Reviews

Practical approach to quantification of mRNA abundance using RT-qPCR, normalization of experimental data and MIQE

M. Yu. Obolenskaya, A. V. Kuklin, R.R. Rodrigez, O. P. Martsenyuk,
K. L. Korneyeva, V. A. Docenko, O. O. Draguschenko

Institute of Molecular Biology and Genetics, NAS of Ukraine
150, Akademika Zabolotnoho Str., Kyiv, Ukraine, 03680
m.obolenska@gmail.com

Reverse transcription and quantitative polymerase chain reaction (RT-qPCR) has become the most common method for characterizing expression patterns of individual mRNAs due to a large dynamic range of linear quantification, high speed, sensitivity, resolution and cost-effectiveness. However, the complexity of the protocol, variability of reagents, an inconsistent quality of biological samples, and the absence of standardized methods of data quantification may produce inconsistent results. In an effort to to standardize ithe procedure and assure high reliability of data, the minimum information for publication of quantitative real-time PCR experiments (MIQE) guidelines was defined and further extended by Prof. Bustin and colleagues (2004). These guidelines have been followed by the development of an XML-based real-time PCR data markup language (RDML) and a RDML data base for consistent reporting of RT-qPCR data. Here we follow the RT-qPCR procedure step by step in compliance with MIQE requirements, local facilities and resources and our own experience in application of RT-qPCR methodology.

K e y w o r d s: RT-qPCR, normalization and standardization of data, MIQE

"Garbage in, garbage out" a teaching mantra by George Fuechsel, an IBM 305 RAMAC technician/instructor, N.Y. 50s XX c. (Early programmers were required to test virtually each program step)

This paper is written to comment step by step the widely accepted conventional experimental protocol of RT-qPCR procedure and to present various currently available methodological solutions at each step of procedure in accordance with mandatory minimum information for publication of quantitative real-time PCR experiments (MIQE) guidelines. The following of these guidelines provide the standardization of procedure, the ability to validate experiments, to consider the conclusion as reliable and to share and compare the results with other colleagues.

Since 1988 when the RT-qPCR was applied for the first time for assaying the transcriptional phenotype of macrophages during normal healing [1] it has become the most common method for characterizing expression patterns of individual mRNAs in biological samples. It has nearly supplanted other similar approaches (e.g., Northern blotting, RNase protection assays) due to its large dynamic range of linear quantification, high speed, sensitivity, resolution and cost-effectiveness. The requirements for RT-qPCR procedure and presentation are formulated in MIQE that is a practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments [2, 3]. In fact these commonsense guidelines are in line with the last year campaign organized by the US National Institutes of Health, Nature and Science against a very low repro-

ducibility for articles published in scientific journals, often as low as 10 – 30 % [4, 5; www.nature.com/news/1.16259]. The reliability of molecular technology and gene expression profiling becomes critically important when it is applicable to diagnostics and predictions for the disease outcome, as well as the choice of the therapy [6].

Here we are not presenting a comprehensive review on RT-qPCR methodology and its widespread application in different fields of contemporary biology as many papers on the subject and detailed guides from the leaders in molecular biological industry may be found in the literature and internet [e.g. 2, 3, 7–15]. Instead we are going to follow the procedure step by step in compliance with MIQE requirements, local facilities and resources along with the currently available at the market using our own experience in application of RT-qPCR methodology. To make the expression results valuable and comparable between runs, real-time RT-qPCR platforms, and between different laboratories worldwide the researchers have to meticulously standardize each step of procedure, starting from the biological sample to the final data analysis.

1. Experimental design

The illustration of different stages of RT-qPCR procedure is based on the experiments conducted in the laboratory of systems biology (Institute of molecular biology and genetics, National Academy of sciences, Kyiv, Ukraine) on the gene expression in rat liver, primary hepatocytes, mouse brain and human placenta at different (patho)physiological states at the steady state and in the course of time [16–20]. Male Wistar rats (200–250 g) at different time points after partial hepatectomy and laparotomy, male BALB/c mice (~ 20 g) and human placenta from midfirst- and third trimesters of gestation were used in the studies. The animals' tissue samples were snap-frozen in liquid nitrogen immediately after collection of organs and tissue dissection; the samples of placental villous tissue were collected immediately after abortion or delivery, rinsed with cold sterile 0.9% NaCl solution and frozen in liquid nitrogen. The samples were stored at –80 °C before processing.

2. RNA isolation

A prerequisite for the performance of RT-PCR is an efficient method for RNA extraction. Currently there are numerous methods that can be used to isolate and purify RNA for RT-qPCR [21]. Here we will focus on the cheapest and more accessible method of RNA extraction based on "TRIzol reagent". The total RNA was isolated from approximately 100 mg of liver tissue, mouse brain, human term placenta and villous tissue from the midfirst trimester of gestation. The samples were ground in liquid nitrogen with mortar and pestle; one ml of homemade "TRIzol reagent" was added and the samples were homogenized using power homogenizer (IKA T10, Cole Parmer LabGen, USA) (for a receipt of homemade "TRIzol reagent" see the Supplement). The liver lysate was additionally passed through 22 G needle five times to shear chromatin. All sequential procedures followed the protocols described in the manual from Thermo Fisher Scientific (USA) for TRIzol® Reagent [http://tools.lifetechnologies.com/content/sfs/manuals/trizol_reagent.pdf]. Briefly, the homogenized samples were processed through separation of aqueous, intermediate and phenol-chloroform phases, containing RNA, DNA and DNA and proteins; correspondingly; RNA precipitation with isopropanol; two washes with 70 % ethanol and dissolving RNA in DEPC-treated water.

