

ARTICLE

# Challenges for the Characterization of Genetically Modified Animals by the qPCR Technique in the Era of Genomic Editing

Ribrio Ivan Tavares Pereira Batista<sup>1,2</sup> Dárcio Ítalo Alves Teixeira<sup>1</sup> Vicente José de Figueirêdo Freitas<sup>1</sup> Luciana Magalhães Melo<sup>1,3\*</sup> Joanna Maria Gonçalves Souza-Fabjan<sup>2\*</sup>

1.Laboratório de Fisiologia e Controle da Reprodução, Faculdade de Veterinária, Universidade Estadual do Ceará (UECE), Av. Dr. Silas Munguba 1700, Fortaleza, CE, 60740-903, Brazil.

2.Faculdade de Veterinária, Universidade Federal Fluminense (UFF), Rua Vital Brazil Filho, 64, Niterói, RJ, 24230-340, Brazil.

3.Molecular Genetics Research Unit, University Center Fametro (Unifametro), Rua Conselheiro Estelita, 500, Fortaleza, CE, 60010-260, Brazil.

ARTICLE INFO

*Article history*

Received: 2 February 2021

Accepted: 12 March 2021

Published Online: 16 April 2021

*Keywords:*

Absolute quantification

Copy number

DNA extraction

Relative quantification

Zygosity

ABSTRACT

Characterization of genetically modified organisms through determination of zygosity and transgene integration concerning both copy number and genome site is important for breeding a transgenic line and the use of these organisms in the purpose for which it was obtained. Southern blot, fluorescence *in situ* hybridization or mating are demanding and time-consuming techniques traditionally used in the characterization of transgenic organisms and, with the exception of mating, give ambiguous results. With the emergence of the real-time quantitative PCR technology, different applications have been described for the analysis of transgenic organisms by determination of several parameters to transgenic analysis. However, the accuracy in quantitation by this method can be influenced in all steps of analysis. This review focuses on the aspects that influence pre-analytical steps (DNA extraction and DNA quantification methods), quantification strategies and data analysis in quantification of copy number and zygosity in transgenic animals.

## 1. Introduction

Technologies for adding exogenous genes to animals have made remarkable progress in recent years and now show promising results in a range of strategies, such as large-scale production of proteins of therapeutic interest

<sup>[1]</sup>, production of experimental models for the study of human and animal diseases, zootechnical improvements, regulation of gene expression studies <sup>[2]</sup>. Since the production of the first transgenic mouse in 1981 <sup>[3]</sup>, significant progress has been made in methods of introducing the gene for transgene in animals. The first transgenic

\*Corresponding Author:

Joanna Maria Gonçalves Souza-Fabjan,

Faculdade de Veterinária, Universidade Federal Fluminense (UFF), Rua Vital Brazil Filho, 64, Niterói, RJ, 24230-340, Brazil;

Email: joannavet@gmail.com

Luciana Magalhães Melo,

University Center Fametro (Unifametro), Rua Conselheiro Estelita, 500, Fortaleza, CE, 60010-260, Brazil;

Email: lucianamelo.2005@gmail.com

approaches were unpredictable, unrepeatably<sup>[4]</sup> and, invariably, resulting in the insertion of exogenous DNA into the host genome in several locations and in several copies. In some cases, this led to disruption of the gene's function, ectopic and overexpression or underexpression of the exogenous gene<sup>[5]</sup>. Conventionally, the characterization of these transgenic animals in terms of copy number, integration site and zygosity, was performed using techniques such as Southern blot, fluorescent *in situ* hybridization (FISH) and mating. However, these methods are technically demanding and time consuming. In addition, except for mating, they give mixed results<sup>[6]</sup>. On the other hand, real-time quantitative PCR (qPCR) has proven to be a reliable, fast and accurate method for determining zygosity<sup>[7]</sup> and copy number<sup>[8]</sup> for transgenic animals.

However, with the evolution of genetic engineering, new methods of producing transgenic animals, called Genetic Editing methods such as zinc finger nucleases (ZFN), transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeats (CRISPRs) have been widely used worldwide for genetic modification of animals<sup>[5,9,10,11]</sup>. Several of these techniques can introduce single nucleotide changes without integrating foreign DNA and, thus, generate organisms with desired phenotypes. Consequently, these organisms can be indistinguishable from their natural counterparts, since the modifications can resemble entirely random mutations, regardless of whether they are spontaneous, chemically induced or by irradiation<sup>[11]</sup>. This required updating the techniques of qPCR, DNA sequencing and DNA hybridization to characterize genetically molded organisms produced with the new genomic editing tools. This review focuses on the aspects that influence pre-analytical steps (DNA extraction and DNA quantification methods), quantification strategies and data analysis in quantification of copy number and zygosity in transgenic animals.

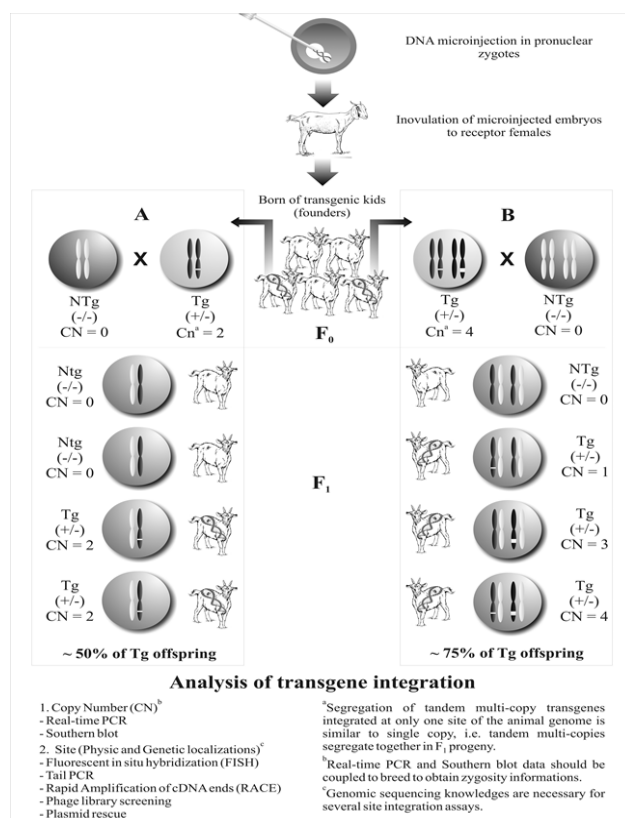
## 2. Analysis of Transgene Integration in Animal Genome

The molecular characterization of the main features that made this a transgenic animal is an essential step. This evaluation aims to characterize the integrity of the inserted sequence in the genome of the specimen, the expression pattern of the transgene to identify its site of insertion (when the transformation is performed by conventional methods), as well as the number of insertions of the expression cassette<sup>[7]</sup>. Molecular characterization of the transgene copy number and the zygosity would

allow inferring the genome receptor stability after gene transformation and the transmission rate of the transgene to generate a F1 (Figure 1).

