Ct shift: a novel and accurate real-time PCR quantification model for direct comparison of different nucleic acid sequences and its application for transposon quantifications

Orsolya Kolacsek [a], Enikő Pergel [a], Nóra Varga [a], Ágota Apáti [a] and Tamás I. Orbán [a,*]

[a] *Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary*

***Correspondence to:**
Tamás I. Orbán
Institute of Enzymology, Research Centre for Natural Sciences,
Hungarian Academy of Sciences
1117 Budapest, Magyar Tudósok krt. 2.
Hungary
Phone: +36 1 382 6638
Fax: +36 1 382 6295
Email: orban.tamas@ttk.mta.hu

Keywords: copy number; *PGBD3*; *Sleeping Beauty*; promoter activity; quantitative PCR; splice variant

Abbreviations: qPCR, quantitative polymerase chain reaction; **Ct**, threshold cycle; **ΔΔCt**, comparative Ct; **RT-PCR**, reverse transcription PCR; ***PGBD3***, PiggyBac derived transposon family member 3; **CS**, Cockayne syndrome; ***CSB***, Cockayne syndrome B; **SB**, Sleeping Beauty; ***Rcor1***, REST corepressor 1; **EGFP**, enhanced green fluorescent protein; **IRDR-L**, inverted repeat direct repeat – left; **IRDR-R**, inverted repeat direct repeat – right; **plst**, plasmid standard; ***RPPH1***, Ribonuclease P RNA component H1; **Hsp90**, heat shock protein 90; **cds**, coding sequence; **RQ**, relative quantity.

Abstract

There are numerous applications of quantitative PCR for both diagnostic and basic research. As in many other techniques the basis of quantification is that comparisons are made between different (unknown and known or reference) specimens of the same entity. When the aim is to compare real quantities of different species in samples, one cannot escape their separate precise absolute quantification. We have established a simple and reliable method for this purpose (Ct shift method) which combines the absolute and the relative approach. It requires a plasmid standard containing both sequences of amplicons to be compared (e.g. the target of interest and the endogenous control). It can serve as a reference sample with equal copies of templates for both targets. Using the ΔΔCt formula we can quantify the exact ratio of the two templates in each unknown sample. The Ct shift method has been successfully applied for transposon gene copy measurements, as well as for comparison of different mRNAs in cDNA samples. This study provides the proof of concept and introduces some potential applications of the method; the absolute nature of results even without the need for real reference samples can contribute to the universality of the method and comparability of different studies.

1. Introduction

Use of real-time quantitative polymerase chain reaction (qPCR) amplifying DNA molecules has become a routine tool in molecular biology to study the abundance of nucleic acid sequences in different samples. While monitoring reaction products in each cycle the basis of quantification is that during the log phase of the reaction, the signal of the product is proportional to the starting material (Livak et al., 1995; Morrison et al., 1998; Freeman et al., 1999). However, in spite of few trials the exact quantities cannot be defined well enough empirically (Liu and Saint, 2002; Ramakers et al., 2003; Schefe et al., 2006); the only thing that can be supposed is that the proportion is probably constant for the same PCR target between different samples (Livak and Schmittgen, 2001). This is why we can make comparisons only between same PCR targets measured from different samples. On this basis two types of quantification approaches are possible: relative (Livak and Schmittgen, 2001; Pfaffl, 2001) and absolute (Morrison et al., 1998; Bustin, 2000) quantifications. Relative quantification is doubly relative because it measures relative content of a target sequence comparing to that in a reference sample, while the ratio of the target is normalized to the ratio of the endogenous control (comparative Ct or ΔΔCt formula). Since the abundance of latter sequence is constant among samples, relative quantification does not require precise quantity input measurements before setting up the qPCR. To investigate changes in gene expression or relative abundance of an mRNA between different types of tissues or cells, relative quantification is easy and adequate for most purposes. However, using relative approach for gene copy quantifications, a reference sample with known copy number is essential. Absolute quantification is based on a calibration curve of an artificial sequence (most often a plasmid) that allows to calculate the exact copy of the target in the starting quantities of dilutions. Ct value of the unknown is compared to the standard curve and copy number is calculated by its

Ct value from the curve. Contrary to relative approach, absolute quantification requires precise quantity input measurements not only for the standard but for all of the samples to be examined.

If one would like to match copy numbers of different sequences in the same specimen – e.g. expression levels of RNA products belonging to the same cellular or genetic pathway – the adequate method is absolute quantification for all sequences. The most frequent and simple standard is a plasmid containing the qPCR target sequence. For the construction of this plasmid, it is sufficient to clone the regular PCR amplicon into any plasmid. Absolute quantification has to be made separately for all sequences that we want to compare. In this case all the errors are added up in the resulting comparison. The best way to minimize the variability between two particular absolute quantifications is to measure them simultaneously using a single plasmid that contains both target sequences to be compared, thus errors of quantity inputs and dilution steps will not vary between them. In spite of this clear advantage, doubly targeted plasmid solution is rarely chosen, e.g. in gene copy number quantification projects, and if so it is used for determining the dilution curves from the same plasmid aliquot (Charrier et al., 2011). The requirement of this unique plasmid standard provided the idea of Ct shift method: since the application of such plasmid *ab ovo* guarantees the determination of correct relation of sequences, we presumed it can be used as a reference for quantifications by the ΔΔCt method.

We applied the Ct shift method successfully for gene copy number measurement purposes, as well as for comparing the abundance of different mRNA splicing variants in cDNA samples. In these scenarios, the reference plasmid contains the target of interest, as well as the sequence of a reference gene (often called the endogenous control), the abundance of which is presumably constant in all samples. This plasmid can be applied either for absolute quantification in parallel standard curve experiments, or as a single calibrator sample containing one copy of the target amplicon and one copy of the endogenous control amplicon. In this way the plasmid standard can serve as a reference sample with equal copy number of templates for the two sequences, and the Ct difference of the amplicons is denoted as the Ct shift parameter. With the use of the ΔΔCt formula referring to the plasmid standard, we can quantify the exact ratio of the two sequences in each sample. An important advantage of the Ct shift method is that similarly to relative quantification, it does not necessarily require precise quantity inputs when setting up qPCR.

As a proof of concept, we started using the Ct shift method to measure *Sleeping Beauty* (SB) transposon copy numbers in transgenic cell cultures and in transgenic animals. Few years ago we developed diagnostic copy number assays for such purposes which required tedious preliminary work of establishing reference samples with known transposon copy numbers (Kolacsek et al., 2011). Applying the Ct shift method we provided evidence that this approach is applicable not only for the routine copy number determinations but can also be easily and reliably extended to new copy number projects, even without the laborious pilot experiments to generate reference samples with known copies.

