4. Designing and validating primers

DESIGN: When designing primers, consideration should be given to the area of the target gene being amplified (target sequence) and the thermodynamic properties of the primers themselves. All of these properties factor into the efficiency and specificity of target gene amplification. Forward and reverse primer pairs for qPCR should amplify unique target sequences between 70-150 bp long, that are areas of low secondary structure (GC content < 60%). The primers themselves should follow the guidelines for good primer design, but briefly, their length is typically 15-20 bp long, their GC content should be approximately 50-55% with low secondary structure (hairpins, loops etc..) but primers should be "anchored" with a G/C at the 5' end of the sequence and their melting temperature is approximately 5°C higher than the annealing temperature (60 °C) in your proposed qPCR run profile (therefore, optimally 65 °C). Keep in mind that all the primers used in a single qPCR reaction should have similar annealing/extension temperatures. There are many good on-line sites for primer design:

Primer BLAST http://www.ncbi.nlm.nih.gov/tools/primer-blast/
Primer 3 http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/
IDT SciTools http://www.idtdna.com/pages/scitools

Primer pair 1

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	GCGTGAAGGCGAATGCCAGC	Plus	20	2691	2710	65.23	65.00	3.00	2.00
Reverse primer	GTAGCCTGTCACGGGCGCTT	Minus	20	2838	2819	65.22	65.00	7.00	2.00
Product length	148								

Sample read-out from "Primer BLAST" primer design program

IT IS CRITICAL TO "BLAST" (nucleotide) THE PROPOSED PRIMER SEQUENCE TO ENSURE SPECIFICITY FOR THE TARGET (amplified) GENE.

Once your primers arrive, if they are lyophilized, it is recommended to make a 100 uM stock and make a working stock (10 uM) from this. The 100 uM stock should be kept at -20°C and opened only to create new working stocks. This will reduce the possibility of contamination.

VALIDATION: It is recommended to purchase the top two predicted primer pairs and perform the validation on both primers. Depending on the availability of your template (i.e. whether your samples are readily obtained from experimental animals or cell culture or whether they are precious human samples) you may wish to validate your primers on a commercially available universal cDNA. For example, for human primers you can validate on Clontech Laboratories Human Universal Reference

Total RNA (cat #639653/54). While there may be some variability in your target gene expression between this cDNA and that obtained from your biopsy/sample, it is "as good as you can get" without wasting your precious samples on validation.

Ideally, the multiple primers you will use for qPCR all have a reasonable efficiency at the same annealing/extension temperature (usually 60°C). You will have factored this into the design of the primer by specifying melt temperatures (between 63-68°C) as one of the variables. Before performing a qPCR reaction and using expensive SYBR Master Mix, it is recommended to first check the specificity of your primers by running an endpoint PCR reaction first and running the products on a DNA acrylamide gel. In this reaction, you should include a no-template control (NTC) to test for contamination of your buffers and solutions as well as the prevalence of primer-dimer formation. You should also perform multiple PCR reactions to determine the optimal concentration of primers to use for amplification (usually between 100-500 nM).

Sample end-point reaction: (50 ul final volume)

40 ul	UPW (ultra pure water)
1 ul	forward primer (10 uM stock)
1 ul	reverse primer (10 uM stock)
5 ul	10X Taq buffer (which includes $MgCl_2$)
1 ul	dNTP mix (2.5 mM stock)
1 ul	template (Universal cDNA or target cDNA – 10-50 ng/ul stock) (OR WATER, IF A NO-TEMPLATE CONTROL)
1 ul	Taq polymerase

PCR Parameters:

95°C 2:00 minutes

40 cycles of:

95°C 60 seconds (denature)

60 °C 30 seconds (anneal)

4 °C hold

Run products on an **acrylamide DNA gel**, prepared as follows:

8.10 ml Ultrapure H_2O

4.17 ml 29:1 Acrylamide:Bis (30%; Biorad cat# 1610156)

0.25 ml 50x TAE solution

62.5µl APS

6.25µl TEMED

Using a SDS-PAGE gel apparatus (i.e. vertical electrophoresis) set up plates and fill entirely with this solution (i.e., there is no stacking/separating gel). Top with either 10 or 15 well comb. Allow polymerization, remove comb, flush wells with water and load PCR sample with 6X loading buffer (for DNA agarose gels NOT sample buffer for protein gels) diluted to 1X. Load DNA ladder according to predicted size of products. The gel requires 30-45 minutes to run (at 120V). When lower loading dye is ¾ of the way down the gel, remove from plates and incubate gel in ethidium bromide bath (250ml H_2O with 5µl Ethidium Bromide stock solution (10mg/ml; Invitrogen cat# 15585-011)) for up to 60 minutes and then wash with water for up to 60 minutes.