The yield of RNA, its purity and integrity are assessed by the conventional methods, particularly in our lab, by spectrophotometry on a NanoDrop ND-2000 device (Life Technologies, USA) and denaturing formaldehyde 1 % agarose electrophoresis with image analysis using Gel-Pro 3.1 software (Media Cybernetics, Inc., USA). The A260/A280 ratio around 2.0 and the A260/A230 ratio around 2.2 evidence the sufficient purity from proteins, phenol, and aromatic compounds or carbohydrates, respectively [22, 23]. The ratio of 28S/18S bands intensities near 2.0 at electrophoregram the integrity of RNA. At this step the RNA may be stored either at – 80 °C or may be precipitated with 3 vol. of 96 % ethanol – 0.1 vol. 3 M NaAc and stored at – 20 °C before the further procedure.

We have to note here that the main disadvantage of general UV spectrophotometry for RNA quantification is low sensitivity (≥ 4 µg/ml) and the contribution of signals from DNA, degraded RNA and salts. The use of an ultrasensitive fluorescent RiboGreen® (RNA detection range: 5 ng/ml – 50 ng/ml) alleviates this problem [24]. The RiboGreen® reagent is nonfluorescent when free in solution; upon binding to RNA, the fluorescence of the RiboGreen® reagent increases more than 1000-fold. The RNA-bound RiboGreen® reagent has an excitation maximum of approximately 500 nm and an emission maximum of approximately 525 nm. The results may be read either with Fluorescence Microplate Reader or with Thermo Scientific NanoDropTM 3300 Fluorospectrometer. RiboGreen® is produced by Invtirogen, sold as a RiboGreen®RNA Reagent (#R11491) and within a kit (#R11490).

The routine assessment of RNA integrity with denaturing formaldehyde agarose electrophoresis is low-sensitive, time-consuming, laborious and hazardous technique. This method using EtBr as a fluorescent intercalating agent requires approximately 200 ng of RNA to make an accurate assessment of its integrity. The amount needed can be reduced by using alternative fluorescent dyes such as SYBR® and SYBR® Green II RNA gel stain (Invitrogen Corporation). However, when RNA quantity is very limited (e.g. RNA from biopsies), agarose gel analysis may not be possible. A major improvement in RNA analysis occurred with the introduction of a system combining microfluidics, capillary electrophoresis, and fluorescence to evaluate both RNA concentration and integrity [25]. There are the Experion automated elecltrophoresis system (Bio-Rad Laboratories, Inc.) and the Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.) at the market. This technology requires very small inputs due to Chip format, allowing user to assay RNA quality in limiting samples. Bio-Rad introduced the RNA quality index (RQI) in Experion software and the Agilent Technologies provide the freely available RNA integrity number (RIN) algorithm software that assess the RNA integrity by the numbers from 1 (highly degraded) to 10 (intatct RNA) [26, 27]. Such method favors the standardization of the whole procedure. The example of electrophoregrams obtained with routine and advanced method is represented in the Fig. S1.

3. Elimination of genomic DNA from RNA samples

The isolated RNA usually contains different amount of genomic DNA, sometimes up to 50 %, that could result in false measurement of RNA concentration by UV spectrophotometry and induce the variability of results. The kits and reagents for elimination of genomic DNA by RNAse free DNAse I treatment are provided by different vendors that propose DNAses active at somewhat different concentration of Tris, $MgCl_2$, $CaCl_2$ or $MnCl_2$ (compare kits from Life Technologies, USA, Cat.# AM1906, Cat.# EN0525 and Cat.# EN0521 and one from Promega, USA, Cat.# M6101) and possessing different enzymatic activities. The vendors' recommendations are usually given in the units of enzyme per reaction, from roughly one unit for 1 µg of RNA (enzyme #EN0525) to two units per 10 µg of RNA (enzyme#AM1907).

Though the time and the temperature during enzymatic treatment are uniform throughout, the ways for the effective DNAse I removal are various. This step is especially critical when the RNA will be used directly to synthesize cDNA according to the recommendations of several producers [http://www.biomart.cn/upload/asset/attachment/fermentas%E4%B8%AD%E5%9B%BD/coa_en0521.pdf]. DNAse I may be removed by routine phenol/chloroform extraction; irreversible enzyme inactivation for 10 min at 65 °C in the presence of 2.5 mM EDTA [28]; conventional precipitation with isopropanol [29] or ethanol and column clean up systems (e.g. GeneJET RNA Cleanup and Concentration Micro Kit, #K0841 from Thermo Fisher Scientific). There is also a novel DNAse Removal Reagent at the market (Ambion's DNA-freeTM, #1906, It sequesters and precipitates DNAse I and cations in minutes at room temperature

and thus RNA cannot be degraded while inactivation of DNAse I by heating in the presence of divalent cations can partially degrade RNA.

The disadvantage of the whole step of DNA elimination is a risk to overestimate the initial amount of RNA for PCR and to inhibit reverse transcriptase if the reaction mixes after DNAse I inactivation are used directly to cDNA synthesis. In our lab we control the concentration and integrity of the RNA after DNAse I treatment and selectively check the efficiency of DNA elimination by the ratio of 18S rDNA products after minus RT – qPCR and 18S rRNA products after RT – qPCR. While before DNAse treatment up to ≥50 % of isolated putative RNA is represented by DNA then after DNAse I treatment the traces of DNA usually comprise < 0.5 % (for details see Supplement).