### 2.1 Copy Number

Transgene copy number, defined as the number of exogenous DNA insert(s) in the genome is a key issue for transgenic studies, since it is directly relevant to the effectiveness of transgenic event and data interpretation<sup>[11]</sup>. Theoretically, a single intact copy may be sufficient to produce the recombinant protein. However, the expression level may be correlated with transgenic gene dosage; a higher copy number may result in increased expression<sup>[12]</sup>. Nevertheless, this observation is not true for all transgenic events, since an exceptionally high copy number may, in fact, result in low expression<sup>[13,14]</sup>. A high copy number may not only affect the expression level but also the genetic stability of the transgenic locus<sup>[15]</sup>, due to failure in the recombination process during meiosis. According to Mahon et al.<sup>[16]</sup>, the insertion of a large number of transgenes in tandem can result in intrachromosomal recombination, deletion, breakage or translocation of the transgenic.



**Figure 1.** Schematic representation of transgenic animal production by pronuclear microinjection, segregation of the transgene in F1 lineage and methods to analyze transgenic integration.

After obtaining and characterizing the founding animal, selective breeding allows the establishment of a stable transgenic line. Normally, when the founder animal has more than one transgene integration site, segregation of the transgenes and the creation of independent strains are recommended<sup>[7]</sup>. However, not always the presence of an intact transgene into the host genome ensures their expression. This is because the expression of a transgene is influenced by its location in genomic DNA, e.g., its position in relation to transcriptional control elements, heterochromatin regions of chromosomes non-transcribed, and other silenced regions<sup>[17]</sup>. This demonstrates the importance of new genomic editing tools in which site-specific modifications are achieved by targeted cleavage of DNA and homologous recombination using ZFN, which are chimeric molecules, composed of a nuclease domain and specifically designed DNA-recognition domains.

The conventional method for transgene copy number determination is Southern blot hybridization. The usage of a restriction enzyme with only one restriction site in the transgene cassette should be chosen to digest the genomic DNA. Thus, the digested nucleic acid will be used for Southern blot hybridization with transgene specific probes. If Southern blot hybridization renders only one band, the transgene copy number should be one<sup>[18-20]</sup>. However, Southern blot hybridization requires a relatively large amount of DNA, labor-intensive and time-consuming<sup>[20]</sup>. Additionally, Southern blot analysis may not be accurate enough to determine copy numbers greater than two<sup>[9-11]</sup>.

The polymerase chain reaction (PCR) is one of the most sensitive methods for detecting the integrated gene in the transgenic animal genome, and thus it can reduce the amount of DNA required for analysis<sup>[9,10]</sup>. Quantitative PCR has been successfully used to determine the copy number of genes into the genome of many species<sup>[8]</sup>. In addition, to improve the accuracy of real-time PCR for this application, the most used methods are the external standard curve-based method and the  $\Delta C_T$  method involving an internal reference gene (Table 1). Ballester et al.<sup>[8]</sup> described a rapid and accurate qPCR-based system to determine transgene copy number in transgenic animals. The authors used the  $2^{-\Delta\Delta CT}$  methods to analyze several mouse lines carrying a goat  $\beta$ -lactoglobulin transgene without the requirement of a control sample previously determined by Southern blot analysis. Instead of a murine DNA, the calibrator was a goat genomic DNA, which was used to amplify both  $\beta$ -lactoglobulin (target) and glucagon (reference) genes. Chandler et al.<sup>[21]</sup> also used the  $2^{-\Delta\Delta CT}$  methods to estimate bacterial ar-

tificial chromosomes (BAC) transgene copy number in mice embryos and lines. They observed accuracy and reproducibility in copy number quantification in several of independent transgenic lines and showed that increased BAC transgene copy number is correlated with increased BAC transgene expression. To determine the correlation of transgene expression with copy number, Kong et al.<sup>[22]</sup> examined the green fluorescent protein (GFP) copy number in ears of newborn and mature transgenic pigs. Interestingly, a decline in copy number was found by both absolute quantitative real-time PCR and Southern blot analysis. In addition, the authors observed a significant correlation between GFP expression levels and copy number in transgenic fibroblast cells. The authors hypothesized that the decline of transgene expression may be due to the loss of copies.

## 2.2 Zygoty

The zygoty of transgenic animals describes the similarity or dissimilarity of the transgene insertion of homologous chromosomes in a specific allelic position. The term is used to describe homozygous event of double insertion of the transgene in the same position of allelic chromosomes, whereas hemizygote describes the insertion position in a single specific allele. Based on the Mendelian inheritance, the transgene transmission to F1 is considered 100% when the animal is homozygous and 50% when hemizygote. However, animals in hemizygotes with multiple insertion position different allelic can produce the rate of 75% transmission of the transgene (Figure 1). This information is always required for effective breeding and colony maintenance. In addition, heterozygous or homozygous status for the transgene has been shown to correlate with gene expression levels and rates transmission of transgene<sup>[7]</sup>. When integration takes place at a single location in the genome, in the cell embryo, without compromising the genes involved in gametogenesis that can induce transmission rate distortion, half of the F1 offspring will be transgenic<sup>[23]</sup>. Normally, the interaction occurs at several different locations in the genome, which may be on the same chromosome or on different chromosomes<sup>[23]</sup>. When two or more interactions occur on the same chromosome, the rate of segregation of the transgenes or the frequency of recombination during meiotics will be determined by the distance between the integration sites<sup>[23]</sup>.

