Habilitation

Livestock Transcriptomics:
Quantitative mRNA Analytics in Molecular Endocrinology and Physiology

submitted to the Faculty Center of Life and Food Sciences,
Technische Universität München - Weihenstephan

Author: Dr. agr. Michael W. Pfaffl

Chair of Physiology, Department of Animal Science, Center of Life and Food Sciences,
Weihenstephaner Berg 3, 85354 Freising, Germany

Michael.Pfaffl@wzw.tum.de
http://www.weihenstephan.de/fml/physio/personen/pfaffl.htm
http://www.weihenstephan.de/gene-quantification/habilitation.html

January 2003
# Table of contents

Table of contents ........................................................................................................... 1

Abstract .......................................................................................................................... 2

Introduction ...................................................................................................................... 2

Livestock transcriptome .................................................................................................... 4

RNA extraction ................................................................................................................. 4

Reverse transcription ........................................................................................................ 6

Comparison of real-time RT-PCR with end-point detection method ................................ 7

Chemistry developed and real-time RT-PCR platforms .................................................. 8

Quantification strategies in kinetic RT-PCR ................................................................. 9

Absolute quantification ................................................................................................. 10

Advantages and disadvantages of external standards ................................................... 12

Application of absolute quantification ........................................................................ 13

Relative quantification ..................................................................................................... 17

Application of relative quantification ........................................................................... 19

How to get rid of unspecific real-time PCR artefacts ................................................... 23

Real-time PCR amplification efficiency ....................................................................... 26

Normalisation of expression results .............................................................................. 29

Real-time PCR data processing ....................................................................................... 33

Data evaluation ................................................................................................................ 33

Automation of the quantification procedure .................................................................. 34

Statistical comparison ...................................................................................................... 35

Conclusion ........................................................................................................................ 36

Acknowledgements ......................................................................................................... 37

References .......................................................................................................................... 38
Abstract

Molecular technologies are currently evolving rapidly in agricultural and veterinary sciences. This results in an immense progress in the accumulation of new data potentially useful for molecular diagnostics in farm animal physiology, immunology, diseases and new breeding strategies. While we are still at the “very beginning” of understanding genomics, transcriptomics and proteomics in relation to animal physiology, this development has dramatically changed our perspectives in research during the last decade. It can be foreseen, that the application of sophisticated rather than simple methods will be necessary for numerous diagnostic questions. One of this highly sophisticated methodologies is the quantitative assessment of target nucleic acids, mostly performed as quantitative polymerase chain reaction (PCR) on DNA level or combined with reverse transcription PCR (RT-PCR) to investigate the transcriptome on RNA level. This review will introduce the state of the art in quantitative RT-PCR using real-time RT-PCR on the field of livestock molecular endocrinology and physiology.

Introduction

To investigate local tissue specific expression even in tissues with low abundances, very sensitive methods are required which allow reliable mRNA quantification. Because of its high sensitivity, RT-PCR is being increasingly applied to quantify physiologically relevant changes in gene expression. RT-PCR has a detection limit 10- to 100-fold better than other methods, e.g. RNA-Protection-Assay or Northern-Hybridisation respectively (1-4). It offers a new dimension in the detection of rare RNA by amplifying a single stranded cDNA after reverse transcription. The RT-PCR quantification technique of choice depends on the target sequence, the expected range of the mRNA amount present in the tissue, the degree of accuracy required, and on the question whether quantification needs to be relative or absolute. Externally standardised RT-PCR with quantification on ethidium bromide stained gels followed by densitometry is widely used, but the degree of accuracy is limited and the quantification is more relative than absolute (4). Today the microarray technology is a powerful technique in order to analyse the gene expression of thousands of genes (up to 20000) in a short time. Presently the microarrays or gene chips of cDNA and oligo-nucleotide type are available from several manufacturers. Problems encountered include inconsistent fidelity, high variability, sensitivity not sufficient for low abundant expressed genes (like some growth factors and their receptors), discrepancy in “fold-changes calculation” and lack of specificity for different isoforms or differentially expressed genes (5-9).

For an exact quantitative measurement of low abundant gene expression only a few PCR methods allow reliable mRNA quantification. The aim of a full quantitative method is to estimate, as exact and reliable as possible, the number or target molecules in the sample. At present the following RT-PCR methods are suitable for such a sensitive quantification.
A) Internally standardised competitive RT-PCR measured by HPLC separation and UV detection (10, 11) or high resolution gel electrophoresis followed by densitometric analysis (12,13): In a competitive RT-PCR, a reference RNA mutant is reverse transcribed and co-amplified in the same reaction tube with the native mRNA sequence of interest. Internally standardised RT-PCR is a very time consuming and laborious technique. It is generally believed to yield the most precise results, because all parameters throughout RT-PCR act on both the analyte and reference mutant.

B) Externally standardised RT-PCR with online-detection using LightCycler SYBR ® Green I (Molecular Probes) technology (14-17): real-time RT-PCR with SYBR ® Green I detection produces reliable and rapid results. Due to the use of an external standard curve, the amplification efficiencies for the calibration curve and the analyt must be equal for accurate quantification.

C) Externally standardised RT-PCR with online-detection using specific hybridisation probes (18-20): This detection format is based on various fluorescence detection formats, e.g. fluorescence resonance energy transfer (FRET).

D) DNA array technologies and real-time RT-PCR: The microarray based screening of tissue specific gene expression and confirmation of putative candidate target genes by kinetic RT-PCR represents a powerful and optimal combination (21, 22). Hereby the advantages of both quantification systems can be added - the high throughput capacity of the microarray platform as well as sensitivity and specificity of the real-time RT-PCR platform (5-9).

Today real-time RT-PCR (or kinetic RT-PCR) is increasingly used because of its high sensitivity, good reproducibility, and wide dynamic quantification range (1-4). The first practical kinetic PCR technology, the 5’-nuclease assay, was established 1993 and combines the exponential PCR amplification of a specific transcript with the monitoring of newly synthesized DNA in each performed PCR cycle (14-17). It is the most sensitive method for the detection and quantification of gene expression levels, in particular for low abundant transcripts in tissues with low RNA concentrations, from limited tissue samples, and for the elucidation of small changes in mRNA expression levels (4, 19, 20, 23, 24). While kinetic RT-PCR has a tremendous potential for analytical and quantitative applications, a comprehensive understanding of its underlying principles is important. Fidelity of real-time RT-PCR is associated with its “true” specificity, sensitivity, reproducibility, robustness and, as a fully reliable quantitative method, it suffers from the problems inherent in RT and PCR, e.g. Amplification of unspecific products, primer-dimers, amplification efficiencies, hetero-duplex formation, etc. (19).

This review analyses the mRNA quantification analytics and quantification strategies in real-time RT-PCR and describes all corresponding markers of a successful real-time RT-PCR. The following aspects are discussed in detail: RNA extraction, reverse transcription, general quantification strategies: absolute and relative quantification, efficiency calculation, normalisation of expression results, data processing and
statistical comparison. Further this review turns into practical considerations with focus on specificity, sensitivity, variability, reproducibility and the experimental design of the experiments. Useful real-time RT-PCR applications in animal molecular endocrinology and molecular physiology are discussed and descriptive examples are shown in figures. The corresponding tables will be shown exclusively in the appendix, where some “papers by Pfaffl et al.” are shown in full length. All chapters described and the corresponding literature is also accessible on the Gene Quantification web page: http://www.wzw.tum.de/gene-quantification/

Livestock transcriptome

RNA extraction

The integrity of purified RNA is critical to all gene expression analysis techniques. The preparation of intact cellular total RNA or pure mRNA is the first critical step in gene quantification. For successful and reliable diagnostic use, real-time RT-PCR needs high quality, DNA free, and undegraded RNA (25, 26). Accurate quantification and quality assessment (30) of the starting RNA sample is particularly important for absolute quantification methods that normalize specific mRNA expression levels against total RNA (‘molecules/g total RNA’ or ‘transcript concentrations/g total RNA’). RNA, especially long mRNA up to 10 kb (25), is easily degraded by cleavage of RNAes during tissue sampling, RNA purification and RNA storage. The source of RNA, sampling techniques (biopsy material, purified somatic cells from complex matrices, e.g. blood and milk, single cell sampling, laser micro-dissection) (27-30) as well as RNA isolation techniques (either total RNA or poly-adenylated RNA isolation techniques) often vary significantly between processing laboratories (26).

In agricultural and veterinary research RNA extracted from mammary gland, especially during lactation period (29, 30), adipose or collagen rich tissues, tissue sampled after long lack time at the slaughterhouse (10-13) often has a lower yield and is of lesser quality, and contains partly degraded RNA sub-fractions, especially the messenger RNA (mRNA) fraction (own unpublished results). Particular RNA extraction techniques can act more effectively on one specific tissue type compared to another one, and result in up to 10-fold variations in total RNA yield (26).

In figure 1 the livestock transcriptome (various cattle tissue total RNA) is characterised and all sub-fractions are described. As shown, most genes are transcribed at very low abundant mRNA levels under 20 copies per
Therefore highly sensitive quantification methods are necessary to perform a precise and reliable quantification of growth factors, hormones, their corresponding receptors and enzymes being important for understanding animal endocrinology and physiology.

<table>
<thead>
<tr>
<th>sub-fraction</th>
<th>content of total RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ribosomal RNA</td>
<td>rRNA</td>
</tr>
<tr>
<td>transfer RNA</td>
<td>tRNA</td>
</tr>
<tr>
<td>messenger RNA</td>
<td>mRNA</td>
</tr>
<tr>
<td>average length</td>
<td></td>
</tr>
<tr>
<td>high abundant</td>
<td>&lt;10 genes</td>
</tr>
<tr>
<td>intermediate abundant</td>
<td>~500-1000 genes</td>
</tr>
<tr>
<td>low abundant</td>
<td>&gt;10000 genes</td>
</tr>
</tbody>
</table>

**Figure 1:** Characterisation of the transcriptome isolated from numerous cattle tissue samples (29, 30).

Isolated RNA may contain tissue enzyme inhibitors that result in reduced RT and PCR reaction efficiencies and generate unreliable and wrong quantification results (25, 26). Most RNA preparations are contaminated with DNA and protein at very low levels. Even high quality commercially obtained RNA contain detectable amounts of DNA (26). While this is not a problem for some applications, the tremendous amplification power of kinetic PCR may result in even the smallest amount of DNA contamination to interfering with the desired “specific amplification”. To confirm the absence of residual DNA either a “minus-RT” or “water control” should always be included in the experimental design. It may be necessary to treat the RNA sample with commercially available RNAse-free DNAse, to get rid of residual DNA.