4. The synthesis of Luc RNA as an external standard

Now the preprocessing of RNA is over and the RNA is ready for the main step of RT-qPCR analysis. To maximally diminish the preprocessing related variability, which affects cDNA synthesis, the exogenous RNA (spike RNA) as an external standard can be used. Here we present our procedure of the external standard synthesis [20]. The fragment of firefly luciferase gene (Luc), irrelevant to the mammalian genomes, was cut out from the pGL3-Basic plasmid (Promega, Wisconsin, USA) by XbaI and HindIII restriction enzymes and cloned into the same restriction sites of the pGEM-3Z vector(Promega, Wisconsin, USA) according to the routine procedures. The pGEM®-3Z Vector is intended for use as a standard cloning vector, as well as for the highly efficient synthesis of RNA in vitro. The obtained product was referred to as pGEM-3ZLuc and was subjected to in vitro transcription. The final mix of 100 μl contained 1 μg of pGEM-3ZLuc DNA linearized by EcoRI, 2 mM NTP, 10 μl 10x IVT Buffer, 50 U of RiboLockTMRNAse inhibitor, 30 U SP6 polymerase (Thermo Fisher Scientific, USA). The reaction lasted for 2 h at 37 °C and was followed by subsequent phenol/chlorophorm extraction of RNA.

The Luc RNA was dissolved in 40 μl of DEPC treated water, its concentration was detected with a NanoDrop ND-2000 device and its integrity and specificity was confirmed by non-denaturing electrophoresis in 1% agarose gel. The main steps of external RNA synthesis are illustrated in Fig. 1. The obtained probe was aliquoted and stored at – 80 °C. The Luc RNA has very high PCR efficiency in the absence of inhibitors (E > 95 %) and does not produce the primer-dimer products (no PD signal in 40 cycles). The amount of the Luc RNA template for RT-qPCR was adapted for genes of interest in a way that both Cq would be in the same range.

There are several ready spikes available at the market, such as Universal RNA Spike from TATAA Biocenter (http://www.tataa.com/products-page/quality-assessment/tataa-universal-rna-spike/) or SPUD from Sigma-Aldrich, which are very effective tools for the quality control throughout entire RT-qPCR experimental workflow [30]. Both have a synthetic sequence that is not present in any known living organism, and mimic eukaryotic mRNA. In the presence of a "clean" sample, the quantification or threshold cycles (Cq, aka Ct) of spike will remain the same as their control whereas in the presence of a contaminated sample, the Cq will shift to higher cycles.

There is another approach to check whether a sample contains inhibitors. For10X difference in concentration of starting template, the difference between Cqs should be approximately equal to PCR slope. The greatly reduced difference in Cqs after 10 to 1 dilutions compared to the expected difference indicates the presence of inhibitors. For more detailed information how to check the presence of inhibitors and to eliminate them see [Pennington R. Dealing with Amplification Inhibitors: Reagent Choice Matters. 2014 (https://www.promega.com)].

5. Design of primers and target sequence

The quantification of target cDNA during real-time PCR is accomplished by the DNA binding fluorescent dye SYBR Green I or TaqMan hydrolysis probes. We describe here the first approach as it is easier to design, faster to set up, more cost-effective

and applicable for estimation of different genes expression from the same cDNA sample if it is obtained with the random primers but not with the target specific ones.

Both primers and target sequence may affect the efficiency and specificity of amplification. A number of free and commercial programs are available for the primers design and control of their specifity: PrimerQuest software (Integrated DNA Technologies, http://eu.idtdna.com/scitools/applications/primerquest/ default.aspx), Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/ primer3plus.cgi) or Primer-Blast (http://www.ncbi.nlm.nih.gov/tools/primer-blast/), OligoAnalyzer 3.1, Primer-BLAST, Vector NTIAdvanceTM10.0) [31]. In silico PCR on UCSC Genome Browser is very useful to check for non-specific binding. There is also an integrative publicly available database for the storage, retrieval and analysis of primer and probe information - RTPrimerDB (http://www.rtprimerdb.org) [15]. It contains primers and probe sequences used in real-

time PCR assays employing popular chemistries (SYBR Green I, Taqman, Hybridisation Probes, Molecular Beacon). It was organized to avoid time-consuming primer design and experimental optimization, and to introduce a certain level of uniformity and standardisation among different laboratories. The database is linked with the reference databases to allow the submission of assays for all genes and organisms officially registered in Entrez Gene and RefSeq. The submission of own tested primers to RTprimerDB is strongly encouraged.

To avoid amplification of genomic DNA the intron spanning primers are designed with at least one incorporated exon-exon junction (Fig. 2). The primers have not to be placed on polymorphic sites and splice forms of RNA. The secondary structure of the RNA fragment that is reverse transcribed may have also a substantial impact on the efficiency of RT-qPCR. Therefore, RNA sequences and primers should be checked with nucleic acid–folding software, MFOLD program [32]. According to MIQE
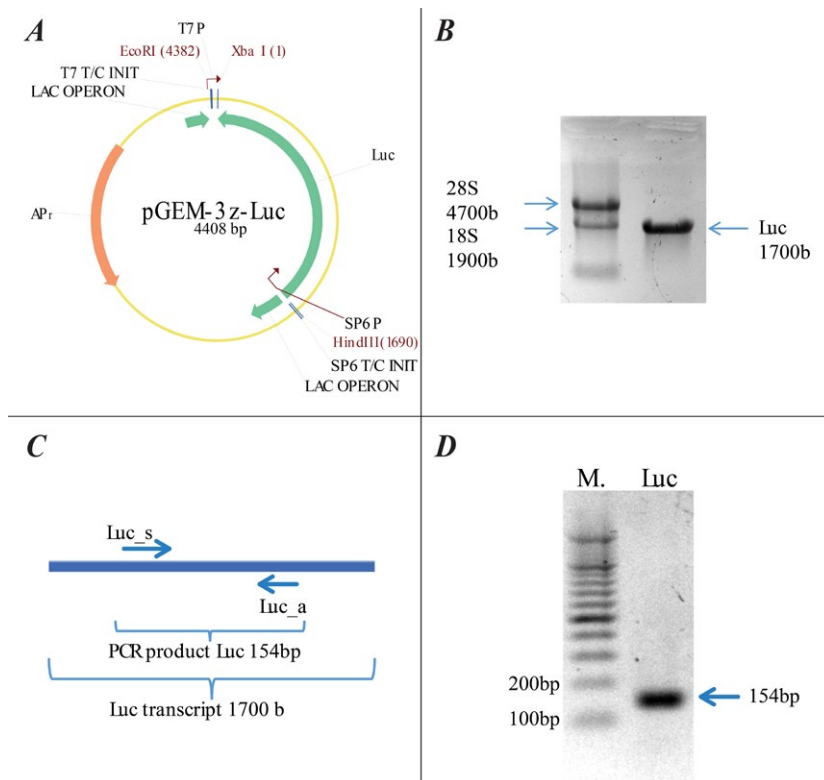

