Journal of Biomolecular Techniques
15:107–111 © 2004 ABRF

Quantitative Real-Time Polymerase Chain Reaction: Methodical Analysis and Mathematical Model

Stefan Wilkening[a] and Augustinus Bader[b]

[a]German Research Centre for Biotechnology, Braunschweig, Germany; [b]Biomedical-Biotechnological Center, Leipzig, Germany

Real-time polymerase chain reaction was established for 16 genes using the LightCycler system to evaluate gene expression in human hepatocytes. During the experiments a large set of data has been obtained. These data have now been evaluated with respect to template stability, accuracy of melting curve analysis, and reproducibility. In addition, the statistical evaluation of the efficiencies of all 16 polymerase chain reactions led to a new mathematical model. To examine template stability, the degradation of mRNA and cDNA was determined at different temperatures. Surprisingly, cDNA, which was obtained by first-strand synthesis, appeared to degrade significantly faster than the respective mRNA. Melting curve analysis is a fast and sensitive method to check for polymerase chain reaction specificity. However, we show that two transcription variants of the glutathione S-transferase 1 gene, with over 100 bp length difference, could not be distinguished by this method. Furthermore, an equation was set up describing the correlation between polymerase chain reaction efficiency and crossing point. This equation can be used to estimate the number of template molecules without having a standard of known concentration. Finally, experimental reproducibility of the real-time polymerase chain reaction was defined.

KEYWORDS: Quantitative real-time polymerase chain reaction, LightCycler, melting curve analysis, polymerase chain reaction efficiency, cDNA stability.

ADDRESS CORRESPONDENCE AND REPRINT REQUESTS TO: Stefan Wilkening, German Cancer Research Center (DKFZ), Molecular Genetic Epidemiology, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany (tel.: 49-6221-421803; fax: 49-6221-421810; e-mail: stefan_wilkening@web.de).

This work was supported by a fellowship from the Dr. Hilmer-Stiftung for Stefan Wilkening.

In a previous report, we described the use of real-time polymerase chain reaction (RT-PCR) to quantify a large number of genes coding for phase I and phase II enzymes of drug metabolism.[1] This was done with human hepatocytes using the LightCycler system and SYBR green I as fluorescent marker. In the current work we analyzed single experimental steps of this method. The key point for quantitative RT-PCR is the quality of the template for reverse transcription (mRNA) and the template for the PCR (cDNA). To get a general idea of the stability of these templates, we determined mRNA and cDNA degradation at different storage temperatures. Thereby we could disprove the assumption of many scientists that DNA in general is more stable than RNA. Furthermore, we demonstrate the limits of melting curve analysis in the LightCycler, showing that two significantly different PCR amplificates are indistinguishable by this method.

For an absolute quantification by RT-PCR, standard curves are normally generated with standards of known concentration. Presupposing that the number of PCR products at the crossing point (CP) is always similar, it is possible to estimate the original number of template molecules only by determining the PCR efficiency. To do so, we set up an equation. Additionally, we determined the reproducibility of RT-PCR with the LightCycler by specifying intra- and interassay variations.

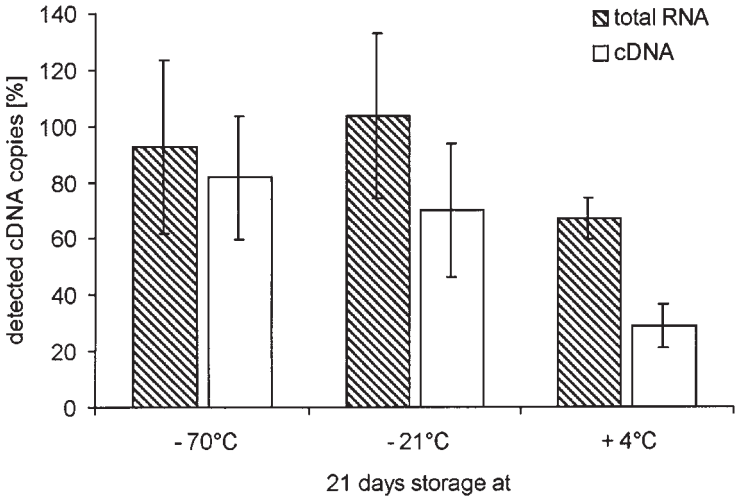

FIGURE 1

Temperature-dependent degradation of mRNA and cDNA. values are given as percent of the molecule number at day 0 (mean ± SD). Total RNA samples for this experiment were isolated from two different cultures of human hepatocytes and were examined for the expression of two genes (glyceraldehyde-3-phosphate dehydrogenase and microsomal epoxide hydrolase), respectively (n = 4). For all quantifications the same stock of reverse transcriptase and SYBR Green was used.