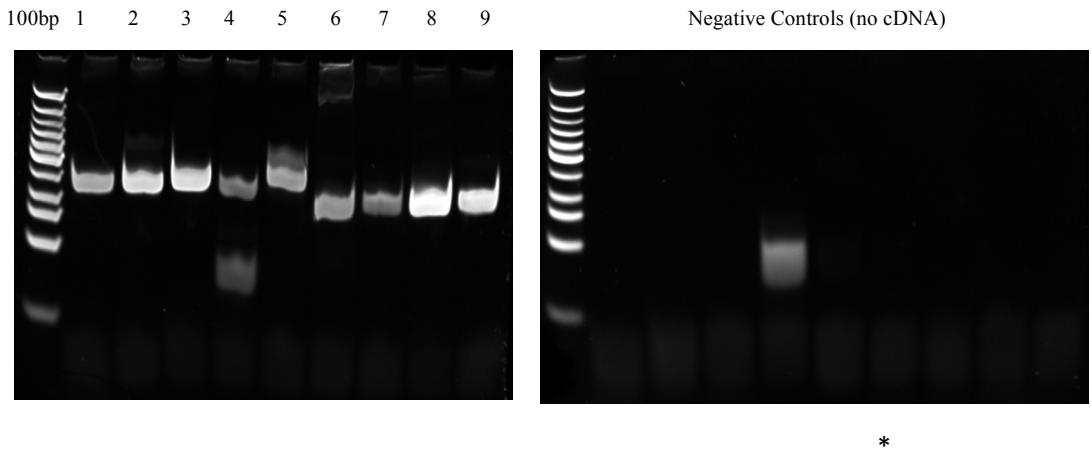


Figure 2. End-point PCR with reference cDNA to confirm amplicon size (A) Reference cDNA (universal human cDNA) was utilized as template for end-point PCR to verify the ability of primers to amplify a SINGLE amplicon of predicted size as well as no background (*) in cDNA negative control.

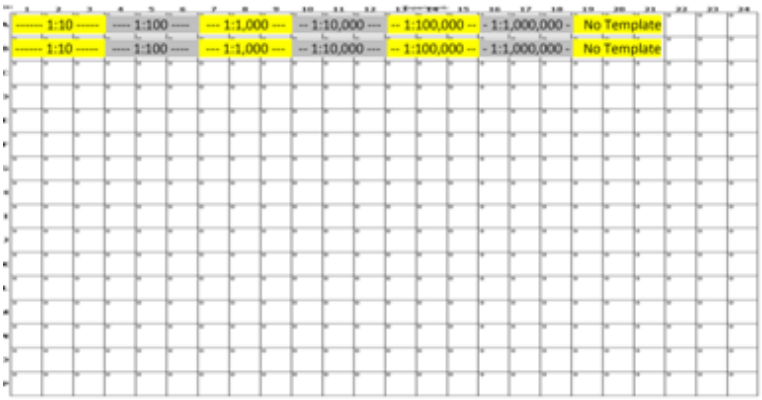
Once you have confirmed that your primers render a single product with no background and the product migrates at the predicted size, you can use this primer set in a qPCR reaction for **melt-curve analysis and generating a standard curve to evaluate PCR efficiency.**

Initially, try only a few primers (i.e. housekeepers such as GAPDH and HMBS) to get used to the equipment, procedures and pipetting. It's not difficult once you get used to it, but pipetting a full 384 well plate is not trivial for a beginner. It is recommended to use a 6 point standard curve, using 1:5 or 1:10 serial dilutions (in triplicate) of your cDNA template along with NTC and NRT (no reverse transcription) controls, each in triplicate.