Fig. 1. Synthesis of Luc RNA as an external standard. *A* – Plasmid pGEM-2z LUC containing the fragment of firefly luciferase gene Luc; XbaI and HindIII are the cloning sites for Luc insert; EcoRI – a site for linearization. *B* – Transcript of the Luc fragment (1700b) loaded next to the total RNA from rat liver. *C* – Scheme of the Luc transcript and location of primers. *D* – Electrophoregramof Luc PCR product: M – 100bp DNA Ladder (Thermo Scientific.
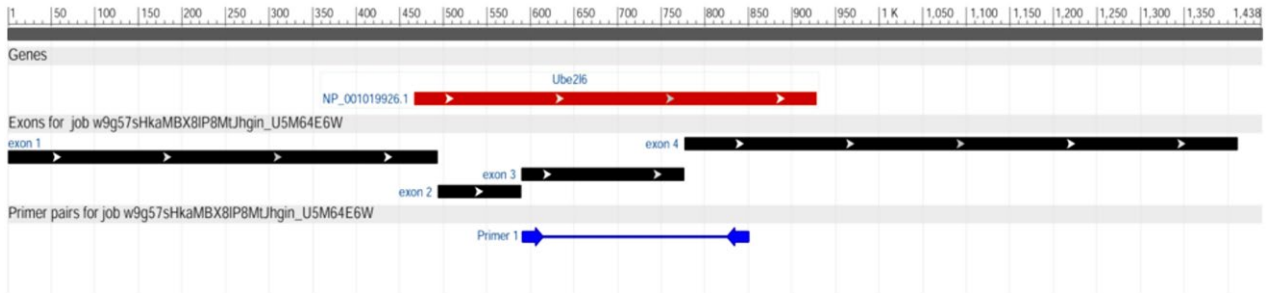
Fig. 2. Location of primers for PCR amplification of Ube 2L6 cDNA.
Note: the amplicon includes the junction between third and fourth exons

the folding structures should be ideally made available to reviewers as well.

We design the primers using Vector NTIAdvance TM10.0 program, check their specificity with Primer-BLASTprogram (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and secondary structure with MFOLD. The target sites for restriction endonucleases within amplicon are found withVector NTI Advance® 11.5.4 software (ThermoFisher Scientific, USA).

All new primers have to be tested by qPCR using a 3-fold serial dilution of cDNA in order to determine the amplification efficiency of the cDNA fragment, to validate the specificity of amplicons and the presence/absence of primer dimers after qPCR [32]. The amplification efficiency for each pair of primers is determined from the linear slope of standard curve; only primers with a standard curve slope between –2.92 and –3.92 (efficiency 100 ± 20 %) are used for further quantification. Efficiency 100% is defined by a slope of − 3.32 and a 2-fold change for each change in Cq of the 10 fold dilution (Fig. S2) [33]. If the slope is below –3.92, then the PCR has poor efficiency. The efficiency of amplification may be optimized by changing the concentration of $MgCl2$ and dNTP, adjusting the temperature and time of annealing. If these procedures do not improve the efficiency of amplification the primers have to be redesigned.

The amplicon specificity is validated by the length of the amplicon and its restriction fragments (For example see Table 1 and Fig. 3). Due to the critical importance of the above mentioned characterristics of primers, the information regarding their sequence, location on cDNA, specific restriction endonucleas-

es for an amplicon, the length of restriction fragments and their experimental approval must be provided in the papers according to the MIQE guidelines. The sequence of amplicon is appreciated by many reviewers [2].

6. Reverse transcription and Real time PCR

The reverse transcription and real time PCR are conducted according to the conventional protocols for specific enzymes. In our lab the first-strand cDNA synthesis from total RNA is carried out with the spike RNA Luc and RiboLock RNase Inhibitor. For everything else the procedure follows the protocol of ReverseTranscriptase producer (in our case RevertAid ReverseTranscriptase, cat. #EP0441, Thermo Fisher Scientific, USA). After termination of the reaction the mixture is diluted with sterile DEPC-treated water to the concentration of 25–50 ng/µl, aliquoted and stored at − 80 °C.
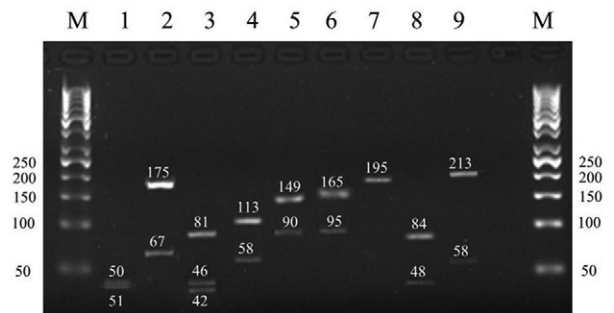

Fig. 3. Specificity of amplicons proved by restriction analysis. Lanes: 1 – Tbp, 2 – 18S, 3 – Ifnα, 4 – Isg15, 5 – Ube1l, 6 – Ube2l6, 7 – Trim25, 8 – Usp18, 9 – Irf7, M-50bp DNA Ladder (Thermo Scientific, USA) Compare experimental results with theoretically predicted (Table 1).

