GAPDH spike RNA as an alternative for housekeeping genes in relative gene expression assay using real-time PCR

Zeenah Weheed Atwan

Abstract

Background and aim of study: qPCR is a robust technique which quantifies the expressions of target genes in relation to reference genes. Stresses such as virus infection or heat shock change expressions of many cellular genes including the reference genes, so the aim was to introduce a constant calibrator to normalize the data to.

Methodology: Constructed glyceraldehyde 3-phosphate dehydrogenase (GAPDH) plasmid was transcribed to GAPDH RNA and used as spike RNA. Spiked RNA samples were subjected to qPCR at different conditions such as virus infection, IFN treatment, or mild heat shock.

The results: Adenovirus hexon in interferon-deficient cells showed different expression levels when data were normalized to GAPDH or 18S. Consistently, hexon expression levels were different in untreated cells under the control or heat-shocked conditions when data were normalized to GAPDH or 18S. Promyelocytic leukemia protein II (PML-II) expression level was lower in HeLa-PML-II-deficient cells (PML-II-Kd) compared to the control when the data were normalized to GAPDH as a reference gene and also in GAPDH RNA spiked, which showed reasonable consistency. More consistent data were obtained when the GAPDH normalizer was added before the step of treating the extracted RNA with DNase compared to add it after the treatment or directly to the qPCR reaction.

Conclusion: The internal controls that were chosen for this study completely changed the experimental results since they were affected with the experimental conditions. However, GAPDH spike RNA level was stable in its amplification at different kinds of stresses. So it can be an alternative for housekeeping gene due to its stability at these different conditions.

Keywords: Spike RNA, qPCR, Relative expression, Housekeeping genes, Universal normalizer

Introduction

More than two decades ago, a video camera was used to monitor the amplification of DNA fragment over a course of thermocycling (Higuchi et al., 1993). The quantitative real-time PCR was firstly used to quantify the levels of cystic fibrosis transmembrane conductance regulator (CFTR) mRNA. Relying on 5′ exonuclease activity of Taq polymerase and tetrachlor-6-carboxy-fluorescein as a reporter dye, an internal probe was designed to detect the CFTR in real time with a parallel control that was detected using specific primer with a different internal DNA probe (Gibson et al., 1996).

Similar to many quantitative assays, qPCR has many troubleshootings either in PCR steps or in the fluorescence problems (Bustin, 2000). To quantify mRNA expression, researchers developed a method using SYBR green or recRNA or recDNA (Pfaffl & Hageleit, 2001).

In order to normalize the expression of any target gene, it is critical to choose the most stable gene which expression unaffected by different conditions or at least the conditions in a certain experiment. For example, GAPDH, β-actin, and 18S RNA are the most common housekeeping or control genes, but in all cases, it is necessary to optimize the conditions to choose the right control before starting. Prior to do the qPCR experiment, it is very important to choose a reference gene that does not change under the experiment conditions.
to ensure that the alteration in gene expression is due to
the target gene not to the control gene (Livak &
Schmittgen, 2001).

qPCR is a very sensitive assay, and it needs sequential
steps starting from RNA or DNA extraction and reverse
transcription and ending with the qPCR (Skern et al.
2005). In addition, there is no reference gene that stays
stable over different types of stresses or cellular treatments
even the very common reference genes such as GAPDH
or β-actin. So these housekeeping genes were re-
introduced as unreliable control genes as their expression
is changing with treatments and alternatives were intro-
duced to the field (Radonić et al., 2004). Virus infection as
one kind of stress interferes with many cellular pathways,
so it changes the regular expression of many genes includ-
ing the housekeeping genes, such as the human cyto-
megalovirus (HCMV) infection. HCMV infection alters
transcription and stimulation of cellular transcription and
post-translational modifications (Kalejta & Shenk, 2002).
So many studies started to detect the most stable refer-
cence gene during infection with different types of viruses
in different cell lines (Radonić et al., 2005).

The reference or control could be an active control
which gives a detectable signal during the qPCR run,
and it could be internal or external; the internal is an
RNA that is extracted in the same sample and the signal
of the target gene is normalized to, while the external
means adding a synthetic spike RNA or DNA to the ex-
tacted RNA and the data are normalized to that foreign
nucleic acid (Livak, 1997).

