

Accurate Measurement of the Relative Abundance of Different DNA Species in Complex DNA Mixtures

SANGKYUN Jeong^{1,*}, HYUNJOO Yu¹, and KARL Pfeifer²

Medical Research Division, Korea Institute of Oriental Medicine, 483 Exporo, Yusong-gu, Daejeon 305-811, Republic of Korea¹ and Laboratory of Mammalian Genes and Development, National Institutes of Child Health and Development, National Institute of Health, Bethesda, MD 20892, USA²

*To whom correspondence should be addressed. Tel. +82 42-868-9314. Fax. +82 42-868-9480.

Email: skjeong@kiom.re.kr

Edited by Naotake Ogasawara

(Received 26 October 2011; accepted 6 January 2012)

Abstract

A molecular tool that can compare the abundances of different DNA sequences is necessary for comparing intergenic or interspecific gene expression. We devised and verified such a tool using a quantitative competitive polymerase chain reaction approach. For this approach, we adapted a competitor array, an artificially made plasmid DNA in which all the competitor templates for the target DNAs are arranged with a defined ratio, and melting analysis for allele quantitation for accurate quantitation of the fractional ratios of competitively amplified DNAs. Assays on two sets of DNA mixtures with explicitly known compositional structures of the test sequences were performed. The resultant average relative errors of 0.059 and 0.021 emphasize the highly accurate nature of this method. Furthermore, the method's capability of obtaining biological data is demonstrated by the fact that it can illustrate the tissue-specific quantitative expression signatures of the three housekeeping genes *G6pdx*, *Ubc*, and *Rps27* by using the forms of the relative abundances of their transcripts, and the differential preferences of *Igf2* enhancers for each of the multiple *Igf2* promoters for the transcription.

Key words: competitive PCR; relative quantitation; transcription

1. Introduction

Gene transcription levels have been measured by methods based on Northern analysis or reverse transcription (RT)–polymerase chain reaction (PCR) assays. These approaches are generally effective for comparing the RNA levels of any gene among many samples. However, these methods are unsuitable for comparing the quantitation data generated in different places or times, primarily due to the paucity of quantitation standardization. Comparing the expression levels of different genes is even more problematic, even though this task is important for understanding the coordinative nature of gene expressions and the relative strength of regulatory elements.¹ Microarray

technology and recent advancements in sequencing technology, such as RNAseq, have opened the way for a thorough examination of the expression of many RNAs; they also facilitate the visualization of relative expression levels of multiple genes.^{2,3} Nonetheless, these assays are yet to be tailored to attain sufficient accuracy or reproducibility between different assay platforms; they are also unsuitable for examining a small number of genes in numerous samples.^{4,5}

The quantitative competitive PCR (qcPCR) method measures the amount of target DNA of interest relative to the amount of an artificially made DNA competitor. With this measurement, the absolute quantity of the target DNA can be determined if the amount of the corresponding competitor is known.^{6,7}

A single-nucleotide difference between two otherwise identical DNAs is the best competition structure that warrants the equal amplification efficiency of these dimorphic DNAs in the PCR. The resultant PCR amplicons have been analysed for the allele ratios in the samples using a wide range of application platforms including Maldi-TOF, pyrosequencing, and real-time DNA melting.^{8–11} However, qPCR approach needs to deal with some inevitable uncertainties. The quantitation of the competitor templates or samples of interest is largely performed in a spectrophotometric manner; that is, the method takes into account the common measurement biases that occur between measurements and between laboratories.

In this study, we present an alternative method of nucleic acid quantitation based on a qPCR approach. With this method, the amount of a given sequence in a mixture such as cDNA or genomic DNA (gDNA) is determined in relation to the amount of another sequence in the mixture. To accomplish this task, we introduced a competitor array as a novel reaction component in the qPCR. The competitor array is an artificially made plasmid DNA in which the competitor templates for the target DNAs in question are arranged with a defined aspect ratio. This structure of the competitor array enables one to infer the quantitative relation between two or more DNA sequences if the relative quantity of each sequence to its respective competitor templates in the competitor array is determined. We also exploited high accuracy and precision of the melting analysis for allele quantitation (MAAQ) method in determining the relative quantity of the target DNA with respect to the competitor template. The MAAQ method utilizes the differential melting patterns of either the fluorescent resonance energy transfer (FRET) probe encompassing on a single-nucleotide variation or restriction fragment length polymorphic DNAs. This method mathematically analyses the melting curve of a sample to extract the constitutional fraction with which each allelic version contributes to the curve.^{10,11} It offers a simple and rapid procedure as it does not require post-PCR purification or additional biochemical reactions while ensuring high accuracy and precision. In addition, the MAAQ analysis can be performed with a relatively less expensive device than other methods. Various exemplary assays are presented to emphasize the validity and versatility of this method in the field of nucleic acid quantitation.