We also show the advantage of the Ct shift method in gene expression studies. Cryptic transcription start sites and alternative splicing often results in several mRNA species produced from a single gene and the ratio among the mRNA variants is certainly crucial for the function of the gene. The Ct shift method offers a good solution for monitoring the levels

of these gene products, and this could be applied for splice variants of certain domesticated transposase genes whose cellular functions have only been partially revealed (Sinzelle et al., 2009) or have not yet been elucidated (Alzohairy et al., 2013; Vogt et al., 2013). Some members of the PiggyBac transposase family express themselves by a gene trapping strategy, and in the absence of an own promoter they can only be transcribed as an alternative mRNA fusion product from the host (Newman et al., 2008). *PGBD3*, one of the PiggyBac derived transposons in the human genome is abundantly expressed in such a fusion form with Cockayne syndrome B (CSB) protein (http://biogps.org/). CSB protein has functions in transcription-coupled nucleotide excision repair (Sarker et al., 2005; Laine and Egly, 2006) and chromatin remodeling (Newman et al., 2006), and mutations of the *CSB* gene are responsible for 70% of CS disease (Laugel et al., 2010). *PGBD3* has been conserved in the 5th intron of *CSB* gene in Primates, and alternative splicing produces a protein fusion product with the N-terminal part of CSB protein (Newman et al., 2008). In addition, using a semi-quantitative method, the authors showed that a cryptic transcript is also expressed starting at the end of exon 5 (Newman et al., 2008). The role of the PGBD3 transposon as part of these transcripts is not known but it is probably beneficial for normal cell physiology. Our first aim in this topic was to reveal the ratio of these mRNA variants and whether this ratio is under cell type specific regulations. In this study while interpreting our results by the Ct shift model we have also verified the reliability of the method also in splicing pattern determinations.

2. Materials and methods

2.1. Cell lines and culturing

Human embryonic kidney cells (HEK-293) and HeLa cells were maintained as described previously (Kolacsek et al., 2014a); the HUES9 embryonic stem cell line was cultured as described earlier (Apati et al., 2008). The establishment and maintenance of the MSCL-2 mesenchymal-like cell line was described in detail previously (Varga et al., 2011).

2.2. Calculation with Ct shift

Ct deviation of two different PCRs – e.g. the target and the endogenous control – with equal input of template we have named as Ct_{shift} and determined by the plasmid standard:

$$\mathbf{Ct_{shift}} = \Delta\mathbf{Ct_P} = \mathbf{Ct_{TP}} - \mathbf{Ct_{EP}} \qquad (1)$$

where P stands for plasmid standard, T for target and E for endogenous control.

In a particular measurement Ct shift is likely a uniform characteristic for all specimens as reflected by all standard points of the plasmid standards (Figure 1A). Thus the value of Ct shift is constant and independent of the plasmid input quantity but only if efficiencies of the target and the endogenous control PCRs are approximately identical which can be tested empirically with the help of the slopes of dilution plots (Kolacsek et al., 2014b). This constant is probably characteristic to an unknown sample containing equal copy of templates.

$$\mathbf{Ct_{shift}} = \mathbf{Ct_{TS}} - \mathbf{Ct_{ES}} \quad (2)$$

where S stands for sample. So with the help of Ct shift we can estimate a theoretical Ct for the target of an unknown sample when it would be present in equal copy number to endogenous control in the specimen.

$$\mathbf{Ct_{T(Theor)S}} = \mathbf{Ct_{ES}} + \mathbf{Ct_{shift}} \quad (3)$$

Now the difference between the real Ct of target and this theoretical Ct can be calculated to determine the fold difference between them, which in fact gives the ratio of target and endogenous control of the unknown.

$$\mathbf{\Delta Ct_{T(Theor)S}} = \mathbf{Ct_{TS}} - \mathbf{Ct_{T(Theor)S}} = \mathbf{Ct_{TS}} - \mathbf{Ct_{ES}} - \mathbf{Ct_{shift}} \quad (4)$$

Fold difference of initial number of the same nucleic acid molecule between reactions can express as an exponential function of the difference of Cts (Livak and Schmittgen, 2001) and we can calculate our theoretical difference with $\mathbf{\Delta Ct_{T(Theor)S}}$.

$$\mathbf{Fold\ difference} = [\mathbf{1} + \mathbf{E}]^{-\Delta \mathbf{Ct\ T(Theor)S}} \quad (5)$$

Where E is PCR-efficiency of the two PCR. When E=1, then:

$$\mathbf{Fold\ difference} = \mathbf{2}^{-(\Delta \mathbf{Ct\ S} - \mathbf{Ct\ shift})} \quad (6)$$

Where: $$\mathbf{\Delta Ct_S} = \mathbf{Ct_{TS}} - \mathbf{Ct_{ES}} \quad (7)$$

2.3. Copy number quantifications of SB transposons

For endogenous controls, genes present in one haploid copy were chosen: the *RPPH1* gene encoding for the H1 RNA subunit of the RNaseP complex for human samples, whereas the *Rcor1* (REST Corepressor 1) gene for rat samples. For construction of plasmid standards, *RPPH1* and *Rcor1* PCR amplicons were gel purified and first cloned into a pGEM®-T vector system (Promega). After verification by sequencing, amplicons were subcloned into a SB-CMV-EGFP transposon containing plasmid separately for *RPPH1* and *Rcor1* (pSB-EGFP-RPPH1, pSB-EGFP-Rcor1). These plasmids can be used as quantification standards since the targeted amplicons, the EGFP transgene, the SB transposon inverted repeat-direct repeat (IRDR) regions, and the control gene are all present in exactly one copy on them. Plasmid standards were linearized with restriction digestion leaving intact amplicon sequences and purified from gel electrophoresis. For absolute quantification measurements input quantity of plasmid was calculated precisely (Kolacsek et al., 2014b), but for the Ct shift method it is sufficient to input them into reactions in a quantity where the endogenous control amplification from the plasmid approximates the amplifications of samples.

Human transgenic cell clones (Kolacsek et al., 2011) and the transgenic rat line (Szebenyi et al., 2015b) were established and genomic DNA was isolated as described earlier.