Results from the zygoty analysis of transgenic animals using qPCR are listed in Table 1. Comparative analysis of zygoty between qPCR and Southern blot analysis in 45 transgenic rats for the human decay-accelerating factor showed ambiguous results when Southern

blot technique was applied. However, this same analysis using qPCR permitted the clear identification of all transgenic animals as homozygous or heterozygous. Mating of homozygous and heterozygous animals, defined by qPCR, could show transgene transmission to the offspring as expected by Mendelian laws [7]. Shitara et al. [24] developed the system to determine the zygosity using only two experimental processes: estimation of the concentration of DNA and SYBR Green PCR analysis. With this method, the authors successfully discriminated homozygous, heterozygous and non-transgenic animals. In this same work, to confirm the accuracy of zygosity determination by this method, blastocysts obtained from superovulated female mice, which had been mated to male mice of the transgenic strains by *in vitro* fertilization and the green fluorescent protein fluorescence (EGFP) was then visualized under an inverted fluorescence microscope. All embryos (70/70) derived from the F1 male, determined to be homozygous by real-time quantitative PCR, showed EGFP. In embryos derived also from the F1 male, which were determined to be heterozygous, about 55% of these embryos (49/89) showed fluorescence. As expected, these results were per-

fectly consistent when the used method was qPCR.

### 3. Real Time Quantitative PCR (qPCR)

The qPCR features such as use of low amounts of template DNA and high specificity due to the high temperature annealing of the primers, compared to low hybridization specificity of the probes in the Southern blot, made this technique a powerful tool in the characterization of transgenic animals. [25]. The high sensitivity is conferred by the exponential nature of the PCR reactions, which enable specific sequences to be detected in samples even if only a few copies are present. The procedure for copy number and zygosity analysis in transgenic animals using qPCR technique can be divided into three steps: pre-qPCR procedures (such as DNA extraction and quantification), qPCR amplifications and post-qPCR procedures (mathematical and statistical data analysis) [26]. Concerning qPCR amplifications, an important issue is the choice of the quantification strategy (with acceptable specificity and sensitivity), which should determine the accuracy of the measurement. The following topics describe these steps

**Table 1.** Technical details in DNA extraction, quantification, and real-time PCR for copy number and zygosity analysis of transgenic animals.

	Copy number				Zygosity			
	Joshi et al. [45]	Ballester et al. [8]	Chandler et al. [21]	Kong et al. [22]	Haurogné et al. [19]	Shitara et al. [24]	Tesson et al. [9]	Ji et al. [48]
Species	Mus musculus	Mus musculus	Mus musculus	Sus domesticus	Mus musculus	Mus musculus	Mus musculus	Danio rerio
Tissue for DNAg extraction	liver and lung	liver	tail	Fibroblast cells	tail	tail	tail	Tail
DNAg extraction method	P.C.	P.C.	P.C.	Universal Genomic DNA Extraction Kit Ver.3.0	phenol-chloroform protocol	phenol-chloroform protocol	phenol-chloroform protocol	DNeasy Tissue Kit (Qiagen)
Quantification method	UV spectrophotometry at 260 nm	N.E.	UV spectrophotometry at 260 nm	--	UV spectrophotometry at 260 nm	UV spectrophotometry at 260 nm	UV spectrophotometry at 260 nm	UV spectrophotometry at 260 nm
Reference gene	$\beta$ -actin	Glucagon	<i>Jun</i> gene	TFRC	Mouse gap junction channel protein alpha 5	--	Rat hypoxanthine phosphoribosyltransferase (rHPRT)	Gene bank AC087105
qPCR chemistry	SYBR Green	TaqMan	TaqMan	SYBR Green	GeneAmp 5700 SYBR Green	SDS 7900 CYBR Green	SDS7700 TaqMan	SDS7700 TaqMan
Normalization	DNA quantification	Endogenous control	DNA quantification	DNA quantification	Endogenous control	Endogenous control	Endogenous control	Endogenous control
Quantification	Absolute	2 <sup>-<math>\Delta\Delta C_t</math></sup>	Absolute	Absolute	2 <sup>-<math>\Delta\Delta C_t</math></sup>	2 <sup>-<math>\Delta\Delta C_t</math></sup>	2 <sup>-<math>\Delta\Delta C_t</math></sup>	Relative standard curve
Calibrator	N.A.	Mus musculus	N.A.	N.A.	Animal of known zygosity	Animal of known zygosity	Animal of known zygosity	Animal of known zygosity
Quantification method of calibrator	N.A.	Southern blot	N.A.	N.A.	--	FISH	Southern blot	FISH

N. A. = Not applicable. P.C.= Protocol based on phenol-chloroform method. P. DNAg extraction kit based on spin technology (TaKaRa).



and the effects in final qPCR accuracy.

### 3.1 Pre-qPCR Procedures – DNA Extraction

The aim of a nucleic acid extraction method is to isolate DNA of suitable integrity, purity and of sufficient quantity for diagnostic applications by qPCR [27]. Obtaining DNA of high quality is paramount for ensuring confidence in all subsequent steps in the process of generating analytical measurements. However, the quality of the template DNA obtained can vary according to the extraction method used, thereby influencing the accuracy of the quantification.

The basic phenol/chloroform extraction buffer consists of 100 mM Tris-HCl pH 8.0, 5 mM EDTA pH 8.0, 200 mM NaCl, 0.2% sodium dodecyl sulfate (SDS) and 0.1 mg/mL proteinase K (molecular biology grade). The phenol/chloroform/isoamyl alcohol (25:24:1) mix is used to remove proteins and polysaccharides and ammonium acetate and ethanol to DNA precipitation [7]. For real-time PCR use, RNA is removed from the nucleic acid preparation with enzymes, as RNase A and RNase T1. However, the prior work of our group demonstrated that this DNA extraction method compromises the efficiency of qPCR [28]. According to published reports [29, 30], phenol/chloroform extracted DNA needs further purification to be used for real-time PCR. The purification of phenol/chloroform extracted DNA with a Genomic Tip 20 column (Qiagen) resulted in a linear calibration curve and produced the expected values [30]. Nevertheless, some studies have succeeded to perform the copy number quantifications and zygosity analysis using this method without purification provided for amplification by real-time PCR [9,10]. Additionally, Sakurai et al. [31] succeeded in the analysis of zygosity in transgenic animals using crude extract obtained from incubation of samples from different tissues of transgenic mice. These controversial results can be attributed to variation in the tissue type used for gDNA extraction.