2. Materials and methods

2.1. Plasmids

The plasmid DNAs containing a single target or competitor template were made by TA cloning of

the corresponding PCR amplicons into pCR2.1 (Invitrogen) so that the inserted fragments could be liberated with *EcoRI* restriction. Competitor template was generated by PCR amplification by applying a mismatch-carrying primer. Plasmid DNAs containing multiple template sequences which were used for an equimolar mixture of the target and competitor template and for the competitor arrays were constructed by a means of a single or repeated subcloning process. For the subcloning, a *SpeI/XbaI* fragment of a plasmid was placed into the *XbaI* or *SpeI* site of another plasmid. Plasmids that contained any form of an inverted repeat were avoided. Further information on the plasmids used in this study is presented in Supplementary Table S5.

2.2. RT and PCR

RT was carried out on 1 μ g of the total RNA with an iScriptTM cDNA synthesis kit (BioRad) in accordance with the manufacturer's instructions. PCR was performed with ~ 0.2 μ l of the reverse transcribed cDNA template, 20 ng of gDNA or 0.1 μ g of plasmid DNA for appropriate analyses and 0.5 μ M of corresponding primer pairs: G6pd-1F and G6pd-7R for *G6pdx*; Ubc-1F and Ubc-4R for *Ubc*; Rps-4F and Rps-2R for *Rps27*; Igf2-E1F1 and Igf2-E4R1 for the *Igf2* P1 transcript; Igf2-E2F1 and Igf2-E4R1 for the *Igf2* P2 transcript; Igf2-E3F1 and Igf2-E4R1 for the *Igf2* P3 transcript; and Igf2-E4F1 (GTACCAATGGGGATCCCAGT) and Igf2-E4R1 for the *Igf2* total transcript. Information on the primer sequences is presented in Supplementary Table S5. The competitor arrays were prepared by cleaving the corresponding plasmid and subsequently diluting it in water containing λ phage DNA (1 ng/ μ l). For the competitive PCR, each template (except for the references) was mixed with its appropriate competitor array and amplified by PCR by using 40 cycles at 94°C for 10 s, 56°C for 20 s, and 72°C for 30 s.

2.3. Melting analysis

The relative ratios of the target DNA to its competitor in the PCR amplicon were determined as previously described. Briefly, 5 μ l of PCR amplicons were mixed with 5 μ l of 10 mM ethylenediaminetetraacetic acid and with 0.4 μ M of FRET probes, Igf2-S1 and Igf2-A1 (see Fig. 1b) for *Igf2*; G6pdx-S, Red640-5'-GCACTG TTGGTGAAGATGTCA-3', and G6pdx-A, 5'-CACACG GATGACCAGCTCATTACG-3'-FITC, for *G6pdx*; Ubc-S, Red640-5'-TACGTCTGTCTTCTCTGTTTGAC-3', and Ubc-A, 5'-GGCTCTTTT TAGATACTGTGGTGAGG-3'-FITC, for *Ubc*; and Rps27-A, 5'-CCTTCTGATGA ATGTGGTGTGGAG-3'-FITC and Rps27-S, Red640-5'-TCTGGGAAGCCACTTTGACAG-3', for *Rps27*. The DNA melting was performed with Roche LightCycler 2.0 under the following conditions: 30 s at 95°C; and