TaqMan® primer pairs with FAM-MGB probes were also detailed previously (Kolacsek et al., 2011; Kolacsek et al., 2013) with the exception of *Rcor1* for which an Applied Biosystems pre-developed gene expression assay (Rn03810960_s1) was used. Quantitative PCRs were performed on a real-time PCR platform (StepOne™ or StepOnePlus™; Applied Biosystems) with the following thermal profile: 95°C 10min, 40 cycles of 95°C 15sec and 60°C 1min. 100ng gDNA was inputted into a 20µl reaction volume measured in triplicate for each target, final concentrations of primers and probes were 900 and 250nM, respectively.

2.4. RT-qPCR for gene expression analyses of CSB gene products

RNA variants from the CSB locus were detected by five different amplicons as shown on Figure 3A. For construction of plasmid standards, first the endogenous control PCR amplicon of Hsp90 and the five target PCR amplicons were gel purified, cloned into separate pGEM®-T vectors, and after verification by sequencing, the target amplicons were subcloned into the Hsp90 containing pGHsp90 pGEM® plasmid separately (pGHsp-tr1svPG3, pGHsp-tr1+2svPG3, pGHsp-tr1svCSB, pGHsp-tr1+2svCSB, pGHsp-PG3cds). Hsp90 containing pGHsp90 were digested with Spe I and Hinc II. Spe I and Zra I sites were utilized to cut out target amplicons. Plasmid standards were processed into qPCRs as described in the former section.

Total RNA was isolated from cells in near confluent 6-well plates using RNeasy Plus Mini kit (QIAGEN). 1µg of DNase I (NEB) treated total RNA was reverse transcribed using random oligonucleotides with the High-Capacity cDNA Reverse Transcription Kit (Life Technologies). Ten-fold diluted cDNAs were inputted into qPCRs in 20µl reaction volume containing 50nM of primers and Power SYBR® Green PCR Master Mix (Life Technologies). Reactions were performed in triplicates for each target in the same platform with same heat profile as for copy number measurements except that melting curve profile was added. For gDNA content control, equivalent DNase I digested RNA was tested in PG3cds PCR and no remarkable amplification was detected. All other PCRs were intron spanning. Primers were designed with Primer-BLAST software at NCBI website; the sequences are as follows (see Figure 3A for CSB locations):
Hsp90 forward 5'-TGGATATCCCATTACTCTTTTTGTG, reverse 5'-TTCTTTTTCTTCTTCTTTGTCTTCCT; tr1svPG3 forward 5'-TGAAGAGTCTGAGTATTTCCCCAC, reverse 5'-TGCTTCTATGCTGTCATCTGTCTC; tr1+2svPG3 forward 5'-AGGTCGGAAAGTGGGAAGATAC, reverse primer is the same as for tr1svPG3; tr1svCSB forward primer is the same as for tr1svPG3, reverse 5'-TTCCTCAGAATCGTCCTCCAG; tr1+2svCSB forward primer is the same as for tr1+2svPG3, reverse primer is the same as for tr1svCSB; PG3cds forward 5'-TCATGGGAGGCGTAGACAGA, reverse 5'-CACGTCGACGAAACTCCAGA.

3. Results

3.1. Validation of the Ct shift method: SB transposon copy numbers in transgenic rats and human cell clones

We provide evidence for the reliability of the Ct shift method first by applying it for copy number determination. Gene copies are considered as corpuscular units, which uniformly present in individual cells in case of a cell clone or in the germline of an animal, offering a good candidate to prove that a particular qPCR quantification method is reliable. We have established a transgenic rat line expressing fluorescent calcium sensor protein which contains an SB-CAG-GCaMP2 transposon cassette in a homozygous form (Szebenyi et al., 2015a; Szebenyi et al., 2015b). Phenotypic analysis combined with transgene copy number determination of founders and progenies of subsequent generations enabled an effective and rapid selection of animals, resulting in a stable homozygous line in a short time schedule of crossings. One form of validation of the Ct shift method is whether we can gain the same copy number results arising from different sequences of the transgene. Analyses were done for left- and right transposon repeat sequences (IRDR-L and IRDR-R) with TaqMan® assays developed earlier. Since our EGFP TaqMan® assay can detect the *GCaMP2* fluorescent calcium sensor sequence, this was also analyzed. Plasmid standard was constructed from a transposon containing plasmid, cloning the *Rcor1* TaqMan® amplicon also into this plasmid, serving as the endogenous control as it is present in one copy per haploid genome.

Evidence for the existence of the Ct shift can be seen in Figure 1A. Measurement of serial dilution of the plasmid standard produced good efficiencies for all sequences. However, applying the same volume of an individual plasmid standard for the reactions, it produced different separate Ct values for all of the analyzed sequences, and these Ct differences were consequent for all of the standard points originated from any of the dilutions; in this way, each standard point represents the same copy input of different sequences. These Ct delays occurring from the same copy input of targets and endogenous control were denominated as Ct shifts and were determined by multiple targeted plasmid standards in this quantification model. Ct shift is unique for a particular target and to some extent, for the actual run. Here we applied the measurement on the progeny derived from the crossing of F1 generation transgenic animals having two hemizygous copies of transgene on separate loci. As they descended from different founders, they contained different integration sites, thus the possible copy numbers of their offspring varies between 0 and 4. We have chosen one progeny for all representing copies and used the plasmid standard for parallel absolute quantification to determine the relative copy numbers – target/Rcor1 – precisely. After counting the total copy inputs of samples from the plasmid standard curves for all targets and the endogenous control (Figure 1A), relative copies were calculated for each target (Figure 1B). As clearly seen, the three different targets present on the same transgene produced very similar results and the multiple targeted plasmid standards provided credible relative copies (*Rcor1* control has 2 endogenous haploid copies, so the relative copy of 0.5 means 1 hemizygous copy per cell). While absolute quantification based on only one sequence and DNA quantities of inputs are often error prone, this standard seems to guarantee correct relations of sequences. Figure 1C represents our formerly used routine copy number quantifications: it shows relative quantification of the same samples normalized to the endogenous control plus to a reference rat gDNA with one transgene copy. In addition, plasmid dilution points (plst) – that were previously used for absolute quantification to determine standard curves – were quantified as unknown samples: these plasmid inputs were determined as they were gDNA samples with 2

copies, proving that plasmid standards can serve as an excellent reference sample mimicking the reference gDNAs with known transgene copies. Therefore, plasmid standards could be used as references for the Ct shift method (Figure 1D), providing exactly the same results as previous quantification models (Figure 1B and 1C).