The total reaction volume per well is 10µl (8 ul master mix plus 2 ul template), following a similar formula for this SAMPLE of master mix calculations:

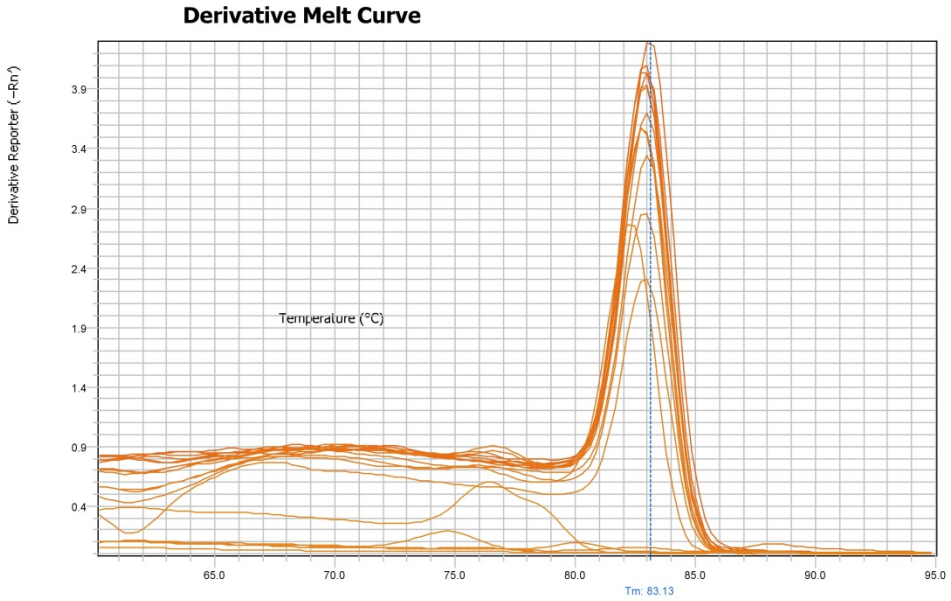
Reagent	[Stock]	Volume/rxn	samples	Master Vol.	Final concentration
SYBR Green Mix	2x	5.00 µl	21	105 µl	1x
Primer mix (FWD/REV)	10µM	0.40 µl	21	8.4 µl	400nM (i.e. determined prior)
H_2O (RNase/DNase free)		2.60 µl	21	54.6 µl	N/A
Total:		**8.00 µl**		**168 µl**	

For serial template (cDNA) dilutions: Commercial human cDNA (Clontech cat #636693) or experimental cDNA was serially diluted 1:10 enough for 6 wells (per concentration): i.e. 2µl of stock cDNA was diluted in 18µl H_2O and then serially in the same manner (6 times) to 1:1,000,000. Each concentration was done in triplicate for each primer 2µl of the cDNA was dispensed into each well. Of note, when dispensing into the microplate, the pipette tip should be placed on the side of the well: the plate will be spun to bring everything together. This is a SAMPLE of what the 384 well plate will look like.

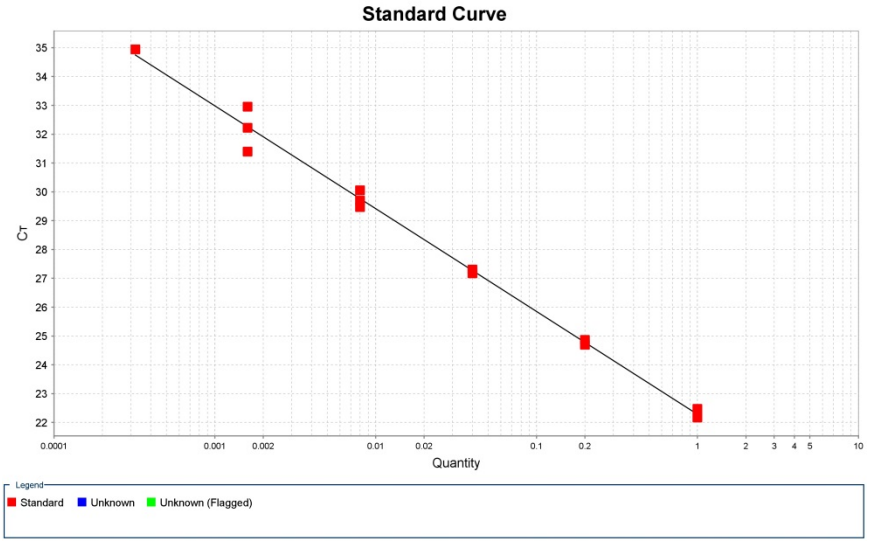


Master mix and template are dispensed with a repeat electronic pipettor (for accuracy) into a 384-well plate. Once the plate is loaded, it must be sealed with an optical cover (ABI 4311971) to prevent evaporative loss. Spin the plate for 2 minutes at 1200rpm and then run the qPCR reaction.