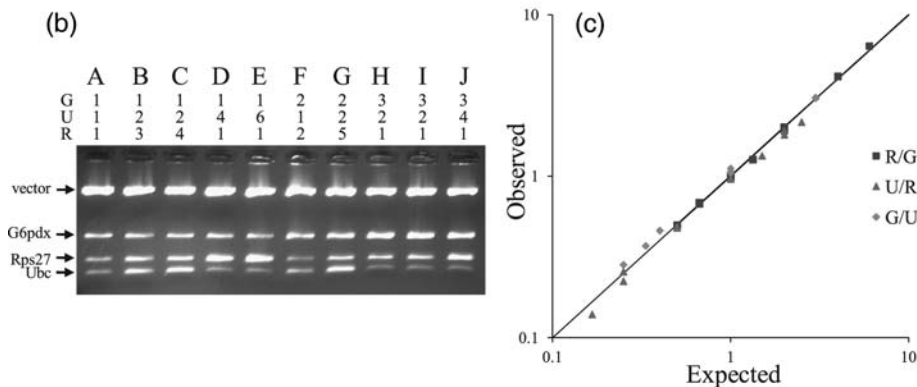
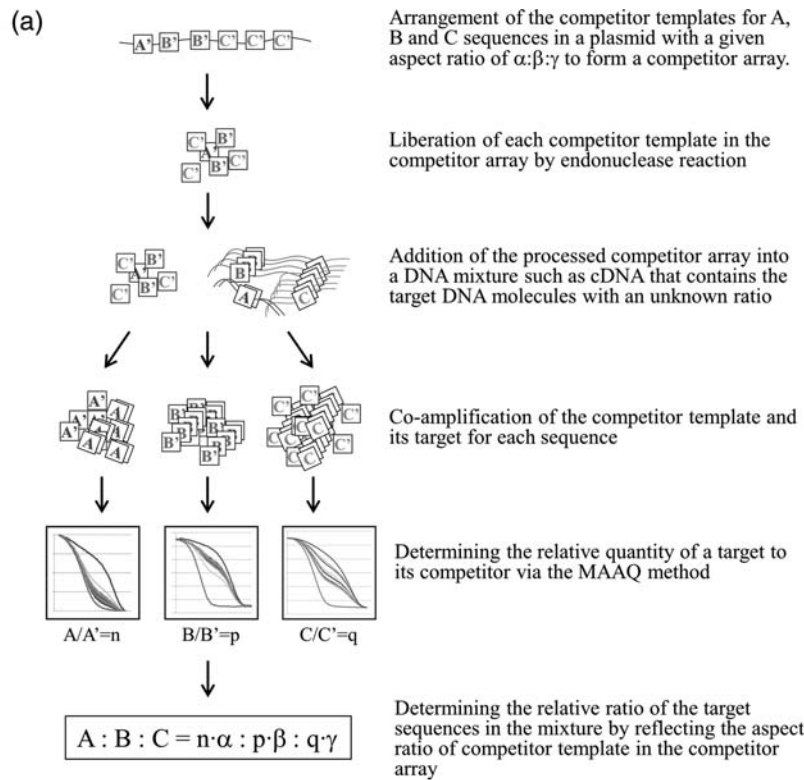


Figure 1. Determination of relative abundances of different DNA sequences. (a) A schematic representation of the entire procedure for determining the relative abundances among three sequences (A, B, and C) in a cDNA sample. The respective competitor templates (A', B', and C') which are arranged in a competitor array with a given aspect ratio of $\alpha:\beta:\gamma$ are liberated by endonuclease restriction, mixed with cDNAs, and subjected to PCR reactions for each sequence of interest. The relative quantities of the target sequences with respect to the corresponding competitor in each amplicon, n, p, and q, as assessed by MAAQ, are fed into a calculation based on the aspect ratio of the competitor template in the competitor array to yield the relative abundances among the target sequences. (b) A set of artificial DNA mixtures constituted by the restriction of the plasmids (A–J) that contain cDNA fragments of three housekeeping genes with variable copy numbers as denoted under the name of the mixtures and fractionated on an agarose gel. The identity of each DNA is denoted on the left. (c) The relative abundances of two cDNA sequences were determined and scatter-plotted against the expected values on the log scale axes.

8 min at either 45 °C (for *Igf2*, *G6pdx*, and *Ubc*) or 44 °C (for *Rps27*) at 0.1 °C stepwise increments in temperature until the probe's melting is completed. The background-subtracted fluorescence values (640/Back-530) from the Red640 dye were extracted with the aid of LightCycler Software 4.05 and then

submitted for calculation as previously reported.¹¹ The relative abundance of two sequences in each sample was determined from the ratio of the T/C values of the two comparing sequences after reflecting the copy numbers of their corresponding competitor templates in the competitor array used.

3. Results

3.1. Determining the relative abundance of DNAs in simple mixtures

Figure 1a depicts the typical procedure for determining the relative abundances of the sequences of interest in a mixture. To demonstrate the feasibility of this quantitation approach, we used it to determine the relative quantity of a cDNA sequence with respect to another sequence in the artificially made cDNA mixtures. A set of cDNA mixtures was composed through the construction of 10 plasmids (A–J); they contained variable numbers of each of the three cDNA fragments of mouse *Ubc*, *Rps27*, and *G6pdx* genes. Plasmids H and I were from independent clones with the same sequence composition. The relative abundances of the sequences in each plasmid are therefore explicitly known (Fig. 1b). Another plasmid DNA, the competitor array CA1, was made to contain a single copy of each competitor template, namely a single-nucleotide substituted version of the respective cDNA sequence. These plasmid DNAs were structured to liberate each cDNA fragment as a single entity upon *EcoRI* digestion (Fig. 1b). Each mixture generated by the *EcoRI* digestion of the corresponding plasmid was individually mixed with the *EcoRI*-restricted CA1 and subsequently subjected to competitive PCR reactions along with three standard templates, namely the target cDNA, the corresponding competitor template, and an equimolar mixture of the target cDNA and its competitor template. One of the fluorescently labelled FRET probes for each cDNA sequence was designed to span the nucleotide of variation between two competing templates so that the probe's melting from the complementary sequences are differentiated depending on the completeness of the base pairing. The melting curves that we generated from the PCR amplicons by applying the respective FRET probes were fed into the MAAQ process to determine the relative quantities of each target cDNA sequence with respect to the competitor template (T/C for short).