We have also tested the Ct shift quantification model in previously established human cell clones (Kolacsek et al., 2011). We have constructed similar plasmid standards that contained the *RPPH1* human endogenous control amplicon. Using the ΔΔCt method with a one copy reference clone, relative quantification of this plasmid standard again mimicked a 2 copy clone (Figure 2A). When we applied the Ct shift method for the same dataset, it resulted in similar copy numbers for all clones (Figure 2B). Moreover, it produced more balanced results for all targets, especially smoothening the discrepancies in the case of the higher copy number clones (such as G2C3).

3.2. Application of the Ct shift method for RT-qPCR: comparison of the complex human CSB-PGBD3 mRNA variant pool

In RT-qPCR application of the Ct shift method, intra-sample relative gene expression measurements of different targets can be achieved using a single reference gene with unique calibrator plasmids for quantifying each target. All calibrator plasmids should contain the particular target-, as well as the common endogenous control amplicon, thus each target is quantified with separate calculations relative to the same endogenous control. Since all targets will be compared precisely to the same control gene, relations between them will also be just as accurate as the relative quantification measurements referring to even different individual plasmid standards.

Here we investigated the relative abundance of diverse mRNA products expressed from the *CSB* gene in four different human cell lines – HEK293, HeLa, HUES9 and MSCL-2. The latter is a mesenchymal stem cell like cell line which was differentiated *in vitro* in our laboratory from the HUES9 embryonic stem cell line (Varga et al., 2011). Figure 3A shows the exon-intron structure of the CSB locus containing the domesticated *PGBD3* transposase gene, and also depicts primer pairs used to differentiate various transcripts and splice variants. We have chosen Hsp90 for endogenous control as it is proven that even the confluence of cells does not alter expression of this gene (Greer et al., 2010). PCR products of the different variants have been cloned into calibrator vectors also containing the *Hsp90* amplicon, to produce individual plasmid standards for each target. As shown previously for copy number quantifications, one proof for reliability is if different assays for the same target produce similar results. With the use of the Ct shift method, this was exactly the case in our RT-qPCR application shown in Figure 3B. cDNAs were tested with two different primer pairings detecting all mRNAs containing *PGBD3* coding sequence. The assays produced very similar results for all cell types though they are targeting different parts of the same mRNA, providing further evidence for the reliability of the Ct shift method (Figure 3B).

The Ct shift method is directly suitable for comparison of normal- and fusion splice variants (svCSB and svPG3) derived from the canonical (tr1) transcription start site (tr1svPG3 vs tr1svCSB, Figure 3C), because these primer pairs detect solely the canonical transcripts. Quantifying their levels revealed that they are expressed at similar levels only in two cell

lines, Hela and MSCL-2. However, when the sum of the canonical and the cryptic promoter activities are measured, higher splice variant expression levels suggest significant cryptic promoter activity (tr1svPG3 vs tr1+2svPG3 and tr1svCSB vs tr1+2svCSB, Figure 3C). These variants (tr1+2) showed similar pattern in cell types such as canonical splice variants but higher levels in MSCL-2 cell line (compare tr1+2svPG3 and tr1+2svCSB, Figure 3C). As signal levels of tr1+2 splice products showed more diversity we have tested the reproducibility of the method using different cDNA inputs in separate measurements (Figure 3D). The data clearly show that regardless of the amount of template applied, the absolute values of relative quantities are reproducible. The measurements also indicate that there is a substantial variation among the expression levels of the PG3 splice variants in the four cell lines examined (Figure 3C and 3D).

Since in our gene expression analyses with the Ct shift method the relative quantities (RQs) are measured in the endogenous control expression unit of *Hsp90*, different RQ signals quantified in the same sample might be subtracted or added to extract more information from the result. Because of the "absolute" nature of resulting signal we can recover the products from the cryptic promoter by simply subtracting the canonical transcript splice variant (tr1 values) levels from the entire transcription activities (tr1+2 values) from Figure 3C (Figure 4A). Utilizing this feature we can reveal real differences between cell types when referring to one of the samples: when normalizing the values to the ones in HEK293 (Figure 4B), the results show that the highest variability is detected in the MSCL-2 cell line. Different calculations also allow us to reflect the activities of the two promoters separately: with the addition of RQs of appropriate splice variants deriving from the same transcript from Figure 4A, total promoter activities can be expressed (Figure 4C). It suggests that in some of the cell types, transcription from the cryptic promoter can be elevated by at least two fold as compared to the canonical promoter activity.

4. Discussion

In this study we provide evidence that the existence of the "Ct shift", featured as the real-time PCR based Ct delay produced from different PCRs of same copy number, can be utilized for reliable quantification measurements. If the molar ratio of two sequences in PCR inputs of the plasmid standard is 1:1, one would expect that Ct values will be the same for both assays, especially when PCR efficiencies are similar and close to 100% for all of them – as proved by the slope of standard curves fitted onto plasmid dilution points. However, Ct values in qPCR measurements will differ in most of sequence cases (see Figure 1A), even with application of the same threshold levels and detection chemistry. Based on our investigations, this deviation shows a partially stochastic nature: Ct shifts represent characteristic values for the particular measurements – this is the basis of our method, but it could show remarkable inter-assay variability. The reason behind the phenomenon is not entirely clear, but the most likely explanation is that the few initial cycles of the latent phase of PCR do not follow the kinetics measured by the reaction efficiency. This kinetic change during cycling may occur due to template change from the initial longer DNA template to the smaller PCR amplicon present in excess in later cycles. Because of these Ct delays, Ct values

of different targets cannot directly be compared, and it is incorrect to represent ΔCt values as a fold difference by these Cts of different PCR target reactions, e.g. to compare the expression levels of different genes even if they were similarly compared to the same control gene.

Using the Ct shift method for three different targets and two different endogenous controls, we have provided evidence that precise gene copy numbers can be accurately determined similarly to traditional quantification models. The calibrator plasmids (containing amplicons of both the target gene and the endogenous control) can serve as excellent references, mimicking the gDNA of real reference samples containing single copies of target and control per haploid genome, and ensuring exact results by the ΔΔCt formula. Using such easy-to-make calibrators, copy number quantification projects do not require the tedious establishment of real reference samples, the copy numbers of which also need to be confirmed with other unrelated techniques. Here we provide evidence that Ct shift can be well characterized by plasmid standards and is also uniformly characteristic for the real samples. It definitely means that it is solely attributable to the particular amplicons independently of sequential or sample environment. So the Ct shift method is perfectly suitable for other qPCR studies regardless of targets and specimens, providing good reaction efficiencies and sample qualities.