Two main factors that compromise PCR amplification are: i) the quality of template DNA in the reaction and ii) the presence of a series of inhibitors (Table 2). The presence of inhibitors in the PCR reaction compromises not only the efficiency of the reaction, but also the reproducibility of the PCR, thus contributing to inaccurate qPCR results. The inhibitory mode of action of some of these compounds may be linked with precipitation and denaturation of DNA or the ability of the polymerase enzyme to bind to magnesium ions [32]. Animal tissue or reagents used in the DNA extraction stage, inhibitors generally modify the kinetics of the PCR reaction, chelating Mg<sup>2+</sup> (a cofactor of DNA polymerases) and / or by binding to template DNA or DNA polymerase [33,34,35]. To overcome this

limitation, commercial kits DNA binding to silica-based matrices, followed by elution, can be used to remove inhibitors and organic solvents, such as chloroform. According to Burkhart et al. [34] and our experiments [28], gDNA obtained by silica matrix-based methods are more efficient for amplification by qPCR. An inhibition test using either internal controls or evaluation of the linearity of the calibration curves should be performed to determine the suitability of the extracted DNA for real-time PCR amplification [28, 30].

**Table 2.** Examples of PCR inhibitors reported and methods to minimize inhibition.

Inhibitors	Description and inhibitory concentration	Methods to minimize inhibition
EDTA	≥ 0.5 mM 1 mM [31]	Reduce the concentration of EDTA to 0.1 mM in the TE buffer or simply use Tris-HCl (10 mM) to bring DNA in solution. DNA can also be brought in pure water (but the DNA cannot be stored for long-term use)
Ethanol	>1% (v/v) [53]	Dry pellet and resuspend
Isopropanol		Dry pellet and resuspend [26]
Proteinase K	1% casein hydrolysate in PCR mixture caused inhibition [31] Proteinase K [33]	Use SDS or guanidinium buffers, proteinase K Non proteinase K based genomic DNA isolation method [33]
Detergents	SDS [53]	Wash with 70% ethanol
Sodium acetate	≥5 mM [53]	Wash with 70% ethanol
Sodium chloride	≥25 mM [54]	Wash with 70% ethanol or use silica-based purification [26]

### 3.2 Pre-qPCR Procedures – DNA Quantification

Prior to qPCR, stock DNA extracts are commonly quantified and diluted so that all reference and test samples contain identical amounts of this nucleic acid. An accurate determination of DNA concentration in a sample is a critical component for analysis of copy number and zygosity by qPCR. In general, DNA quantification prior to qPCR increases confidence in negative PCR results, where insufficient target DNA could otherwise be interpreted as a false-negative. Two principal methods used for DNA quantification are UV spectrophotometry (absorbance 260 nm – A260) and fluorometry (Table 2). However, the spectrophotometric methods of quantifying DNA do not discriminate intact DNA, RNA and free nucleotides, [36] on the other hand, fluorescent dyes, are highly specific for double-stranded DNA. [37] According to Shokere et al. [36] A260 and fluorescent-dye methods of quantifying intact genomic DNA provide relatively concordant DNA quantification values. However, the quantification values differ significantly for an identical DNA extract that has been

degraded with its non-degraded counterpart. This study revealed that A260 values overestimate by an average of 20.3% ( $\pm 6.1$ ) and fluorescent–dye methods underestimate by an average of 145.8% ( $\pm 6.0$ ). The DNA concentration of PCR-amplifiable intact DNA extracts. Furthermore, when fluorescent–dye methods of DNA quantification were compared with A260 methods, an average percent difference of 10.1% ( $\pm 6.3$ ) was reported for intact genomic DNA, but a much more significant percent difference of 152% ( $\pm 10.3$ ) was reported in degraded genomic DNA. [36]

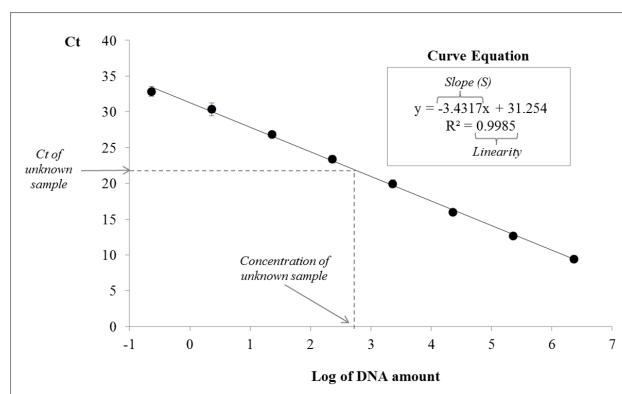
### 3.3 qPCR Amplifications – Quantitative Strategies

Target nucleic acids can be quantified using either absolute or relative quantification. The absolute quantification determines the absolute amount of target, whereas relative quantification determines the ratio between the gene target amounts in two samples (an unknown and another previous quantified, named calibrator) [38]. For relative strategies of quantification, in general, is necessary to amplify an endogenous reference gene (usually an appropriate housekeeping gene) in the two samples of DNA, to normalize the data [39]. All qPCR methods demand to plot standard curves for mathematical validation of gene amplifications. A standard curve is generated using a dilution series of at least five different concentrations of the DNA template. [40] The most important parameters calculated with the plots are: 1) linearity (Pearson correlation coefficient,  $R^2$ ) – must be greater than 0.96 and it is also important to make PCR reactions with DNA amounts that are within the linear range of amplification. [28] 2) Efficiency (E) – must be close or equal to 1.0 (ideal values are between 1.1 and 0.9). [41] When all these requirements are fulfilled, both methods can be successfully used to estimate the number of copies or zygosity in transgenic animal. The work of our group demonstrated that high error rate ranging from 11-177% in absolute quantification, when these requirements are not met [28].