However, unspecific side reactions of the DNAse often result in RNA degradation (own unpublished results). It is always necessary to remove the DNAse prior to any RT or PCR step. Furthermore, the design of the PCR product should incorporate at least one exon-exon splice junction to allow a product obtained from the cDNA to be distinguished on electrophoresis from genomic DNA contamination (4, 10-13, 19, 20). Processed pseudogenes [e.g. β-actin, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), 18S rRNA] can be present and lead to confusion in data interpretation. In addition, intron-lacking pseudogenes (e.g. β-actin) with equal sequence length to endogenous mRNA have been described (31-37). They prevent a distinction between products originating from genomic DNA versus mRNA, which poses a significant problem on qualitative and quantitative gene quantification. Therefore, various housekeeping genes must be tested or
multiplex assays of reference genes as internal controls for the assessment of RNA and cDNA quality must be performed (38-42).

**Reverse transcription**

The second step in quantitative RT-PCR is the production of a single-stranded (ss) complementary DNA copy (cDNA) of the RNA through the reverse transcriptase (RT) and its dynamic range, sensitivity and specificity are of prime consideration for a successful kinetic RT-PCR assay (43). For many quantitative applications, MMLV H\(^{-}\)RT is the enzyme of choice (44-47), as its cDNA synthesis rate is up to 50-fold greater than that of AMV (own unpublished results). Newly available thermo stable RNAse H\(^{-}\)RT maintains its activity up to 70°C, thus permitting increased specificity and efficiency of first primer annealing (4, 44). However, this enzyme may be less robust than more conventional ones as it appears to be more sensitive to inhibitors present in RNA preparation (44-47).

The RT step is the source of most of the variability in a kinetic RT-PCR experiment and for each enzyme the specific reaction conditions has to be optimised. Salt contamination, alcohol, phenol and other inhibitors carried over from the RNA isolation process can affect the apparent RT efficiency (19, 48-52). Another source of variability is the choice of priming method used to initiate cDNA synthesis, which can be either target gene-specific or non-specific. Target gene specific primers work well in conjunction with elevated RT reaction temperatures to eliminate spurious transcripts (53, 54). The same anti-sense (reverse primer) is used for the subsequent PCR assay in conjunction with the corresponding gene-specific sense primer (forward primer). However, the use of gene-specific primers necessitates a separate RT reaction for each gene of interest. It cannot be assumed that different reactions have the same cDNA synthesis efficiency; the result can be high variability during multiple RT reactions.

To circumvent this high inter-assay variations in RT, target gene unspecific primers, e.g. random-hexamer, -octamer or -decamer primers, can be used and a cDNA pool can be synthesized. Similarly, poly-T oligonucleotides (consisting solely of 16-25 deoxythymidine residues) can anneal to the poly-adenylated 3’- (poly-A) tail found on most mRNA (19, 47). The cDNA pools synthesized with unspecific primers can be split into a number of different target-specific kinetic PCR assays. This maximizes the number of genes that can be assayed from a single cDNA pool, derived from one small RNA sample. Therefore the gene expression results are directly comparable between the applied assays, at least within one and the same RT pool. In conclusion, a rank order of RT efficiency can be shown for the applied different primers for ONE specific gene (own unpublished results):

\[
\text{random hexamer primers} > \text{poly-dt primer} > \text{gene specific primer}
\]

Importantly, not only RNA quantity and quality, but also yield and quality of cDNA can be highly variable. Certainly, there is evidence that cDNA yield from sequences near the 5’-end of partially degraded mRNA is significantly less than from sequences near the poly-A tail and assays aimed at identifying RNA degradation...
are being developed (3, 25, 51, 56). Thus, reliable internal quality control of cDNA synthesis is essential. Controls are generally performed by PCR amplification of reference genes, mostly common housekeeping genes GAPDH, albumin, actins, tubulins, cyclophilin, microglobulins, 18S ribosomal RNA (rRNA) or 28S rRNA (40, 52-57). The reference genes used as well as the expression levels vary between different laboratories, and only few of them have been critically evaluated (see Normalisation).

Comparison of real-time RT-PCR with end-point detection method

The efficacy of kinetic RT-PCR is measured by its specificity, low background fluorescence, steep fluorescence increase and high amplification efficiency, and high level plateau (58). Typically the PCR can be divided in four characteristic phases as shown in figure 2 (59):

1st phase is hidden under the background fluorescence, where an exponential amplification is expected;
2nd phase with exponential amplification, that can be detected above the background;
3rd phase with linear amplification efficiency and a steep increase of fluorescence;
4th phase or plateau phase, defined as the attenuation in the rate of exponential product accumulation, which is seen concomitantly in later cycles (58-60);

![Figure 2](image)

**Figure 2:** The four characteristic phases of PCR, evaluated by real-time PCR fluorescence acquisition (59, 60).
The amount of amplified target is directly proportional to the input amount of target only during the exponential phase of PCR amplification. Hence the key factor in the quantitative ability of kinetic RT-PCR is that it measures the product of the target gene within that phase (24, 59, 61-64). Since data acquisition and analysis are performed in the same tube, this increases sample throughput, reduces carryover contamination and removes post-PCR processing as a potential source of error (65).

In contrast, during the plateau phase of the PCR there is no direct relation of "DNA input" to "amplified target"; hence classical RT-PCR assays have to be stopped at least in linear phase (61, 62, 66). The exponential range of amplification has to be determined for each transcript empirically by amplifying equivalent amounts of cDNA over various cycles of the PCR or by amplifying dilutions of cDNA over the same number of PCR cycles (24). Amplified RT-PCR end product is later detected by ethidium bromide gel staining, radioactivity labelling, fluorescence labelling, high performance liquid chromatography, southern blotting, densitometric analysis, or other post amplification detection methods. This step wise accumulation of post-PCR variability leads to semi-quantitative results with high intra-assay (around 30-40%) and inter-assay variability (around 50-70%) in end-point detection assays, including competitive RT-PCR, performed earlier (10-13, 67-71). Finally, whereas real-time methods have a dynamic range of greater than eight orders of magnitude (60, 72), the dynamic range of the endpoint assays is at best two orders of magnitude (10-13).

**Chemistry developed and real-time RT-PCR platforms**

The third critical step in kinetic RT-PCR is the correct detection chemistry. Two general methods for the quantitative detection of the amplicon became established: specific double strand (ds) DNA binding agents or gene specific fluorescent probes (3, 14-16, 72-75), e.g. based on fluorescence resonance energy transfer (FRET) (18, 61, 62). The best-know probe-based system is ABI TaqMan (16, 17, 19, 20, 75, 76) which makes use of the 5'-exonuclease activity of Taq polymerase to quantify target sequences in the samples. Probe hydrolysis separates fluorophore and quencher and results in an increased fluorescence signal, called "Förster type energy transfer" (77, 78). The alternative is a non-sequence specific fluorescent intercalating ds DNA binding dye e.g. SYBR Green I (Molecular Probes) or ethidium bromide (79). For single PCR product reactions with well designed primers, SYBR Green I can work extremely well, with spurious non-specific background only showing up in very late cycles (4, 63, 64, 80, 85). Among the real-time detection chemistry, SYBR Green I and TaqMan assays produced comparable dynamic range and sensitivity, while SYBR Green I detection was more precise and produced a more linear decay plot than the TaqMan probe detection (24). For more detailed information about real-time PCR detection chemistry on real-time platforms, e.g. using hybridisation probes, molecular beacons or scorpions, please visit the following link: http://www.wzw.tum.de/gene-quantification/chemistry.html.

The PCR machines differ in sample capacity, up to 96-well and 384-well standard format, others process only 32 samples and require specialized glass capillaries, excitation method (lasers, others broad spectrum light sources with various filters), and fluorescence acquisition channels. There are also platform-specific
Quantification strategies in kinetic RT-PCR

Two strategies can be performed in real-time RT-PCR: The levels of expressed genes may be measured by absolute or relative quantitative real-time RT-PCR (figure 3). Absolute quantification relates the PCR signal to input copy number using a calibration curve, while relative quantification measures the relative change in mRNA expression levels. The reliability of an absolute real-time RT-PCR assay depends on the condition of ‘identical’ amplification efficiency for both the native target and the calibration curve in RT reaction and in following kinetic PCR (75-77). This problem was also evident in competitive RT-PCR, where an ‘identical’ amplification efficiencies was a necessary precondition for a reliable RT-PCR (10-13, 29, 30). Nowadays a relative quantification is easier to perform than absolute quantification because there is no need for a calibration curve. It is based on the expression levels of a target gene versus a housekeeping gene (reference gene or control gene) and in theory is adequate for most purposes in animal sciences to investigate physiological changes in gene expression levels (4, 69, 72-74). The units used to express relative quantities are irrelevant, and the relative quantities can be compared across multiple real-time RT-PCR experiments (1, 3, 41, 42).

Figure 3: Quantification strategies in real-time PCR (http://www.wzw.tum.de/gene-quantification/).
Absolute quantification

Calibration curves allow the generation of highly specific, sensitive and reproducible data (3, 4, 41, 42, 60). However, the external calibration curve model has to be thoroughly validated as the accuracy of absolute quantification in real-time RT-PCR depends entirely on the accuracy of the standard material. Standard design, production, determination of the exact standard concentration and stability over long storage time is not straightforward and can be problematic. The dynamic range of the calibration curve performed can be up to nine orders of magnitude from $<10^1$ to $>10^{10}$ start molecules, depending on the applied purity of the standard material. The calibration curves used in absolute quantification can be based on known concentrations of DNA standard molecules, e.g. recombinant plasmid DNA (recDNA) (figure 4), in vitro transcribed recombinant RNA (recRNA), genomic DNA, RT-PCR product, commercially synthesized big oligo-nucleotide (3, 4, 41, 42, 83). Stability and reproducibility in kinetic RT-PCR depends on the type of standard used. Cloned recDNA and genomic DNA are very stable and generate highly reproducible standard curves even after a long storage time, in comparison to freshly synthesized recRNA standard material.