Application of the Ct shift method for RT-qPCR studies is exemplified by our measurements of the splice variants from the human CSB-PGBD3 locus in different cell lines. The complexity of the mRNA pool is provided by the alternative splicing of the domesticated *PGBD3* transposase as part of a fusion CSB mRNA product, as well as by a cryptic promoter activity in addition to the canonical transcriptional start site of the *CSB* gene. The reliability of our method was also shown when the same RNA sequence region (PGBD3) was targeted with different PCRs, and the intra- and intersample splice variant patterns could also be reproducibly determined. Because of the "absolute" nature of Ct shift resulting signal, the method also allows to calculate indirectly determined transcript levels by subtraction or addition of different target RQs, in this case differentiating between promoter activities and alternative splicing. Using this approach, we could show remarkable variability of the cryptic promoter usage among cell types (Figure 4C), and the comparable expression of the "host" (CSB) mRNA and the CSB-PGBD3 fusion product (Figure 3C), further supporting the importance of the domesticated PGBD3 expression, despite its as yet undetermined function (Newman et al., 2008). It is clearly seen that higher variance is observed among cell types for the PGBD3 splice variants than normal CSB splice variants; moreover, we can suspect important function of the PGBD3 splice product from the cryptic promoter in the MSCL-2 fibroblast cell line, as its expression is significantly higher than in the other cell lines (Figure 4B). In fact, an autonomous PGBD3 protein product corresponding to this splice variant has been already detected by western blots (Newman et al., 2008), underlining its potential function. Although we also detected a significant level of the "normal" CSB splice product from this cryptic promoter, it is currently unknown if it is translated. The final conclusion of this part of our study is that the newly developed Ct shift method is readily applicable to decipher the fine-tuned complex regulation of mRNA patterns, such as the human CSB-PGBD3 locus, and could help to point toward potential functional studies, in this case examining the regulation of these mRNA variants in fibroblast like cell lines, such as the examined MSCL-2 cell line.

Up to now, several principally different technologies have been developed for nucleic acid quantifications. Some new methodologies, such as next generation sequencing, are powerful tools to analyze large datasets, but despite their applicability for quantitative comparisons, they are not optimal for quantifying small sample sizes, such as determining single transgene copy numbers or splice variant variability of a given locus in a small scale. Another powerful technology is droplet digital PCR which performs absolute quantitation without the assumption of equal primer efficiencies, and in principle it offers high precision in a relatively large dynamic range. Although it seems ideal for the quantitative measurement of nucleic acids in very low concentrations, it is often difficult to optimize the reaction conditions without *a priori* knowledge of the approximate concentration range, not to mention the demand for a relatively expensive machinery (Weaver et al., 2010; Majumdar et al., 2015). Thus, traditional qPCR remains substantial to gain precise results in cases of low throughput analyses, and our newly developed Ct shift method can become a useful contributor in the field of nucleic acid quantification providing a more absolute nature of expression measurements. This model allows direct comparison of target and endogenous control expressing the target level in quantity units of the control, so different measurements or studies become more comparable even for independent targets when same endogenous control is used in these assays. This model combines absolute and relative quantification approaches: similarly to absolute quantification, it uses artificially made standards with known copy number content. However, since these plasmid standards contain both the target and the control amplicons with known copies, they can be directly applied as a reference sample regardless of precise quantity input. With the application of this modified ΔΔCt method we can avoid the traditional absolute quantification measurements, and avoid the tedious establishment of known copy controls, such as for transgene copy number determinations. The Ct shift method might not be limited to the two most frequent quantification forms demonstrated here but may also find potential use in quality control areas or in healthcare, e.g. determination of microbiological or foreign load of any biological material.

Acknowledgement

We would like to thank Kornélia Némethy for excellent technical assistance and Dr. Douglas Melton for originally providing the HUES9 cell line. Tamás I. Orbán is a recipient of the János Bolyai Scholarship of the Hungarian Academy of Sciences. This work was supported by the Hungarian Scientific Research Fund (OTKA) K112112 grant.

References

Alzohairy, A.M., Gyulai, G., Jansen, R.K. and Bahieldin, A., 2013. Transposable elements domesticated and neofunctionalized by eukaryotic genomes. Plasmid 69, 1-15.

Apati, A., Orban, T.I., Varga, N., Nemeth, A., Schamberger, A., Krizsik, V., Erdelyi-Belle, B., Homolya, L., Varady, G., Padanyi, R., Karaszi, E., Kemna, E.W., Nemet, K. and Sarkadi, B., 2008. High level functional expression of the ABCG2 multidrug

transporter in undifferentiated human embryonic stem cells. Biochim Biophys Acta 1778, 2700-9.
Bustin, S.A., 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J Mol Endocrinol 25, 169-93.
Charrier, S., Ferrand, M., Zerbato, M., Precigout, G., Viornery, A., Bucher-Laurent, S., Benkhelifa-Ziyyat, S., Merten, O.W., Perea, J. and Galy, A., 2011. Quantification of lentiviral vector copy numbers in individual hematopoietic colony-forming cells shows vector dose-dependent effects on the frequency and level of transduction. Gene Ther 18, 479-87.
Freeman, W.M., Walker, S.J. and Vrana, K.E., 1999. Quantitative RT-PCR: pitfalls and potential. Biotechniques 26, 112-22, 124-5.
Greer, S., Honeywell, R., Geletu, M., Arulanandam, R. and Raptis, L., 2010. Housekeeping genes; expression levels may change with density of cultured cells. J Immunol Methods 355, 76-9.
Kolacsek, O., Erdei, Z., Apati, A., Sandor, S., Izsvak, Z., Ivics, Z., Sarkadi, B. and Orban, T.I., 2014a. Excision efficiency is not strongly coupled to transgenic rate: cell type-dependent transposition efficiency of sleeping beauty and piggyBac DNA transposons. Hum Gene Ther Methods 25, 241-52.
Kolacsek, O., Izsvák, Z., Ivics, Z., Sarkadi, B. and Orbán, T.I., 2014b. Quantitative analysis of DNA transposon-mediated gene delivery: the Sleeping Beauty system as an example, Genomics III - Methods, Techniques and Applications. iConcept Press Ltd., pp. 97-123.
Kolacsek, O., Krizsik, V., Schamberger, A., Erdei, Z., Apati, A., Varady, G., Mates, L., Izsvak, Z., Ivics, Z., Sarkadi, B. and Orban, T.I., 2011. Reliable transgene-independent method for determining Sleeping Beauty transposon copy numbers. Mobile DNA 2, 5.
Kolacsek, O., Krizsik, V., Schamberger, A., Erdei, Z., Apati, A., Varady, G., Mates, L., Izsvak, Z., Ivics, Z., Sarkadi, B. and Orban, T.I., 2013. Correction: Reliable transgene-independent method for determining sleeping beauty transposon copy numbers. Mob DNA 4, 11.
Laine, J.P. and Egly, J.M., 2006. Initiation of DNA repair mediated by a stalled RNA polymerase IIO. EMBO J 25, 387-97.
Laugel, V., Dalloz, C., Durand, M., Sauvanaud, F., Kristensen, U., Vincent, M.C., Pasquier, L., Odent, S., Cormier-Daire, V., Gener, B., Tobias, E.S., Tolmie, J.L., Martin-Coignard, D., Drouin-Garraud, V., Heron, D., Journel, H., Raffo, E., Vigneron, J., Lyonnet, S., Murday, V., Gubser-Mercati, D., Funalot, B., Brueton, L., Sanchez Del Pozo, J., Munoz, E., Gennery, A.R., Salih, M., Noruzinia, M., Prescott, K., Ramos, L., Stark, Z., Fieggen, K., Chabrol, B., Sarda, P., Edery, P., Bloch-Zupan, A., Fawcett, H., Pham, D., Egly, J.M., Lehmann, A.R., Sarasin, A. and Dollfus, H., 2010. Mutation update for the CSB/ERCC6 and CSA/ERCC8 genes involved in Cockayne syndrome. Hum Mutat 31, 113-26.
Liu, W. and Saint, D.A., 2002. A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics. Anal Biochem 302, 52-9.
Livak, K.J., Flood, S.J., Marmaro, J., Giusti, W. and Deetz, K., 1995. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. Genome Res 4, 357-62.
Livak, K.J. and Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402-8.
Majumdar, N., Wessel, T. and Marks, J., 2015. Digital PCR modeling for maximal sensitivity, dynamic range and measurement precision. PLoS One 10, e0118833.