The **melt-curve** performed at the end of the qPCR cycles is important in confirming the specificity of the primer annealing. The curve should display a single peak with no shouldering (see below).

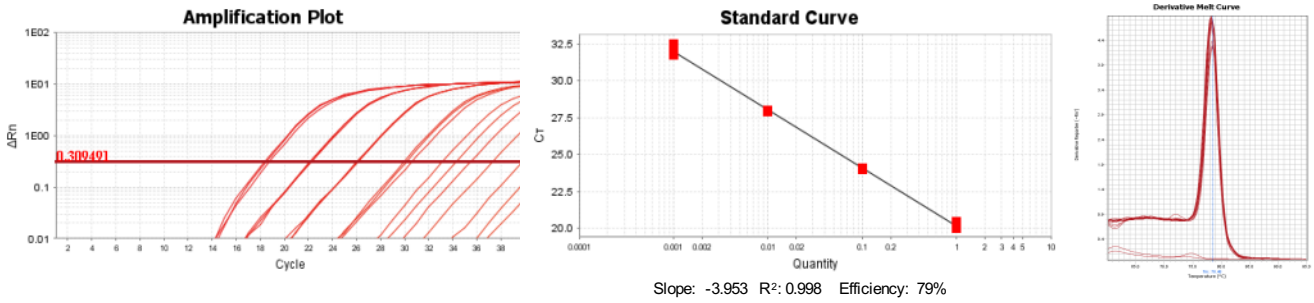


The standard curve is used to calculate **the efficiency of the qPCR reaction**, or how well the polymerase is able to amplify the target cDNA given the primer sequences and the reaction conditions. In a perfectly efficient (100%) reaction, there is a product doubling every cycle, but reaction efficiencies can be either higher or lower than 100% due to inhibitory factors in the cDNA preparation, ineffective primer binding, secondary structure in the target gene, amplicon or primers, old reagents and enzyme or sub-optimal annealing/extension temperatures. The standard curve will also inform you of the optimal template concentration to use so that your target genes amplify within the optimal range (Ct 18-25).



Slope: -3.566 Y-inter: 22.282 R^2:0.992 Eff%: 90.737

The efficiency is calculated by plotting the template quantity vs. the Ct value. The slope of the (linear regression) best fit line reflects the efficiency of the reaction. The reaction efficiency, should always be between 90-110% and it is important that the efficiencies of all the primers used in the qPCR reaction are approximately equal for accurate comparison. The R^2 value represents how well the data fit the regression line or whether the efficiencies are different at any point in the range of template concentrations. This R^2 value should be as close to 1.0 as possible with an acceptable range between 0.98- 1.0. If the value is significantly lower, check the variability between your technical replicates and discard any data points that fall outside of 0.5 Ct of each other. For example, with the plot above, the triplicates for the highest concentration of cDNA (22.174, 22.471, 24.349) were assessed and the outlier (24.349) was omitted from the analysis. For standard curves in which a 1:10 dilution is performed, the highest and lowest concentrations of the standard curve may have to be eliminated for accuracy as the Ct values may start "falling off the curve" with increasing dlution. For example, take this validation scenario using a GAPDH primer:



Assessment: (i) optimal dilution of cDNA (Ct 18-25) with first 3 dilutions

(ii) First four sets of replicates are good, but then variability is high when cDNA is too dilute

Can repeat experiment and do standard curve with lower dilution (1:5) to be within dynamic range

(iii) Calculated efficiency (79%) is borderline acceptable; should be 80-110%

(iv) $R^2 = 0.998$ which is good correlation

(v) Melt curve good – single product, no shouldering

Once your primers have been validated for efficiency, R_2, and specificity, you MUST ensure that you select a proper housekeeping gene(s) for your experimental system. Again, reference genes normalize the data by correcting for differences in starting quantities of cDNA and therefore, CANNOT CHANGE IN EXPRESSION LEVEL (Ct VALUE) BETWEEN EXPERIMENTAL CONDITIONS/TIME POINTS. The final validation step, therefore, is to compare the Ct value at a given concentration of cDNA across all your different

experimental conditions. It is also useful to note that often, one housekeeping gene is not enough for accurate determination of gene expression changes. Vandesompele et al. (2002) wrote an excellent article outlining that the geometric mean of multiple carefully selected housekeeping genes was the most accurate normalization factor. There are also many good on-line algorithms to assess the variability in your housekeepers expression levels for proper normalization; these can be found on the Gene Quantification website (**http://www.gene-quantification.info/**).

