Guidelines for PCR

The invention of the polymerase chain reaction (PCR) by K. Mullis and co-workers in 1985 revolutionized molecular biology and molecular medicine. Major research areas, such as biomarker discovery, gene regulation, and cancer research, are challenging today's PCR technologies with more demanding requirements. These include the need for increased throughput, higher assay sensitivity, and reliable data analysis. Assay development and evaluation, reproducibility of data, and time to result are still major problems encountered by researchers.

PCR amplification is performed routinely and thousands of PCR protocols have been developed, yet researchers still encounter technical difficulties with PCR experiments and often fail to obtain specific amplification products. Although there are several different challenges (e.g., smearing, low yield, and nonspecific amplification), there are two main reasons for PCR failure or poor results: the specificity of the reaction and template secondary structure.

PCR is both a thermodynamic and an enzymatic process. Successful real-time PCR requires amplification and detection under optimal conditions and each reaction component can affect the result. The annealing step is critical for high PCR specificity. When primers anneal to the template with high specificity, this leads to high yields of specific PCR products and increases the sensitivity of the amplification reaction. However, due to the high primer concentration in the reaction, primers will also hybridize to non-complementary sequences with mismatches. If the primers anneal to the template sequence with low specificity, amplification of nonspecific PCR products and primer–dimers may occur. Competition in the amplification reaction between these artifacts and the desired PCR product may reduce the yield of the specific product, thereby reducing the sensitivity and linear range of the real-time reaction. Low PCR specificity can significantly affect quantitative PCR particularly when using SYBR Green for detection. As SYBR Green binds to any double-stranded DNA sequence, primer–dimers and other nonspecific PCR products will generate a fluorescent signal. This reduces the overall sensitivity of the assay and also leads to inaccurate quantification of the transcript of interest. Factors critical for high specificity in PCR include primer design and the reaction chemistry used.

Back to top

PCR primer design

Optimal primer sequences and appropriate primer concentrations are essential for maximal specificity and efficiency in PCR. The table, Guidelines for the design and use of primers provides an overview of primer design and use for standard and multiplex PCR, as well as one-step RT-PCR. Molar conversions can be found in the table Molar conversions for PCR primers.

Guidelines for the design and use of primers	Standard PCR	Multiplex PCR	One-step RT-PCR
Length	18–30 nt	21–30 nt	18–30 nt
GC content	40–60%	40–60%	40–60%
T_m information	The T_m of all primer pairs should be similar	The T_m of all primer pairs should be similar. For optimal results, the $T_m$should be 60–88°C	The T_m of all primer pairs should be similar. The T_m should not be lower than the temperature of the reverse transcription (e.g., 50°C)
Estimating optimal annealing temperature	5°C below the calculated T_m	5–8°C below the calculated T_m (when greater than 68°C) or 3–6°C below the calculated T_m (when 60–67°C)	5°C below the calculated T_m

Location	–	–	To prevent detection of gDNA: Primer hybridizes to the 3' end of one exon and the 5' end of the adjacent exon. Alternatively, the primer hybridizes to a flanking region that contains at least one intron. If only the mRNA sequence is known, choose primer annealing sites that are 300–400 bp apart.
Concentration, A_{260} unit equivalence	20–30 µg	20–30 µg	20–30 µg

Molar conversions for PCR primers Primer length	pmol/µg	20 pmol
18mer	168	119 ng
20mer	152	132 ng
25mer	121	165 ng
30mer	101	198 ng

The following points should be considered when designing PCR primers and are common to all types of PCR:

- T_m calculation: 2°C x (A+T) + 4°C x (G+C)
- Avoid complementarity in the 2–3 bases at the 3' end of the primer pairs
- Avoid mismatches between the 3' end of the primer and the template
- Avoid runs of 3 or more Cs or Gs at the 3' end of the primer
- Avoid complementarity within primers and between the primer pair
- Avoid a T as ultimate base at the 3' end
- Ensure primer sequence is unique for the template sequence
- Use a concentration of 0.1–1.0 µM of each primer. For many applications, a primer concentration of 0.2 µM will be sufficient

Lyophilized primers should be dissolved in a small volume of distilled water or TE to make a concentrated stock solution. Prepare small aliquots of working solutions containing 10 pmol/µl to avoid repeated thawing and freezing. Store all primer solutions at –20°C. Primer quality can be checked on a denaturing polyacrylamide gel; a single band should be seen.

Back to top

PCR conditions

The primer and Mg^{2+} concentration in the PCR buffer and annealing temperature of the reaction may need to be optimized for each primer pair for efficient PCR. In addition, PCR efficiency can be improved by additives that promote DNA polymerase stability and processivity or increase hybridization stringency, and by using strategies that reduce nonspecific primer–template interactions (1). Use of high-purity reagents is also essential for successful PCR, especially for amplification of rare templates, for example, single copy genes in genomic DNA or pathogenic viral

DNA sequences in genomic DNA isolated from an infected organism.
Inclusion of control reactions is essential for monitoring the success of PCR reactions. Wherever possible, a positive control should be included to check that the PCR conditions used can successfully amplify the target sequence. As PCR is extremely sensitive, requiring only a few copies of target template, a negative control containing no template DNA should always be included to ensure that the solutions used for PCR have not become contaminated with the template DNA.

PCR setup should be performed in a separate area from PCR analysis to ensure that reagents used for PCR do not become contaminated with PCR products. Similarly, pipets used for analysis of PCR products should never be used for setting up PCR.

Primer annealing specificity and PCR buffers

In PCR, annealing occurs between the primers and complementary DNA sequences in the template. Primer annealing must be specific for successful amplification. Due to the high concentration of primers necessary for efficient hybridization during short annealing times, primers can anneal to non-complementary sequences. Amplification of products from nonspecific annealing competes with specific amplification and may drastically reduce the yield of the specific product.

The success of PCR largely depends on maintaining a high ratio of specific to nonspecific annealing of the primer molecules. Annealing is primarily influenced by the components of the PCR buffer (in particular the cations) and annealing temperature. Special cation combinations can maintain high primer annealing specificity over a broad range of annealing temperatures. This eliminates the need for optimization of annealing temperatures for each individual primer–template system and also allows the use of non-ideal PCR assays with different primer annealing temperatures.

Annealing temperature

The optimal primer annealing temperature is dependent on the base composition (i.e., the proportion of A, T, G, and C nucleotides), primer concentration, and ionic reaction environment.

Magnesium ion concentration

Magnesium ions are a critical DNA polymerase cofactor necessary for enzyme activity. Mg^{2+} binds to DNA, primers, and nucleotides contained in the amplification reaction. The Mg^{2+} concentration is generally higher than that of dNTPs and primers, and some optimization may be necessary for different template and primer concentrations. A higher than optimal concentration of Mg^{2+} can stabilize nonspecific binding and is often indicated by decreased yields of specific PCR products and the appearance of background smear or other PCR artifacts.

PCR additives

Various PCR additives or enhancers are available for improving PCR results. It is claimed that these reagents relieve secondary DNA structure (e.g., in GC-rich regions or in long amplification products), lower template melting temperature, enhance enzyme processivity, stabilize DNA polymerases, or prevent attachment of polymerases to plasticware.

Commonly used PCR additives include dimethyl sulfoxide (DMSO), bovine serum albumin (BSA), and glycerol.

Back to top

Guidelines for degenerate primer design and use

PCR primer sequences are often deduced from amino acid sequences if the exact nucleotide sequence of their target is unknown. However, because of the degeneracy of the genetic code, the deduced sequences may vary at one or more positions. A common solution in these cases is to use a degenerate primer, which is a mixture of similar primers that have different bases at the variable positions. Using degenerate primers can lead to difficulties optimizing PCR assays: within a degenerate primer mixture only a limited number of primer molecules are complementary to the template; the melting temperature (T_m) of primer sequences may vary significantly; and the sequences of some primers can be complementary to those of others. For these reasons, amplification conditions are required that

minimize nonspecific primer–template and primer–primer interactions. The following guidance may help when designing and using degenerate primers.

Primer sequence:

- Avoid degeneracy in the 3 nucleotides at the 3' end, i.e., if possible use Met- or Trp-encoding triplets at the 3' end
- To increase primer–template binding efficiency, reduce degeneracy by allowing some mismatches between the primer and template, especially towards the 5' end, but not the 3' end
- Try to design primers with less than 4-fold degeneracy at any given position.