Supplementary Table S1 presents the T/C values of each cDNA sequence (*G6pdx/C*, *Ubc/C*, and *Rps27/C*) in the cDNA mixtures; the quantities were determined from the quadruplicate PCR reactions. The abundances of a cDNA sequence relative to another (*Rps27/G6pdx*, *G6pdx/Ubc*, and *Ubc/Rps27*) in each mixture were then obtained from the ratios of the respective T/C values of the two sequences in question. Figure 1c shows a scattered plot of the observed versus expected values of the relative abundances; the plot highlights the accuracy of this assay. The average relative error for this assay is 0.059, demonstrating that our qPCR approach can accurately

assess the abundance of a sequence in relation to another sequence in a simple mixture.

3.2. Determining the relative abundance of different DNAs in complex mixtures

We subsequently considered whether our approach would be reliable in the analysis of complex samples such as those that are likely to be encountered in a laboratory. To address this matter, we measured the copy number ratios of the autosomal *Ubc* and the X-linked *G6pdx* sequences in male and female gDNAs. We chose gDNA because it is a particularly complex mixture of sequences and provides a stringent test for our assay. The *Ubc/G6pdx* ratios were expected to be 1 and 2 for females and males, respectively. About 20 ng of each liver gDNA of five females and five males was mixed with 0.05 pg of CA1 and subjected to PCR reactions for two sequences. We then applied the MAAQ. Five XY samples have an average ratio of 1.978, and five XX samples have a ratio of 1.000 with an average relative error of 0.021. These results strongly confirm that this method can measure the relative abundances of the target DNAs, even in a highly complex mixture such as gDNA (Table 1).

3.3. Determining the tissue-specific expression signatures via the forms of relative abundance of cDNAs

Because of the proven accuracy and reliability of our method, we used it to obtain two types of biological data. The first type of data is the relative expression abundances among three housekeeping genes in different mouse tissues. A preliminary assessment using the serial dilutions of the CA1 on the tissue cDNAs indicates that the abundances of the *Ubc* and *Rps27* transcripts are in an equivalent range whereas the abundance of *G6pdx* transcript is far less than those of the two transcripts (more than a hundredfold in some cases, data not shown). To accurately cover this wide range of expression levels, we made an additional competitor array, CA2, to carry the copies of *G6pdx*, *Ubc*, and *Rps27* sequences at a ratio of 1:7:7. The PCR amplicons of each sequence were generated by using a mixture of each cDNA and CA2 as a template. Subsequent analysis revealed the T/C values and the relative abundances among the target sequences (Supplementary Table S2). Figure 2a visualizes the relative abundances of two subjected sequences in the cDNAs of the liver, heart, and brain of three female adult mice. These results demonstrate the tissue-specific expression signatures of these three genes with forms of their relative transcript abundances; they also demonstrate the relatively small

Table 1. The relative abundance of nucleic acid species can be reliably calculated in gDNA samples

	Female					Male				
	F1	F2	F3	F4	F5	M1	M2	M3	M4	M5
G6pdx										
γG6pdx	0.601	0.512	0.518	0.510	0.543	0.378	0.385	0.214	0.326	0.400
SD	0.008	0.007	0.010	0.013	0.010	0.020	0.015	0.011	0.012	0.015
G6pdx/CA1	1.504	1.048	1.074	1.039	1.187	0.607	0.625	0.273	0.484	0.666
Ubc										
γUbc	0.599	0.511	0.519	0.506	0.548	0.561	0.553	0.336	0.49	0.567
SD	0.008	0.007	0.017	0.008	0.006	0.006	0.010	0.023	0.011	0.014
Ubc/CA1	1.492	1.044	1.081	1.024	1.213	1.276	1.238	0.506	0.960	1.310
Ubc/G6pdx ^a	0.992	0.996	1.006	0.985	1.022	2.102	1.981	1.857	1.985	1.967
Obs/Exp	0.992	0.996	1.006	0.985	1.022	1.051	0.990	0.929	0.992	0.984

A sample of gDNA was prepared from the kidneys of 10 adult mice (five females: F1–F5 and five males: M1–M5). The relative abundance of the *Ubc* and *G6pdx* genes was then determined as described in the text. The terms γG6pdx and γUbc are the proportions of the target sequences within the competitively amplified amplicons, and G6pdx/CA1 and Ubc/CA1 are the relative ratios of the *G6pdx* and the *Ubc* sequences with respect to their corresponding competitors, respectively; the standard deviation (SD) is based on from three measurements; the term Obs/Exp refers to the ratios of the observed per expected values.