MATERIALS AND METHODS

RNA Isolation and Reverse Transcription

Total RNA was isolated from primary human hepatocytes[1] using 50 µL Trizol reagent (Invitrogen, Karlsruhe, Germany) per square centimeter of the culture dish. This method yielded an average of 30 µg total RNA from 10^6 cells. To avoid loss and degradation of RNA, no DNase digestion was done. First-strand cDNA synthesis was performed in a total volume of 20 µL, where 5 µg of total RNA were reverse transcribed using 200 units of Superscript II RNase H$^-$ reverse transcriptase (Invitrogen), 5 pmol oligo$(dT)_{18}$ primer, and 40 units RNase inhibitor (RNasin, Promega, Madison, WI, USA).

Quantification of cDNA by RT-PCR

To avoid amplification of genomic DNA, primers were designed to be intron-spanning. Primer sequences and conditions used for quantitative RT-PCR are described in detail elsewhere.[1] Briefly, quantification of cDNA was performed in the LightCycler (Roche, Mannheim, Germany) using the DNA Master SYBR Green I kit (Roche) and TaqStart Antibody (Clontech, Heidelberg, Germany) to avoid unspecific amplification reactions. For each gene a standard curve was generated with serial dilutions of plasmids containing the respective PCR product. These plasmids were linearized by endonuclease digestion to avoid supercoiled structure. Amplification curves from linearized plasmids arose about three cycles earlier than from circular plasmides (data not shown). To quantify gene expression, PCR was performed with cDNA probes and one dilution (10^6 copies) of the respective plasmid as standard. The expression was then normalized against the expression of glyceraldehyde-3-phosphate dehydrogenase. PCR specificity was proved by sequencing, gel electrophoresis, and melting curve analysis, respectively.

RESULTS

Comparison of mRNA and cDNA stability

To estimate template stability, we determined how much mRNA and single-stranded cDNA degraded after 3 weeks at different temperatures (stored in H_2O, pH 8.3). As expected, the highest level of degradation was observed at the highest temperature (Fig. 1). Interestingly, mRNA appeared to be more stable than cDNA. Even at 4°C, 65% of the initially quantified molecules could be detected whereas only about 30% from could be detected from cDNA.

Melting Curve Analysis

The primer pair we used for glutathione *S*-transferase 1 *(GSTM1)* span a variable region of the gene. With these primers two PCR products could be amplified, representing two alleles: *GSTM1* variant 1 (accession no.: NM_000561) and variant 2 (accession no.: NM_146421). This variants are clearly distinguished by gel electrophoresis (Fig. 2). However, PCR products of both variants show exactly the same melting curve in the LightCycler. This demonstrates that melting curve analysis is not always suited to detect different PCR products. A total of seven livers were examined where the overall allele frequency was 64% for *GSTM1* variant 1 and 36% for *GSTM1* variant 2.

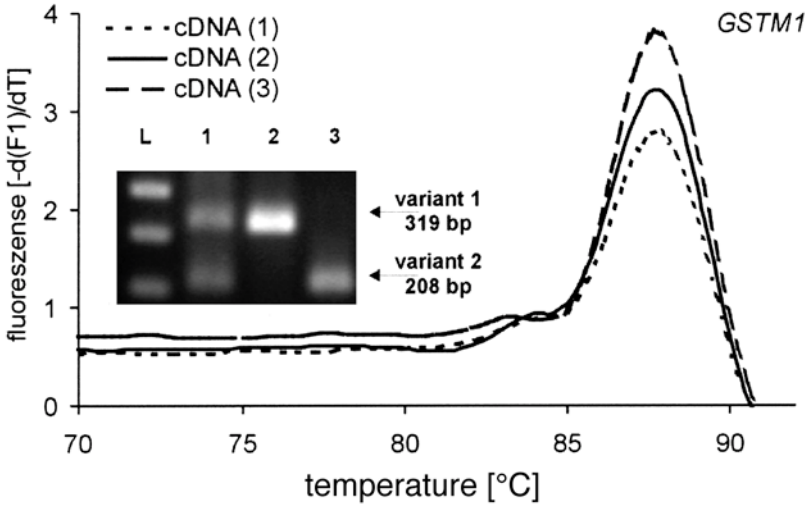

FIGURE 2

LightCycler melting curve analysis from the amplification of the glutathione *S*-transferase 1 *(GSTM1)* gene. The melting curves are displayed as first negative derivative of the fluorescence versus the temperature. Thus, a peak can be seen at the melting temperature. The gel bands represent two allele variants of the *GSTM1* gene ("L" stands for 100 bp leader). In cDNA-1 both alleles are present, in cDNA-2 only variant 1 and cDNA-3 only variant 2 (lacking an 111 bp exon). However, no differences could be seen in the melting curves of the respective PCR products.