Primer concentration:

- Begin PCR with a primer concentration of 0.2 µM
- In case of poor PCR efficiency, increase primer concentrations in increments of 0.25 µM until satisfactory results are obtained

Back to top

Amplification of long PCR products

Amplification of PCR products longer than 3–4 kb is often compromised by nonspecific primer annealing, suboptimal cycling conditions, and secondary structures in the DNA template. Lengthy optimization is often necessary, by varying factors such as cycling conditions, primer and dNTP concentrations, and special additives.

Optimizing cycling conditions

While depurination is usually not a problem in standard PCR, it can significantly influence the amplification of longer PCR fragments. This is because longer templates are proportionally more depurinated than shorter ones. For this reason, very short denaturation steps of only 10 seconds give higher yields and no background smearing compared to denaturation steps of 30 seconds or 1 minute (which leads to PCR failure; see figure Effect of cycling conditions). Extensive depurination is also observed during the final extension step. Therefore, using a lower extension temperature of 68°C instead of 72°C dramatically improves yield of longer amplification products.

Ideal cycling conditions for longer PCR products are given in the table Cycling conditions for amplifying longer PCR products.

* 5°C below Tm of primers.

Cycling conditions for amplifying longer PCR products Step	Time/cycles	Temperature
Initial activation step	2 min	95°C
3-step cycling		
Denaturation	10 s	94°C
Annealing	1 min	50–68°C*
Extension	1 min/kb	
Number of cycles	40 cycles	68°C
End of PCR cycling	Indefinite	4°C

Optimizing PCR additives

Secondary structures such as hairpin loops, which are often caused by GC-rich template stretches, interfere with efficient amplification of long PCR products. This problem can be overcome by adding reagents that modify the melting behavior of DNA to help resolve secondary structures at lower temperatures.

Optimizing 3' to 5' exonuclease activity

Taq DNA Polymerase introduces more errors into the PCR product while copying the template than do so-called proofreading DNA polymerases. Once a mismatch occurs during synthesis, *Taq* DNA polymerase will either extend the mismatched strand or fall off the template strand, leading to mutated or incomplete PCR products, respectively. Although this does not generally affect PCR efficiency when amplifying shorter PCR fragments, amplification of longer PCR products can be significantly impaired by mismatches introduced during DNA synthesis.

Proofreading DNA polymerases contain an inherent 3' to 5' exonuclease activity that removes base-pair mismatches. Adding a small amount of a proofreading DNA polymerase to the PCR mixture therefore significantly improves the amplification efficiency of longer PCR products.

Back to top

Enzymes used in PCR

Several types of thermostable DNA polymerases are available for use in PCR, providing a choice of enzymatic properties, see table DNA polymerases used in PCR.

Taq DNA polymerase, isolated from the eubacterium Thermus aquaticus, is the most commonly used enzyme for standard end-point PCR. The robustness of this enzyme allows its use in many different PCR assays. However, as this enzyme is active at room temperature, it is necessary to perform reaction setup on ice to avoid nonspecific amplification.

A number of modifications of the original “PCR polymerase” — *Taq* DNA polymerase — are now available for different downstream application needs, such as hot-start, single-cell, high-fidelity, or multiplex PCR. With an average error rate of 1 in 10,000 nucleotides, *Taq* DNA polymerase and its variants are less accurate than thermostable enzymes of DNA polymerase family B. However, due to its versatility, *Taq* DNA polymerase is still the enzyme of choice for most routine applications and when used with a stringent hot-start, is suitable for several challenging PCR applications.

Adapted from reference 2.

DNA polymerases used in PCR Enzyme properties	DNA polymerase family A	DNA polymerase family B
Available enzymes	*Taq* DNA polymerase	Proofreading enzymes
5'–3' exonuclease activity	+	–
3'–5' exonuclease activity	–	+
Extension rate (nucleotides/second)	~150	~25
Error rate (per bp/per cycle)	1 in 10^3/10^4	1 in 10^5/10^6
PCR applications	Standard, hot-start, reverse transcription, real-time	High fidelity, cloning, site-directed mutagenesis
A-addition	+	Sometimes

Hot-start DNA polymerase

When amplification reaction setup is performed at room temperature, primers can bind nonspecifically to each other, forming primer–dimers. During amplification cycles, primer–dimers can be extended to produce nonspecific products, which reduces specific product yield. For more challenging PCR applications, the use of hot-start PCR is crucial for successful specific results. To produce hot-start DNA polymerases, *Taq* DNA polymerase activity can be inhibited at lower temperatures with antibodies or, more effectively, with chemical modifiers that form covalent bonds with amino acids in the polymerase. The chemical modification leads to complete inactivation of the polymerase until the covalent bonds are broken during the initial heat activation step. In contrast, in antibody-mediated hot-start procedures, antibodies bind to the polymerase by relatively weak non-covalent forces, which leaves some polymerase molecules in their active state. This sometimes leads to nonspecific primer extension products that can be further amplified during PCR. These products appear as smearing or incorrectly sized fragments when run on an agarose gel.

High-fidelity DNA polymerase

Unlike standard DNA polymerases (such as *Taq* DNA polymerase), high-fidelity PCR enzymes generally provide a 3' to 5' exonuclease activity for removing incorrectly incorporated bases. High-fidelity PCR enzymes are ideally suited to applications requiring a low error rate, such as cloning, sequencing, and site-directed mutagenesis. However, if the enzyme is not provided in a hot-start version, the 3' to 5' exonuclease activity can degrade primers during PCR setup and the early stages of PCR. Nonspecific priming caused by shortened primers can result in smearing on a gel or amplification failure — especially when using low amounts of template. It should be noted that the proofreading function often causes high-fidelity enzymes to work more slowly than other DNA polymerases. In addition, the A-addition function required for direct UA- or TA-cloning is strongly reduced, resulting in the need for blunt-end cloning with lower ligation and transformation efficiency.

Back to top

PCR cycling

In theory, each PCR cycle doubles the amount of amplicon in the reaction. Therefore, 10 cycles multiply the amplicon by a factor of ~1000 and so on.

Each PCR cycle consists of template denaturation, primer annealing and primer extension. If the temperatures for annealing and extension are similar, these two processes can be combined. Each stage of the cycle must be optimized in terms of time and temperature for each template and primer pair combination.

After the required number of cycles has been completed (see table Guidelines for determining the number of PCR cyclesfor further information), the amplified product may be analyzed or used in downstream applications.

Guidelines for determining the number of PCR cycles Amount of 1 kb DNA fragment	Amount of *E. coli*DNA	Amount of human DNA	Number of single-copy targets	Number of PCR cycles
0.0.1–0.11 fg	0.05–0.56 pg	36–360 pg	10–100	40–45
0.11–1.1 fg	0.56–5.56 pg	0.36–3.6 ng	100–1000	35–40
1.1–5.5 fg	5.56–278 pg	3.6–179 ng	1×10^3–5×10^4	30–35
>5.5 fg	>278 pg	>179 ng	$>5 \times 10^4$	25–35

Back to top

Commonly used terms in PCR

Basic terms used in data analysis are given below. For more information on data analysis, refer to the recommendations from the manufacturer of your real-time cycler. Data are displayed as sigmoidal-shaped

amplification plots (when using a linear scale), in which the fluorescence is plotted against the number of cycles (see figure Typical amplification plot).

Before levels of nucleic acid target can be quantified in real-time PCR, the raw data must be analyzed and baseline and threshold values set. When different probes are used in a single experiment (e.g., when analyzing several genes in parallel or when using probes carrying different reporter dyes), the baseline and threshold settings must be adjusted for each probe.

Furthermore, analysis of different PCR products from a single experiment using SYBR Green detection requires baseline and threshold adjustments for each individual assay.