![Calibration curve](image)

**Figure 4:** Characteristics of an external androgen receptor (AR) calibration curve on recDNA basis: detection limit, quantification limit and quantification range using SYBR Green I technology (83).

Furthermore, the longer templates derived from recDNA and genomic DNA mimic the average native mRNA length of about 2 kb better than shorter templates derived from RT-PCR product or oligo-nucleotides. They are more resistant against unspecific cleavage and proofreading activity of polymerase during reaction setup and in kinetic PCR (own unpublished results). One advantage of the shorter templates and commercially available templates is an accurate knowledge of its concentration and length. Second shorter templates avoid the very time consuming production process of standard material: standard synthesis,
purification, cloning, transformation, plasmid preparation, linearisation, verification and exact
determination of standard concentration (3, 4, 42).

DNA based calibration curves are subject to the PCR step only, unlike the “unknown mRNA samples” that
must first be reverse transcribed. This increases the potential for variability of the RT-PCR results and the
amplification results may not be strictly comparable with the results from the unknown samples. However,
the problem of sensitivity of the RT-PCR to small variations in the reaction setup is always lurking in the
background as a potential drawback to this simple procedure. Therefore, quantification with external
standards requires careful optimisation of its precision (replicates in the same kinetic PCR run – intra-assay
variation) and reproducibility (replicates in separate kinetic PCR runs – inter-assay variation) in order to
understand the limitations within the given application (3, 4, 60, 70).

A recombinant RNA (recRNA) standard that was synthesized in vitro from a cloned RT-PCR fragment in
plasmid DNA is one option (4, 7, 10-13, 84). However, identical RT efficiency, as well as real-time PCR
amplification efficiencies for calibration curve and target cDNA will have to be tested and confirmed, if the
recDNA is to provide a valid standard for mRNA quantification (4). This is because only the specific
recRNA molecules are present during RT and the kinetics of cDNA synthesis are not like those in native
RNA (the unknown sample) that also contain a high percentage of natural occurring sub-fractions, e.g.
ribosomal RNA (rRNA, ~80%) and transfer RNA (tRNA, 10-15%), shown in figure 1. These missing RNA
sub-fractions can influence the cDNA synthesis rate and in consequence RT efficiency rises and calibration
curves are then overestimated in gene quantification (3, 4, 85, 86). To compensate for background effects
and mimic a natural RNA distribution like in native total RNA, total RNA isolated from bacterial or insect
cell lines, can be used. Alternatively commercially available RNA sources can be used as RNA background,
e.g. poly-A RNA or tRNA, but they do not represent a native RNA distribution over all RNA sub-fractions
(3, 4, 42). Earlier results suggest, that a minimum of RNA background is generally needed and that it
enhances RT synthesis efficiency rate. Low concentrations of recRNA used in calibration curves should
always be buffered with moderate concentrations of unspecific background or carrier RNA, otherwise the
low amounts can be degraded easily by RNase. Very high background concentrations had a more significant
suppression effect in RT synthesis rate and in later real-time PCR efficiency (4).

No matter how accurately the concentration of the standard material is known, the final result is always
reported relatively compared to a defined unit of interest: e.g. copies per defined ng of total RNA, copies per
genome (6.4 pg DNA), copies per cell, copies per gram of tissue, copies per ml blood, etc. (3, 42, 57, 70). If
absolute changes in copy number are important then the denominator still must be shown to be invariable
across the comparison. The quality of gene quantification data cannot be better than the quality of the
denominator itself. Any denominator variation will obscure real changes, produce artificial changes and
wrong quantification results. Careful use of controls is critical to demonstrate that the choice of denominator
was a wise one (42). Under certain circumstances, absolute quantification models can also be normalized
using suitable and unregulated references or housekeeping genes (see normalisation).
Advantages and disadvantages of external standards

External standard quantification is the method of choice for the nucleic acid quantification, independent of any hardware platform used. The specificity, sensitivity, linearity and reproducibility allow absolute and accurate quantification of molecules even in tissues with low mRNA abundance (<100 mRNA molecules) and a detection down to a few molecules (<10 mRNA molecules) (4, 29, 30, 62). The dynamic range of an optimally validated and optimised external standardized real-time RT-PCR assay can accurately detect target mRNA up to nine orders of magnitude or a billion-fold range with high assay linearity (Pearson correlation coefficient; r>0.99) (4, 29, 30, 62). In general a mean intra-assay variation of 10-20% (n = 7x3) and a mean inter-assay variation of 15-30% (n = 7x7) (figure 5) on molecule basis (maximal 2-4% variability on crossing point [CP] basis, respectively) is realistic over the wide dynamic range, as shown for estrogen receptor (ER) kinetic RT-PCR, performed on the LightCycler (83, 87-89).

**Figure 5:** Mean intra-assay (n = 7x3) and inter-assay variation (n = 7x7) of estrogen receptor (ER) alpha real-time RT-PCR assay, using externally standardised recDNA calibration curve (73-75, 83).

At high (> 10^7) and low (< 10^3) template copy input levels the assay variability is higher than in the range between the two (4, 49, 51). At very low copy numbers, under 20 copies per tube, the random variation due to sampling error (Poisson’s error law) becomes significant (59, 62-64). A recDNA calibration curve model can quantify precisely only cDNA molecules derived from the RT step; it says nothing about the conversion to cDNA of the mRNA molecules present in the native total RNA sample. Variability in cDNA synthesis efficiency during reverse transcription must always be kept in mind. Therefore, a recRNA calibration curve model has the advantage that both RNA templates undergo parallel RT and real-time PCR steps. However, a direct comparison suggests that the recDNA quantification model shows higher sensitivity, exhibits a larger quantification range, has a higher reproducibility, and is more stable than the recRNA model (4). Furthermore, recDNA external calibration curves exhibit lower variation (intra-assay variation < 0.7%; inter-assay variation < 2.6% on CP basis) than the recRNA model (< 2.7% and < 4.5%, respectively). Clearly, the
RT step has a profound affect on the overall result obtained from an RT-PCR assay and more thorough consideration of RT efficiency is needed. The main disadvantage of external standards is the lack of internal control for RT and PCR inhibitors. All quantitative PCR methods assume that the target and the sample amplify with similar efficiency (3, 90). The risk with external standards is that some of the unknown samples may contain substances that significantly reduce the efficiency of the PCR reaction in the unknown samples. As discussed, sporadic RT and PCR inhibitors or different RNA/cDNA distributions can occur. A dilution series can be run on the unknown samples and the inhibitory factors can often be diluted out, causing a non-linear standard curve (62, 79).

“Do we need to run a calibration curve in each run?” (62) and “Do we need a calibration curve at all?” (81, 82, 91) are frequently posed questions, together with “What about the reproducibility between the runs?” (http://www.idahotec.com/LightCycler_u/lectures/quantification_on_lc.htm). Repeated runs of the same standard curve give minor variations of a 2-3% in the slope (from which the real-time PCR efficiency is calculated) and about 10% in the intercept of calibration curve. Since the variation in the standard curve correlates with variation in the unknowns, a detection of a 2-fold difference over a wide range of target concentrations is possible (62). The slope of the calibration curve is more reproducible than the intercept, hence only a single standard point will be required to “re-register” a previously performed calibration curve level for the new unknowns. The curve can be imported into any run, as done in the LightCycler software (81, 82). Never changing variations and 100% reproducibility are the big advantages of such a calibration curve import, but there are also disadvantages as variations of reagents, primers and probe (sequence alterations and fluorescence intensity), day-to-day or sample-to-sample variations will not be covered in this “copy and paste” approach. Since these affect PCR efficiency, such an approach can introduce significant errors into the quantification.

Application of absolute quantification

The absolute quantification of specific transcripts was applied in molecular endocrinology. We have examined the tissue specific mRNA expression of estrogen receptors (ER) ERα and ERβ in various bovine tissues using real-time RT-PCR (83, 87-89). Knowledge of the distribution and regulation of ERβ in various tissues of ruminants are missing at this time. The available publications about ERβ expression in bovine species are limited to the cattle reproductive organs and to the sheep hypothalamus (92-95). However, a more detailed study of the tissue distribution of both ER subtypes is essential to continue investigations of their regulation and physiological function. It is well known that steroids lead to an increased synthesis of specific proteins (96) and it is proposed that estradiol can stimulate via ERα its own receptor expression at least in the uterus (97). Goal of the study was to evaluate the deviating tissue sensitivities and the influence of the estrogenic active preparation RALGRO on the tissue specific expression and regulation of both ER subtypes. RALGRO contains Zeranol (α-Zearalanol), a derivative of the mycotoxin Zearalenon, showing
strong estrogenic and anabolic effects, and exhibits all symptoms of hyper-estrogenism in particular reproductive and developmental disorders. Eight heifers were treated over 8 weeks with multiple dose implantations.

Both real-time RT-PCR were ERα and ERβ product specific, and effective PCR amplification kinetic was shown by high PCR efficiency per cycle. Assay sensitivities were confirmed by detection limits down to 10 ssDNA molecules and linear quantification ranges between $10^2$ to $10^9$ molecules. Intra- and inter-assay variation of <19% to <30%, respectively, were determined over the entire quantification range (83, 87-89). The advantage of a high temperature fluorescence acquisition in the 4th segment (60) during the amplification program results in reliable and sensitive ER subtype specific quantification with high linearity (Pearson correlation coefficient; r>0.995) over seven orders of magnitude. High temperature fluorescence acquisition melts the unspecific PCR products at 82°C and 87°C respectively, eliminates the non-specific fluorescence signal derived from primer dimers and ensures an accurate quantification of the desired products. Expression results indicate the existence of both ER subtypes in all 15 investigated tissues and quantified mRNA expression results are shown below as mean values with bi-directional error bars (figure 6.).