Table 1. **Characteristics of primers and amplicons [Kuklin et al., 2015]**

mRNA (Refseq)	Primers	Amplicon location, bp	Restriction analysis of amplicon	
			Enzyme	Restricts, bp**
Tbp (NM_001004198.1)	F 5'- TCAGTCCAATGATGCCTTACG - 3' R 5'- CTGCTGCTGCTGTCTTTGTT - 3'	348–448	Hpy 1881	50; 51
18 S (NR_046237.1)	F 5'- GTTCCGACCATAAACGATGC-3' R5'- CGCTCCACCAACTAAGAACG -3'	1078–1341	*Hinf*I	175; 67; 44
*Ifna**	F 5' - CTGCTGTCTAGGATGTGACCTGC -3' R 5' - TTGAGCCTTCTGGATCTGCTG - 3'	57–225	HinfI	81; 46; 42
Isg15 (NM_001106700.1)	F 5' - CCTCTGAGCATCCTGGTGAG- 3' R5' - CAGTGGCTCTTT GTCCTCCA - 3'	376–546	PvuII	58; 113
Ube1l (NM_001106856)	F 5' - GGGCCTGGGAGTTAGGGATAATGG- 3'' R 5' - CGTCCACCCTGGAGAAGAAGTCGT - 3'	1492–1730	Ear I	90; 149
Ube2L6 (NM_001024755.1)	F 5' -ACCAACTTCCCTATCGCCTCAAGG- 3' R 5' - GAGGTCAGCTAGTTCCAAACGCACA- 3'	591–850	Bgl II	95; 165
Trim25 (NM_001009536.1)	F 5' - CGCAAATGTTCCAGGCACAACC- 3' R 5' -CATCCTCCAGTGCTTTGCTCGCT - 3'	521–725	Rsal	10; 195
Usp18 (NM_001014058.1)	F 5' - ATACAACGTGC CATTGTTTGTCC- 3' R 5' - TCGGTCCAGATTGT GAACAGATC- 3'	496–627	EarI	48; 84
Irf7 (NM_001033691.1)	F 5'- GCTGAGCGAAGAGGCTGGAAGA - 3' R 5'- CCAGAAAGCAGAGGGCTTGG - 3'	650–918	HaeII	58; 113

Note: For Ifna indicated with an asterisk (*), the primers are common for Ifna1, NM_001014786.1; Ifna2, NM_001271218.1;Ifna4, NM_001106667.1 and Ifna16l, XM_575856.1. Double asterisks (**) denote the length ofr theoretically predicted length of restricts.

The amplification of individual mRNA is performed in triplicate with real-time PCR at different types of Real-Time PCR Detection Systems. We use Bio-Rad CFX96 Real-Time PCR Detection System (Bio-Rad Laboratories Ltd., USA). The mandatory requirement includes the presence on each plate the samples of interest and 'No Template' control in each RT- q PCR plate. Also, there has to be a set of corresponding diluted standards (amplicons in our case) for the standard curve if the absolute quantification is used.

The specificity of amplification is confirmed with the same device by the melt-curve analysis (Fig. S3). For the methodological details see Supplement.

7. Absolute *vs* relative quantification

There are two strategies to quantify the level of gene expression – by absolute or relative RT-qPCR. The first one relates on a calibration curve and the obtained results are represented in amount of mRNA in ng, nanomoles or copies per one unit of total RNA,

per cell or per g of tissue while the second reflects the relative changes in the mRNA expression level between the samples and the results are represented by relative units.

The standard curve method determines the input mRNA levels of a gene by relating the PCR signal to a standard curve. It requires the production of a high quality standard and a high quality RNA. A purified specific RT-PCR product, recombinant DNA or recombinant RNA, may serve as standards. The production of the specific RT-PCR product or amplicon by ordinary RT-PCR reaction with subsequent purification is rather simple but the standard may be not very stable therefore its integrity has to be checked from time to time. The production of recDNA is much more time consuming and cost effective, however, it is more stable than a short RT-PCR product. The recRNA best of all mimics the RT-PCR procedure of natural RNA but the standard is very unstable and its preparation is rather complicate and ex-

pensive. The standard curve method has several advantages: representation of dilutions for the standard curve at each PCR plate provides the routine validation for methodology; it substantially simplifies all calculations and comparison of the expression levels between different mRNAs in the same sample or between different samples [12]; it does not require a laborious work for validation of reference genes; the results have the absolute values and thus are easily comparable between different laboratories worldwide. The disadvantage of the method is a requirement for production of high quality standards and more stringent requirements for maximal standardization of the whole procedure (especially the high standards for RNA quality taken for RT – qPCR). An example of using the standard curve for assessment of mRNA abundance is represented in the Fig. S2.

Relative quantification relates the PCR signal of the target transcript in one sample to that of another sample chosen as a control or calibrator. The choice of calibrator depends on the gene expression experiments. It may be an untreated sample in the experiment with treatment; a control sample in the study of some disease; a sample at time zero in the time course study of gene expression; a sample for one tissue that is compared with other tissue.

The relative quantification is mostly applied to steady–state conditions and an endogenous presumably invariant reference transcript is used to diminish the variability between the samples. Occurring background interferences retrieved from the extracted tissue components and the cDNA synthesis efficiency relate to target and internal reference similarly so all the variations preceding PCR are compensated by the calculations after amplification. The relative quantification method is dependent on the determination of real time amplification efficiency of the target and reference templates. To date, two types of relative quantification models are available. In the first one the correction of amplification efficiencies is not required if they are near 100 % and within 5% of both target and reference template. But it is required in the second model where the amplification efficiencies of target and reference templates differ for more than 5 % of

each other values [12, 13, 34]. The advantages of the relative quantification method consist of no need to produce the standard material, to optimize and to validate the standard curve. The disadvantages of the method relate to the necessity to validate the internal standards, low standardization of the method and the difficulties to compare the results obtained by different research groups as the references are usually chosen according to the own preferences of each researcher. Frequently, to find the invariable referent gene is problematic.