Baseline: The baseline is the noise level in early cycles, typically measured between cycles 3 and 15, where there is no detectable increase in fluorescence due to amplification products. The number of cycles used to calculate the baseline can be changed and should be reduced if high template amounts are used or if the expression level of the target gene is high (see figure Baseline and threshold settings). To set the baseline, view the fluorescence data in the linear scale amplification plot. Set the baseline so that growth of the amplification plot begins at a cycle number greater than the highest baseline cycle number. The baseline needs to be set individually for each target sequence. The average fluorescence value detected within the early cycles is subtracted from the fluorescence value obtained from amplification products. Recent versions of software for various real-time cyclers allow automatic, optimized baseline settings for individual samples.

Background: This refers to nonspecific fluorescence in the reaction, for example, due to inefficient quenching of the fluorophore or the presence of large amounts of double-stranded DNA template when using SYBR Green. The background component of the signal is mathematically removed by the software algorithm of the real-time cycler.

Reporter signal: Fluorescent signal that is generated during real-time PCR by either SYBR Green or a fluorescently labeled sequence-specific probe.

Normalized reporter signal (Rn): This is the emission intensity of the reporter dye divided by the emission intensity of the passive reference dye measured in each cycle.

Passive reference dye: On some real-time cyclers, the fluorescent dye ROX serves as an internal reference for normalization of the fluorescent signal. It allows correction of well-to-well variation due to pipetting inaccuracies, well position, and fluorescence fluctuations.

Threshold: The threshold is adjusted to a value above the background and significantly below the plateau of an amplification plot. It must be placed within the linear region of the amplification curve, which represents the detectable log-linear range of the PCR. The threshold value should be set within the logarithmic amplification plot view to enable easy identification of the log-linear phase of the PCR. If several targets are used in the real-time experiment, the threshold must be set for each target.

Threshold cycle (C_T) or crossing point (Cp): The cycle at which the amplification plot crosses the threshold (i.e., there is a significant detectable increase in fluorescence). C_T can be a fractional number and allows calculation of the starting template amount.

ΔC_T value: The ΔC_T value describes the difference between the C_T value of the target gene and the C_T value of the corresponding endogenous reference gene, such as a housekeeping gene, and is used to normalize for the amount of template used:

- $\Delta C_T = C_T$ (target gene) – C_T (endogenous reference gene)
- $\Delta\Delta C_T$ value: The $\Delta\Delta C_T$ value describes the difference between the average ΔC_T value of the sample of interest (e.g., stimulated cells) and the average ΔC_T value of a reference sample (e.g., unstimulated cells). The reference sample is also known as the calibrator sample and all other samples will be normalized to this when performing relative quantification:
- $\Delta\Delta C_T$ = average ΔC_T (sample of interest) – average ΔC_T (reference sample)

Endogenous reference gene: This is a gene whose expression level should not differ between samples, such as a housekeeping gene (3). Comparing the C_T value of a target gene with that of the endogenous reference gene allows normalization of the expression level of the target gene to the amount of input RNA or cDNA (see above section about ΔC_Tvalue). The exact amount of template in the reaction is not determined. An endogenous reference gene corrects for possible RNA degradation or presence of inhibitors in the RNA sample, and for variation in RNA content, reverse-transcription efficiency, nucleic acid recovery, and sample handling. For selection of the optimal reference gene(s), algorithms have been developed which allow the choice of the optimal reference, dependent on the experimental set-up (4).

Internal control: This is a control sequence that is amplified in the same reaction as the target sequence and detected with a different probe (i.e., duplex PCR is carried out). An internal control is often used to rule out failure of amplification in cases where the target sequence is not detected.

Calibrator sample: This is a reference sample used in relative quantification (e.g., RNA purified from a cell line or tissue) to which all other samples are compared to determine the relative expression level of a gene. The calibrator sample can be any sample, but is usually a control (e.g., an untreated sample or a sample from time zero of the experiment).

Positive control: This is a control reaction using a known amount of template. A positive control is usually used to check that the primer set or primer–probe set works and that the reaction has been set up correctly.

No template control (NTC): This is a control reaction that contains all essential components of the amplification reaction except the template. This enables detection of contamination due to contaminated reagents or foreign DNA, e.g., from previous PCRs.

No RT control: RNA preparations may contain residual genomic DNA, which may be detected in real-time RT-PCR if assays are not designed to detect and amplify RNA sequences only. DNA contamination can be detected by performing a no RT control reaction in which no reverse transcriptase is added.

Standard: This is a sample of known concentration or copy number used to construct a standard curve.

Standard curve: To generate a standard curve, C_T values/crossing points of different standard dilutions are plotted against the logarithm of input amount of standard material. The standard curve is commonly generated using a dilution series of at least 5 different concentrations of the standard. Each standard curve should be checked for validity, with the value for the slope falling between –3.3 to –3.8. Standards are ideally measured in triplicate for each concentration. Standards which give a slope differing greatly from these values should be discarded.

Efficiency and slope: The slope of a standard curve provides an indication of the efficiency of the real-time PCR. A slope of –3.322 means that the PCR has an efficiency of 1, or 100%, and the amount of PCR product doubles during each cycle. A slope of less than –3.322 (e.g., –3.8) is indicative of a PCR efficiency <1. Generally, most amplification reactions do not reach 100% efficiency due to experimental limitations. A slope of greater than –3.322 (e.g., –3.0) indicates a PCR efficiency which appears to be greater than 100%. This can occur when values are measured in the nonlinear phase of the reaction or it can indicate the presence of inhibitors in the reaction.

The efficiency of a real-time PCR assay can be calculated by analyzing a template dilution series, plotting the C_T values against the log template amount, and determining the slope of the resulting standard curve. From the slope (S), efficiency can be calculated using the following formula: PCR efficiency (%) = (10(–1/S) – 1) x 100

Back to top

Real-time PCR

Real-time PCR and RT-PCR (also known as quantitative or qPCR) allow accurate quantification of starting amounts of DNA, cDNA, and RNA targets. Fluorescence is measured during each cycle, which greatly increases the dynamic range of the reaction, since the amount of fluorescence is proportional to the amount of PCR product. PCR products can be detected using either fluorescent dyes that bind to double-stranded DNA or fluorescently labeled sequence-specific probes.

Back to top

What is SYBR Green PCR?

The fluorescent dye SYBR Green I binds all double-stranded DNA molecules, emitting a fluorescent signal of a defined wavelength on binding (see figure SYBR Green principle). The excitation and emission maxima of SYBR Green I are at 494 nm and 521 nm, respectively, allowing use of the dye with any real-time cycler. Detection takes place in the extension step of real-time PCR. Signal intensity increases with increasing cycle number due to the accumulation of PCR product. Use of SYBR Green enables analysis of many different targets without having to synthesize target-specific labeled probes. However, nonspecific PCR products and primer–dimers will also contribute to the fluorescent signal. Therefore, high PCR specificity is required when using SYBR Green.

Back to top

What is probe-based PCR?

Fluorescently labeled probes provide a highly sensitive method of detection, as only the desired PCR product is detected. However, PCR specificity is also important when using sequence-specific probes. Amplification artifacts such as nonspecific PCR products and primer–dimers may also be produced, which can result in reduced yields of the desired PCR product. Competition between the specific product and reaction artifacts for reaction components can compromise assay sensitivity and efficiency. The following probe chemistries are frequently used.

TaqMan probes: sequence-specific oligonucleotide probes carrying a fluorophore and a quencher moiety. The fluorophore is attached at the 5' end of the probe and the quencher moiety is located at the 3' end. During the combined annealing/extension phase of PCR, the probe is cleaved by the 5'–3' exonuclease activity of Taq DNA polymerase, separating the fluorophore and the quencher moiety. This results in detectable fluorescence that is proportional to the amount of accumulated PCR product.

FRET probes: PCR with fluorescence resonance energy transfer (FRET) probes uses 2 labeled oligonucleotide probes that bind to the PCR product in a head-to-tail fashion. When the 2 probes bind, their fluorophores come into close proximity, allowing energy transfer from a donor fluorophore to an acceptor fluorophore. Therefore, fluorescence is detected during the annealing phase of PCR and is proportional to the amount of PCR product. As the FRET system uses 2 primers and 2 probes, good design of the primers and probes is critical for successful results.