^aAverage relative error, 0.021.

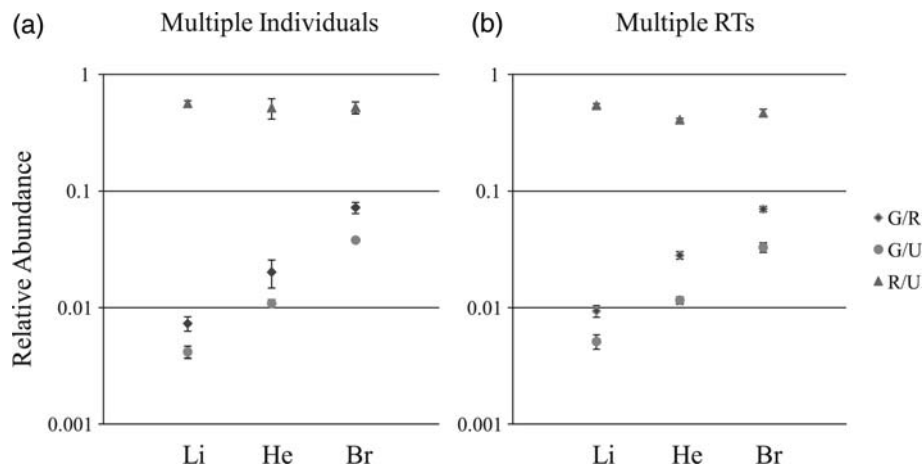


Figure 2. Relative abundances among the transcripts of three housekeeping genes that display tissue-specific expression signatures. (a) A plot of the relative expression ratios of two of the three housekeeping genes in three adult tissue samples as determined from three individuals. The error bar denotes the standard deviations among the individuals. (b) A plot of the relative expression ratio as determined from multiple RT reactions of a single individual; the plot highlights the precision of the assay. The error bar denotes the standard deviations among multiple RT reactions.

individual variations [which have a 0.125 coefficient of variation (CV)].

Figure 2b shows the relative abundances among the transcripts of the three tissues of a single individual. The values were averaged from four measurements. We determined three of the measurements from three independent RT reactions by adapting another version of competitor array, CA3, with the respective copy ratio of 1:9:9 for the sequences of *G6pdx*, *Ubc*, and *Rps27*. The CV of these measurements (0.079) confirms that our quantitation method has a high level of precision.

3.4. Measuring the relative promoter strength of multiple promoters of a gene

We tried to obtain the relative transcriptional activity of the multiple *Igf2* promoters in embryonic tissue. The utilization of multiple promoters for transcription is not rare in a eukaryotic system.¹² The usage rate of each promoter could be the subject of developmental or tissue-specific regulations. If the promoter strength is linearly correlated with the cellular level of the transcript it drives, our assay scheme is well suited for assessing the relative strength among the promoters.

This type of assessment provides insight into the relative promoter activity of the given samples and confirms the potential of our approach as a versatile quantitation method. Transcriptional regulation of the murine *Igf2* gene is relatively well characterized with respect to the role and position of *cis*-regulatory elements, including promoters and enhancers.¹³ Fetal and neonatal *Igf2* transcription is derived from multiple promoters, P1, P2, and P3, through appropriate interactions between regulatory sequences; these types of interactions provide a good model system for comparing the transcriptions of multiple promoters and specific interactions.¹⁴

We conducted an assay in which each promoter-specific transcription was compared with the total level of *Igf2* transcription in two embryonic tissues, the liver and muscle with endodermal and mesodermal origins, respectively. The primers were designed to span two consecutive exons to prevent any potential noise from gDNA contamination. The promoter specificity of the PCR amplicon was given by placing a part of the forward primers on the promoter-specific exons (Fig. 3a). The competitor templates for the promoter-specific amplification were made by placing only a single-nucleotide substitution on exon E4; they were then arranged into a plasmid to constitute the competitor array CA4. Each promoter-specific competitor was also used as the competitor for the entire *Igf2* transcript by applying a new forward PCR primer placed on the 5'-end of the shared exon E4 (Fig. 3a). In this way, CA4 can be exploited in two different competitions with an aspect ratio of 1:3. One competition is for the promoter-specific transcript; the other is for the total transcript. Each cDNA was added to the CA4 and subsequently subjected to four separate PCR amplifications: three of the PCR amplifications were specific for the promoter-specific transcript and the other was used for the total transcript. After determining the T/C values of each PCR reaction, we obtained the relative abundance of each promoter-specific transcript with respect to the total *Igf2* transcript; the results are presented in Fig. 3b and Supplementary Table S3.