**Figure 6:** Tissue specific estrogen receptor (ER) alpha (ERα) and ERβ mRNA expression cluster from 15 heifers. Transcripts were measured on molecule basis with real-time RT-PCR in 25 ng total RNA (n = 8) and quantified results are shown as mean with bi-directional error bars (s.e.m.) (89).
All tissue exhibit a specific ER\(\alpha\) and ER\(\beta\) expression pattern and regulation under RALGRO treatment. With increasing Zeranol concentrations a significant down-regulation of ER\(\alpha\) mRNA expression could be observed in jejunum (\(p<0.001\)) and kidney medulla (\(p<0.05\)). Due to the given results of the mainly non-significant relationship (except in kidney-medulla and jejunum) between Zeranol treatment and ER\(\alpha\) and ER\(\beta\) expression levels all expression data within one tissue were pooled. The derived mean expression concentrations and variations are characteristic for all investigated tissues and the relation of both ER subtypes result in a tissue specific expression cluster. In future, the presented ER\(\alpha\)/ER\(\beta\) cluster with its tissue specific variations will lead us to more detailed studies for a better understanding of estrogen receptor regulation and physiology. Data show high ER\(\alpha\) expression rates in the classical estrogen sensitive tissues and a wide ER\(\alpha\)/ER\(\beta\) ratio (89). ER\(\beta\) transcripts were mainly expressed on lower level. The dominant ER\(\alpha\) expression (ratio of 2.6) could also be shown in various brain regions, including hypothalamus, pituitary gland and 6 different brain lobes (own unpublished results). No expression of ER\(\alpha\) could be observed in bovine endothelial cell preparations, whereas the ER\(\beta\) transcripts were expressed very low abundant (own unpublished results). The dominant role of ER\(\alpha\), especially in uterus explains why it was the first cloned ER as most purification and cloning and attempts were based on uterine tissue (89). Also in the muscular tissues both subtypes could be detected with a very wide expression ratio, especially in muscle parts which are involved in cattle allometric growth like neck and shoulder muscularity. The molecular basis for this sexually dimorphic muscle growth pattern might be attributed to relatively higher sensitivities to sexual steroids in this muscles. Above that, the present study implies that local differences in ER\(\alpha\) expression might be one promoting factor for higher growth velocities and play a key role in allometric growth. The localisation and dominant expression of ER\(\beta\) in both kidney regions and in the jejunum leads to the hypothesis that ER\(\beta\) plays a dominant role in these tissues. ER\(\alpha\) was already detected earlier in the bovine gastrointestinal tract (67), but ER\(\beta\) might be the major actor in the absorptive processes. However, any notations on physiological direct effects of estrogens on gastrointestinal tissues and kidney remain speculative, but there are some indications, that estrogens might influence the calcium transport.

These expression data support the hypothesis, that the ER\(\beta\) may have different biological functions than ER\(\alpha\), especially in kidney and the jejunum (89). In a second study, in ten bovine gastrointestinal tract compartments, using optimised primer pairs of numerous steroid receptors [androgen receptor (AR), ER\(\alpha\), ER\(\beta\) and progesterone receptor (PR)] the data examined above were confirmed (83). It reconfirms the bovine gastrointestinal tract as target of steroid hormone action.

In a further collaborative study with A. Didier and A. Tichopad (76) we investigated the bovine prion mRNA expression using externally calibrated highly sensitive externally standardized real-time RT-PCR with LightCycler instrument. Total RNA was isolated from 16 different tissues, nine regions of the central neural system (CNS) and seven peripheral organs (spleen, lymph node, thymus, liver, muscle, kidney, lung). PrP\(^c\) mRNA copy numbers could be determined in all tissues under study. In agreement with prior studies high
mRNA levels were found in neocortex and cerebellum, calculated on the basis of mg tissue. As shown in figure 7 the lymphatic organs showed at least as high expression levels of prion mRNA as overall in the brain. Lowest expression was detected in kidney. Results of our study provide insight into the involvement of different organs in pathogenesis with respect to prion mRNA expression. LightCycler technology is currently considered the most precise method for nucleic acid quantification and showed to be powerful tool for further studies on prion diseases pathogenesis (80).

![Figure 7: Mean PrP\\(^c\) mRNA copy numbers in various bovine tissue on mg tissue basis (76). Error lines indicate SD.](image)

To quantify all representative members of the somatotropic axis in newborn calves absolute quantification method was used (72). This study was performed in cooperation with J. Blum and T. Mircheva-Georgieva (University of Berne). The somatotropic axis is mainly involved in the growth physiology and tissue differentiation in numerous animal tissues during pre-natal and post-natal development. The factors can act in an endocrine as well as in a para-/autocrine way. To obtain highly accurate and reliable quantitative results in a real-time RT-PCR a highly defined calibration curve is needed. We designed and developed nine different calibration curves, based on recDNA plasmid standards and established them on a constant real-time PCR platform (LightCycler) for the following factors: growth hormone receptor (GHR), insulin-like growth factor (IGF)-1, IGF-1 receptor (IGF-1R), IGF-2, IGF-2 receptor (IGF-2R), insulin receptor (INSR), and IGF-binding proteins (IGF-BP) 1, 2 and 3. The newly elucidated sequences of IGF-2 receptor in cattle
(see appendix) was published in the sequences databases EMBL (European Molecular Biology Library; http://www.ebi.ac.uk/) and GenBank (http://www.ncbi.nlm.nih.gov/Entrez/index.html). Developed assays were applied in the LightCycler system on bovine ileum and liver total RNA and showed high specificity and sensitivity of quantification. All assays had a detection limit of under 35 recombinant DNA molecules present in the capillary (see table in appendix). The SYBR Green I determination resulted in a reliable and accurate quantification with high test linearity (Pearson correlation coefficient r > 0.99) over seven orders of magnitude from <10^2 to >10^8 recombinant DNA start molecules and an assay variation of maximal 5.3%. Applicability of the externally standardised exact quantification method was shown by analysing mRNA levels in newborn calves: mRNA concentrations per gram tissue of mRNA of IGF-1, IGF-1R, IGF-2, IGF-2R, GHR, INSR, and IGFBP-1, -2 and -3 were all different in liver and ileum and the traits exhibited individual differences (72). Further experimental trials and investigation on the somatotropic axis are in progress in Berne (workgroup J. Blum) and Weihenstephan.

Relative quantification

Relative quantification determines the changes in steady-state mRNA levels of a gene across multiple samples and gives a result relative to the levels of an internal control RNA. This reference gene is often a housekeeping gene and can be co-amplified in the same tube in a multiplex assay or can be amplified in a separate tube (figure 3) (62, 73, 74). Therefore, relative quantification does not require standards with known concentrations and the reference can be any transcript, as long as its sequence is known (3, 42). Relative quantification is based on the expression levels of a target gene versus a reference gene and in many experiments is adequate for investigating physiological changes in gene expression levels. To calculate the expression of a target gene in relation to an adequate reference gene various mathematical models are established. Calculations are based on the comparison of the distinct cycle determined by various methods, e.g. crossing points (CP) or threshold values (Ct) at a constant level of fluorescence; or CP acquisition according to established mathematic algorithm. To date several mathematical models that calculate the relative expression ratio have been developed. Relative quantification models without efficiency correction are available and published (equations 1-2) (20, 76, 90).

\[
R = 2^{-\Delta\Delta CP} = 2^{-[\Delta CP_{sample} - \Delta CP_{control}]} \\
R = 2^{-\Delta CP}
\]
equation 1
equation 2

More recently developed mathematical models with kinetic PCR efficiency correction, like in equations 3-6 (81, 82, 98-101), allow for a more precise quantification of the relative expression levels. Further, the
available models allow for the determination of single transcription difference between ONE control and only ONE sample, assayed in triplicates (n =1/3), e.g. LightCycler Relative Quantification Software (42), or Q-Gene (102). For a group wise comparison for up to 100 samples the software application REST and REST-XL (99, 100) are relevant. In REST the relative expression ratio of a target gene is computed, based on its calculated real-time PCR efficiencies (E) or alternatively on the static efficiency of E = 2, and the CP difference (ΔCP) of one unknown sample (treatment) versus one control (ΔCP_control - treatment), shown in equation 3. Using REST and REST-XL the relative calculation procedure is based on the MEAN CP of the experimental groups (equation 4) (http://www.wzw.tum.de/gene-quantification/rest.html).

$$\text{ratio} = \frac{(E_{target})^{\Delta \text{CP}_{target} (\text{control} - \text{sample})}}{(E_{ref})^{\Delta \text{CP}_{ref} (\text{control} - \text{sample})}}$$

equation 3

$$\text{ratio} = \frac{(E_{target})^{\Delta \text{CP}_{target} (\text{MEAN control} - \text{MEAN sample})}}{(E_{ref})^{\Delta \text{CP}_{ref} (\text{MEAN control} - \text{MEAN sample})}}$$

equation 4

In these models the target gene expression is normalized by a non regulated reference gene expression, e.g. derived from classical and frequently described housekeeping genes (53-56). The crucial problem in this relative approach is that the most common reference gene transcripts from “so-called housekeeping genes”, whose mRNA expression can be regulated and whose levels vary significantly with treatment or between individuals (103-106). However, relative quantification can generate useful and biologically relevant information when used appropriately.

$$\text{ratio} = \frac{(E_{ref})^{\text{CP}_{sample}}}{(E_{target})^{\text{CP}_{sample}}} \div \frac{(E_{ref})^{\text{CP}_{calibrator}}}{(E_{target})^{\text{CP}_{calibrator}}}$$

equation 5

$$\text{ratio} = \frac{\text{conc}_{(target\ sample)} / \text{conc}_{(ref\ sample)} * \text{MF}}{\text{conc}_{(target\ cal.)} / \text{conc}_{(ref\ cal.)} * \text{CF}}$$

equation 6

The validity, the variability and the reproducibility of the relative quantification concept using REST was tested in numerous experiments (99, 100). On the basis of the previously published mathematical model (equation 3 and 4) the relative expression ratios on the basis of group means for metallothionein gene versus reference gene GAPDH expression in a zinc deficiency experiment and tests the group ratio results for significance. Herein different cDNA input concentrations were tested (up to 3-fold differences = ± 300%) to
mimic these huge RT variations. It resulted in no significant changes of relative expression ratio evaluated by REST. Also the reproducibility of the developed mathematical model used was given, based on the exact determination of real-time amplification efficiencies and low LightCycler CP variability documented. No significant differences between cDNA starting concentration on expression ratio could be found using relative quantification model (99, 100).