Dyes used for fluorogenic probes in real-time PCR: For real-time PCR with sequence-specific probes, various fluorescent dyes are available, each with its own excitation and emission maxima (see table Dyes commonly used for quantitative, real-time PCR. The wide variety of dyes makes multiplex, real-time PCR possible (detection of 2 or more amplicons in the same reaction), provided the dyes are compatible with the excitation and detection capabilities of the real-time cycler used, and the emission spectra of the chosen dyes are sufficiently distinct from one another. Therefore, when carrying out multiplex PCR, it is best practice to use dyes with the widest channel separation possible to avoid any potential signal crosstalk.

Other probes: Many probe suppliers have developed their own proprietary dyes. For further information, please refer to the web pages of the respective suppliers.

* Emission spectra may vary depending on the buffer conditions.

Dyes commonly used for quantitative, real-time PCR Dye	Excitation maximum (nm)	Emission maximum (nm)*
Fluorescein	490	513
Oregon Green	492	517
FAM	494	518

SYBR Green I	494	521
TET	521	538
JOE	520	548
VIC	538	552
Yakima Yellow	526	552
HEX	535	553
Cy3	552	570
Bodipy TMR	544	574
NED	546	575
TAMRA	560	582
Cy3.5	588	604
ROX	587	607
Texas Red	596	615
LightCycler Red 640 (LC640)	625	640
Bodipy 630/650	625	640
Alexa Fluor 647	650	666
Cy5	643	667
Alexa Fluor 660	663	690
Cy 5.5	683	707

Back to top

PCR quantification

Target nucleic acids can be quantified using either absolute quantification or relative quantification.

Absolute quantification determines the absolute amount of target (expressed as copy number or concentration), whereas relative quantification determines, as the first step of analysis, the ratio between the amount of target and the amount of a control (e.g., an endogenous reference molecule, usually a suitable housekeeping gene). Subsequently, this normalized value can then be used to compare, for example, differential gene expression in different samples.

What is absolute quantification?

Use of external standards enables the level of a gene to be given as an absolute copy number. For gene expression analysis, the most accurate standards are RNA molecules of known copy number or concentration. Depending on the sequence and structure of the target and the efficiency of reverse transcription, only a proportion of the target RNA in the RNA sample will be reverse transcribed. The cDNA generated during reverse transcription then serves as template in the subsequent real-time PCR. The use of RNA standards takes into account the variable efficiency of reverse transcription.

A standard curve (plot of C_T values/crossing points of different standard dilutions against log of amount of standard) is generated using a dilution series of at least 5 different concentrations of the standard (see figure, Absolute quantification). The amount of unknown target should fall within the range tested. Amplification of the standard dilution series and of the target sequence is carried out in separate wells. The C_T values of the standard samples are determined. Then, the C_T value of the unknown sample is compared with the standard curve to determine the amount of target in the unknown sample. It is important to select an appropriate standard for the type of nucleic acid to be quantified. The copy number or concentration of the nucleic acids used as standards must be known. In addition, standards should have the following features:

- Primer and probe binding sites identical to the target to be quantified

- Sequence between primer binding sites identical or highly similar to the target sequence
- Sequences upstream and downstream from the amplified sequence identical or similar to the “natural” target
- Equivalent amplification efficiencies of standard and target molecules

RNA standards for absolute quantification

RNA standards can be created by cloning part or all of the transcript of interest into a standard cloning vector. The insert can be generated by RT-PCR from total RNA or mRNA, or by PCR from cDNA. The cloning vector must contain an RNA polymerase promoter such as T7, SP6, or T3. Ensure that in vitro transcription of the insert leads to generation of the sense transcript. After in vitro transcription, plasmid DNA must be removed completely with RNase-free DNase, since residual plasmid DNA will lead to errors in spectrophotometric determination of RNA concentration and will also serve as a template in the subsequent PCR. Furthermore, ensure that the RNA used as a standard does not contain any degradation products or aberrant transcripts by checking that it migrates as a single band in gel or capillary electrophoresis.

After determination of RNA concentration by spectrophotometry, the copy number of standard RNA molecules can be calculated using the following formula:

(X g/µl RNA / [transcript length in nucleotides x 340]) x 6.022 x 10^{23} = Y molecules/µl

An alternative to the use of in vitro transcripts as RNA standards is the use of a defined RNA preparation (e.g., from a cell line or virus preparation), for which the absolute concentration of the target has already been determined.

DNA standards for absolute quantification

Plasmid DNA: The most convenient way to create a DNA standard is to clone a PCR product into a standard vector. Advantages of this method are that large amounts of standard can be produced, its identity can be verified by sequencing, and the DNA can easily be quantified by spectrophotometry. Plasmid standards should be linearized upstream or downstream of the target sequence, rather than using supercoiled plasmid for amplification. This is because the amplification efficiency of a linearized plasmid often differs from that of the supercoiled conformation and more closely simulates the amplification efficiency of genomic DNA or cDNA.

After spectrophotometric determination of plasmid DNA concentration, the copy number of standard DNA molecules can be calculated using the following formula:

(X g/µl DNA / [plasmid length in base pairs x 660]) x 6.022 x 10^{23} = Y molecules/µl

PCR fragment: A PCR product containing the target sequence can also be used as a DNA standard. We recommend including at least 20 bp upstream and downstream of the primer binding sites of the amplicons. Copy number is calculated using the formula for plasmid DNA (see above), replacing “plasmid length” with the length of the PCR product.

Genomic DNA: If the target of interest is present in only 1 copy per haploid genome and amplification of pseudogenes and/or closely related sequences can be excluded, genomic DNA can also be used as a DNA standard for absolute quantification. The copy number of the target present in the genomic DNA can be directly calculated if the genome size of the organism is known. For example, the genome size (haploid) of Mus musculus is 2.7 x 109 bp, a molecular weight of 1.78 x 10^{12} Daltons.

1.78 x 10^{12} g of genomic DNA corresponds to 6.022 x 10^{23} copies of a single-copy gene.

1 µg of genomic DNA corresponds to 3.4 x 10^{5} copies of a single-copy gene.

What is relative quantification?

In relative quantification, the ratio between the amounts of a target gene and a control gene (e.g., an endogenous reference gene present in all samples) is determined. This ratio is then compared between different samples. In gene expression analysis, housekeeping or maintenance genes are usually chosen as an endogenous reference. The

target and reference gene are amplified from the same sample, either separately or in the same reaction (duplex, real-time PCR). The normalized value is determined for each sample and can be used, for example, to compare differential expression of a gene in different tissues or to compare gene expression between siRNA-transfected cells and untransfected cells. However, the expression level of the endogenous reference gene must not vary under different experimental conditions or in different states of the tissue (e.g., “stimulated” versus “unstimulated” samples). When gene expression levels are compared between samples, the expression level of the target is referred to as being, for example, 100-fold higher in stimulated cells than in unstimulated cells. The quantification procedure differs depending on whether the target and the endogenous reference gene are amplified with comparable or different efficiencies.

Determining amplification efficiencies

The amplification efficiency of 2 genes (target A and target B) can be compared by preparing a dilution series for both genes from a reference RNA or cDNA sample. Each dilution series is then amplified in real-time one-step or two-step RT-PCR and the CT values obtained are used to construct standard curves for target A and target B. The amplification efficiency (E) for each target can be calculated according to the following equation:

$E = 10^{(-1/S)} - 1$ (where S = slope of the standard curve)

To compare the amplification efficiencies of the 2 target sequences, the C_T values of target A are subtracted from the $C_T$values of target B. The difference in C_T values is then plotted against the logarithm of the template amount (see figure Efficiency comparison). If the slope of the resulting straight line is <0.1, amplification efficiencies are comparable.

Different amplification efficiencies

Amplification efficiencies of the target gene and the endogenous reference gene are usually different since efficiency of primer annealing, GC-content of the sequences to be amplified, and PCR product size usually vary between the 2 genes. In this case, a standard curve needs to be prepared for the target gene as well as for the endogenous reference gene, for example, using total RNA prepared from a reference cell line (calibrator or reference sample).

Due to differences in PCR efficiency, the resulting standard curves will not be parallel and the differences in C_T values of the target and the reference will not be constant when the template amounts are varied (see figure Different PCR efficiencies).