Using external standard RNA is very important to
compensate the function of the internal control as a
constant calibrator (Jiang et al., 2011). Such artificial
molecules are either made as aliens where they are made
as pool sequences which do not belong to any known
genome and cloned in a vector or they could be refer-
ce known-sequence gene that is cloned and added to
the test samples artificially during the RNA extraction
step. Despite that this RNA will be ideal for
normalization as it will not be affected by treatment as
the internal control does, it did not come from the same
cells or tissues so it will need to be validated beforehand
(Cronin et al., 2004; Huggett et al., 2005). Based on the
need to build unbiased conclusions when analyzing gene
expression by real-time PCR, the plan was to analyze the
gene expression of some target genes under different
types of stresses such as mild heat shock and adenovirus
infection. So the aim was to analyze the expression of
hexon (late adenovirus gene) and promyelocytic
leukemia protein isoform II (PML-II) (interferon-regu-
lating gene) under normal or slightly higher temperature
normalized to selected housekeeping genes or to
GAPDH spike RNA.

Materials and methods
Construction of GAPDH DNA plasmid
Genomic GAPDH insert was constructed from a subcul-
ture of HeLa cells which was used as a DNA source.
DNA was extracted by column-based purification tech-
nique which was provided by Sigma-Aldrich GenEl-
luteTM Mammalian Genomic DNA Miniprep Kit
(G1N70). DNA samples were extracted by following the
manufacturer’s instructions, and the eluted DNA was
stored at −20 °C. A fragment of 97 bp was amplified by
GAPDH-specific primers using an Eppendorf Master
Cycler Gradient Thermo-cycler and Taq DNA polymer-
ase. The conventional PCR reaction mixture was pre-
pared as follows:

| Reagent                  | Concentration/amount |
|--------------------------|----------------------|
| 1 DNA template           | 50 ng/µl             |
| 2 Forward and reverse primers | 1 µl of 20 mM     |
| 3 dNTPs                  | 1 µl of 20 mM        |
| 4 MgCl2                  | 50 mM                |
| 5 H2O                    | Completed to 50 µl    |

The thermal cycle was set as follows: 94 °C for 3.5 min,
30 s at 60 °C, and 7 min at 70 °C. Using QIAquick PCR
purification kit, the amplified GAPDH was further
purified according to the manufacturer’s directions.
Ligation was achieved by mixing 2× pGEMT rapid
ligation buffer (PGEM R-T Easy Vector Promega) at a
ratio of 3:1 insert to vector in a total volume of 10 µl,
according to the manufacturer’s instructions. Fifty percent
of the mixture was used to transform 100 µl of E. coli
competent DH5α cells and plated on. White-blue selec-
tion was achieved by preparing XGAL/IPTG LB ampicill-
in plates by mixing 40 µl of 20 mg/ml XGAL with 40 µl
of 100 mM IPTG and spread on the surface of the sterile
LB-Amp agar plates. Positive colonies (white) were
grown in 5 ml of LB broth containing 100 µg/ml ampicill-
in for 16 h at 37 °C. The growth then used to purify
GAPDH plasmids and purified using QIAprep Spin
Miniprep Kit and finally sent to GATC Biotech to deter-
mine correct DNA sequence (Atwan, 2016).

GAPDH plasmid linearization
The reaction mixture was prepared by mixing 4 µl of
NEB 10× buffer, 2 µl of 500 ng/µl of GAPDH plasmid,
and 1 µl of ECORI restriction enzyme, and the reaction
was completed to 10 µl by H2O. The mixture then left at
37 °C in water bath for 1 h. The digestion was stopped
by adding 0.5 µl of 0.5 M EDTA, 1 µl of 3 M Na acetate,
and 20 µl of ethanol. The mixture then cooled at −20 °C
for 30 ′, and the linearized DNA then precipitated by
14K micro-centrifuge for 15 ′. Supernatant was removed,
and the tube was re-spun again for few seconds to remove the residual liquid phase; then, DNA was re-suspended in H₂O at a concentration of 1000 ng/μl.

**In vitro transcription**

All the kit reagents were briefly centrifuged before using, and all of them except the 10× reaction buffer were left on ice; then, the reaction mixture was prepared as in the table below:

| The reagent                  | Amount |
|------------------------------|--------|
| ATP solution                 | 2 µl   |
| CTP solution                 | 2 µl   |
| GTP solution                 | 2 µl   |
| UTP solution                 | 2 µl   |
| 10× reaction buffer          | 2 µl   |
| Linearized GAPDH plasmid     | 1 µg   |
| RNA polymerase enzyme mix    | 2 µl   |
| H₂O                          | 20 μl  |

The reaction was mixed thoroughly and briefly micro-centrifuged and was left at 37 for 8 h.