Application of relative quantification

In the following cooperation with W. Windisch (University of Vienna), the influence of Zinc on the relative mRNA expression of zinc transporter proteins was investigated (107). Zinc (Zn) is a small, hydrophobic, highly positive charged ion (Zn2+), which can not cross biological membranes by passive diffusion (108). Therefore, specialized mechanisms are required in the organism for both the Zn uptake and the Zn release. These active transport processes via Zn-binding ligands require energy for Zn transport and their presence can significantly affect Zn transport into the cells. The importance of Zn in cell physiology is related mainly to its intracellular involvement into enzyme catalysis, protein structure, protein-protein interactions, and protein oligo-nucleotide interactions (109-111). The accumulation of Zn in the cell is a sum of influx and efflux processes via transporter proteins (112), like the four Zn transporters (ZnT1 – ZnT4), the divalent cation transporter 1 (DCT1) and of storage processes mainly bound to metallothionein (MT) (113). To study the effect of Zn deficiency on mRNA expression levels adult rats were used as an animal model. Feed intake was restricted to 8 g/d containing 2 µg Zn/g fortified with pure phytate in Zn deficiency rats (114) and 58 µg Zn/g in controls (n=7). At day 1, 2, 4, 7, 11, 16, 22, and 29 of Zn deficiency, 3 animals each were euthanised (n = 24). Zn deficiency was evident from reduced plasma Zn, plasma alkaline phosphatase activity and severe mobilization of Zn from tissue stores (mainly skeleton), while feed intake and body weight remained unaffected. Tissues representing Zn absorption (jejunum, colon), Zn storage and utilization (muscle, liver), and Zn excretion (kidney) were retrieved. Real-time reverse transcription (RT) polymerase chain reaction (PCR) assays were developed and a relative quantification on the basis of GAPDH mRNA expression was applied. Assays allowed a relative and accurate quantification of mRNA molecules with a sufficiently high sensitivity and repeatability. All known Zn transporter subtypes were found to be expressed in the tissues. Expression patterns and reactions to Zn deficiency were specific for the tissue analysed. Expression results imply that some transporters are expressed constitutively, whereas others are highly regulated in tissues responsible for Zn homeostasis. The most distinct changes of expression levels were shown in colon which can therefore be postulated as a highly Zn sensitive tissue. MT was down-regulated in all tissues, in parallel with intracellular Zn status, and is therefore a potent candidate gene for Zn deficiency.

This study provides the first comparative view of gene expression regulation and fully quantitative expression analysis of all known Zn transporters in a non-growing adult rat model. In view of the data provided the developed RT-PCR assay developed herein allows a relative and accurate quantification of Zn transporters and MT mRNA molecules with a sufficiently high sensitivity even for tissues with low mRNA
abundance. The expression results indicate the existence of Zn transporter subtypes in various rat tissues, their different expression pattern and their tissue specific regulation under Zn deficiency treatment. The results show that all transporters and MT have unique expression patterns (figure 8). Colon is a very Zn sensitive tissue in view to the expression results. Expression results imply that some transporters are expressed constitutively, whereas others are highly regulated in tissues responsible for Zn homeostasis. In all tissues MT expression level reflects the intracellular Zn status best. In comparison to control group MT mRNA was down-regulated in all tissues. MT subtype 1 and 2 mRNA expression is a potent candidate as a marker gene for Zn deficiency (107).

Figure 8: Relative metallothionein (MT) expression on the basis of GAPDH expression (= 1.00). The exponential relationship in liver and colon of zinc deficiency and MT mRNA expression is illustrated by the linear regression mean values (n = 3/7) on logarithmic scale. (r = Pearson correlation coefficient; p-value of correlation) (107).

From the same experimental setup (107, 114) of Zn deficiency rats, a microarray screening of 1001 genes (PIQOR™ array) was performed in liver and jejunum, in collaboration with MEMOREC Stoffel GmbH (medical molecular research cologne, Köln, Germany). The PIQOR™ system allows the parallel identification and quantification of thousands of genes from two different samples (e.g. diseased versus normal tissue or samples of physiologically treated animals versus treated animals) (7-9). For verification of candidate genes found in array experiments, quantitative reverse transcription - polymerase chain reaction (RT-PCR) on a real-time platform represents a suitable tool (21, 22, 115). Thus during the recent years, real-
time RT-PCR using SYBR Green I technology is more and more used to quantify physiologically changes in gene expression (3, 4) and to verify gene expression results derived from microarrays (21, 22, 115).

The effect of zinc deficiency on a mRNA expression levels in liver and jejunum of adult rats was analysed. After experimental Zn deficiency, jejunum and liver were retrieved and total RNA was extracted. Tissue specific expression pattern were quantified by microarray analysis and verified via real-time RT-PCR. A relative quantification was performed with the newly developed Relative Expression Software Tool© (100) on 35 candidate genes which showed a highly differential expression. RNAs from liver and jejunum of control samples were reverse transcribed and Cy3 (green fluorescence) labelled, RNAs from Zn deficiency samples were reverse transcribed and Cy5 (red fluorescence) labelled. A false colour overlay from the hybridisation of liver samples on a PIQOR™ array is shown. The spots of the strongest down-regulated genes metallothionein subtype 1 and 2 (MT-1 and MT-2) as well as up-regulated gene interleukins receptor type 6 beta (IL-6R-beta) are indicated in figure 9.

![Image of gene expression pattern](image)

**Figure 9:** Representative example of a gene expression pattern captured as an image of a cDNA-array hybridised with Cy3-labelled control sample (green fluorescence) and Cy5-labelled sample (Zn deficiency in red fluorescence). Each of the 1001 cDNAs is spotted either in quadrant A and B. Four replicates for each cDNAs are spotted, resulting in four A and B quadrants, respectively. A magnification for the most up-regulated (MT-1 and MT-2) and down-regulated (IL-6R-beta) gene transcripts is shown.

From the 1001 genes present on the microarray 457 genes in liver and 566 genes in jejunum were found to be expressed with signal intensities in at least one of the two channels for Cy3 or Cy5 greater than two-fold above negative controls. Genes were combined in classes (0.1-fold expression width), between 1-fold and 2-fold expression ratio and over 2-fold expression in wider ranges. A three parametric Gaussian regression was calculated for each tissue (figure 10) on the bases of a logarithmical conversion (10 log) of the median expression ratios of each class. This resulted in high regression coefficient ($r_{\text{Liver}} = 0.854, p < 0.0001; r_{\text{Jejunum}} = 0.853, p < 0.0001$) and in a normal distribution of both frequency datasets. A 95% confidential interval was calculated for the x-fold expression ratio ($x$), two sided from the mean ($\mu$) according to the Gaussian distribution ($\mu - 1.96 \times \text{standard deviation} < \mu < \mu + 1.96 \times \text{standard deviation}$) and the two sided 2.5%
significance level borders were defined ($p < 0.05$). For liver and jejunum the following 95% confidence intervals were calculated: $-1.82\text{-fold} < x < +1.74\text{-fold}$, $-1.71\text{-fold} < x < +1.61\text{-fold}$, respectively. According to the derived cut offs 85 candidate genes were selected in total (figure 10).

In liver 10 genes were up-regulated higher than 1.74-fold. The mean variation of the calculated expression ratio was 6.3%, calculated from the signal ratios of 4 cDNA spots per array experiment. In liver 23 genes were down-regulated more than 1.82-fold with a variation of 8.6%. In jejunum 25 genes were up-regulated higher than 1.61-fold under zinc deficiency (variation = 4.9%), and 27 genes were down-regulated 1.71-fold with an average variation of 7.1%.

Figure 10: Frequency and level of down- or up-regulation of regulated genes of microarray experiments in liver and jejunum of Zn deficiency rats. Frequency plot of both tissue expression pattern exhibit a three parametric Gaussian distribution ($p <0.0001$). Mean ($\mu$) and boarders of confidential interval are indicated ($\mu \pm 1.96$ times the standard deviation of the Gaussian distribution). Significantly different expressed genes ($p<0.05$) were selected outside the 95% confidential interval. Lines indicate an approximation of 95% interval in liver and jejunum.

The expression results indicate the existence of individual expression pattern in liver and jejunum and their tissue specific regulation under Zn deficiency. In addition, in jejunum a number of B-cell related genes could be demonstrated to be suppressed at Zn deficiency. In liver, MT-1 and MT-2 genes could be shown to be dramatically repressed and therefore represent putative markers for Zn deficiency. Expression results imply
that some genes are expressed constitutively, whereas others are highly regulated in tissues responsible for Zn homeostasis. The combination of microarray analysis and RT-PCR allowed a high effective screening over 1001 genes and afterwards a high quantitative relative quantification of mRNA molecules with high sensitivity and reproducibility.

Data achieved demonstrate that the combination of microarray and real time RT-PCR experiments represents a powerful approach, that summarizes the advantages of both quantification systems - high throughput of the microarray and sensitivity of the real-time RT-PCR. The results demonstrate the feasibility and utility of both methodologies to genome wide exploration of gene expression patterns. But, normalisation in array experiments as well as in kinetic RT-PCR is a general problem and needs high input of further improvements of new concepts, to gain more comparability and reliability in gene expression analysis (21, 22, 115). The expression results indicate the existence of individual expression pattern in liver and jejunum and their tissue specific regulation under Zn deficiency. Jejunum represents a very Zn sensitive tissue with regard to the expression results of immunological relevant genes. Our results imply that some genes are expressed constitutively, whereas others are highly regulated in tissues responsible for Zn homeostasis. Finally, MT subtype 1 and 2 represent potent candidate genes as markers for Zn deficiency.