Morrison, T.B., Weis, J.J. and Wittwer, C.T., 1998. Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. Biotechniques 24, 954-8, 960, 962.

Newman, J.C., Bailey, A.D., Fan, H.Y., Pavelitz, T. and Weiner, A.M., 2008. An abundant evolutionarily conserved CSB-PiggyBac fusion protein expressed in Cockayne syndrome. PLoS Genet 4, e1000031.

Newman, J.C., Bailey, A.D. and Weiner, A.M., 2006. Cockayne syndrome group B protein (CSB) plays a general role in chromatin maintenance and remodeling. Proc Natl Acad Sci U S A 103, 9613-8.

Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29, e45.

Ramakers, C., Ruijter, J.M., Deprez, R.H. and Moorman, A.F., 2003. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett 339, 62-6.

Sarker, A.H., Tsutakawa, S.E., Kostek, S., Ng, C., Shin, D.S., Peris, M., Campeau, E., Tainer, J.A., Nogales, E. and Cooper, P.K., 2005. Recognition of RNA polymerase II and transcription bubbles by XPG, CSB, and TFIIH: insights for transcription-coupled repair and Cockayne Syndrome. Mol Cell 20, 187-98.

Schefe, J.H., Lehmann, K.E., Buschmann, I.R., Unger, T. and Funke-Kaiser, H., 2006. Quantitative real-time RT-PCR data analysis: current concepts and the novel "gene expression's CT difference" formula. J Mol Med (Berl) 84, 901-10.

Sinzelle, L., Izsvak, Z. and Ivics, Z., 2009. Molecular domestication of transposable elements: from detrimental parasites to useful host genes. Cellular and molecular life sciences : CMLS 66, 1073-93.

Szebenyi, K., Furedi, A., Kolacsek, O., Csohany, R., Prokai, A., Kis-Petik, K., Szabo, A., Bosze, Z., Bender, B., Tovari, J., Enyedi, A., Orban, T.I., Apati, A. and Sarkadi, B., 2015a. Visualization of Calcium Dynamics in Kidney Proximal Tubules. J Am Soc Nephrol 26, 2731-40.

Szebenyi, K., Furedi, A., Kolacsek, O., Pergel, E., Bosze, Z., Bender, B., Vajdovich, P., Tovari, J., Homolya, L., Szakacs, G., Heja, L., Enyedi, A., Sarkadi, B., Apati, A. and Orban, T.I., 2015b. Generation of a Homozygous Transgenic Rat Strain Stably Expressing a Calcium Sensor Protein for Direct Examination of Calcium Signaling. Sci Rep 5, 12645.

Varga, N., Vereb, Z., Rajnavolgyi, E., Nemet, K., Uher, F., Sarkadi, B. and Apati, A., 2011. Mesenchymal stem cell like (MSCl) cells generated from human embryonic stem cells support pluripotent cell growth. Biochem Biophys Res Commun 414, 474-80.

Vogt, A., Goldman, A.D., Mochizuki, K. and Landweber, L.F., 2013. Transposon domestication versus mutualism in ciliate genome rearrangements. PLoS Genet 9, e1003659.

Weaver, S., Dube, S., Mir, A., Qin, J., Sun, G., Ramakrishnan, R., Jones, R.C. and Livak, K.J., 2010. Taking qPCR to a higher level: Analysis of CNV reveals the power of high throughput qPCR to enhance quantitative resolution. Methods 50, 271-6.

Figure legends

Figure 1. Validation of the Ct shift method with transgenic copy number quantification of progeny of founder rats established by the Sleeping Beauty transposon based gene delivery. (A) Standard curves of absolute quantification. Red points are plasmid standard dilutions,

blue ones are transgenic rat measurements. Copy inputs estimated from the plasmid quantities are on the X axis, resulting Cts are on the Y axis. Efficiencies calculated from the slopes are: Rcor1 (endogenous control) – 95%, EGFP – 96%, IRDR-L – 96%, IRDR-R – 98%. (B) Relative copy numbers by absolute quantification were calculated from the Ct values of targets and *Rcor1*. *Rcor1* represents 1 copy per haploid genome so 2 copies per cell, 0.5 copy of a target means 1 copy per cell (hemizygous for the locus). Previously determined copy numbers are shown in brackets. (C) Relative quantification of the same dataset with comparative Ct method (ΔΔCt). The one copy F2/4 rat sample was used as a reference sample, the three plasmid standard points (plst) were measured as a 2-copy-gDNA sample. (D) Relative quantification of the same dataset with the Ct shift method using one of the plasmid standard points as a reference sample.
Means of at least 3 independent measurements are shown, error bars represent 95% confidence intervals.