8. Data normalization

Several strategies have been proposed for normalizing RT-qPCR data. They range from the calculation of individual mRNA content per unit of sample, per unit of investigated RNA, either total or its fraction, to using an internal housekeeping gene as a reference. These approaches are not mutually exclusive and can be incorporated into a protocol at many stages. The relative merits of different normalization strategies are explicitly represented in multiple reviews [35–39].

The choice of the normalization strategy strictly depends on the goal of experiment and the essence of biological sample(s) under the study whether they are in the steady or dynamical state. Calculations of mRNA amount per total RNA and per internal reference gene(s) are applicable to steady state situation. In dynamical systems the qualitative and quantitative composition of total RNA changes in the course of time. So it is difficult to find out the internal reference gene with constant concentration particularly when the rRNA : mRNA ratio varies.

In our lab we explore the gene expression in two dynamical systems – regenerating rat liver and human placenta in the midfirst and third trimesters of gestation. We have shown that the concentration of 18S rRNA and TBP mRNA encoding TATA-binding protein, both widely used as internal reference templates, gradually increase during the first 12 h after partial hepatectomy and gradually decrease during the same time after sham operation [20]. Another group of researchers propose different pairs of reference templates for eight types of hepatic cells in re-

generating liver [40] Given that isolated RNA reflects the composition of liver RNA in vivo we present the target mRNA abundance in copies per mass unit of total liver RNA [16, 20]. Somewhat similar situation is observed with placental samples at different time points of gestation. The abundance of 18S rRNA and YWHAZ template, recommended as placental internal standards [38], is invariable either in the first or in the third trimester of gestation, however, the values are nearly two times less in the third trimester in comparison to the first one making these templates inconsistent for inter-trimesters comparison. So we also use the absolute values in copies of mRNA per mass of total RNA [17, 18].

To mitigate the variability during preprocessing of RNA we additionally normalize the data according to Luc RNA recovery. We added an external Luc RNA spike together with total RNA to the reaction of reverse transcription. The 100 % value of Luc recovery corresponds to the copies of amplified Luc RNA in the absence of cDNA of interest. The recovery of less than 85 % of Luc RNA, when it was processed together with the sample RNA points to the presence of inhibitors in the RNA sample. Such samples are either additionally purified or RNA isolation is repeated (Table S1). Ideally, it is possible to include the additional artificial RNA at the very beginning of procedure to follow all the steps of manipulation with the target mRNAs.

During the last decade the captious exploration of the reference genes reliability has completely changed the opinion about their applicability. The most commonly used housekeeping genes in époque of Northern blot and semiquantitative examination of mRNA level like glyceraldehyde-3-phosphate dehydrogenase (GADPH), β-actin and 18S ribosomal RNA became inadequate in most cases of RT-qPCR due to their variability in different tissues, presence of pseudogenes that may misleads the results, inconsistency of their expression under different treatments or in comparison of control with disease *etc.* [37–39, 42, 43]. Moreover it is considered that the choice of housekeeping genes is highly specific for particular experiment [37].

To facilitate and standardize the selection of reference genes, a lot of methods have been developed. The geNorm VBA applet for Microsoft Excel calculates a gene expression normalization factor for each tissue sample based on the geometric mean of a user-defined number of reference genes [44]. There are already geNorm Kits at the market (https://genorm.cmgg.be/). BestKeeper (www.gene-quantification.de/bestkeeper.html) also selects the least variable genes using the geometric mean but uses the raw data instead of the data converted to a copy number as geNorm[13]. NormFinder (http://www.multid.se/genex/hs410.htm) measures the variation and ranks the potential reference genes according to the grouping of samples, such as untreated/treatment, *etc.* [45].

The standardized exchange of qPCR data is already a reality

The demand for a standardized qPCR data format stimulated the emergence of two innovations. A set of guidelines that describe the minimum information necessary for evaluating qPCR experiments, MIQE, was proposed by Professor S. Bustin with colleagues [2], and the XML-based Real-Time PCR Data Markup Language (RDML) was developed by the RDML consortium [15]. The MIQE provides this approach with a checklist that contains 85 parameters to assure quality results. The detailed practical approach to RT-qPCR—publishing data that conform to the MIQE guidelines may be found in the mini review by Dr. S.Taylor with colleagues [11] and contain the mandatory and elective parameters. The RDML was created to enable the straightforward exchange of raw fluorescence data free of smoothing or baseline subtraction, as they allow quality control, the evaluation and reevaluation of the validity of conclusions if new approaches become available. Recently the new RDML version 1.2 was published [46], which includes the open source editor RDML-Ninja (http://sourceforge.net/projects/qpcr-ninja/) and the database RDMLdb (http://www.rdmldb.org).The editor allows researcher to visualize, edit and validate the RDML files. The database RDMLdb is an online repository for RDML

files. Authors may upload their RDML files into this database and provide the matching IDs in the article, as it is customarily done for the microarray and RNA sequencing. This new RDML infrastructure is the foundation for the standardized qPCR data exchange among scientists and research groups, and for the meta-analysis of gene expression assessed by RT-qPCR in foreseeable future.

The accurate following the protocol of the RT-qPCR workflow with maximal standardization of each step is critical for producing the reliable, comparable and reproducible results. The extension of obtained results via specified open access databases is highly appreciated for dissemination of new knowledge and strengthening the contacts within scientific community for the further scientific achievements.