How to get rid of unspecific real-time PCR artefacts

Real-time assays using SYBR Green I can easily reveal the presence of primer dimers, which are the product of non-specific annealing and primer elongation events (73). These events take place as soon as PCR reagents are combined. During PCR, formation of primer dimers competes with formation of specific PCR product, leading to reduced amplification efficiency and a less successful specific RT-PCR product (116). To distinguish primer dimers from the specific amplicon a melting curve analysis can be performed in numerous available quantification software, shown in figure 11 for the LightCycler platform (81, 82). The pure and homogeneous RT-PCR product produce a single, sharply defined melting curve with a narrow peak (60).

In contrast, the primer dimers melt at relatively low temperatures and have broader peaks, at 80°C (117). To avoid primer dimer formation an intensive primer optimisation is needed, by testing multiple primer pair by cross-wise combinations (71, 118). Multiple optimisation strategies have been developed and are published (60, 119-122). The easiest and most affective way to get rid of any dimer structures, at least during the quantification procedure, is to add an additional 4th segment to the classical three segmented PCR procedure:

- **1st segment** with denaturation at 95°C;
- **2nd segment** with primer annealing at 55-65°C;
- **3rd segment** with elongation at 72°C;
- **4th segment** with fluorescence acquisition at elevated temperatures (51, 64, 103).
Figure 11: Melting curves of insulin-like growth factor-1 (IGF-1) real-time RT-PCR products from multiple species. Melting temperatures of IGF-1 products are between 88.7°C (Callithrix jacchus), 87.7°C (Sus scrofa), 89.9°C (Ovis aries), 90.5°C (Bos taurus) and for primer-dimers are lower than 82°C. The 4th segment during the amplification program melts the unspecific LightCycler PCR products at 85°C and eliminates any non-specific fluorescence signals (51). Callithrix jacchus samples were kindly donated by R Einspanier and A. Einspanier in collaboration with the German Primate Center in Göttingen.

The advantages of a high temperature fluorescence acquisition during amplification are a more reliable quantification with less variation in a wider quantification range (60, 72, 83). The fluorescence acquisition in 4th segment is performed mainly in the range of 80-87°C, eliminates the non-specific fluorescence signals derived by primer dimers or unspecific minor products and ensures accurate quantification of the desired product. High temperature quantification keeps the background fluorescence and the “no template control” fluorescence under 2-3% of maximal fluorescence at plateau (58, 60, 72). Especially in SYBR® Green I determination sensitive transcript quantification with high linearity (correlation coefficient r = 0.99) over eight orders of magnitude (10^2 to 10^9 RNA start molecules) can be achieved (figure 12 B). In contrast, a conventional determination at 72°C results in a truncated quantification range (r = 0.99) over only four orders of magnitude (10^5 to 10^9 RNA start molecules, figure 12 A).
Figure 12: The effect of fluorescence acquisition in 4th segment at elevated temperatures (60). SYBR® Green I acquisition at 72°C in the 3rd segment (figure 12 A, upper graph) and 85°C in the 4th segment (figure 12 B, lower graph) from $10^1$ to $10^9$ recRNA start molecules and one “negative control”. Both online quantifications were done in one and the same LightCycler experiment within the same capillaries.
Real-time PCR amplification efficiency

Individual samples generate different and individual fluorescence histories in kinetic RT-PCR. The shapes of amplification curves differ in the steepness of any fluorescence increase and in the absolute fluorescence levels at plateau depending on background fluorescence levels. The PCR efficiency has a major impact on the fluorescence history and the accuracy of the calculated expression result and is critically influenced by PCR reaction components. Efficiency evaluation is an essential marker in gene quantification procedure (59, 63, 64, 123-125). Constant amplification efficiency in all samples compared is one important criterion for reliable comparison between samples. This becomes crucially important when analysing the relationship between an unknown sequence versus a standard sequence, which is performed in all relative quantification models. In experimental designs employing standardisation with housekeeping genes, the demand for invariable amplification efficiency between target and standard is often ignored, despite the fact that corrections have been suggested (99, 100, 124, 125). A correction for efficiency, as performed in efficiency corrected mathematical models (equation 3-6), is strongly recommended and results in a more reliable estimation of the „real expression ratio“ compared to no efficiency correction. Small efficiency differences between target and reference gene generate false expression ratio, resulting in over- or under-estimation of the „real” initial mRNA amount. Difference in PCR efficiency (ΔE) of 3% (ΔE = 0.03) between target gene and reference gene generate falsely calculated differences in expression of 47% in case of E_{target} < E_{ref} and 209% in case of E_{target} > E_{ref} after 25 performed cycles. This gap will increase dramatically by higher efficiency differences ΔE = 0.05 (28% and 338%, respectively) and ΔE = 0.10 (7.2% and 1083%, respectively) and higher cycle number performed (62, 100). Therefore efficiency corrected quantification corrections should be included in the automation and calculation procedure in relative quantification models (59). The assessment of the exact amplification efficiencies of target and reference genes must be carried out before any calculation of the normalized gene expression is done. LightCycler Relative Expression Software (72, 77), Q-Gene (91), REST and REST-XL software applications (99, 100) allow the evaluation of amplification efficiency plots. A separate determination of real-time PCR efficiency in triplets for every tissue and each performed transcript is necessary (82, 99, 100, 102). Different tissues exhibit different PCR efficiencies, caused by RT inhibitors, PCR inhibitors and by variations in the total RNA fraction pattern extracted (64).

Several methods are described in the literature or were developed by Pfaffl et al. (59, 60, 63, 64) to calculate real-time PCR efficiency (http://www.wzw.tum.de/gene-quantification/efficiency.html):

A) Real-time PCR efficiency calculation from the slopes of the calibration curve, according to the equation: \( E = 10^{\frac{1}{-\text{slope}}} \), as shown in figure 13 (14, 62, 76). Determination of efficiency should be evaluated in a pool of all starting RNA to accumulate all possible „negative impacts“ on kinetic PCR efficiency. Usually, real-time PCR efficiency vary with high linearity (\( r > 0.989 \)) from \( E = 1.60 \) to maximal values up to \( E = 2.10 \) for cDNA input ranges from a few pg to 75 ng cDNA input. Typically, the relationship
between CP and the logarithm of the starting copy number of the target sequence should remain linear for up to five orders of magnitude in the calibration curve as well as in the native sample RNA (4, 72, 99). This calculation method results, in some cases, in efficiencies higher than \( E > 2.0 \), which is practically impossible in the PCR amplification theory. But as shown in given results they are highly reproducible and constant within one transcript and tissue. This probably indicates that this efficiency calculation method is not optimal and overestimates the “real efficiency”.

**Figure 13:** Determination of real-time PCR efficiencies from the slopes of the calibration curve (method A), according to the equation: \( E = 10^{-\frac{1}{\text{slope}}} \) (5, 53, 65) of “reference gene” glutathione transferase (Gst), and two “target genes”: tryptophan operon (TyrA) and aspartate transcarbamylase (PyrB) (99). CP cycles versus cDNA (reverse transcribed total RNA) concentration input (log scale) were plotted to calculate the slope (mean ± SD; \( n = 3 \)).

**B)** Efficiency calculation from the fluorescence increase in 3rd linear phase (as shown in figure 2) of each logarithmic fluorescence history plot. The investigator has to decide which cycle number to include in the analysis and plot an linear regression (similarly to the “Fit Point Method” regression), where the slope of the regression line represents the PCR efficiency. Here efficiencies between \( E = 1.35 \) and \( E = 1.60 \) are realistic and differ dramatically from the results above (60, 58, 126). This efficiency calculation method might underestimate the “real efficiency”, because data evaluation is made in linear phase near the plateau, where reaction trends to get restrictive (58).

**C)** Efficiency calculation on the basis of all fluorescence data points (starting at cycle 1st up to the last cycle), according to a sigmoidal or logistic curve fit model. The advantage of such model is that all data points will be included in the calculation process. No background subtraction is necessary (59, 63, 64, 123-125). Slope value is “nearly” identical to method B and only measured at the point of inflexion at absolute
maximum fluorescence increase of “first derivate maximum” (FDM) \((1.35 < E < 1.60)\). But the derived slope parameters generated by the full sigmoidal or full logistic (figure 14) models are not directly comparable with the „real PCR efficiency“. This method is easy to perform and a good estimator for the maximum curve slope with high correlation coefficient \((r > 0.99)\) and level of significance \((p < 0.001)\) \((59, 63, 64)\).

**D)** Efficiency calculation from the fluorescence increase only in the 2nd real exponential phase, according to a polynomial curve fit, as described earlier \(Y_n = Y_0 (E)^n\), where \(Y_n\) is fluorescence acquired at cycle \(n\), and \(Y_0\) initial fluorescence, so called ground fluorescence \((59, 126-128)\). This phase around the “Second Derivate Maximum” (SDM) exhibit a real exponential amplification behaviour \((59)\) (figure 14). Here in the exponential part the PCR reaction kinetic is still under „full power“ with no restrictions \((62, 63, 64, 82)\). In this method the calculation is performed on each reaction kinetic plot and the amplification efficiency can be determined exactly. They range from \(E = 1.75\) to \(E = 1.90\), hence are between the other methods.

![Figure 14](image)

**Figure 14:** Plot of fluorescence observations from LightCycler (Roche Diagnostics). Forty observations give a sigmoid trajectory that can be described by full data fit (four-parametric logistic model). Ground phase can be well linearly regressed. FDM and SDM denote position of first and second derivate maximum within full data fit.

Which efficiency calculation method (estimation A to D) is „the correct one“ and which one shows the realistic real-time PCR kinetic and thereby is highly reproducible has to be evaluated in further trials.
**Normalisation of expression results**

Data normalisation in real-time RT-PCR is a further major step in gene quantification analysis (3, 42, 99, 100, 129). The reliability of any relative RT-PCR experiment can be improved by including an invariant endogenous control in the assay to correct for sample to sample variations in RT-PCR efficiency and errors in sample quantification. A biologically meaningful reporting of target mRNA copy numbers requires accurate and relevant normalisation to some standard and is strongly recommended in kinetic RT-PCR. But the quality of normalized quantitative expression data cannot be better than the quality of the normalizer itself. Any variation in the normalizer will obscure real changes and produce artifactual changes (3, 41, 42).