Figure 2. The Ct shift method is applied for copy number quantification of different human cell clones, formerly created by the Sleeping Beauty transposon system. (A) Relative quantification with the ΔΔCt method using a previously established one copy reference clone, and the human *RPPH1* as the endogenous control. Values represent the average results of the three different assays, plasmid standard was measured as a clone with 2 copies. (B) Relative quantification with the Ct shift method using a single plasmid standard (plst) as a reference. A correction factor of 2 was applied to gain copies per cell results.
Means of at least 3 independent measurements are shown, error bars represent 95% confidence intervals.

Figure 3. Applying the Ct shift method for RT-qPCR. (A) Structure of the human *CSB* gene locus containing the domesticated *PGBD3* transposase, and primer pairs used to differentiate splice variants and promoter activities. Arrows depicted as "trx 1" indicates transcription from the canonical gene promoter, whereas "trx 2" shows transcription starting from a cryptic promoter located in exon 5. pA: polyA signal; "TTAA" sequences are the original target site duplications of the domesticated *PGBD3* transposase; Sv stands for splice variant. (B) Validation of the Ct shift method for gene expression quantification in four human cell lines. Results of two different primer pairs detecting separate parts of PGBD3 mRNA. Signal levels of plasmid standards were set as 1. (C) Quantification measurements differentiating all variants expressed from the CSB gene. Tr1 stands for the canonical, whereas tr1+2 for the sum of canonical and cryptic promoter activity. (D) Testing the reproducibility of the Ct shift method. RQs from two separate measurements (/1 and /2) with different cDNA inputs derived from 20ng or 40ng total RNA, respectively.
Representative experiments are shown, error bars indicate mean values ± 95% confidence intervals.

Figure 4. Quantification of canonical and cryptic promoter products of *CSB* gene in four cell lines. (A) Splice variants produced from the canonical and cryptic promoter. RQs of splice variants from the cryptic promoter were calculated by subtracting splice variant levels accounting for solely the canonical promoter activity (tr1) from measurements accounting for

the sum of the appropriate splice variants produced from the two promoters (tr1+2) in Figure 3C. (B) Splice variant RQ values normalized to the levels measured in HEK293 cell line (C) Comparison of individual products of the two promoters in *CSB* gene. It was calculated by the addition of RQs of two possible splice variants produced from the appropriate promoters separately from Figure 4A.
Representative experiments are shown, error bars indicate mean values ± 95% confidence intervals.

