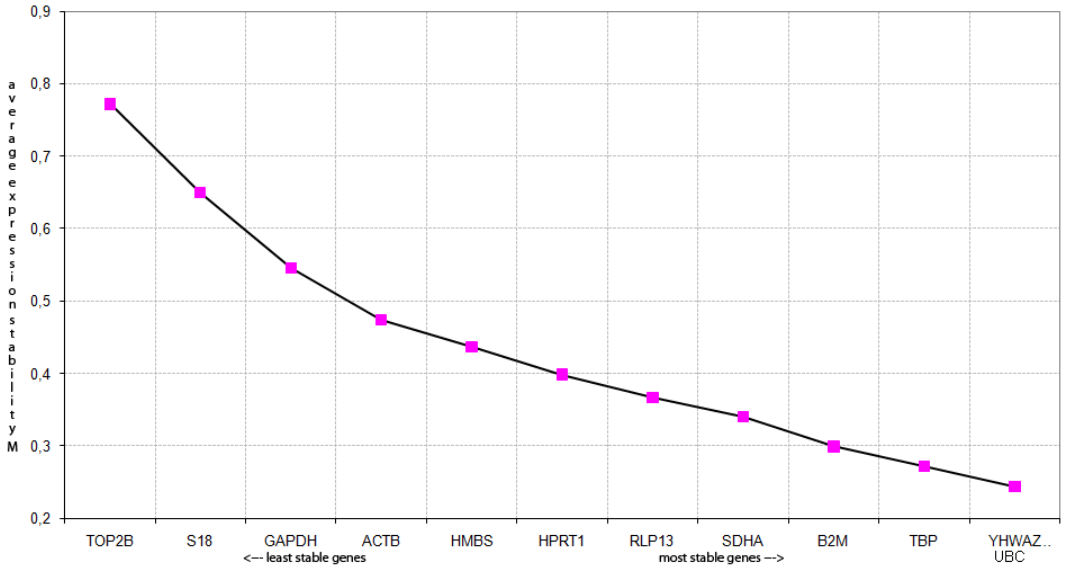


Figure 2b.



ACCEPTED MANUSCRIPT

Figure 2c.



Accepted Manuscript