Real-time RT-PCR-specific errors in the quantification of mRNA transcripts are easily compounded with any variation in the amount of starting material between the samples, e.g. caused by sample-to-sample variation, variation in RNA integrity, RT efficiency differences and cDNA sample loading variation (26, 27, 42, 99). This is especially relevant when the samples have been obtained from different individuals, different tissues and different time courses, and will result in the misinterpretation of the derived expression profile of the target genes. Therefore, normalisation of target gene expression levels must be performed to compensate intra- and inter-kinetic RT-PCR variations (sample-to-sample and run-to-run variations).

Data normalisation can be carried out against an endogenous unregulated reference gene transcript or against total cellular DNA or RNA content (molecules/g total DNA/RNA and concentrations/g total DNA/RNA). Normalisation according the total cellular RNA content is increasingly used, but little is known about the total RNA content of cells or even about the mRNA concentrations. The content per cell or per gram tissue may vary in different tissues in vivo, in cell culture (in vitro), between individuals and under different experimental conditions. Nevertheless, it has been shown that normalisation to total cellular RNA is the least unreliable method (3, 42, 86, 87). It requires an accurate quantification of the isolated total RNA or mRNA fraction by optical density at 260 nm (43), Agilent Bioanlyser 2100, or Ribogreen RNA Quantification Kit. Alternatively the rRNA content has been proposed as an optimal and stable basis for normalisation, despite reservations concerning its expression levels, transcription by a different RNA polymerase and possible imbalances in rRNA and mRNA fractions between different samples (42, 55, 103, 130-132).

To normalize the absolute quantification according to a single reference gene, a second set of kinetic PCR reactions has to be performed for the invariant endogenous control on all experimental samples and the relative abundance values are calculated for internal control as well as for the target gene. For each target gene sample, the relative abundance value obtained is divided by the value derived from the control sequence in the corresponding target gene. The normalized values for different samples can then directly be compared.

The objective of the following study, in collaboration with D. Schams and B. Berisha, was to demonstrate the mRNA expression of estrogen receptor α (ERα), ERβ and progesterone receptor (PR) in bovine ovarian follicles and in corpus luteum during estrous cycle and pregnancy (87). To quantify steroid receptor transcripts, sensitive and reliable real-time RT-PCR absolute quantification methods were developed.
Expression results were normalised on the basis of GAPDH mRNA expression, which was constant over the experimental trial (estrous cycle and pregnancy).

The investigated ovarian steroid hormones estrogen and progesterone fulfil a number of important functions related to reproduction by endocrine mechanisms of action. In addition to acting as hormones on structures remote from the ovary, the steroids produced by follicle or corpus luteum cells also act locally within the follicles or corpora lutea in which they are produced as paracrine/autocrine agents, acting on or within the cells in which they are produced. The main ovarian events studied for steroid involvement have been folliculogenesis, steroidogenesis, ovulation, corpus luteum formation and function. Estradiol-17β is a most active estrogen in the ovary, and is synthesized and secreted by granulosa cells in antral follicles, especially pre-ovulatory dominant follicles and in corpus luteum.

The mRNA expression of ERα and ERβ mRNA in theca interna tissue (TI) (lower pg/µg RNA) increased continuously and significantly during final growth of follicles, with much higher levels for ERα. The mRNA expression of ERα and ERβ in granulosa cells (GC) (fg/µg RNA) increased continuously during follicle growth (see figures 15 A, B, C).

The expression of mRNA for PR in follicles (lower fg/µg RNA) increased continuously to maximum level in pre-ovulatory follicles with a significant change only in TI. The highest mRNA expression for ERα (fg/µg RNA) was detected in corpus luteum (CL) during the early luteal phase, following by a significant decrease of expression during the mid, late, and regression phases. In contrast, ERβ mRNA expression is relatively high during the early stage, decreased during the late early and mid luteal phase, and increased significantly again during the late luteal phase and after CL regression. During pregnancy, low levels of ERα and ERβ mRNA expression (<25 fg/µg RNA) with no significant changes were measured. No significant alterations in PR mRNA expression (levels <13 fg/µg RNA) during the estrous cycle and pregnancy in bovine CL was found. The results suggest an autocrine and/or paracrine role of steroid receptors in the regulation of final follicle growth and corpus luteum formation and function (87).
Here a central questions arise: “What is the appropriate reference gene for an experimental treatment and investigated tissue?” (3, 42, 133).

Commonly used housekeeping genes, e.g. GAPDH, albumin, actins, tubulins, cyclophilin, microglobulins, 18S rRNA or 28S rRNA (40, 52-54) may be suitable for reference genes, since they are present in all nucleated cell types and are necessary for basic cell survival. The mRNA synthesis of housekeeping genes is considered to be stable in various tissues, even under experimental treatments. However, numerous treatments and studies have already shown that above mentioned housekeeping genes are regulated and vary under experimental conditions (103-106, 134). It remains up to the individual investigator to choose a reference gene that is best for reliable normalisation in their particular experimental setting. In addition, the
endogenous control should be expressed at roughly the same CP level as the target gene (3, 40, 42, 62). At the same CP level, reference and target experience the same condition and real-time RT-PCR kinetics with respect to polymerase activation (heat activation of polymerase), reaction inactivation, stochastic relation between target and primer concentration, and reaction end product inhibition by the generated RT-PCR product.

It has to be emphasized again, that the choice of housekeeping or lineage specific genes is critical. For a number of commonly used reference genes processed pseudogenes have been shown to exist, e.g. β-actin or GAPDH (31-37). These pseudogenes may be responsible for specific amplification products in an mRNA-independent fashion and result in specific amplification even in the absence of intact mRNA (31-33, 91). It is vital to develop universal, artificial, stable, internal standard materials that can be added prior to the RNA preparation to monitor the efficiency of RT as well as the kinetic PCR, respectively (3, 42). Usually more than one housekeeping gene should be tested in a multiple correlation analysis and its behaviour summarized to a housekeeping gene index called BestKeeper© (software tool and publication for Nucleic Acids Research is in preparation). According to this BestKeeper© index, which is based on the weighted expression of at least three housekeeping genes, a more reliable basis of normalisation in relative quantification can be postulated.

The weighted index (using 18S, ubiquitin, GAPDH and beta-actin as housekeeping genes) was calculated for normalisation in three studies, together with R. Bruckmaier, T. Inderwies, T. Neuvians and M. Reist (University of Berne). Subject of the first study was the detection and quantification of mRNA expression of α- and β-adrenergic receptor subtypes in the bovine mammary gland of dairy cows (118). Herein two new subtypes were elucidated in cattle, the β-adrenergic receptor subtype 1B and 2C.

In the second collaborative study together with the University of Berne further nine 5-hydroxytryptamine receptor subtypes were elucidated in cattle intestinal tract (135). Serotonin (5-Hydroxytryptamine) is involved in a wide range of physiological functions and pathological states in humans. There is evidence that serotonergic pathways are also involved in gastrointestinal (GI) motility disorders in ruminants such as abomasal displacement or cecal dilatation/dislocation. This study aimed to develop and validate real-time PCR assays for quantitative mRNA analysis of 5-HT receptor subtypes in bovine tissues. Because the bovine 5-HT receptor nucleotide sequences were completely unknown before in cattle, multiple species (human, mouse, and rat) comparisons of nucleotide sequences were done and primers used for bovine cDNA amplification were derived from human or mouse sequences in highly homologous regions (135).

In the third study, in cooperation with T. Neuvians and D. Schams (136), a normalisation was applied on the weighted housekeeping genes index using 18S, ubiquitin, GAPDH and beta-actin. There a new isoforms of the bovine insulin receptor in the corpus luteum was elucidated in quantitative studies. All newly elucidated sequences (summarised in the appendix) were published in the sequences databases EMBL (European Molecular Biology Library; http://www.ebi.ac.uk/) and GenBank (http://www.ncbi.nlm.nih.gov/Entrez/index.html).
Today there is increasing appreciation of a more reliable normalisation in relative quantification. Recently
the software tool GeNorm was established for the evaluation of housekeeping genes expression levels (40).
GeNorm allows for an accurate normalisation of real-time quantitative RT-PCR data by geometric averaging
of multiple internal control genes (http://allserv.rug.ac.be/~jvdesomp/genorm/). The GeNorm VBA applet
for Microsoft Excel determines the most stable housekeeping genes from a set of 10 tested genes in a given
cDNA sample panel, and calculates a gene expression normalisation factor for each tissue sample based on
the geometric mean of a user defined number of housekeeping genes. The normalisation strategy used in
GeNorm is a prerequisite for accurate kinetic RT-PCR expression profiling, which opens up the possibility of
studying the biological relevance of small expression differences (40).

Real-time PCR data processing

Data evaluation

The next step in reliable gene quantification using real-time RT-PCR is the data evaluation. The calculation
unit in real-time PCR is a sample specific and characteristic crossing points (CP). For CP determination
various fluorescence acquisition methodologies are possible. The “Fit Point Method” and “Threshold Cycle
Method” measure the CP at a constant fluorescence level (76, 81). These constant threshold methods assume
that all samples have the same cDNA concentration at the point were the fluorescence signal significantly
increases in 2nd to 3rd phase over the background fluorescence (figure 2). Measuring the level of
background fluorescence can be a challenge in real-time PCR reactions with significant background
fluorescence variations, caused by drift-ups and drift-downs over the course of the reaction. Averaging over
a drifting background will give an overestimation of variance and thus increase the threshold level (62, 76,
82). The threshold level can be calculated by fitting the intersecting line upon the ten-times value of ground
fluorescence standard deviation. This acquisition mode can be easily automated and is very robust (76). In
the “Fit Point Method” the user discards the uninformative background points, excludes the plateau values
by entering the number of log-linear points, and then fits a log-line to the linear portion of the amplification
curves. These log lines are extrapolated back to a common threshold line and the intersection of the two lines
provides the CP value. The strength of this method is its extreme robustness. The weakness is that it can not
be easily automated and requires a lot of user input (62, 81). “Fit Point Method” or “Threshold Cycle
Method” can be used on all available platforms with different evaluation of background variability.