Guidelines for relative quantification with different amplification efficiencies:

- Choose an appropriate endogenous reference gene whose expression level does not change under the experimental conditions or between different tissues.
- Prepare a dilution series (e.g., 5-fold or 10-fold dilutions) of a cDNA or RNA control sample to construct standard curves for the target and reference.
- Perform real-time PCR/RT-PCR.
- Determine the C_T values for the standards and the samples of interest.
- Construct standard curves for both the target and reference by plotting C_T values (Y-axis) against the log of template amount or dilution (X-axis).
- Calculate the amount of target and reference in the samples of interest using their C_T values and the corresponding standard curve.
- To calculate the normalized amount of target, divide the amount of target by the amount of reference (if replicate reactions were performed, use the average value).
- Define the calibrator sample and compare the relative expression level of the target gene in the samples of interest by dividing the normalized target amounts by the value of the calibrator.

Comparable amplification efficiencies

If the amplification efficiencies of the target gene and the endogenous reference gene are comparable, one standard curve for the reference gene is sufficient. The differences in C_T values of the target and the reference will be constant when the amounts of template are varied (see figure Same PCR efficiencies). The amounts of target and reference in an unknown sample are calculated by comparing the C_T values with the standard curve for the reference gene.

Guidelines for relative quantification with comparable amplification efficiencies:

- Choose an appropriate endogenous reference gene whose expression level does not change under the experimental conditions or between different tissues.
- Prepare a dilution series (e.g., 5-fold or 10-fold dilutions) of a cDNA or RNA control sample to construct a standard curve for the endogenous reference gene only.
- Perform real-time PCR/RT-PCR.
- Determine the C_T values for the standards and the samples of interest.
- Construct a standard curve for the endogenous reference gene by plotting C_T values (Y-axis) against the log of template amount or dilution (X-axis).
- Calculate the amount of target and reference in the samples of interest using their C_T values and the standard curve.
- To calculate the normalized amount of target, divide the amount of target by the amount of reference (if replicate reactions were performed, use the average value).
- Define the calibrator sample and compare the relative expression level of the target gene in the samples of interest by dividing the normalized target amounts by the value of the calibrator.

Comparative method or $\Delta\Delta C_T$ method of relative quantification

An alternative approach is the comparative or $\Delta\Delta C_T$ method, which relies on direct comparison of C_T values. The preparation of standard curves is only required to determine the amplification efficiencies of the target and endogenous reference genes in an initial experiment. In all subsequent experiments, no standard curve is required for quantification of the target sequence. If amplification efficiencies are comparable, amounts of target are simply calculated by using C_T values as described below.

First of all, the ΔC_T value for each sample is determined by calculating the difference between the C_T value of the target gene and the C_T value of the endogenous reference gene. This is determined for each unknown sample as well as for the calibrator sample.

- ΔC_T (sample) = C_T target gene – C_T reference gene
- ΔC_T (calibrator) = C_T target gene – C_T reference gene

Next, the $\Delta\Delta C_T$ value for each sample is determined by subtracting the ΔC_T value of the calibrator from the ΔC_T value of the sample.

- $\Delta\Delta C_T = \Delta C_T$ (sample) – ΔC_T (calibrator)

If the PCR efficiencies of the target gene and endogenous reference gene are comparable, the normalized level of target gene expression is calculated by using the formula:

- Normalized target gene expression level in sample = $2^{-\Delta\Delta CT}$

However, if the PCR efficiency is not the same between the target gene and endogenous reference gene, this method of quantification may lead to inaccurate estimation of gene expression levels.

The error is a function of the PCR efficiency and the cycle number and can be calculated according to the formula:

- Error (%) = $[(2^n/(1+E)^n) \times 100)] - 100$ (where E = efficiency of PCR; n = cycle number)

Therefore, if the PCR efficiency is only 0.9 instead of 1.0, the resulting error at a threshold cycle of 25 will be 261%. The calculated expression level will be 3.6-fold less than the actual value.

Tip: The $\Delta\Delta C_T$ method should only be chosen if the PCR efficiency of the target gene and endogenous reference gene are the same, or if the difference in expression levels is sufficiently high to tolerate the resulting error. However, errors can be corrected by using efficiency-corrected calculation programs, such as the Relative Expression Software Tool (REST; see reference 5).

Guidelines for relative quantification using $\Delta\Delta C_T$ method:

- Perform a validation experiment to determine the PCR efficiency for the target and reference (see Determining amplification efficiencies).
- Perform real-time RT-PCR for the target and reference with RNA derived from different samples.
- Determine the ΔC_T value by subtracting the endogenous reference gene C_T value from the target gene C_T value for each sample.
- Define the calibrator sample and determine the $\Delta\Delta C_T$ value by subtracting the calibrator ΔC_T value from the ΔC_T value of each sample.
- Calculate the normalized level of target expression relative to the calibrator by using the formula $2^{-\Delta\Delta CT}$.

Endogenous reference genes

For relative quantification of gene expression, it is important to choose a suitable gene to use as a reference (see table Housekeeping genes commonly used as endogenous references). The expression level of the reference gene must not vary under experimental conditions, or in different states of the same tissue or cell line (e.g., “disease” versus “normal” samples). The expression level of the reference RNA should also be approximately the same as the RNA under study. Reference RNA commonly used for relative quantification includes β-actin, β-2-microglobulin, peptidylprolyl isomerase A, and GAPDH mRNAs, and also 18S rRNA. β-actin mRNA is ubiquitously expressed and was one of the first RNAs to be used as a reference sequence. However, its transcription levels may vary and the presence of pseudogenes may mean that genomic DNA is detected during real-time PCR, leading to inaccuracies in quantification. GAPDH is a housekeeping gene commonly used as a reference for quantification of gene expression. GAPDH mRNA levels may vary between individuals, at different stages of the cell cycle, and following treatment with different drugs, making GAPDH unsuitable as a reference in some systems. As 18S rRNA is not an mRNA, its expression levels in the cell may not accurately reflect the cellular mRNA population. Therefore, a combination of genes may provide the most reliable reference for quantification studies.

* “+” indicates relative abundance of the transcripts.

Housekeeping genes commonly used as endogenous references Gene	Human gene symbol	Mouse gene symbol	Relative expression level in humans*	Relative expression level in mice*
18S ribosomal RNA	RRN18S	Rn18s	++++	++++
Actin, beta	ACTB	Actb	+++	+++
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH	Gapdh	+++	+++
Phosphoglycerate kinase 1	PGK1	Pgk1	+++	+++
Peptidylprolyl isomerase A	PPIA	Ppia	+++	+++
Ribosomal protein L13a	RPL13A	Rpl13a	+++	+++
Ribosomal protein, large, P0	RPLP0	–	+++	–
Acidic ribosomal phosphoprotein PO	–	Arbp	–	+++
Beta-2-microglobulin	B2M	B2m	++ to +++	++ to +++

Tyrosine 3-monooxygenase/tryptophan5-monooxygenase activation protein, zeta polypeptide	YWHAZ	Ywhaz	++ to +++	+
Succinate dehydrogenase complex, subunit A, flavoprotein (Fp)	SDHA	Sdha	++	+
Transferrin receptor	TFRC	Tfrc	++	+
Aminolevulinate, delta-, synthase 1	ALAS1	Alas1	+	+
Glucuronidase, beta	GUSB	Gusb	+	+
Hydroxymethylbilane synthase	HMBS	Hmbs	+	++ to +++
Hypoxanthine phosphoribosyltransferase 1	HPRT1	Hrpt1	+	+
TATA box binding protein	TBP	Tbp	+	+
Tubulin, beta	TUBB	–	+	–
Tubulin, beta 4	–	Tubb4	–	+

Back to top

PCR controls

No-template control

A no-template control (NTC) allows detection of contamination of the PCR reagents. An NTC reaction contains all real-time PCR components except the template. Detection of a positive signal in an NTC reaction indicates the presence of contaminating nucleic acids.

Positive control

A positive control may be necessary, for example, when amplifying a new target sequence to confirm whether the primer set or primer–probe set works. A positive control can be an absolute standard, which is a nucleic acid template of known copy number that provides quantitative information. Absolute standards, such as a nucleic acid from an established cell line, a plasmid containing cloned sequences, or in vitro transcribed RNA, are commercially available or can be generated in the lab. A positive control can also be a known positive sample, which is usually a substitute for an absolute standard and used only to test for the presence or absence of a target.