Correlation Between PCR Efficiency and CP

The efficiency of a PCR can be deduced by the slope of the standard curve according to equation:[2]

$$E = 10^{[-1/\text{slope}]} \quad (1)$$

The maximum efficiency possible in PCR is 2—every PCR product is replicated every cycle. The minimum value is 1, corresponding to no amplification. If the efficiency of a PCR is relatively high, the respective amplification curve should arise earlier, and the CP should lay at fewer cycles than a PCR with a low efficiency. To prove if this is true, we plotted PCR efficiencies against CPs (Fig. 3). This was done for PCRs of 16 genes, where 10^6 copies of the respective DNA standard were used as starting concentration. The number of copies of the PCR product at the CP (N_{CP}) can be calculated by the equation:[2]

$$N_{CP} = N_0 = E^{CN}, \quad (2)$$

where N_0 is the starting concentration and CN the cycle number at the CP. This way we defined the average N_{CP} value from all 16 PCRs which was $1.12 \times 10^{10} \pm 3.66 \times 10^9$. The calculated number of copies of a PCR product at the CP (N_{CP}) comes very close to the value of 10^{10} given by Rasmussen[2]. To predict the CP for a given PCR efficiency, Eq. 2 can be linearized to:

$$CN = \ln(N_{CP}/N_0)/\ln E. \quad (3)$$

If N_{CP} is 1.12×10^{10} and N_0 is 10^6, then the equation is:

$$CN = 9.32/\ln E \quad (4)$$

Most of our empirical values lay precisely on the dotted graph which is reflected by this equation (Fig. 3). The lengths of the PCR products (ranging from 250 to 560 bp) did not appear to have any influence on the correlation between PCR efficiency and CP. Resolving Eq. 2 to N_0, an approximate value of the initial copy number (N_0) can be calculated if the PCR efficiency is known and 1.12×10^{10} is entered for N_{CP}:

$$N_0 = 1.12 \times 10^{10}/E^{CN}. \quad (5)$$

Sensitivity and Accuracy of RT-PCR

To prove sensitivity and accuracy of real-time PCR in the LightCycler, intra- and interassay variations were determined for different template concentrations (Table 1). Intra-assay variation was determined in three repeats within one LightCycler run. Interassay variation was determined from three runs on three different days using the same cDNA sample. Additionally, variation within three different reverse transcriptions was determined. As expected, at higher template concentrations lower variations were observed.

DISCUSSION

Here, we observed cDNA to degrade faster than mRNA. There are two possible explanations for this finding: (1) After reverse transcription, a cDNA–RNA complex is formed, which might be more vulnerable to nucleases than the RNA with its complex secondary structure; (2) the buffer composition after reverse transcription enhances degradation. Further studies need to be performed to clarify the reason for the fast