#### 3.3.1 Absolute Quantification

Absolute quantification (Figure 2) can be achieved by a relation of the  $C_T$  measurement to a standard curve that can be obtained by diluting a standard DNA sample (as a plasmid) with the transgene sequence for which the exact DNA concentration and molecular weight is known [42]. The  $C_T$  values can thereby be related to a distinct number of plasmids and with the knowledge of the molecular weight of the haploid animal genome, the number of molecules represented by a certain amount of animal DNA can be estimated [28, 43]. For this, it is necessary to construct

a standard curve using serial dilutions of at least five different concentrations. The amount of unknown target should fall within the range tested. Amplification of the standard dilution series and the target sequence is carried out in separate wells. The  $C_T$  values of the standard samples are determined. Then, the  $C_T$  value of the unknown sample is compared with the standard curve to determine the amount of target in the unknown sample. A given number of animal DNA molecules yield the same  $C_T$  value as the same number of plasmids, if all molecules contain one copy of the transgene (i.e., if the animal is homozygous). [28] For heterozygous animal with only half of the molecules containing the expression cassette, the  $C_T$  value will count for half of the number of plasmids. [28] This method was validated by Schmidt and Parrott [43]. The advantages of this method are that large amounts of standard that can be produced, its identity can be verified by sequencing and DNA can easily be quantified by spectrophotometry or fluorometry. Plasmid standards should be linearized since the amplification efficiency of a linearized plasmid often differs from that supercoiled conformation and more closely simulates the amplification efficiency of genomic DNA. Additionally, due to variations in inhibitor levels of qPCR between tissues is recommended also the use of plasmids mixed with genomic DNA samples of non-transgenic animals to delineate the curves to simulate possible interferences of each tissue [44].



**Figure 2.** Typical standard curve and principles of absolute quantification strategy. The  $C_T$  values were plotted versus DNA amounts used for qPCR amplifications. The slope of the tendency curve achieved by linear regression is used to determine the efficiency of qPCR. The Pearson correlation coefficient ( $R^2$ ) is the linearity and should be close to the unit (or 100%). Determination of a sample of interest (unknown sample) is performed by extrapolating the  $C_T$  value on the standard curve.

The gDNA can also be used to construct the standard curve for absolute quantification. However, in this situation, analysts need to certify the presence of only one

copy per haploid genome of the target DNA and the exclusion of closely related pseudogenes and / or sequences from amplification. The main advantage of this approach to quantification is that there is no need to use a calibrator (a sample for which the copy number is exactly known, typically from a Southern blot). However, the accuracy of this strategy is directly associated with the precision in DNA quantification. As discussed above, currently, the main DNA quantification methods are spectrometry (A 260 nm) and fluorometry, which can, respectively, overestimate and underestimate the real amounts of DNA. Consequently, these matters will reflect in the  $C_T$  values, compromising the result [35]. This imprecision in the result can be even more pronounced when the plasmids are used to construct the standard curve, due to the small mass of this deoxyribonucleic acid (small errors in the quantification reflect in large variations in the  $C_T$ ). However, several groups have used this method successfully for the characterization of transgenic animals (Table 2). Using, this strategy to estimate the number of copies of transgenic mice for hG-CSF, we observed an accuracy of 100% when DNA was quantified with fluorometer and qPCR all requirements have been met [28].

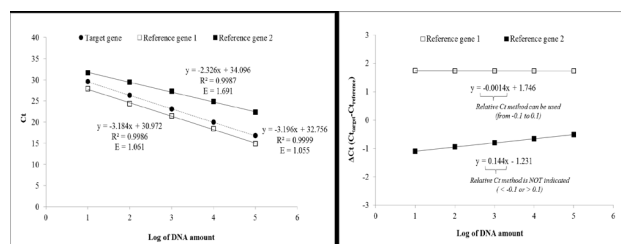
### 3.3.2 Relative Quantification

The comparative  $C_T$  is the most used relative quantification method for several purposes, including transgene analysis [10, 44]. However, the validation of comparative  $C_T$  for transgene copy quantification requires a previous comparison between standard curves plotted for the transgene (target) and the reference gene (Figure 3). Thus, the efficiency ( $E = 10^{(-1/Slope)}$ ) of both transgene and reference gene amplifications must be highly similar. A simple way to determine what relative quantification method can be applied is to plot  $\Delta C_T$  values (calculated as the difference between target and reference gene  $C_T$ s) versus log of DNA amounts. The comparative  $C_T$  method can be used if the slope of this plot is between -0.1 and 0.1. However, if the slope is out of this range, the indicated method should be the relative quantification by standard curves. Another limitation of using this strategy to copy number quantification and zygosity analysis is referred to the accuracy of the results of the quantification calibrator. The main justification for the use of real-time PCR in the characterization of transgenic animals is the increased reliability of results obtained in relation to conventional methods, as Southern blot, and FISH [8,9,10]. However, copy number and zygosity analysis in calibrator animals are usually performed by conventional methods. In this sense, any errors that may have occurred in the characterization of this animal will be impressed in real-time PCR quantification

of target animals.

### 3.3.3 Relative Quantification by Standard Curves

The characterization of transgenic animals using these methods requires the use of a reference sample, with the number of copies or the known zygosity, and an endogenous control gene [46]. The amount of the target is determined from the standard curve of the transgene and an endogenous control. Normalization is performed by dividing the equivalent dilution of the transgene by the equivalent dilution of the endogenous control [47]. Standard dilution equivalents without a unit require a sample to serve as a calibrator. A good calibrator for quantifying copy number is a homozygous animal with one copy of the transgene. Samples with half the normalized dilution equivalent used as the calibrator are heterozygous; samples with the same normalized dilution equivalent are homozygous for one copy, and so on. The quantification procedure differs depending on whether the target and the endogenous reference gene are amplified with comparable or different efficiencies. This strategy has been used mainly to determine the zygosity of transgenic animals [25, 47, 48].



**Figure 3.** Determination of real-time PCR efficiencies from the slopes of the calibration curve. To compare the amplification efficiencies of the 2 target sequences, the  $C_T$  values of reference gene 1 are subtracted from the  $C_T$  values of target gene. The difference in  $C_T$  values is then plotted against the logarithm of the template amount. If the slope of the resulting straight line is  $< 0.1$ , amplification efficiencies are comparable.