Supplementary material

Details of the analysis are provided in the Supplement http://biopolymers.org.ua/content/32/3/161/

Acknowledgements

REFERENCES

1. *Rappolee DA, Mark D, Banda MJ, Werb Z.* Wound macrophages express TGF-alpha and other growth factors in vivo: analysis by mRNA phenotyping. *Science.* 1988;**241**(4866): 708–12.
2. *Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT.* The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem.* 2009;**55**(4):611–22.
3. *Bustin SA, Beaulieu JF, Huggett J, Jaggi R, Kibenge FS, Olsvik PA, Penning LC, Toegel S.* MIQE précis: Practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. *BMC Mol Biol.* 2010;**11**:74.
4. *McNutt M.* Journals unite for reproducibility. *Science.* 2014;**346**(6210):679.
5. *Leek JT, Peng RD.* Statistics. What is the question? *Science.* 2015;**347**(6228):1314–5.
6. *Bustin SA.* Molecular medicine, gene-expression profiling and molecular diagnostics: putting the cart before the horse. *Biomark Med.* 2008;**2**(3):201–7.
7. *Bustin SA, Benes V, Nolan T, Pfaffl MW.* Quantitative real-time RT-PCR--a perspective. *J Mol Endocrinol.* 2005;**34**(3): 597–601.
8. *Bustin S, Dhillon HS, Kirvell S, Greenwood C, Parker M, Shipley GL, Nolan T.* Variability of the reverse transcription step: practical implications. *Clin Chem.* 2015;**61**(1):202–12.
9. *Joyce C.* Quantitative RT-PCR. A review of current methodologies. *Methods Mol Biol.* 2002;**193**:83–92.
10. *Martin LA, Smith TJ, Obermoeller D, Bruner B, Kracklauer M, Dharmaraj S.* RNA purification. *In: Molecular Biology Problem Solver: A Laboratory Guide.* Ed Alan S. Gerstein, 2001 by Wiley-Liss, Inc. Ch.8: 197–224.
11. *Taylor S, Wakem M, Dijkman G, Alsarraj M, Nguyen M.* A practical approach to RT-qPCR-Publishing data that conform to the MIQE guidelines. *Methods.* 2010;**50**(4):S1–5.
12. *Pfaffl MW.* Quantification strategies in real-time PCR. *In: A-Z of quantitative PCR* Ed.: S.A. Bustin. International University Line (IUL) La Jolla, CA, USA. 2004; Chaper 3: 87–112.
13. *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**(6):509–15.
14. *Lefever S, Hellemans J, Pattyn F, Przybylski DR, Taylor C, Geurts R, Untergasser A, Vandesompele J; RDML consortium.* RDML: structured language and reporting guidelines for real-time quantitative PCR data. *Nucleic Acids Res.* 2009;**37**(7):2065–9.
15. *Lefever S, Vandesompele J, Speleman F, Pattyn F.* RTPrimerDB: the portal for real-time PCR primers and probes. *Nucleic Acids Res.* 2009;**37**(Database issue):D942–5.
16. *Perepelyuk MM, Fedorchenko DB, Rybalko SL, Obolenskaya MYu.* Interferon α expression in the rat liver after partial hepatectomy. *Biopolym Cell 2006;22(4):276–282.*
17. *Korneyeva KL, Rodrigues RR, Ralchenko SV, Martunovska OV, Frolova AO, Martsenyuk OP, Manzhula LV, Melnyk VT, Shkoropad OY, Obolenska MYu.* Expression of genes, encoding the enzymes of cysteine metabolism in human placenta in the first and the third trimesters of gestation. *Ukr Biochem J.* 2016;**8**(1):88–98.
18. *Korneyeva KL, Rodrigues RR, Ralchenko SV, Vakulenko AV, Manzhula LV, Melnik VT, VereshchakOYu, Obolenskaya-MYu.* The expression of genes encoding the key enzymes of folate dependent metabolism in human placenta in first and third trimesters of uncomplicated pregnancy. *Perinatologiya I pediatriya.* 2014; **4**(60):24–30.
19. *Kuklin AV, Tokovenko BT, Obolenska MYu.* Evaluation criteria of rat hepatocytes transcriptome analysis under the influ-

ence of interferon alpha by DNA microarray. *Biopolym Cell.* 2013; **29**(6): 21–6.

20. *Kuklin AV, Poliezhaieva TA, Zhyryakova IO, Ogryzko VV, Obolenskaya MYu.* Expression of ISGylation related genes in regenerating rat liver. *Biopolym Cell.* 2015; **31**(5):351–61.

21. *Deng MY, Wang H, Ward GB, Beckham TR, McKenna TS.* Comparison of six RNA extraction methods for the detection of classical swine fever virus by real-time and conventional reverse transcription-PCR. *J Vet Diagn Invest.* 2005;**17**(6):574–8.

22. *Sambrook J, Fritsch EF, Maniatis T.* Molecular Cloning: A Laboratory Manual. *NY: Cold Spring Harbor Laboratory, 1989.* 2nd ed.: vol. 3.

23. *Wilfinger WW, Mackey K, Chomczynski P.* Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. *Biotechniques.* 1997;**22**(3):474–6, 478–81.

24. *Jones LJ, Yue ST, Cheung CY, Singer VL.* RNA quantitation by fluorescence-based solution assay: RiboGreen reagent characterization. *Anal Biochem.* 1998;**265**(2):368–74.

25. *Imbeaud S, Graudens E, Boulanger V, Barlet X, Zaborski P, Eveno E, Mueller O, Schroeder A, Auffray C.* Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Res.* 2005;**33**(6):e56.

26. *Mueller O, Lightfoot S, and Schroeder A.* RNA Integrity Number (RIN)—standardization of RNA quality control. 2004. Agilent Application Note, Publication Number-5989-1165EN, 1–8.

27. *Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, Ragg T.* The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol.* 2006;**7**:3.