No RT control

A no RT control, where real-time RT-PCR is carried out without reverse transcriptase, should be included when performing gene expression analysis. For viral load monitoring, a no RT control may be necessary, depending on the sample type and the life cycle of the virus species detected. Since reverse transcription cannot take place, a no RT control reaction allows detection of contaminating DNA, such as DNA from viral sequences integrated into the host genome. Contaminating DNA in RNA samples can be removed by DNase treatment before starting RT-PCR.

Internal controls

An internal, positive control can be used to test for the presence of PCR inhibitors. A duplex reaction is carried out, where the target sequence is amplified with one primer–probe set, and a control sequence (i.e., the internal, positive control) is amplified with a different primer–probe set. The internal, positive control should be at a high enough copy number for accurate detection. If the internal, positive control is detected, but the target sequence is not, then this indicates that the amplification reaction was successful and that the target sequence is absent (or at too low a copy number to be detected).

Several factors can generate a false negative result, such as errors in sample extraction or thermocycler malfunction. Assay failure due to PCR or RT-PCR inhibition is the most common cause.

The most practical approach to control for the presence of inhibitors is to include an Internal Positive Control, or Internal Control (IC). This IC is simultaneously extracted and amplified (or only amplified) in the same tube with the

pathogen target, and should always be combined with an external positive control to prove the functionality of the reaction mix for amplification of the target. This combination rules out inhibition, among other malfunctions, and confirms that a negative result is truly negative.

Not all internal controls are the same (see table Features of internal controls), and each IC concept has value for specific applications. Endogenous ICs occur naturally in test specimens, such as a sequence of the host genome (e.g., ß-actin) or from normal microflora genomes (e.g., 16s). Exogenous ICs, on the other hand, are spiked into samples either during nucleic acid extraction or before PCR amplification.

Exogenous ICs can be homologous, where an artificial template is constructed with the same primer binding sites as the targeted pathogen sequence. Although the same primer set is used for target and IC, the sequence differs, enabling differentiation of pathogen and IC amplicons with different probes. Heterologous ICs, on the other hand, are designed with their own primers and probe.

Endogenous and exogenous homologous ICs carry the risk of impairing detection sensitivity for the pathogen target due to competition for reaction components. For example, a high starting amount of an endogenous IC template can impair assay sensitivity. This high starting amount can result from variations in the sample type or sampling technique. In the case of RNA applications, the high starting amount can also be due to enhanced expression levels of the IC due to disease-related cellular pathology. In the case of exogenous homologous ICs, using the same primers to amplify both target and IC leads to primer competition. Additionally, both endogenous and homologous ICs involve tedious IC design, and their use is restricted to a few applications or even individual assays.

In the context of process safety and workflow simplification, exogenous heterologous ICs are the most informative and flexible. The amount of IC template spiked into a sample is defined and consistent, and unrestricted design options enable optimization of IC properties. Only heterologous ICs allow for a design and setup that prevents competition for PCR components, and heterologous ICs are suited as universal controls, thereby making their implementation in new assays easy.

Features of internal controls Feature	Exogenous homologous	Exogenous heterologous	Endogenous
Universal use in multiple assays	No	Yes	No
Serves as control for purification procedure	Yes	Yes	Yes
Differentiates purification errors from amplification errors	Yes	Yes	No
Template quantities are defined and consistent	Yes	Yes	No
Non-competitive internal control design	No	Yes	Yes

Back to top

Types of PCR

Multiplex PCR

Multiplex PCR employs different primer pairs in the same reaction for simultaneous amplification of multiple targets. This type of PCR often requires extensive optimization of annealing conditions for maximum amplification efficiency of the different primer–template systems and is often compromised by nonspecific PCR artifacts. A stringent hot-start procedure and specially optimized buffer systems are absolutely crucial for successful multiplex PCR.

Compared with standard PCR systems using only 2 primers, an additional challenge of multiplex PCR is the varying hybridization kinetics of different primer pairs. Primers that bind with high efficiency could utilize more of the PCR reaction components, thereby reducing the yield of other PCR products. This often results in unamplified DNA sequences and absence of expected PCR products. Commercial PCR kits are available that are specifically designed to overcome the challenges of multiplex PCR and it is recommended that, where possible, such a kit is used.

Long-range PCR

PCR products of up to 4 kb can be routinely amplified using standard PCR protocols using *Taq* DNA polymerase. However, amplification of PCR products longer than 4 kb often fails without lengthy optimization. Reasons for failure include nonspecific primer annealing, secondary structures in the DNA template, and suboptimal cycling conditions — all factors which have a greater effect on the amplification of longer PCR products than on shorter ones. Preventing DNA damage, such as DNA depurination, is of particular importance for amplification of long PCR products, as a single DNA lesion within the template is sufficient to stall the PCR enzyme. DNA damage during PCR cycling can be minimized with specific buffering substances that stabilize the pH of the reaction. Commercial PCR kits are available that are specifically designed to overcome the challenges of long-range PCR, for example, by using an optimized mixture of *Taq* DNA polymerase and proofreading enzymes, and it is recommended that, where possible, such a kit is used.

Single-cell PCR

Single-cell PCR provides a valuable tool for genetic characterization using a limited amount of starting material. By flow cytometry or micromanipulation, individual cells of interest can be isolated based on cell-surface markers or physical appearance. Amplification of low-abundance template molecules — which can be as low as one or two gene copies — requires a PCR system that is highly efficient, specific, and sensitive. Again, commercial PCR kits are available that are specifically designed for single-cell PCR.

Fast-cycling PCR

Faster PCR amplification enables increased PCR throughput and allows researchers to spend more time on downstream analysis. The demand for reducing time-to-result is met by the recent development of faster PCR techniques. Fast PCR can be achieved using new thermal cyclers with faster ramping times or through innovative PCR chemistries that allow reduced cycling times due to significantly shortened DNA denaturation, primer annealing, and DNA extension times. Fast-cycling PCR reagents must be highly optimized to ensure amplification specificity and sensitivity.

Methylation-specific PCR (MSP)

MSP enables the methylation status of target DNA to be determined after sodium bisulfite treatment. The method requires two sets of primers to be designed: one set that anneals to unchanged cytosines (i.e., methylated in the genomic DNA) and one set that anneals to uracil resulting from bisulfite treatment of cytosines not methlyated in the genomic DNA. Amplification products derived from the primer set for unchanged sequences indicates the cytosines were methylated and thus protected from alteration (6).

Stringent and highly specific PCR conditions must be used to avoid nonspecific primer binding and the amplification of PCR artifacts. This is particularly important as the conversion of unmethylated cytosines to uracils reduces the complexity of the DNA and increases the likelihood of nonspecific primer–template binding.

Hot start PCR

See Hot-start DNA polymerase for more information.

High-fidelity PCR

See High-fidelity DNA polymerase for more information.

RAPD: Rapid amplified polymorphic DNA analysis

RAPD is a PCR-based tool enabling the study of organisms at the molecular level. It uses small, nonspecific primers to amplify seemingly random regions of genomic DNA. Successful primer pairs produce different banding profiles of PCR products between individuals, strains, species, etc., when analyzed using an agarose gel.

In RAPD, the primers are only ~10 bases long. As a result, annealing temperatures required are <40°C.

RACE: Rapid amplification of cDNA ends

RACE is a variant of RT-PCR and is a procedure for amplification of nucleic acid sequences from a messenger RNA template between a defined internal site and unknown sequences at either the 3' or the 5' -end of the mRNA. RACE only requires the knowledge of a short sequence within the mRNA of interest. It is often used for cloning the remainder of incomplete cDNAs. There are two techniques:

- 5' RACE — amplifies 5' cDNA ends
- 3' RACE — amplifies 3' cDNA ends

The first step is common to both types of RACE and involves the conversion of RNA to single-stranded cDNA using a reverse transcriptase. The second steps are unique to each type of RACE; although each generates information that may yield the full-length cDNA sequence.

Because RACE uses an “anchor site” within the mRNA as a point of reference, it is sometimes known as “anchored PCR”.