The reaction was stopped by adding 115 µl of H₂O and 15 µl of ammonium acetate stop solution.

The RNA was purified using column kits provided by Promega SV total RNA isolation system z3100, and according to the manufacturer’s instructions, the purified RNA was then eluted with nuclease-free water at a concentration of 600 ng/μl measured by NanoDrop.

**Stable transduced HeLa cells**

Permanent PML-II depleted cells, control empty vector EV, or GFP cells were produced by genetic modification using the lentiviral system to deliver the anti-PML-II, EV, or GFP constructs into HeLa cell genome to investigate the function of PML-II gene in adenovirus life cycle (Atwan, 2016).

**Heat shock protein 70 (HSP70)**

Hsp70 mRNA in HeLa cells was transiently interfered with HSP70 siRNA (Atwan, 2016) to reduce the gene expression of HASA1 and HSPB1.

**qPCR assay**

PML-II, EV, or GFP cells were plated on at a density of 1 × 10⁵ cell/ml; 24 h later, both types of cells were infected with adenovirus type 5 wild type with multiplicity of infection of 5. Twenty hours later, cells were washed with PBS, then were lysed and used as a source of adenovirus-infected RNA which was extracted by Promega SV total RNA isolation system z3100, and according to the manufacturer’s instructions, the purified RNA was then eluted with a nuclease-free water. The RNAs were DNase treated and reverse transcribed to cDNA by GoScript™ reverse transcription system A5000 and according to the manufacturer’s protocol. Then, qPCR reaction was prepared by mixing 0.1 µl of 20 µM hexon and IFN forward and reverse primers, 10 µl of SYBR green reaction mix, and 0.3 µl (Agilent). For hexon and IFN gene expression, the adenovirus cDNA were used at a concentration of 5 ng/μl as the DNA template; the qPCR experiments were achieved by Stratagene MX3005P light cycler (Agilent Technologies) under 95 °C for 3 min, then cycles of 1 min at 95 °C, 30 s at 55 °C, and 30 s at 95 °C thermal cycles. Cycle threshold and baseline values were determined using the machine software, and viral or cellular products were subjected to dissociation curve analysis to ensure primer specificity; all the primers which were used in this work are listed in Table 1.

**Table 1 qPCR primer sequences**

| Primer  | Forward sequence | Reverse sequence | Reference |
|---------|------------------|------------------|-----------|
| gGAPDH  | CCCACACACATGCACT | CCTAGTCCCCAGGCCTTT | Krog et al. 2007 |
| H       | TACC             | GATT             | Chen 2014  |
| PML-II  | AGGCAGAGGAAGCGGT | GGCTCCATGCAGGGATT | Chen 2014  |
| TGT     | CCCATTTTCCTTTGCTTC | TGCCTTTGTAGGCTTC | Chen 2014  |
| ISG56   | GCCATTTGTCTTCCTTC | TGCCCCCTTGAGGCTTC | Chen 2014  |
| GAPDH   | GGTCGGAGTCAACGGA | CCAGCATCGCCCCACTT | Vestergaard et al. 2010 |
| 18S     | ACCTGCTGACGGTCAACGGA | CCTAGGAGGGACACGTT | Menager et al. 2009 |
| HSP70   | AGGCCAACAAGACTAC | TCGGCCTCCGGTCTGAC | Tanaka et al. 2007 |
| Hexon   | CGCTGGACATGACTTGT | GAACGGTGTCGCCAGGTA | Schreiner et al. 2013 |
The data analyses were calculated as Cts after subtracting housekeeping gene Ct values from the gene of interest.

**Results**

**Observation**

One of the early observations during working with qPCR assays is the variation in the internal control amplification with the treatment. Genetically modified cells, and PML-II depleted and control cells GFP were used to analyze adenovirus gene expression. In a try to compare the expression of hexon, one of the late gene expressed proteins in adenovirus infected cell. Hexon expression showed dramatic variability in its mRNA level using different internal controls. In the first experiment, the hexon expression which is the target gene was expressed 3 Cts earlier in PML-II depleted cells compared to GFP control cells (Fig. 1) which is very expected as PML-II protein is regulating the interferon type I response (Chen et al., 2015). Despite this dramatic change in adenovirus hexon, when data were normalized to GAPDH as an internal control, the expression of hexon was roughly the same in both types of cell lines (Fig. 1).