### 3.3.4 Relative Quantification by Comparative $C_T$ Method

While requiring an endogenous control and a calibrator, differs from the relative standard method by relying on equal PCR efficiencies with the transgene and the endogenous control genes. The preparation of standard curves is only required to determine the amplification efficiencies of the transgene and endogenous control genes in an initial experiment. In all subsequent experiments, no standard curve is required for quantification of the target sequence. According to Livak and Schmittgen [40], if all amplicons amplify with the same efficiency, the difference

$\Delta C_T$  between the  $C_T$  for the transgene ( $C_{Tt}$ ) and the  $C_T$  for the endogenous control ( $C_{Te}$ ) is constant, independent of the amount of chromosomal DNA ( $\Delta C_T = C_{Tt} - C_{Te}$ ). As for quantification with relative standards, a calibrator is a homozygous one-copy animal. Thus, all samples with the same  $\Delta C_T$  as the calibrator contain one copy of the transgene. More generally, the ratio of the initial amount of transgene in the sample ( $X_s$ ) to the initial amount of transgene in the calibrator ( $X_{cal}$ ) can be calculated as follows ( $X_s/X_{cal} = (1+E)^{-\Delta\Delta CT}$ ), where:  $\Delta\Delta Ct = \Delta C_{Ts} - \Delta C_{Tcal}$ . Whereas for copy number calculation  $\Delta\Delta C_{Ts}$  will be zero (one-copy animals) or negative (multi-copy animals), zygosity analysis should yield  $\Delta\Delta C_{Ts}$  of zero (homozygous) or one (heterozygous). As long as the efficiencies for transgene and endogenous control are the same, calculations with  $E < 1$  are also possible. The  $2^{\Delta\Delta CT}$  method is simple to apply because DNA concentrations do not have to be measured. Its utility has been demonstrated for animal copy number determination<sup>[8]</sup> and zygosity analysis in animals<sup>[7]</sup>.

### 3.5 Post-qPCR Procedures – Mathematical and Statistical Considerations

Due to the high variation in CT values, it has been proposed that the limit for determining the copy number of the transgene by the qPCR technique is two-fold differences. According to Bubner and Baldwin<sup>[43]</sup>, when the standard deviation of the CT values for all samples and amplicons is greater than 0.3, the interpretation of the transgene copy numbers will be compromised, dividing the inability to detect differences twice. Mason et al.<sup>[51]</sup> reported that only about 70% of qPCR-based transgene copy determination results could be verified by Southern blot analysis. In addition, another aspect that affects the determination of the copy number by qPCR is the lack of complete statistical analysis and adequate models, capable of testing the hypotheses<sup>[12]</sup>. Normally, hypothesis tests were not invoked for the transgene copy numbers, in addition, a predetermined P value and the confidence levels of the estimate were not specified. These aspects, especially the confidence levels, are important for determining the number of copies of the transgene, since it defines the precision and sensitivity of the assay<sup>[52]</sup>. Due to the limitations of statistical procedures, the results of the analysis are often ambiguous and without clear confidence intervals. The confidence intervals help to establish the reliable interval for the estimate of  $\Delta\Delta CT$ , while the value of P determines the level of significance<sup>[53]</sup>. In the statistics program, all P values are derived from the null hypothesis test that  $\Delta\Delta CT$  is equal to 0. Therefore, a small P value indicates that  $\Delta\Delta CT$  is significantly different from

0, which demonstrates a significant effect. Despite this, some studies have shown accuracy in the characterization of transgenics. Haurogné et al.<sup>[49]</sup> reported success in estimating the copy number for CT differences close to 1, when they used low concentrations of genomic DNA. In summary, small fluctuations in the initial conditions of a PCR assay led to a large fluctuation in the amount of the product, which is expressed in CT (or equivalent) values. Thus, apparently small standard deviations of the CT values (between 0.3 and 1) are amplified in the analysis because a difference of CT of one represents a difference of twice the initial value. In addition to the variability in transgene measurements, the variability in the measurement of endogenous control must also be considered<sup>[50]</sup>.

## 4. Conclusions

In summary, this review allows us to conclude that real-time PCR is a powerful tool for the characterization of transgenic animals, especially for copy number determination and zygosity analysis. This quantitative technique has the potential to become a widespread tool in animal transformation research, because it helps to characterize the lines, to infer or to explain transgene expression levels and to drive the reproductive managements for livestock establishment. However, the choice of the quantitative PCR method must be accompanied by appropriate validations, ensuring that the measurements are correct and adjusted to the experimental conditions (genes, primers, templates, temperatures). Finally, specific transgenic animal lines that are considered for research or commercial release probably will require the confirmation of both copy number and zygosity data by independent methods, such as Southern blot and mating, respectively. Hence, real-time PCR is a tool that complements rather than replacing traditional procedures.

## References

- [1] Batista RITP, Melo CH, Souza-Fabjan JM, Teixeira DI, Melo LM, Freitas VJ. Phenotypic features of first-generation transgenic goats for human granulocyte-colony stimulation factor production in milk. *Biotechnol Lett.* 2014 36(11): 2155-62. DOI: 10.1007/s10529-014-1588-0.
- [2] Houdebine LM. Use of transgenic animals to improve human health and animal production. *Reprod Domest Anim.* 2005 40(4): 269-81. DOI: 10.1111/j.1439-0531.2005.00596.x.
- [3] Gordon JW, Ruddle FH. Integration and stable germ line transmission of genes injected into mouse pronuclei. *Science.* 1981, 214(4526): 1244-6.