The *Igf2* transcription in the liver necessitates the involvement of an endoderm enhancer, whereas the muscle expression largely relies on a mesoderm enhancer. Our results confirm that minimal transcription activity occurs on promoter P1, regardless of which enhancer is involved; specifically, only 6–7% of the total transcript is attributed to promoter P1 activity in both the liver and muscle. On the other hand, the two types of tissue with developmentally different origins have different preferences with regard to the utilization of promoters P2 and P3: compared with P3, P2 has more transcription in the muscle but less in the liver. This difference indicates that the relative strength of the two promoters depends on the type

of enhancers used. Moreover, the results confirm that this method is highly reproducible; the standard deviations of the measurements of three independent RT reactions are minute (ranging from 0.051 to 0.001, CV = 0.033).

If exon E4 is spliced solely for the transcripts derived from the three promoters addressed here, the relative transcription levels of the three promoters should add up to 1. The fact that the aggregate transcription levels for the muscle samples are 1.010 and 1.008 is consistent with this notion. On the other hand, the aggregate transcription levels for the liver samples are 0.920 and 0.938. This difference may be attributed to the possible existence of one or more transcripts that contain exon E4 in the liver.

4. Discussion

In this study, we presented a qPCR method that can be used to measure the relative abundances among the sequences of interest in any DNA sample. Because of the technical difficulty of obtaining a direct quantitative comparison between two unlinked sequences, we exploited the absolute relative of their respective competitor templates which are arranged in the competitor array. In addition, the high accuracy of the MAAQ method was exploited for the quantitation of the fractional ratios of the competitively amplified DNAs. The relative abundances of two different sequences were consequently obtained in a series of plasmid constructs and gDNAs. These abundances turned out to be very accurate; the average relative errors of 0.059 and 0.021, respectively, confirm the reliability and practical feasibility of this approach.

The quantitation scheme was further validated by its successful acquisition of biological data. We compared three housekeeping genes in terms of the abundances of their transcripts for multiple types of mouse tissue and analysed *Igf2* expression in terms of the proportion of the promoter-specific transcript in the total *Igf2* transcripts. These results efficiently illustrate the tissue-specific expression signatures of the three housekeeping genes and the relative transcriptional strength of the *Igf2* promoters.

With an exploitation of the high level of accuracy and precision, our method can potentially be used as a form of standardized quantitation. That is, the expression level of any gene of interest can be determined relative to the expression of a single gene or a combination of several reference genes in the same sample. Thus, our method can produce a stand-alone value for the transcription level of a gene in a particular sample. Furthermore, this value can readily be compared with other results that are