The whole biological experiment was repeated under the same conditions except the housekeeping gene normalizer. The GAPDH was replaced this time with 18S. In a clear consistency, the hexon mRNA expression appeared 3 Cts earlier than its expression in the control cells (Fig. 2). In contrast to the first experiment, there was a dramatic increase in hexon expression in PML-II depleted cells compared to the control after normalizing data to 18S (Fig. 2).

**Normalization of qPCR data under physical stress in control cells**

Applying different kinds of treatment for the cell culture and quantifying the viral gene expression afterwards needed pre-optimization to consider results that welcome out of the qPCR assay. Both PML-II depleted and control cells were exposed to 40 for about 3 h then infected with adenovirus type 5 for 20 h before extracting RNA from the 2 types of cells and under treatment and normal conditions. Again, the data varied drastically depending on the chosen housekeeping gene. When hexon expression was normalized to GAPDH, it showed higher level at stress conditions compared to the control (Fig. 3), while it...
showed much lower expression at stress conditions compared to the control when the data were normalized to 18S as the reference gene (Fig. 4).

Normalization to 18 in Fig. 4 is more accurate than in Fig. 3 as the internal control in both treated and nontreated samples as its Cts values changed within narrow range between 0.1 and 0.5 which seemed very stable compared to the GAPDH which varied in about 14 Cts. Additionally, the 18S normalized results showed good consistency with Atwan’s (2016) conclusion that the mild stress decreases adenovirus late gene expression which was confirmed by protein analysis.

**Spiking RNA samples with synthetic spike RNA**

In order to compare the stability of GAPDH mRNA expression, the PML-II mRNA level was measured in both PML-II depleted cells compared to the empty vector control cells. As it is expected, the PML-II depleted cells will have lower level of PML-II mRNA in comparison with control cells. In fact, this was the case in both unspiked and spiked samples despite the latter showed more dramatic effect as a consequence of PML-II mRNA depletion (Fig. 5a, b).

**Stage of adding the spike RNA**

One of the important practical questions that need to be answered is at what stage one should contaminate the samples under investigation? To answer such question, it is important to try different stages of RNA extraction to rule out the right stage to work through. In this experiment, the spike RNA was added at the beginning of RNA extraction and before DNase treatment or after DNase treatment or added directly to the qPCR reaction well to investigate the level of interferon stimulated gene 56 (ISG56) in EV control cell line in the presence or absence of heat shock protein 70 (HSP70). Adding the spike RNA before DNase treatment showed difference in about 2 CTs in GAPDH message, and window of variance became narrower when spike RNA was added after DNase treatment and before reverse transcription step (Fig. 6). The GADH message changed between treated and untreated samples in about 0.8 Ct value. The last group of samples was contaminated with spike RNA.
Fig. 5 PML-II expression in PML-II depleted and EV control cells. Cells were plated at a density of $1 \times 10^5$/ml for 24 h. RNA samples were extracted either directly or spiked with 1 ng of GAPDH RNA or left as a regular extracted RBA. Total PML-II RNA was quantified by RT-qPCR using 10 ng/well of its (cDNA) as a template. Data were calculated by the $\Delta\Delta$Ct method, using GAPDH as internal control, and then normalized to the control cell value which was set to be 1. 

- **a** Normalized to GAPDH mRNA.
- **b** Normalized to spike GAPDH mRNA.

Fig. 6 ISG56 expression EV control cells. Cells were plated at a density of $1 \times 10^5$/ml for 24 h. HSP70 was depleted transiently from EV control cells (Atwan, 2016). RNA samples were extracted and spiked either directly and before DNase treatment or spiked or after DNase treatment before reverse transcription or spike RNA was added to the well of qPCR reaction after reverse transcribed separately. Total ISG56 RNA was quantified by RT-qPCR using 10 ng/well of its (cDNA) as a template. Data were calculated by the $\Delta\Delta$Ct method, using GAPDH as internal control, and then normalized to the control cell value which was set to be 1. 

- **a** Normalized to GAPDHmRNA.
- **b** Normalized to spike GAPDH mRNA.
when added to the well of qPCR reaction which showed no differences in the mRNA level which was considered as the control (Fig. 6).