- DOI: 10.1126/science.6272397.
- [4] Capecchi MR. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat Rev Genet.* 2005 Jun;6(6):507-12.  
DOI: 10.1038/nrg1619.
- [5] Ishii T. Reproductive medicine involving genome editing: its clinical and social conundrums. In: *Precision Medicine for Investigators, Practitioners and Providers*, Academic Press, eds. Joel Faintuch, Salomao Faintuch, pp 419-429, 2020.
- [6] B. Tinkle, C. Bieberich, and G. Jay, Molecular approaches involved in mammalian gene transfer: Analysis of transgene integration. In: *Transgenic Animal Technology: A Laboratory Handbook*, Pinkert CA, eds. Academic Press, San Diego, pp 221-234, 1994.
- [7] Tesson L, Heslan JM, Ménoret S, Anegon I. Rapid and accurate determination of zygosity in transgenic animals by real-time quantitative PCR. *Transgenic Res.* 2002, 11(1):43-8.  
DOI: 10.1023/a:1013928600442.
- [8] Ballester M, Castelló A, Ibáñez E, Sánchez A, Folch JM. Real-time quantitative PCR-based system for determining transgene copy number in transgenic animals. *Biotechniques.* 2004, 37(4): 610-3.  
DOI: 10.2144/04374ST06.
- [9] Gaj T, Sirk SJ, Shui SL, Liu J. *Genome-Editing Technologies: Principles and Applications*. Cold Spring Harb Perspect Biol. 2016, 8(12):a023754.  
DOI: 10.1101/cshperspect.a023754.
- [10] Guha TK, Wai A, Hausner G. Programmable Genome Editing Tools and their Regulation for Efficient Genome Engineering. *Comput Struct Biotechnol J.* 2017, 12;15:146-160.  
DOI: 10.1016/j.csbj.2016.12.006.
- [11] Grohmann L, Keilwagen J, Duensing N, Dagand E, Hartung F, Wilhelm R, Bendiek J, Sprink T. Detection and Identification of Genome Editing in Plants: Challenges and Opportunities. *Front Plant Sci.* 2019, 12;10:236.  
DOI: 10.3389/fpls.2019.00236.
- [12] Yuan JS, Burris J, Stewart NR, Mentewab A, Stewart CN Jr. Statistical tools for transgene copy number estimation based on real-time PCR. *BMC Bioinformatics.* 2007 Nov 1;8 Suppl 7(Suppl 7):S6.  
DOI: 10.1186/1471-2105-8-S7-S6. PMID: 18047729; PMCID: PMC2099498.
- [13] Grosveld F, van Assendelft GB, Greaves DR, Kollias G. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell.* 1987, 51(6):975-85.  
DOI: 10.1016/0092-8674(87)90584-8.
- [14] Bieberich C, Scangos G, Tanaka K, Jay G. Regulated expression of a murine class I gene in transgenic mice. *Mol Cell Biol.* 1986, 6(4):1339-42.  
DOI: 10.1128/mcb.6.4.1339.
- [15] Li J, Brunner AM, Meilan R, Strauss SH. Stability of transgenes in trees: expression of two reporter genes in poplar over three field seasons. *Tree Physiol.* 2009, 29(2): 299-312.  
DOI: 10.1093/treephys/tpn028.
- [16] Mahon KA, Overbeek PA, Westphal H. Prenatal lethality in a transgenic mouse line is the result of a chromosomal translocation. *Proc Natl Acad Sci U S A.* 1988, 85(4):1165-8.  
DOI: 10.1073/pnas.85.4.1165.
- [17] Garrick D, Fiering S, Martin DI, Whitelaw E. Repeat-induced gene silencing in mammals. *Nat Genet.* 1998 Jan;18(1):56-9.  
DOI: 10.1038/ng0198-56. PMID: 9425901.
- [18] Koike S, Taya C, Aoki J, Matsuda Y, Ise I, Takeda H, Matsuzaki T, Amanuma H, Yonekawa H, Nomoto A. Characterization of three different transgenic mouse lines that carry human poliovirus receptor gene--influence of the transgene expression on pathogenesis. *Arch Virol.* 1994, 139(3-4): 351-63.  
DOI: 10.1007/BF01310797.
- [19] Ida-Hosonuma M, Iwasaki T, Taya C, Sato Y, Li J, Nagata N, Yonekawa H, Koike S. Comparison of neuropathogenicity of poliovirus in two transgenic mouse strains expressing human poliovirus receptor with different distribution patterns. *J Gen Virol.* 2002, 83(Pt 5):1095-1105.  
DOI: 10.1099/0022-1317-83-5-1095.
- [20] Deatly AM, Taffs RE, McAuliffe JM, Nawoschik SP, Coleman JW, McMullen G, Weeks-Levy C, Johnson AJ, Racaniello VR. Characterization of mouse lines transgenic with the human poliovirus receptor gene. *Microb Pathog.* 1998, 25(1):43-54.  
DOI: 10.1006/mpat.1998.0212.
- [21] Chandler KJ, Chandler RL, Broeckelmann EM, Hou Y, Southard-Smith EM, Mortlock DP. Relevance of BAC transgene copy number in mice: transgene copy number variation across multiple transgenic lines and correlations with transgene integrity and expression. *Mamm Genome.* 2007, 18(10):693-708.  
DOI: 10.1007/s00335-007-9056-y.
- [22] Kong Q, Wu M, Huan Y, Zhang L, Liu H, Bou G, Luo Y, Mu Y, Liu Z. Transgene expression is associated with copy number and cytomegalovirus promoter methylation in transgenic pigs. *PLoS One.* 2009, 18;4(8):e6679.  
DOI: 10.1371/journal.pone.0006679.
- [23] Palmiter RD, Wilkie TM, Chen HY, Brinster RL.