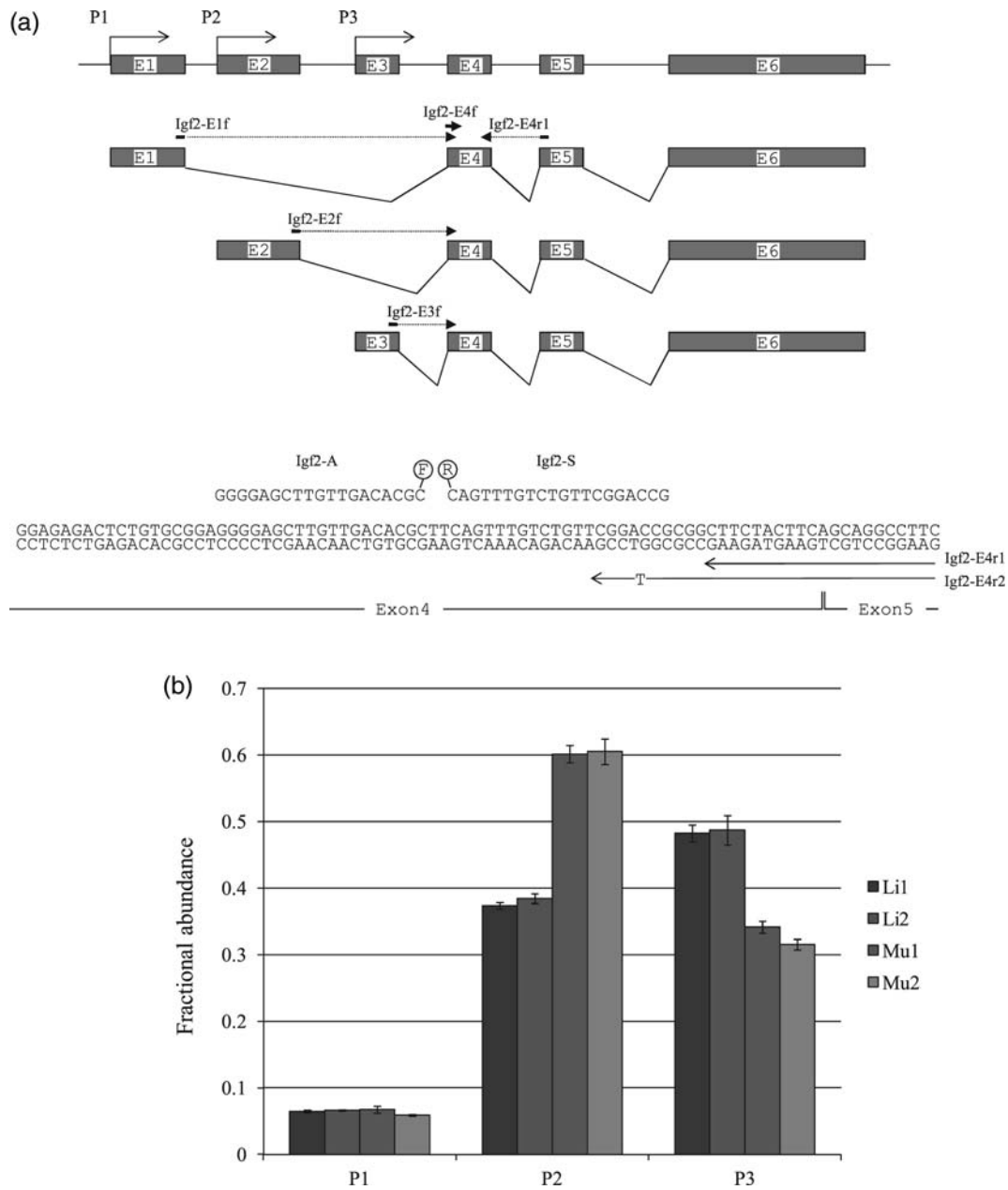


Figure 3. Promoter-specific transcriptions as measured in relation to the total transcription level. (a) The fetal *Igf2* expression as derived from the promoters P1, P2, and P3. Each exon (E1, E2, and E3) constitutes a promoter-specific part of *Igf2* mRNA that is spliced to the shared exon E4. The primers (arrows with dotted line in the middle) were designed to span two consecutive exons except for one forward primer, Igf2-E4F, which is placed on the 5'-most part of exon E4 to account for the total transcription of *Igf2* and the common reverse primer. The detailed structure of the exon E4–exon E5 junction region regarding the positions of the reverse primers Igf2-E4R1 and Igf2-E4R2 and the FRET probes Igf2-A and Igf2-S is depicted in the lower part of the figure. (b) The relative quantities of the promoter-specific transcription with respect to the total transcription yield the relative abundance of each transcript in the samples. The relative transcript levels, which were measured from muscle tissue and liver tissue, represent the mesodermal and endodermal origins, respectively.

independently generated by the same method at any place or time. In this regard, *Ubc* and *Rps27* can be considered as reference genes because their relative transcript abundances were relatively constant in all the tested tissues. Our data also demonstrate that the assay remains robust across a wide range of expression levels (e.g. the *G6pdx/Ubc* value in liver is

0.005). This range can be further extended by applying a competitor array that contains the respective competitor templates with a larger aspect ratio.

For the purpose of sharing and exchanging the quantitation data of the transcripts, various applications have been developed to present the quantities on an absolute scale such as copy numbers per

cell.^{15,16} Miura *et al.*¹⁶ reported a qPCR approach with which the transcription level of thousands of yeast genes were described as copy numbers per cell. Kanno *et al.*¹⁵ also presented a method to convert the microarray data or quantitative PCR data to copy numbers per cell. Both approaches utilize a series of standard RNAs which is spiked into the sample of interest in order to provide a way of calibration of inter-experiment differences or normalization of quantity data in the experimental sets. Although these approaches are effective in providing the structural architecture of transcriptome abundances on an absolute scale, it still requires standardization of standard RNAs in order to cross-refer those data to others that are independently generated. In this circumstance, it would be more practical to present the transcription level of a gene in relation to that of another in the same sample as we did. Many approaches including above two examples and others that are focused on defining copy number variation in the genomes are indeed capable of delineating the abundances of multiple sequences in the view of their relativity.^{17,18} However, the accuracy and the precision that these approaches can confer are not as high as those our method can, or were not addressed probably due to disregard of presenting the relative abundances of intergenic sequences.