Discussion
Clearly, it is not possible to find a reference gene that is stable under different conditions and treatments. The experiments showed instability in the expression of housekeeping genes that were selected in this study under different conditions. This is in consistent with publications since 16 reference genes were tested to determine the most stable genes in control kidney and sever or mild kidney tissue damage; only 4 of them were stable (Cui et al., 2009). Also, Among the 7 reference genes were tested in preeclampsia in diabetes, TATA box binding protein (TBP) and succinate dehydrogenase (SDHA) were the most stable ones (Meller et al., 2005). However, SDHA was registered a one of the most stable genes during early gonad development in mouse (Svingen et al., 2009). The same recommendation emerged from a study which concluded that at different canine articular tissues or its related diseases, it is critical to choose the most stable reference genes (Ayers et al., 2007).

An early observation was made when the data were changing and giving inconsistent results and the threshold values of GAPDH altered over multiple repeated experiments. In an agreement with what was mentioned by (Barber et al., 2005; Dheda et al., 2004), GAPDH as an internal control is drastically variable in 72 types of human tissues which showed that its level could be less in 15 times in breast cells than its level in skeletal muscle. In addition, it also varied from person to person despite its level was measured in the same tissue type (Barber et al., 2005). Furthermore, GAPDH variable expression gave false negative results re-IL4 expression levels in tuberculosis patients (Dheda et al., 2004). 18S is another internal control which was chosen to do the same relative assay instead of GAPDH, since its RNA subunits are the most stable housekeeping gene during different types of nutritional stresses in plants (Chen et al., 2016) or during long-term exposure to salts (Wang et al., 2014).

The study analyzed the data when they were normalized to either GAPDH or 18S to compare the conclusion in both cases. Unfortunately, adenovirus hexon expression drastically changed in the same cell line or in two different cell lines at both GAPDH or 18S normalizations.

GAPDH is preferable to be used as an internal control more than 18S in some kinds of tissues such as adipocyte or pre-adipocytes as it showed the most stable expression compared to a set of housekeeping genes which were selected to investigate the effect of treatment with hormones (Gorzelnia et al., 2001); however, none of them showed the same conclusion. Due to this inconsistency in analyzing data, the need emerged to find an alternative calibrator that is not affected with the treatment conditions in this study. Adding spike RNA or DNA to samples before extraction step to work as pseudo-housekeeping gene makes it possible to normalize the expression of target gene to the external RNA efficiently. The alternative proved that it could be a reliable calibrator especially after DNase treatment, and this agrees with using the foreign RNA in relative gene expression in blastocysts in preimplantation of embryos (Ivan Bower et al., 2007).

Conclusion
GAPDH spike RNA could be a good alternative to normalize the relative expression data when it is added during RNA extraction and after DNase treatment.

Abbreviations
IGF-I: Insulin-like growth factor-1; PML-II-Kd: HeLa-PML-II-deficient cells; BACT: Beta actin; CFTR: Cystic fibrosis transmembrane transduction regulator; EV: Empty vector cells; GAPDH: Glyceraldehyde 3-phosphate dehydrogenase; GFP: Green fluorescent protein; HCMV: Human cytomegalovirus; HSP70: Heat shock protein 70; IFN: Interferon; IL4: Interleukin-4; IPTG: Isopropyl β-D-thiogalactopyranoside; ISG56: Interferon stimulated gene 56; LB-Amp: Lauria broth ampicillin; PML-II: Promyelocytic leukemia protein II; SDHA: Succinate dehydrogenase; TBP: TATA box binding protein; XGAL: 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Acknowledgements
I would like to acknowledge professor Keith Leppard in helping me conduct the work in adenovirus laboratory at the School of Life Sciences, University of Warwick, Coventry, UK. Thanks to Dr. Kirsten Bentley, School of Biology, St Andrews, UK, and Dr. Andrew Woodman, Department of Biochemistry and Molecular Biology, Pennsylvania State University, USA, for providing GFP and EV plasmids.

Authors’ contributions
I have generated all the data by myself at adenovirus laboratory in the School of Life Sciences, University of Warwick, Coventry, UK. The author(s) read and approved the final manuscript.

Funding
The fund was awarded by higher committee for education development in Iraq (HCED Iraq).

Availability of data and materials
The data are available for any further communication,

Ethics approval and consent to participate
I confirm that the work did not include any human or animal subject.

Consent for publication
I agree to submit the paper to the bulletin of the national research center; it is an original paper and does not infringe on any copyright, and also currently, it is not under any consideration by another journal and has no conflict.

Competing interests
The author declares that they have no competing interests.