In situ PCR

In situ PCR is a PCR reaction that occurs inside the cell on a slide, thus combining the sensitivity of PCR or RT-PCR amplification with in situ hybridization. In situ PCR allows cellular markers to be identified and further enables the localization to cell-specific sequences within cell populations, such as tissues and blood samples. Therefore, it is a powerful tool in applications such as the study of disease progression.

Fresh or fixed cells or tissue samples can be used in the procedure, although preparation of the sample is critical to the result, with fixation having a direct influence on PCR signal. The procedure is suitable for use with radiolabeled, fluorescently labeled or biotin-labeled nucleic acid probes.

The PCR process is essentially the same as a standard PCR, but with some modified reaction conditions (e.g., Mg^{2+}concentration).

Differential display PCR

Differential display PCR is based on RT-PCR and is used to compare and identify differences in mRNA (and therefore gene) expression patterns between two cell lines or populations.

In this technique, first-strand cDNA synthesis is primed with an anchored primer complementary to ~13 nucleotides of the poly(A) tail of mRNA and the adjacent 2 nucleotides of the transcribed sequence. After reverse transcription and amplification, amplified products are visualized using gel electrophoresis. The banding patterns observed can be compared to identify differentially expressed cDNAs in the 2 populations.

Invented in the 1990s, the technique fast became a key tool in gene expression analysis. However, it has been more recently superseded by RNA-seq, microarrays, and qRT-PCR.

Back to top

Guidelines for RT-PCR

When performing real-time RT-PCR, the primers and the enzyme for reverse transcription must be carefully chosen. The primers should allow reverse transcription of all targets of interest, and the reverse transcriptase should yield cDNA amounts that accurately represent the original RNA amounts to ensure accurate quantification. In addition, the effects of the components of the RT reaction on subsequent real-time PCR must be minimized.

To perform PCR using RNA as a starting template, the RNA must first be reverse transcribed into cDNA in a reverse transcription (RT) reaction. RT and PCR can be carried out either sequentially in the same tube (one-step RT-PCR) or separately (two-step RT-PCR). One-step RT-PCR requires gene-specific primers.

Back to top

Two-step and one-step RT-PCR

Real-time RT-PCR can take place in a two-step or one-step reaction (see figure Comparison of two-step and one-step RT-PCR and table Advantages of different RT-PCR procedures). With two-step RT-PCR, the RNA is first reverse transcribed into cDNA using oligo-dT primers, random oligomers, or gene-specific primers. An aliquot of the reverse-transcription reaction is then added to the real-time PCR. It is possible to choose between different types of RT primers, depending on experimental needs. Use of oligo-dT primers or random oligomers for reverse transcription means that several different transcripts can be analyzed by PCR from a single RT reaction. In addition, precious RNA samples can be immediately transcribed into more stable cDNA for later use and long-term storage.

In one-step RT-PCR — also referred to as one-tube RT-PCR — both reverse transcription and real-time PCR take place in the same tube, with reverse transcription preceding PCR. This is possible due to specialized reaction chemistries and cycling protocols (see Conditions for one-step RT-PCR). The fast procedure enables rapid processing of multiple samples and is easy to automate. The reduced number of handling steps results in high reproducibility from sample to sample and minimizes the risk of contamination since less manipulation is required.

Advantages of different RT-PCR procedures

Procedure	Advantages
Two-step RT-PCR	Multiple PCRs from a single RT reaction Flexibility with RT primer choice Enables long-term storage of cDNA
One-step RT-PCR	Easy handling Fast procedure High reproducibility Low contamination risk

Back to top

Choice of RT primers

The choice of primers for reverse transcription depends on whether one-step or two-step RT-PCR is being carried out. In one-step RT-PCR, the downstream PCR primer is also the primer for reverse transcription. Therefore, one-step RT-PCR is always performed with gene-specific primers. In two-step RT-PCR, 3 types of primers, and mixtures thereof, can be used for reverse transcription: oligo-dT primers (typically 13–18mers), random oligomers (such as hexamers, octamers, or nonamers), or gene-specific primers (see table “Suitability of primer types for RT-PCR”). If oligo-dT primers are used, only mRNAs will be reverse transcribed starting from the poly-A tail at the 3' end. Random oligomers will enable reverse transcription from the entire RNA population, including ribosomal RNA, transfer RNA, and small nuclear RNAs. Since reverse transcription is initiated from several positions within the RNA molecule, this will lead to relatively short cDNA molecules. In comparison, gene-specific primers allow reverse transcription of a specific transcript.

A universal priming method for the RT step of real-time two-step RT-PCR should allow amplification and detection of any PCR product regardless of transcript length and amplicon position, and achieve this with high sensitivity and reproducibility.

Suitability of primer types for RT-PCR

Application	Recommended type of primer
RT-PCR of specific	Gene-specific primer gives highest selectivity and only the RNA molecule of choice will

transcript	be reverse transcribed
RT-PCR of long amplicon	Oligo-dT or gene-specific primers
RT-PCR of an amplicon within long transcript	Gene-specific primers, random oligomers, or a mixture of oligo-dT primers and random nonamers are recommended so that cDNA covering the complete transcript is produced

Back to top

Conditions for two-step RT-PCR

Effect of RT volume added to two-step RT-PCR

In two-step RT-PCR, the addition of the completed reverse-transcription reaction to the subsequent amplification reaction transfers not only cDNA template, but also salts, dNTPs, and RT enzyme. The RT reaction buffer, which has a different salt composition to that of the real-time PCR buffer, can adversely affect real-time PCR performance. However, if the RT reaction forms 10% or less of the final real-time PCR volume, performance will not be significantly affected. Use of 3 µl of RT reaction in a 20 µl PCR (i.e., 15% of the final volume) can lead to significant inhibition of real-time PCR. We recommend testing dilutions of the RT reaction in real-time PCR to check the linearity of the assay. This helps to eliminate any inhibitory effects of the RT reaction mix that might affect accurate transcript quantification.

Effect of RNA secondary structure

RNA secondary structure can affect RT-PCR in several ways. Regions of RNA with complex secondary structure can cause the reverse transcriptase to stop or dissociate from the RNA template (see figure Effects of complex secondary structure on RT-PCR: RT effects).

The truncated cDNAs, missing the downstream primer-binding site, are then not amplified during PCR. Alternatively, the reverse transcriptase can skip over looped-out regions of RNA, which are then excluded from the synthesized cDNA. In the PCR step, these cDNAs with internal deletions are amplified and appear as shortened PCR products. Ideally, the reverse transcriptase should not be affected by RNA secondary structure and should be capable of reverse transcribing any template, without the need for reaction optimization.

With **high GC content**, the tight association of RNA:DNA hybrids can interfere with primer binding during PCR and prevent DNA polymerases from progressing (see figure Effect of high GC content on RT-PCR: PCR effects). RNase H removes RNA in RNA:DNA hybrids to allow primer binding and second-strand DNA synthesis. RNase H digestion has been previously shown to improve RT-PCR yield and to be required for amplification of some sequences, even as short as 157 bp (7).

Back to top

Conditions for one-step RT-PCR

The ideal reverse transcriptase for one-step RT-PCR should also exhibit the same properties as those described above for reverse transcriptases for two-step RT-PCR. However, one of the main problems in one-step RT-PCR is the inhibitory effect of the reverse transcriptase on the PCR step, which can lead to increased C_T values and thus reduced sensitivity and specificity when compared with two-step RT-PCR.

Back to top

RT-PCR primer design

A critical factor in RT-PCR is the selection of appropriate primers for maximal efficiency and specificity. Primer specificity is affected by a number of factors, including sequence, primer location, and the RT-PCR system used. General primer-design rules for PCR are also applicable in RT-PCR to avoid mispriming and primer–dimer formation (see PCR primer design). These effects are even more pronounced in RT-PCR, where cDNAs produced during reverse transcription are more susceptible to nonspecific priming due to their single-stranded nature. Nonspecific

priming in RT-PCR reduces the sensitivity of the process, leading to reduced yields of specific products or failure of the RT-PCR altogether.

To avoid amplification of contaminating genomic DNA, primers for RT-PCR should be designed so that one half of the primer hybridizes to the 3' end of one exon and the other half to the 5' end of the adjacent exon (see figure RT-PCR primer design). Such primers will anneal to cDNA synthesized from spliced mRNAs, but not to genomic DNA.