For pre-validated human, mouse and rat housekeeper (i.e. GAPDH, HMBS, HPRT, Cyclophilin A, G6PD) primer sequences see Appendix I below or visit the following sites hosted at the University of Ghent and Harvard (**http://medgen.ugent.be/rtprimerdb/, http://pga.mgh.harvard.edu/primerbank/**). Remember, you it is still a good idea to validate these "pre-validated" primers in your experimental system before you use them!

Appendix I: Validated Housekeeping Genes Used for qPCR (with accession numbers)

- ALWAYS "BLAST" (nucleotide) your primer sequences before ordering to check!)
- Also check for more validated qPCR primers in: http://pga.mgh.harvard.edu/primerbank/index.html

HUMAN:

hGAPDH (Glyceraldehyde-3-phosphate dehydrogenase - NM_002046):

Forward 5'-CAATGACCCCTTCATTGACC-3'

Reverse 5'-GACAAGCTTCCCGTTCTCAG-3'

hG6PD (Glucose-6-phosphate dehydrogenase - **NM_000402.3**):

Forward 5'-GAGGCCGTGTACACCAAGAT-3'

Reverse 5' -TCAGGGAGCTTCACGTTCTT-3'

hHMBS (Hydroxymethylbilane synthase - **BC019323.1**):

Forward 5'-TGCAACGGCGGAAGAAAA-3'

Reverse 5'- AGCTGGCTCTTGCGGGTAC-3'

hHPRT1 (Hypoxanthine guanine phosphoribosyl transferase1 - **NM_000194**):

Forward 5' – CCTGGCGTCGTGATTAGTGAT-3'

Reverse 5' – AGACGTTCAGTCCTGTCCATAA - 3'

MOUSE:

mGAPDH (Glyceraldehyde-3-phosphate dehydrogenase - **NM 008084.2**):

Forward 5’ - AGGTCGGTGTGAACGGATTTG-3’

Reverse 5’ -TGTAGACCATGTAGTTGAGGTCA-3’

mG6PD (Glucose-6-phosphate dehydrogenase - **NM 008062**):

Forward 5’ – CACAGTGGACGACATCCGAAA –3’

Reverse 5’ - AGCTACATAGGAATTACGGGCAA-3’

mHMBS (Hydroxymethylbilane synthase - **NM 013551.2**):

Forward 5’-AAGGGCTTTTCTGAGGCACC–3’

Reverse 5’-AGTTGCCCATCTTTCATCACTG -3’

mHPRT1 (Hypoxanthine guanine phosphoribosyl transferase1 - **NM_013556**):

Forward 5’ – TCAGTCAACGGGGGACATAAA – 3’

Reverse 5’ – GGGGCTGTACTGCTTAACCAG – 3

RAT :

rGAPDH (Glyceraldehyde-3-phosphate dehydrogenase – **AF_106860**):

Forward 5’ - GTG CAG TGC CAG CCT CGT C –3’

Reverse 5’ - GGC AGC ACC AGT GGA TGC AG – 3’

rHPRT (Hypoxanthine guanine phosphoribosyl transferase1 - **XM_343829**):

Forward 5’ - GCC GAC CGG TTC TGT CAT – 3’

Reverse 5’- TCA TAA CCT GGT TCA TCA TCA CTA ATC – 3’

rCycloB (Peptidylprolyl isomerase B (cyclophilin B) - **NM_022536**):

Forward 5’ - GGG CTC CGT TGT CTT CCT TT – 3’

Reverse 5’ - GAC TTT AGG TCC CTT CTT CTT ATC GTT – 3’

rHMBS (Hydroxymethylbilane synthase - **NM_013168**):

Forward 5’ - GGC TCA GAT AGC ATG CAA GAG A – 3’

Reverse 5’- TGG ACC ATC TTC TTG CTG AAC A – 3’