On the other hand, application of our method for the standardized quantitation is yet to address certain problems that are potentially caused by the differences in the cDNA conversion efficiency of individual RNAs and between the segments of a given RNA. The different reaction conditions or types of enzymes used for RT consequently produce variations in the relative quantity of given sequences in the cDNA pool. Nonetheless, our approach provides a way of examining the extent to which the relative abundance of those sequences varies on account of the different reaction conditions; it also provides a way of determining the best conditions for minimizing such variations.

With simplicity, speed, and accuracy, our quantitation approach provides a straightforward paradigm for comparing gene expression levels across cell types, developmental stages, and even different species. Most importantly, with this approach, experiments performed in different laboratories or at different times can be correctly referenced to one another.

Supplementary Data: Supplementary Data are available at www.dnaresearch.oxfordjournals.org.

Funding

This work was supported by the National Research Foundation of Korea grant (NRF, No.20110027739)

and a Korea Institute of Oriental Medicine grant (KIOM, No.K11070) funded by the Korea government (MEST).

References

- Cheung, V.G. and Spielman, R.S. 2009, Genetics of human gene expression: mapping DNA variants that influence gene expression, *Nat. Rev. Genet.*, **10**, 595–604.
- Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. 1995, Quantitative monitoring of gene expression patterns with a complementary DNA microarray, *Science*, **270**, 467–70.
- Marioni, J.C., Mason, C.E., Mane, S.M., Stephens, M. and Gilad, Y. 2008, RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays, *Genome Res.*, **18**, 1509–17.
- Tan, P.K., Downey, T.J., Spitznagel, E.L. Jr, et al. 2003, Evaluation of gene expression measurements from commercial microarray platforms, *Nucleic Acids Res.*, **31**, 5676–84.
- Wang, Y., Barbacioru, C., Hyland, F., et al. 2006, Large scale real-time PCR validation on gene expression measurements from two commercial long-oligonucleotide microarrays, *BMC Genomics*, **7**, 59.
- Zentilin, L. and Giacca, M. 2007, Competitive PCR for precise nucleic acid quantification, *Nat. Protoc.*, **2**, 2092–104.
- Siebert, P.D. and Larrick, J.W. 1992, Competitive PCR, *Nature*, **359**, 557–8.
- Ding, C. and Cantor, C.R. 2003, A high-throughput gene expression analysis technique using competitive PCR and matrix-assisted laser desorption ionization time-of-flight MS, *Proc. Natl. Acad. Sci. USA*, **100**, 3059–64.
- Chang, H.W. and Shih Ie, M. 2005, Digital single-nucleotide polymorphism analysis for allelic imbalance, *Methods Mol. Med.*, **103**, 137–41.
- Jeong, S., Hahn, Y., Rong, Q. and Pfeifer, K. 2007, Accurate quantitation of allele-specific expression patterns by analysis of DNA melting, *Genome Res.*, **17**, 1093–100.
- Yu, H., Koo, I. and Jeong, S. 2009, Relative quantitation of restriction fragment length polymorphic DNAs via DNA melting analysis provides an effective way to determine allele frequencies, *Genomics*, **94**, 355–61.
- Davuluri, R.V., Suzuki, Y., Sugano, S., Plass, C. and Huang, T.H. 2008, The functional consequences of alternative promoter use in mammalian genomes, *Trends Genet.*, **24**, 167–77.
- Kaffer, C.R., Grinberg, A. and Pfeifer, K. 2001, Regulatory mechanisms at the mouse Igf2/H19 locus, *Mol. Cell Biol.*, **21**, 8189–96.
- Rotwein, P. and Hall, L.J. 1990, Evolution of insulin-like growth factor II: characterization of the mouse IGF-II gene and identification of two pseudo-exons, *DNA Cell Biol.*, **9**, 725–35.
- Kanno, J., Aisaki, K., Igarashi, K., et al. 2006, Per cell normalization method for mRNA measurement by quantitative PCR and microarrays, *BMC Genomics*, **7**, 64.

16. Miura, F., Kawaguchi, N., Yoshida, M., et al. 2008, Absolute quantification of the budding yeast transcriptome by means of competitive PCR between genomic and complementary DNAs, *BMC Genomics*, **9**, 574.
17. Qiao, Y., Liu, X., Harvard, C., et al. 2007, Large-scale copy number variants (CNVs): distribution in normal subjects and FISH/real-time qPCR analysis, *BMC Genomics*, **8**, 167.
18. Iwao-Koizumi, K., Maekawa, K., Nakamura, Y., et al. 2007, A novel technique for measuring variations in DNA copy-number: competitive genomic polymerase chain reaction, *BMC Genomics*, **8**, 206.