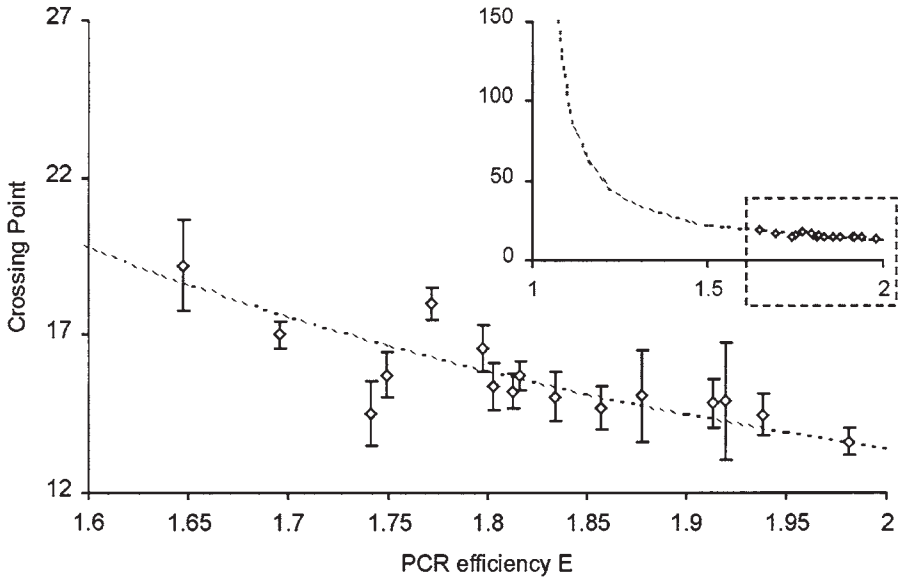

FIGURE 3

Correlation between PCR efficiency and crossing point. PCR efficiencies of 16 PCRs were plotted against cycle number of the respective crossing points (mean ± SD; n = 5). The initial template concentration was 10^6 copies of the respective DNA standard per 20-μL reaction. The crossing point was determined by the "second derivative maximum method." The average length of the PCR products was 325 bp. The *dotted line* describes the theoretical relation between PCR efficiency "E" and the cycle number at the crossing point "CN" according to the equation: CN = 9.32 / ln E, which assumes a constant number of 1.12×10^{10} DNA copies at the crossing point. The small graph in the upper right corner shows the entire graph with a marked section representing the range of the bigger graph.

degradation of cDNA. Because of the observed instability, we avoided longer storage or frequent thawing of cDNA samples. Considering our findings, we would recommend a "one-tube" RT-PCR, if only a small number of genes is to be analyzed. However, for experiments like ours, where a large number of different genes is examined, it is more practical to separate the RT and the PCR steps ("two-tube RT-PCR"). Here, reverse transcription was done with oligo-dT primers so that cDNA represents the whole transcriptome. In this way, the cDNA can be used for all examined genes and RNA does not have to be thawed for every quantification.

In studies where dilutions of total RNA are used to generate the standard curve, often PCR efficiencies above 2 (or slopes lower than −3.32) are reported.[3–5] This suggests that low-concentrated mRNAs are proportionally better reverse transcribed than mRNAs with high concentrations and that standard curves from mRNA dilutions are not linear. To obtain RNA

TABLE 1

Reproducibility of LightCycler RT-PCR

Initial copies of cDNA per PCR	Variation (n = 3, based on copy numbers) within:		
	Capillaries of one run (%)	Runs on different days (%)	Reverse transcriptions (%)
10^1–10^2	14.56	20.60	29.07
10^2–10^3	10.14	15.77	28.66
10^3–10^4	8.94	11.64	13.96

dilutions in which the total amount of transcribed mRNA molecules is the same, the dilution might be done in RNA from another species whose mRNA sequences do not interfere with the cDNA sequence of interest.

The melting curve analysis is an exact and fast method to check PCR specificity. However, in some cases, PCR products of different lengths have the same melting temperature and therefore are undistinguishable by melting curve analysis, as shown here by the example of two variants of the *GSTM1* gene. Probably the missing exon in the short *GSTM1* variant does not have a great impact on the temperature-dependent stability of the PCR product. These results demonstrate that melting curve analysis should not be used exclusively for the evaluation of PCR products to avoid overlooking unspecific PCR byproducts.

On the basis of a large set of data, an equation was set up that described the correlation between CP and the efficiency of a specific PCR in the LightCycler. This equation assumes that the number of PCR amplificates at the CP (calculated by the second derivative maximum method) is always similar. The theoretical data obtained with this equation correlated well with the empirical data (Fig. 3). With this equation, the absolute number of template molecules can be estimated in the absence of standards with known concentration.

ACKNOWLEDGMENTS

The authors thank Katja Müller and Manfred Wirth for the friendly support with the use of the LightCycler system (GBF, Braunschweig, Germany).

REFERENCES

1. Wilkening S, Stahl F, Bader A. Comparison of primary human hepatocytes and hepatoma cell line hepg2 with regard to their biotransformation properties. *Drug Metab Dispos* 2003;31**:**1035–42.
2. Rasmussen R. Quantification on the LightCycler. In Meuer S, Wittwer C, and Nakagawara K (eds): *Rapid Cycle Real-time PCR, Methods and Applications*. Springer Press, Heidelberg, 2001:21–34.
3. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001;29**:**E45–E45.
4. Bustin SA, Gyselman VG, Williams NS, et al. Detection of cytokeratins 19/20 and guanyl cyclase C in peripheral blood of colorectal cancer patients. *Br J Cancer* 1999;79**:**1813–20.
5. Lekanne Deprez RH, Fijnvandraat AC, Ruijter JM, et al. Sensitivity and accuracy of quantitative real-time polymerase chain reaction using SYBR green I depends on cDNA synthesis conditions. *Anal Biochem* 2002;307**:**63–9.