The problems of defining a constant background for all samples within one run, sample-to-sample
differences in variance and absolute fluorescence values lead to develop a new acquisition modus according
to mathematical algorithms. In the LightCycler software the “Second Derivative Maximum Method” is
performed, where CP is automatically identified and measured at the maximum acceleration of fluorescence
(62, 81). The kinetic fluorescence histories of individual curves are different. They show individual
background variability (1st phase), exponential and linear growth of fluorescence (2nd and 3rd phase), and finally reaction specific plateau values (4th phase), as shown in figure 2. The amplification reaction and the kinetic fluorescence history over various cycles is obviously not a smooth and easy function. The mathematical algorithm on which the “Second Derivative Maximum Method” in Roche Molecular Biochemicals software (82) is based is unpublished. But it is possible to fit sigmoidal- and polynomial-curve models, with high significance (p<0.001) and coefficient of correlation (r>0.99), which can be differentiated and the 2nd derivate maximum can be estimated (59, 63, 64, 124, 125). This increase in the rate of fluorescence increase, or better called the acceleration of the fluorescence signal, slows down at the beginning of the 3rd linear phase. Therefore the cycle where the 2nd derivative is at its maximum is always between 2nd exponential and 3rd linear phase (59).

Automation of the quantification procedure

Automation of quantification with any kind of calibration curve using “Fit Point Method”, “Threshold Cycle Method” or “Second Derivative Maximum Method” needs the input of individual settings, e.g. threshold level, noise band, import an existing standard curve, and the corresponding concentration of the used standard material. However, although relative expression is performed according to several established mathematical models (equations 1-6), up to now relative quantification software has been commercially available only from Roche Molecular Biochemicals “LightCycler Relative Quantification Software” (http://www.LightCycler-online.com/le_sys/soft_ind.htm#quant). The software allows a comparison of maximal triplets (n = 3), of a target versus a calibrator gene, both corrected via a reference gene expression and calculates on the basis of the median of the performed triplets. Real-time PCR efficiency correction is possible within the software and calculated from the calibration curve slope, according to the established equation (62) \[ E = 10^{[-1/slope]} \], ranging from \( E = 1.0 \) (minimum value) to \( E = 2.0 \) (theoretical maximum and efficiency optimum). As shown in equation 6 a given correction factor (CF) and a multiplication factor (MF) have to be attended in the equation calculation process (82).

Recently it was not possible to perform a reliable group-wise calculation of the relative expression ratios and a subsequent statistical comparison of the results by a statistical test with more than three repeats or more than three samples. This has changed with the development of new software tools were established, REST and REST-XL, both Excel®-based and programmed in Visual Basic for Applications (99, 100). Both compare two treatment groups, with multiple data points in sample group versus control group, and calculates the relative expression ratio between them. Four target genes with up to 100 data points can be calculated in REST-XL. The mathematical model used is based on the MEAN crossing point deviation between sample and control group of up to four target genes, normalized by the MEAN crossing point deviation of a reference gene (equation 4). Normalisation via endogenous control can be performed according to the users demand, but it is recommended to compensate intra- and inter-RT-PCR variations (130). Therefore the requirement for high reproducibility of RT and RT efficiency is not „that important“.
any more. The cDNA input concentration variation up to ± 3-fold was evaluated to mimic these huge RT variations and resulted in no significant changes of relative expression ratio (99, 100). Specific amplification efficiencies of four target gene genes can be estimated and included in the correction of the quantification ratio. If no real-time PCR efficiency assessment is performed, REST assumes an optimal efficiency of E = 2.0. The main advantage of the software tool is the subsequent statistical test. REST tests the group differences for significance with a newly developed randomisation test - Pair Wise Fixed Reallocation Randomisation Test©. Variation depends only on CP variation of the investigated transcripts and remains stable between 3% and 12%.

Nevertheless, successful application of real-time RT-PCR and REST depends on a clear understanding of the practical problems. Therefore a coherent experimental design, application and validation of the individual real-time RT-PCR assay remains essential for accurate and fully quantitative measurement of mRNA transcripts (http://www.wzw.tum.de/gene-quantification/rest.html).

Recently a second software tool, named Q-Gene, was developed (102). Q-Gene manages and expedites the planning, performance, and evaluation of quantitative real-time PCR experiments, as well as the mathematical and statistical analysis, storage, and graphical presentation of the data. An efficiency correction is possible. The Q-Gene software application is a tool to cope with complex quantitative real-time PCR experiments at a high-throughput scale (96-well and 384-well format) and considerably expedites and rationalizes the experimental setup, data analysis, and data management while ensuring highest reproducibility (http://www.biotechniques.com/softlib/qgene.html).

**Statistical comparison**

Bio-informatics and bio-statistics on real-time RT-PCR experiment data is a new subject and a new challenge in gene quantification analysis. This is because the coordination of the experiments and the efficient management of the collected data has become an additional major hurdle for kinetic RT-PCR experiments. The main challenge remains the evaluation and the mathematical and statistical analysis of the enormous amount of data gained by this technology, as these functions are not included in the software provided by the manufacturers of the detection systems (76, 81). Normally the statistical data analysis in gene quantification, independent of block, competitive or real-time RT-PCR experiments, are all performed on the basis of classical standard parametric tests, such as analysis of variance or t-tests (137). Parametric tests depend on assumptions, such as normality of distributions, whose validity is unclear. In absolute or relative quantification analysis, where the quantities of interest are derived from ratios and variances can be high, normal distributions might not be expected, and it is unclear how a parametric test could best be constructed (100).

Only two free available software packages support statistical analysis of expression results: Q-Gene (102) and REST (100). The Q-Gene Statistics Add-In is a collection of several VBA programs for the rapid and menu-guided performance of frequently used parametric and nonparametric statistical tests. To assess the
significance level difference between any two groups expression values, it is possible to perform a paired or an unpaired Student’s test, a Mann-Whitney U-test, or Wilcoxon signed-rank test (137). In addition, the Pearson’s correlation analysis can be applied between two matched groups of expression values. Furthermore, all statistical programs calculate the mean values of both groups analysed and their difference in percent.

Permutation or randomisation tests are a useful alternative to more standard parametric tests for analysing experimental data (138, 139). They have the advantage of making no distributional assumptions about the data, while remaining as powerful as more standard tests, and is instead based on one we know to be true: that treatments were randomly allocated (138). The randomisation test is conducted as follows: A statistical test is based on the probability of an effect as large as that observed occurring under the null hypothesis of no treatment effect. If this hypothesis is true, the values in one treatment group were just as likely to have occurred in the other group. The randomisation test repeatedly and randomly reallocates the observed values to the two groups, and notes the apparent effect (expression ratio in REST) each time. The proportion of these effects which are as great as that actually observed in the experiment gives us the p-value of the test (http://www.bioss.ac.uk/smart/unix/mrandt/slides/frames.htm).

The REST software package makes full use of the advantages of a randomisation test (100). In the applied two sided Pair Wise Fixed Reallocation Randomisation Test for each sample, the CP values for reference and target genes are jointly reallocated to control and sample groups (= pair wise fixed reallocation), and the expression ratios are calculated on the basis of the mean values. In practice, it is impractical to examine all possible allocations of data to treatment groups, and a random sample is drawn. If 2000 or more samples are taken, a good estimate of p-value (standard error < 0.005 at p = 0.05) will be obtained (138, 139). Randomisation tests with a pair wise reallocation are seen as the most appropriate approach for this type of application. They are more flexible than non-parametric tests based on ranks (Mann-Whitney, Kruskal-Wallis etc.) And do not suffer a reduction in power relative to parametric tests (t-tests, ANOVA etc.). They can be slightly conservative (i.e. Type I error rates lower than the stated significance level) due to acceptance of randomisations with group differences identical to that observed, but this mainly occurs when used with discrete data, which gene expression data are not, and small sample sizes (138, 139).

**Conclusion**

The recent advances in gene quantification strategies, assay optimisation, assay validation, fluorescence and data processing developed at the Institute of Physiology in Weihenstephan have led to the development of various assays whereby mRNA transcripts can be quantified more precisely in very short time. Numerous gene expression assays were established for a reliable mRNA quantification, including 37 newly elucidated genes mainly in *Bos taurus* and *Ovis aries*. The benefits in terms of increased sensitivity, reduced variability, reduced risk of contamination, increased throughput by automation, and meaningful data interpretation are
obvious. If done properly, kinetic RT-PCR will be the most powerful method for quantifying cellular mRNA on a few molecule level. The quantification strategy used should be designed according to the demand, but must be highly optimised and precisely validated. In the future there is a need for greater standardization of the applied assays to make the expression results comparable between runs, between real-time RT-PCR platforms and between different laboratories worldwide.

The achieved innovations during the last years in gene transcript quantification will help in future to understanding the complex relation between genomics, transcriptomics, and physiology in agricultural, veterinary and medical sciences.

Acknowledgements

I would like to express my acknowledgement of thanks to my supervisor Prof. Dr. H. H. D. Meyer for initiation of this work and the opportunity to work at the Institute of Physiology in Weihenstephan.

I am very grateful to all my colleagues at the institute PD Dr. R. Bruckmaier, Prof. Dr. S. Schams, PD Dr. R. Einspanier, Dr. B. Berisha, Dr. A. Didier, Dr. A. Daxenberger, Dr. I. G. Lange, Dr. S. Wittmann, Dr. M. Hageleit, Dt. T. Neuvians, Dipl. Ing. agr. C. Prgomet, Dipl. Biol. A. Tichopad, Msc. A. Dzidic, for encouraging my research in various ways.

I would like to thank D. Tetzlaff and A. Sachsenhauser for their technical assistance and all the other lab members for the good working atmosphere.

Furthermore the author thanks all external persons for the fruitful collaboration in various animal experiments, laboratory work and intensive discussion over the recent years Prof. Dr. L. Dempfle, Prof. Dr. J. Bauer (Department of Animal Science), Prof. Dr. W. Windisch (University of Vienna), Prof. Dr. J. Blum, Dr. M. Reist, Dr. T. Mircheva-Georgieva, Dr. H. Hammon (University of Berne), Prof. Dr. H. Sauerwein (University of Bonn) and Dr. G. Horgan (BIOSS in Scotland).
References

31. Mose, M., and Galliwitz, D., Structure of two human beta-actin related processed genes, one of which is located next to a simple repetitive sequence. EMBO, 2: 757–761, 1983.