28. *Wiame I, Remy S, Swennen R, Sági L.* Irreversible heat inactivation of DNase I without RNA degradation. *Biotechniques.* 2000;**29**(2):252–4, 256.

29. *Chomczynski P, Sacchi N.* The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat Protoc.* 2006;**1**(2):581–5.

30. *Nolan T, Hands RE, Ogunkolade W, Bustin SA.* SPUD: a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations. *Anal Biochem.* 2006;**351**(2):308–10.

31. *Thornton B, Basu C.* Real-time PCR (qPCR) primer design using free online software. *Biochem Mol Biol Educ.* 2011;**39**(2):145–54.

32. *Zuker M.* Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 2003;**31**(13):3406–15.

33. *Yun JJ, Heisler LE, Hwang II, Wilkins O, Lau SK, Hyrcza M, Jayabalasingham B, Jin J, McLaurin J, Tsao MS, Der SD.* Genomic DNA functions as a universal external standard in quantitative real-time PCR. *Nucleic Acids Res.* 2006;**34**(12):e85.

34. *Livak KJ, Schmittgen TD.* Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 2001;**25**(4):402–8.

35. *Bustin SA.* Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol.* 2000;**25**(2):169–93.

36. *Bustin SA.* Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J Mol Endocrinol.* 2002;**29**(1):23–39.

37. *Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A.* Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques.* 2004; **37**(1):112–4, 116, 118–9.

38. *Huggett J, Dheda K, Bustin S, Zumla A.* Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun.* 2005;**6**(4):279–84.

39. *Kozera B, Rapacz M.* Reference genes in real-time PCR. *J Appl Genet.* 2013;**54**(4):391–406.

40. *Wang GP, Xu CS.* Reference gene selection for real-time RT-PCR in eight kinds of rat regenerating hepatic cells. *Mol Biotechnol.* 2010;**46**(1):49–57.

41. *Meller M, Vadachkoria S, Luthy DA, Williams MA.* Evaluation of housekeeping genes in placental comparative expression studies. *Placenta.* 2005;**26**(8–9):601–7.

42. *Lion T.* Current recommendations for positive controls in RT-PCR assays. *Leukemia.* 2001;**15**(7):1033–7.

43. *Barber RD, Harmer DW, Coleman RA, Clark BJ.* GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiol Genomics.* 2005;**21**(3):389–95.

44. *Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 2002;**3**(7): RESEARCH0034.

45. *Andersen CL, Jensen JL, Ørntoft 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**(15):5245–50.

46. *Ruijter JM, Lefever S, Anckaert J, Hellemans J, Pfaffl MW, Benes V, Bustin SA, Vandesompele J, Untergasser A; RDML consortium.* RDML-Ninja and RDMLdb for standardized exchange of qPCR data. *BMC Bioinformatics.* 2015;**16**:197.

До питання про визначення концентрації індивідуальних мРНК, нормалізацію експериментальних даних і мінімальну інформацію, необхідну для їх публікації

М. Ю. Оболенська, А. В. Куклін, Р. Р. Родрігес, О. П. Марценюк, К. Л. Корнєєва, В. А. Доценко, О. О. Драгущенко

Реакція зворотної транскрипції і ланцюгової полімеризації в реальному часі (ЗТ-кПЛР) стала найбільш уживаним методом для характеристики профілів експресії індивідуальних мРНК через можливості оцінки концентрацій в широкому діапазоні,

відносної швидкості реакції, чутливості, роздільності і відносно невеликої вартості. Однак, багатоступеневий характер реакції, різні реактиви, різна якість біологічних зразків і відсутність стандартних приписів для проведення реакції приховують небезпеку отримати викривлені результати. Для стандартизації кожного з етапів методу і підвищення надійності результатів проф. С. Бустіним із співробітниками [2004] була розроблена методична інструкція, що містила мінімальну інформацію, яка необхідна для публікації результатів, отриманих за допомогою ЗТ-кПЛР. Крім того, RDML консорціумом на основі XML (розширювана мова розмічання) створені спеціальна мова RDML і база даних RDML для збору і аналізу результатів ЗТ-кПЛР експериментів. В цій статті ми описуємо весь процес ЗТ-кПЛР по етапах згідно вимог методичної інструкції MIQE і нашим власним досвідом у застосуванні цього методу.

К вопросу об определении концентрации индивидуальных мРНК с помощью ОТ-кПЦР, нормализации экспериментальных данных и минимальной информации, необходимой для их публикации

М. Ю. Оболенская, А. В. Куклин, Р. Р. Родригес,
О. П. Марценюк, К. Л. Корнеева, В. А. Доценко,
Е. О. Драгущенко

Реакция обратной транскрипции и количественной цепной полимеризации (ОТ-кПЦР) стала наиболее используемым методом для характеристики профиля экспрессии индивидуальных мРНК благодаря широкому диапазону измеряемых концентраций, малой затратности по времени исполнения, чувствительности, разрешающей способности и относительно небольшой стоимости. Однако, многоступенчатый характер реакции, разнообразие используемых реактивов, разное качество биологических образцов и отсутствие стандартных подходов для количественной оценки результатов таит опасность получить искаженные результаты. Для максимальной стандартизации каждого из этапов реакции и повышения надежности результатов проф. С. Бустиным с сотрудниками [2004] была разработана методическая инструкция с указанием минимальной информации (MIQE), необходимой для публикации данных, которые были получены с помощью ОТ-кПЦР. Кроме того, RDML консорциумом на основе XML (расширяемый язык разметки) разработан специальный язык RDML и создана база данных RDML для сбора и анализа результатов ОТ-кПЦР экспериментов. В этой статье мы описываем поэтапно весь процесс ОТ-кПЦР в соответствии с требованиями методической инструкции MIQE и нашим опытом в области применения этого метода.