To detect amplification of contaminating DNA, RT-PCR primers should be designed to flank a region that contains at least one intron. Products amplified from cDNA (no introns) will be smaller than those amplified from genomic DNA (containing introns). Size difference in products is used to detect the presence of contaminating DNA.

If only the mRNA sequence is known, choose primer annealing sites that are at least 300–400 bp apart. It is likely that fragments of this size from eukaryotic DNA contain splice junctions. As explained in the previous point, such primers may be used to detect DNA contamination.

In summary, the following factors should be considered when designing primers for RT-PCR:

- Annealing temperature can affect RT-PCR efficiency and sensitivity.
- High primer concentrations can cause mispriming and primer–dimer formation.
- A stringent hot start is essential for optimal RT-PCR performance.
- Primer design in RT-PCR allows differentiation of signals from RNA and contaminating DNA. For best results, DNA-free RNA should be used in order to avoid competition of DNA in RT-PCR.

Back to top

Enzymes used in RT-PCR

RT-PCR allows the analysis of RNA using a combination of reverse transcription and PCR. cDNA is synthesized from RNA templates using reverse transcriptases — RNA-dependent DNA polymerases normally isolated from a variety of retroviral sources (e.g., from Avian Myeloblastosis Virus [AMV] or Moloney murine leukemia virus [MMLV]).

Although thermostable DNA polymerases such as Tth DNA polymerase also exhibit reverse transcriptase activity under specific conditions, these enzymes are not as efficient for reverse transcription as mesophilic reverse transcriptases.

The single-stranded cDNA produced by reverse transcription is more susceptible to nonspecific primer annealing at lower temperatures than double-stranded DNA (e.g., genomic DNA). Nonspecific annealing can result in poor amplification specificity which, especially when combined with limiting cDNA quantity or low transcript abundance, leads to reduced sensitivity and poor reproducibility. Amplification specificity is crucial for successful RT-PCR and is best achieved by combining innovative buffer solutions with specially modified reverse transcriptases and hot-start PCR.

Back to top

Multiplex PCR and RT-PCR

In multiplex, real-time PCR, several genomic DNA targets are quantified simultaneously in the same reaction. Multiplex, real-time RT-PCR is a similar method, allowing simultaneous quantification of several RNA targets in the same reaction. The procedure can be performed either as two-step RT-PCR or as one-step RT-PCR.

Multiplex PCR and RT-PCR offer many advantages for applications such as gene expression analysis, viral load monitoring, and genotyping. The target gene(s) as well as an internal control are co-amplified in the same reaction, eliminating the well-to-well variability that would occur if separate amplification reactions were carried out. The internal control can be either an endogenous gene that does not vary in expression between different samples (e.g., a housekeeping gene; see table Housekeeping genes commonly used as endogenous references) or an exogenous nucleic acid. For viral load monitoring, the use of an exogenous nucleic acid as internal control allows the following

parameters to be checked: the success of sample preparation, the absence of inhibitors, and the success of PCR. Multiplex analysis ensures high precision in relative gene quantification, where the amount of a target gene is normalized to the amount of a control reference gene. Quantification of multiple genes in a single reaction also reduces reagent costs, conserves precious sample material, and allows increased throughput.

Multiplex PCR and RT-PCR are made possible by the use of sequence-specific probes that are each labeled with a distinct fluorescent dye and an appropriate quencher moiety. This means that the emission maxima of the dyes must be clearly separated and must not overlap with each other. In addition, reactions must be carried out on an appropriate real-time cycler that supports multiplex analysis (i.e., the excitation and detection of several non-overlapping dyes in the same well or tube).

Back to top

Whole transcriptome amplification (WTA)

Whole transcriptome amplification (WTA) allows amplification of entire transcriptomes from very small amounts of RNA, enabling unlimited analyses by real-time RT-PCR. WTA of RNA samples can be achieved by reverse transcription and cDNA ligation prior to multiple displacement amplification (MDA).

When only nanogram amounts of an RNA sample are available, the number of real-time RT-PCR analyses that can be performed is limited. This problem can be resolved by WTA. With this technology, all mRNA transcripts in an RNA sample are replicated to provide microgram amounts of cDNA template, enough for unlimited real-time PCR analysis and stable archiving.

WTA techniques

To ensure reliable results in real-time PCR, it is essential that the WTA method provides unbiased and accurate amplification of the whole transcriptome. This means that the sequence and the relative abundance of each transcript should be preserved after WTA, otherwise false results in gene expression analysis will occur. The principle of WTA is shown in the figure Schematic representation of whole transcriptome amplification.

As reverse transcription takes place using a mix of random and oligo-dT primers, a cDNA library covering all transcript sequences, including both 5' and 3' regions, is prepared. Subsequent ligation of the cDNA followed by MDA using a uniquely processive DNA polymerase generates amplified cDNA that preserves the transcript representation of the original RNA sample. This is critical for accurate gene expression analysis.

When carrying out WTA, it is important to consider both the amount of starting material (i.e., the number of cells or the amount of RNA) and the copy number of the transcripts of interest. The table Transcript representation in different cell amounts shows the relationship between the amount of starting material and transcript representation (note that this is only a guide: the number of transcripts per given amount of starting material can vary). In starting material where the copy number of a transcript is 10 or less (highlighted in bold in the table, stochastic problems will occur (i.e., the unequal distribution of a very low number of transcripts in a highly dilute solution). This may result in underrepresentation of the low-copy transcript at the start of WTA. Special consideration should be given to mosaic transcripts, which are derived from genes that are expressed only in a subset of cells in tissues. Since these transcripts are not present in every cell, they will not be accurately represented in low amounts of starting material (i.e., 1–10^2 cells).

Reliable WTA depends on the copy number of the transcripts. 10 ng of RNA corresponds to about 500 cells, and even low-copy transcripts are well represented in this amount of RNA. Using lower amounts of RNA or a very limited number of cells means that the starting material could have a partial representation or an absence of low-copy transcripts.

* Complete representation of all transcripts.
† Stochastic problems for mosaic transcripts.
‡ Stochastic problems for low-copy and mosaic transcripts.
§ Stochastic problems for low-copy transcripts and loss of mosaic transcripts.

Transcript representation in different cell amounts Parameter	10^3 cells*	10^3 cells†	10 cells‡	1 cell§
Amount of RNA (ng)	20	2	0.2	0.02
No. of high-copy transcripts	10^7	10^6	10^5	10^4
No. of medium-copy transcripts	10^5	10^4	10^3	10^2
No. of low-copy transcripts	10^3	10^2	10	1
No. of mosaics transcripts	10^2	10	1	0

Back to top

DNA contamination

Removal of DNA contamination

Contamination of RNA samples with trace amounts of genomic DNA can interfere with real-time RT-PCR quantification if the PCR primers used are also able to amplify genomic DNA sequences. To avoid the negative effects of genomic DNA contamination, careful primer design is required (see Primer design). If this is not possible, RNA samples should be treated with DNase I to digest contaminating DNA.

Detecting DNA contamination in RT-PCR

Use of appropriate controls will enable the detection of any contaminating DNA in the RT-PCR. Reactions should be set up with and without the reverse transcriptase. The presence of a product in the absence of the reverse transcriptase indicates contamination.

Back to top

References

Bustin, S.A., ed. (2004) A-Z of Quantitative PCR. La Jolla, CA: International University Line.

Cited references

1. Ausubel, F.M., et al. (1991) Current Protocols in Molecular Biology. New York: John Wiley and Sons.
2. Pavlov, A.R. et al. (2004) Recent developments in the optimization of thermostable DNA polymerases for efficient amplifications. Trends in Biotechnology **22**, 253.
3. Thellin, O. et al. (1999). Housekeeping genes as internal standards: use and limits. J. Biotechnol. **75**, 291.
4. Vandesompele, J., et al. (2002) Genome Biol. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. **3**, RESEARCH0034.
5. Pfaffl, M.W., Horgan, G.W., and Dempfle, L. (2002) Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res. **30**, e36.
6. Derks, S. et al. (2004) Methylation-specific PCR unraveled. Cell Oncol. **26**, 291.
7. Tacke, E. et al (1995) Transposon tagging of the maize Glossy2 locus with the transposable element En/Spm. Plant J. **8**, 907.