- Transmission distortion and mosaicism in an unusual transgenic mouse pedigree. *Cell*. 1984, 36(4):869-77. DOI: 10.1016/0092-8674(84)90036-9.
- [24] Shitara H, Sato A, Hayashi J, Mizushima N, Yonekawa H, Taya C. Simple method of zygosity identification in transgenic mice by real-time quantitative PCR. *Transgenic Res*. 2004, 13(2):191-4. DOI: 10.1023/b:trag.0000026084.32492.eb.
- [25] Guénin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C, Gutierrez L. Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *J Exp Bot*. 2009, 60(2):487-93. DOI: 10.1093/jxb/ern305.
- [26] Burkhardt HJ. Standardization and quality control of PCR analyses. *Clin Chem Lab Med*. 2000 Feb;38(2):87-91. DOI: 10.1515/CCLM.2000.014. PMID: 10834394.
- [27] Terry CF, Harris N, Parkes HC. Detection of genetically modified crops and their derivatives: critical steps in sample preparation and extraction. *J AOAC Int*. 2002 May-Jun;85(3):768-74. PMID: 12083273.
- [28] Batista RI, Luciano MC, Teixeira DI, Freitas VJ, Melo LM, Andreeva LE, Serova IA, Serov OL. Methodological strategies for transgene copy number quantification in goats (*Capra hircus*) using real-time PCR. *Biotechnol Prog*. 2014, 30(6): 1390-400. DOI: 10.1002/btpr.1946.
- [29] Demeke T, Ratnayaka I, Phan A. Effects of DNA extraction and purification methods on real-time quantitative PCR analysis of Roundup Ready soybean. *J AOAC Int*. 2009, 92(4): 1136-44. DOI: 10.1093/jaoac/92.4.1136
- [30] P. Corbisier, W. Broothaerts, S. Gioria, H. Schimmel, et al., Toward Metrological Traceability for DNA Fragment Ratios in GM Quantification. 1. Effect of DNA Extraction Methods on the Quantitative Determination of Bt176 Corn by Real-Time PCR. *Journal of Agricultural and Food Chemistry*, vol. 55, no. 9, 3249-3257, 2007.
- [31] Sakurai T, Kamiyoshi A, Watanabe S, Sato M, Shin-do T. Rapid zygosity determination in mice by SYBR Green real-time genomic PCR of a crude DNA solution. *Transgenic Res*. 2008, 17(1):149-55. DOI: 10.1007/s11248-007-9134-7.
- [32] Rossen L, Nørskov P, Holmstrøm K, Rasmussen OF. Inhibition of PCR by components of food samples, microbial diagnostic assays and DNA-extraction solutions. *Int J Food Microbiol*. 1992 Sep;17(1):37-45. DOI: 10.1016/0168-1605(92)90017-w.
- [33] Wilson IG. Inhibition and facilitation of nucleic acid amplification. *Appl Environ Microbiol*. 1997 Oct;63(10):3741-51. DOI: 10.1128/AEM.63.10.3741-3751.1997. PMID: 9327537; PMCID: PMC168683.
- [34] Burkhart CA, Norris MD, Haber M. A simple method for the isolation of genomic DNA from mouse tail free of real-time PCR inhibitors. *J Biochem Biophys Methods*. 2002 Jul 31;52(2):145-9. DOI: 10.1016/s0165-022x(02)00052-0. PMID: 12204418.
- [35] Demeke T, Jenkins GR. Influence of DNA extraction methods, PCR inhibitors and quantification methods on real-time PCR assay of biotechnology-derived traits. *Anal Bioanal Chem*. 2010 Mar;396(6):1977-90. DOI: 10.1007/s00216-009-3150-9. Epub 2009 Sep 30. PMID: 19789856.
- [36] Shokere LA, Holden MJ, Jenkins GR. Comparison of fluorometric and spectrophotometric DNA quantification for real-time quantitative PCR of degraded DNA. *Food Control* 2009, 20(4): 391–401. DOI.org/10.1016/j.foodcont.2008.07.009.
- [37] Rye HS, Dabora JM, Quesada MA, Mathies RA, Glazer AN. Fluorometric assay using dimeric dyes for double- and single-stranded DNA and RNA with picogreen sensitivity. *Analytical Biochemistry* 1993, 208(1): 144–150. DOI: 10.1006/abio.1993.1020.
- [38] Wong ML, Medrano JF. Real-time PCR for mRNA quantitation. *Biotechniques*. 2005 Jul;39(1):75-85. DOI: 10.2144/05391RV01. PMID: 16060372..
- [39] Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. *Biotechnol Lett*. 2004, 26(6):509-15. DOI: 10.1023/b:bile.0000019559.84305.47.
- [40] K.J. Livak, and T.D. Schmittgen, "Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C(T)}$  method" *Methods*, vol. 25, no. 4, pp. 402-408, 2001.
- [41] C.J. Smith, and A.M. Osborn, "Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology", *FEMS Microbiology Ecology* vol. 67, no. 1, pp. 6-20, 2009.
- [42] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-(\Delta\Delta C(T))}$  Method. *Methods*. 2001, 25(4):402-8. DOI: 10.1006/meth.2001.1262.
- [43] Bubner B, Baldwin IT. Use of real-time PCR for determining copy number and zygosity in transgenic plants. *Plant Cell Rep*. 2004, 23(5):263-71.

- DOI: 10.1007/s00299-004-0859-y.
- [44] Schmidt M, Parrott W. Quantitative detection of transgenes in soybean [*Glycine max* (L.) Merrill] and peanut (*Arachis hypogaea* L.) by real-time polymerase chain reaction. *Plant Cell Rep.* 2001, 20(5):422-428.  
DOI: 10.1007/s002990100326.
- [45] Joshi M, Keith Pittman H, Haisch C, Verbanac K. Real-time PCR to determine transgene copy number and to quantitate the biolocalization of adoptively transferred cells from EGFP-transgenic mice. *Bio-techniques.* 2008, 45(3):247-58.  
DOI: 10.2144/000112913. PMID: 18778249.
- [46] Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 2001, 29(9):e45.  
DOI: 10.1093/nar/29.9.e45.
- [47] Pfaffl MW. Quantification strategies in real-time PCR. In: *A-Z of Quantitative PCR* (Bustin SA,eds.). IUL Biotechnology Series, International University Line, La Jolla, CA, 87-120, 2004.
- [48] Ji W, Zhou W, Abruzzese R, Guo W, Blake A, Davis S, Davis S, Polejaeva I. A method for determining zygosity of transgenic zebrafish by TaqMan real-time PCR. *Anal Biochem.* 2005, 344(2):240-6.  
DOI: 10.1016/j.ab.2005.06.046.
- [49] Haurogné K, Bach JM, Lieubeau B. Easy and rapid method of zygosity determination in transgenic mice by SYBR Green real-time quantitative PCR with a simple data analysis. *Transgenic Res.* 2007,16(1):127-31.  
DOI: 10.1007/s11248-006-9024-4.
- [50] Bubner B, Gase K, Baldwin IT. Two-fold differences are the detection limit for determining transgene copy numbers in plants by real-time PCR. *BMC Biotechnol.* 2004 Jul 13;4:14.  
DOI: 10.1186/1472-6750-4-14.
- [51] Mason G, Provero P, Vaira AM, Accotto GP. Estimating the number of integrations in transformed plants by quantitative real-time PCR. *BMC Biotechnol.* 2002, 2:20.  
DOI: 10.1186/1472-6750-2-20.
- [52] J Yuan JS, Wang D, Stewart CN Jr. Statistical methods for efficiency adjusted real-time PCR quantification. *Biotechnol J.* 2008, 3(1):112-23.  
DOI: 10.1002/biot.200700169.
- [53] Yuan JS, Reed A, Chen F, Stewart CN Jr. Statistical analysis of real-time PCR data. *BMC Bioinformatics.* 2006, 22;7:85.  
DOI: 10.1186/1471-2105-7-85.