Figure 1
Click here to download high resolution image

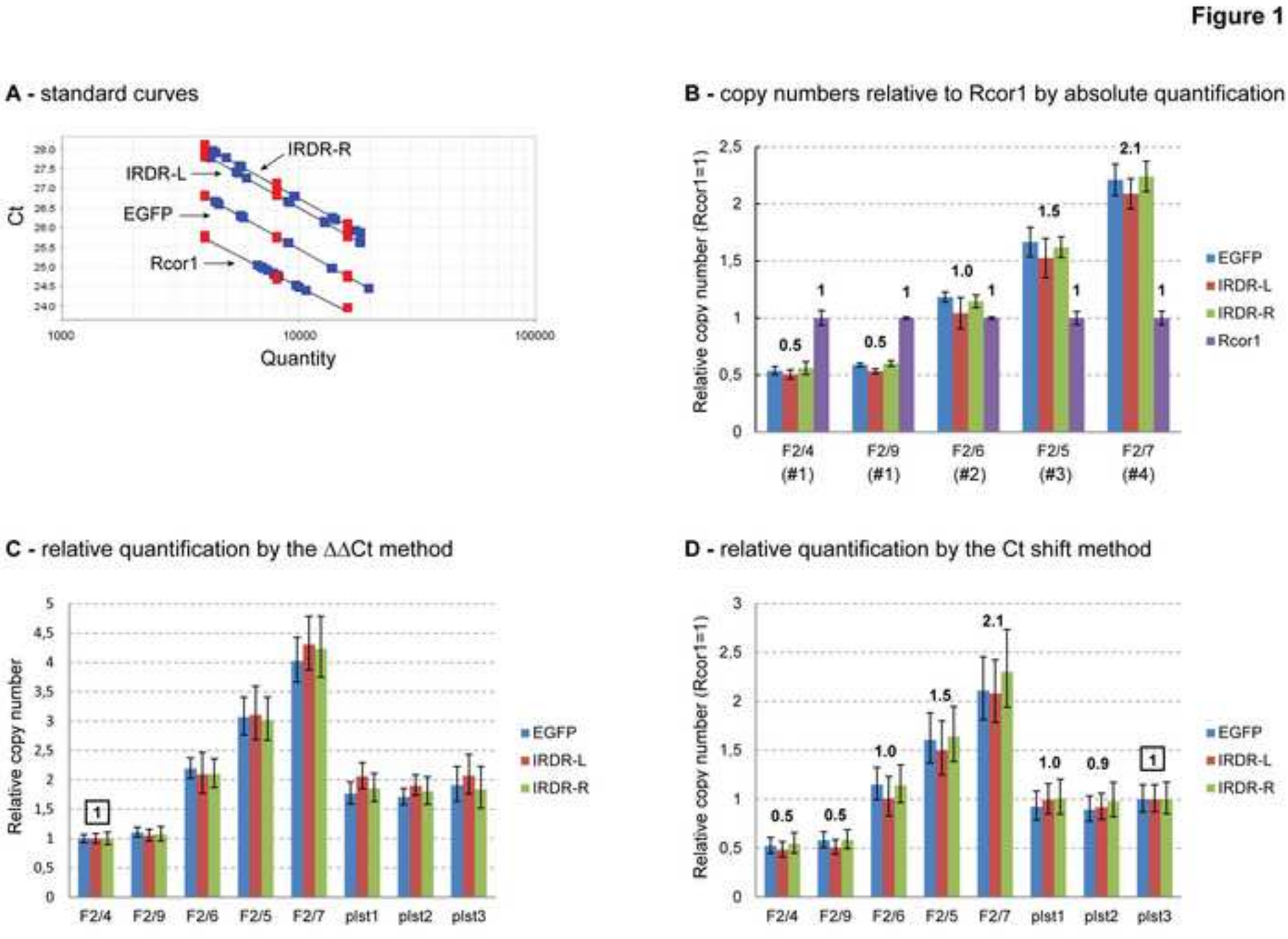

Figure 2

Click here to download high resolution image

Figure 2

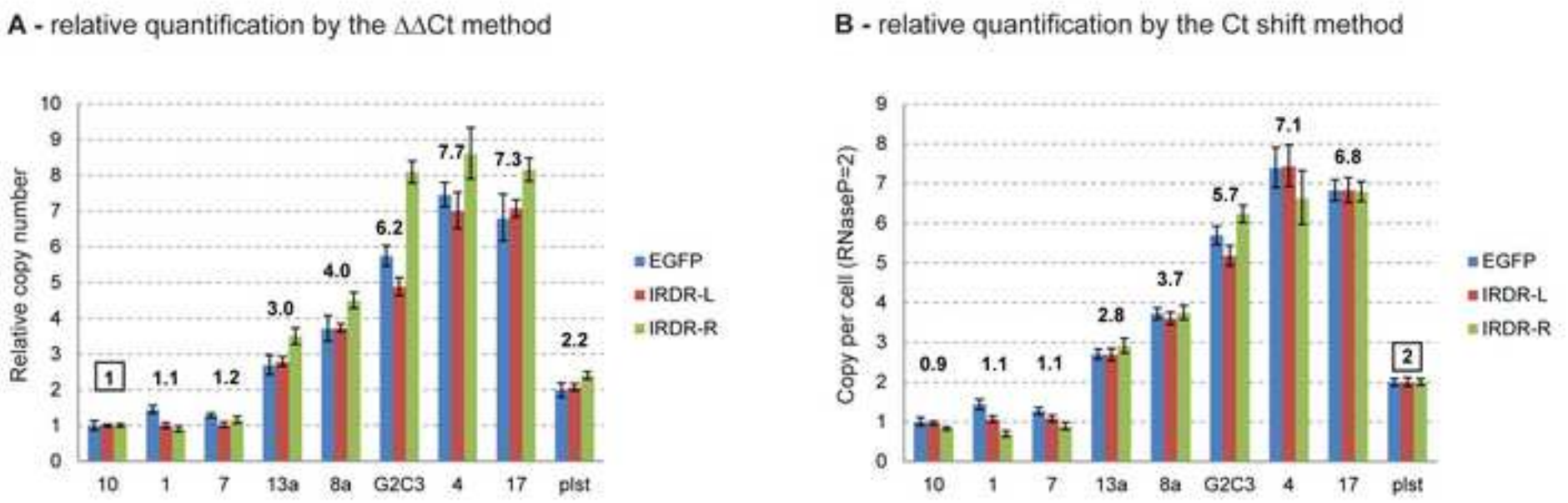

Figure 3

Click here to download high resolution image

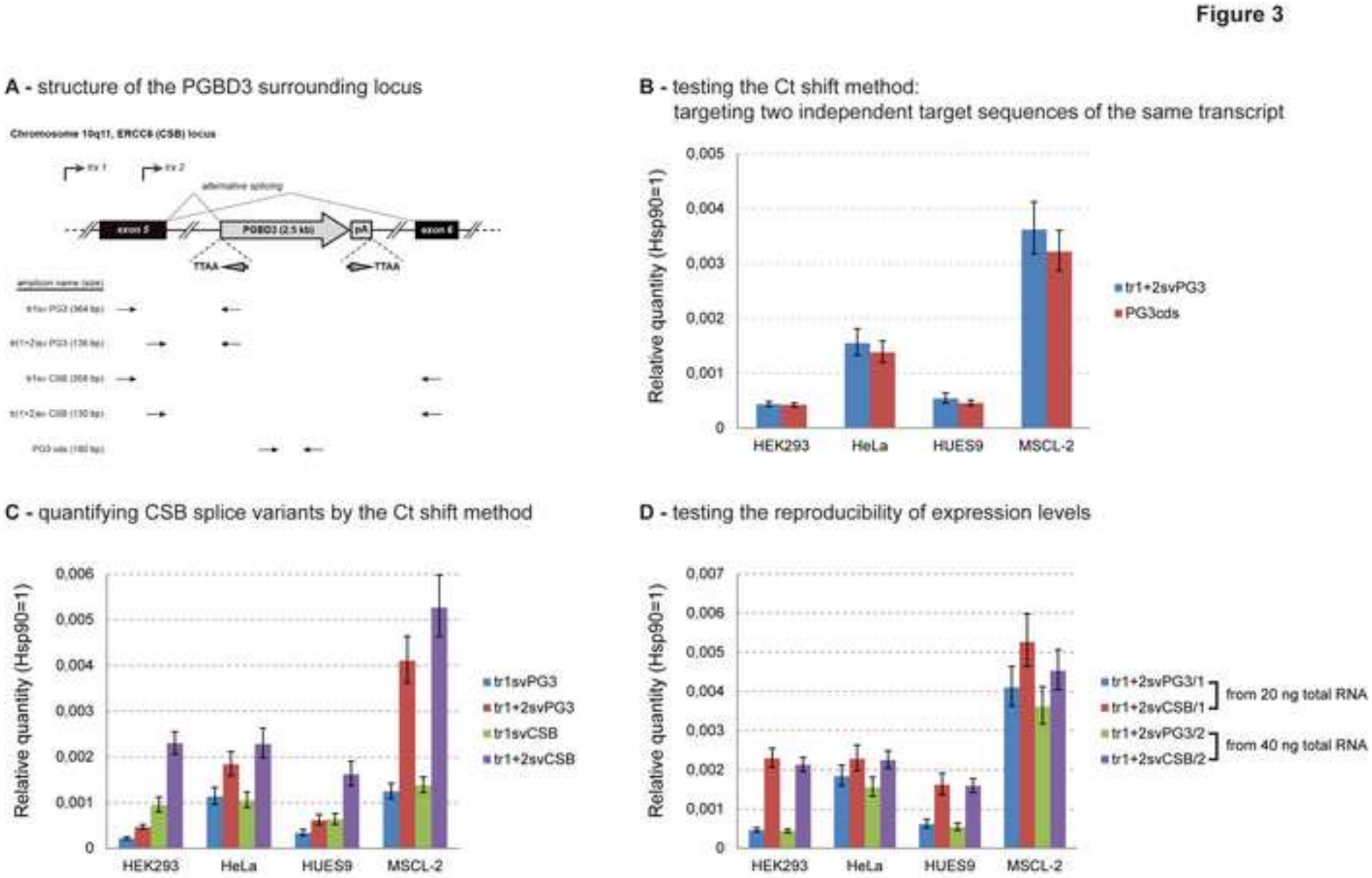

Figure 4

Click here to download high resolution image