

## Relative Gene Expression Workflow

### 1) Isolate sample.

Using RNA<sup>later</sup>® immediately following sample collection protects RNA from degradation by endogenous RNases. It also prevents gene expression profiles from changing prior to RNA isolation. To learn more, click [here](#).

### 2) Choose sample prep method.

Get help choosing the best sample prep approach by clicking [here](#).

### 3) Isolate RNA and assess quality / concentration.

Some helpful technical guidelines for achieving both of these goals can be found [here](#).

### 4) Choose a fluorescent chemistry: either SYBR® Green I or 5'-nuclease (TaqMan®).

For a quick comparison between the two chemistries, click [here](#).

### 5) Choose gene targets / order Assays.

If you're working in one of several commonly studied species *and* using TaqMan® chemistry, go [here](#) to learn how to search for, select, and order from AB's 1,000,000+ bioinformatically validated, highly specific **Pre-developed Assays**.

Need of a TaqMan® Assay that AB doesn't have already? click [here](#) to learn how to submit a context sequence to us for **Custom Assay design**.

To view various possible **endogenous control** Assays, go [here](#).

For quantifying individual **miRNA targets** or other small RNAs, click [here](#).

If you use SYBR® Green I chemistry, or if you already have your own TaqMan® primer and probe designs, go [here](#) to start the **individual oligo** ordering process.

**Important note:** If you are designing your own assays, or you are submitting sequence to AB for a Custom design, you'll need to do thorough bioinformatics on your sequences. For a guide on how to do this, please read Appendix B of [this document](#).

*To see information on all of AB's Assay products, including pre-plated and array card-based solutions for higher throughput GX studies, go to [www.allgenes.com](http://www.allgenes.com).*

### 6) Choose either singleplex or multiplex PCR option, as well as an RT-PCR chemistry (either one-step or two-step).

See page 34 of [this guide](#) for a brief background on the differences. (Note: multiplex PCR requires much more optimization than singleplex.)

**7) Choose RT-PCR Reagents.**

To find the most appropriate reagents, use the [Real-Time PCR Decision Tree](#).

**8) Validate Choice of a Normalizer Gene. (See [Validation #1 document](#).)**

This validation ensures that your choice for a normalizer gene is appropriate.

**9) Validate Dynamic Range of Reverse Transcription. (See [Validation #2](#).)**

This validation ensures that the amount of RNA that you're adding to your reverse transcription reactions is not so high as to introduce target-specific bias into the RT step.

**10) Validate Assay Efficiencies. (See [Validation #3](#).)**

This validation helps one decide between the ddCt and the Relative Standard Curve Methods.

**11) Set up experimental reaction plate and perform run / analyze data.**

In real-time, run triplicate wells for each of your unknown cDNA samples (converted via RT using an appropriate amount of RNA, as defined in step 9) with both the normalizer and target Assay(s). If using the relative standard curve method, remember to also include dilution curves with each Assay.

For detailed discussions on all facets of running an experiment, including software, please click on the appropriate link (according to your instrument and experiment type) and review the indicated chapters.

7500 series: click [here](#) to access the Relative Quantification guide.

Relative Standard Curve Method: review chapters 2-5.

Comparative Ct (ddCt) Method: review chapters 6-9.

StepOne™ series: click [here](#) to access RQ guide.

(RSCM and ddCt chapter numbers – same as above.)

7900:

Relative Standard Curve Method: click [here](#).

Comparative Ct (ddCt) Method: click [here](#) and see p.p, 23-66; or, watch [this video](#).

**12) Troubleshoot data, if necessary.**

AB offers an excellent on-line tool for troubleshooting real-time data. Just click [here](#) and click on the image that most closely represents the problem you're having.

AB also has a Technical Support Hotline: just call 800-762-4001.

**Note:** For a more abbreviated, product-centric workflow, please click [here](#).

## SUPPLEMENTARY GUIDES

### Validation #1: Choosing a stable normalizer for gene expression

When performing a relative gene expression experiment, the choice of an appropriate normalizer (a.k.a. endogenous control) gene is essential for several reasons. Perhaps most notably, differences in the amount of starting template from sample to sample will affect target gene values. However, when data (in the form of Cts) are collected simultaneously for both the target gene and a *gene that is expressed at a consistent level among the various sample types under study*, expression changes for the target can be calculated irrespective of starting mass.

For example, if one were using the ddCt method of quantification, the first normalization step would look something like this:

	<u>Ct (target gene)</u>	<u>Ct (normalizer)</u>	<u>dCt</u>
Untreated	27.5	19.0	8.5
Treated	28.5	20.0	8.5

Here, a dCt value is calculated for each sample by subtracting the normalizer Ct from the target Ct. Note that were we to look solely at the target gene's values, we would be tempted to conclude that the expression in the two samples differed by two-fold ( $28.5 - 27.5 = a 1 \text{ Ct difference}$ ). However, the normalizer gene's one-Ct difference between the untreated and treated samples tells us that there is actually a two-fold difference in the *concentration* of starting cDNAs. Thus, the target's expression is equal in these two samples. (For more information on the calculations for the ddCt method, as well as for the relative standard curve method, please see the Guide to Performing Relative Quantification or similar reference.)

Clearly, the dCt values that one calculates are only trustworthy if the normalizer gene's expression is truly stable. If one were to select a normalizer gene whose expression instead varied by, say, two-fold between the sample types under study, final fold calculations would presumably be in error by this same two-fold factor. Our suggestion is for each researcher **to validate her chosen normalizer in a series of test samples by showing that its expression is consistent when equal amounts of starting template are used.**

In a perfect world, perhaps we could always start with exactly the same amount of template for each sample; doing so might, in theory, even allow us to bypass using a normalizer gene altogether. However, this approach would require extreme care on our part in the quantification of every sample and would likely prove cumbersome. What's more, reviewers for scientific journals almost universally expect to see a normalizer gene for each experiment, and are not likely to accept a paper that neglected to include one.

**What follows is a step-by-step procedure for testing the stability of a single normalizer in one's experimental samples.** Just be aware that alternative methods of normalization are advocated by some members of the scientific community, including the use of multiple normalizers as an effort to minimize the error of any single choice. Others argue that normalization based on starting template mass – which lies at the heart of the approach described here – may not always be appropriate, depending upon the cell types being studied and the experimental question being asked. As these are somewhat complex considerations beyond the scope of this discussion, further reading may be necessary.

Here is the step-by-step process for our recommended validation:

- 1) Choose 3-4 representatives for each distinct sample type (treatment, tissue source, time point, etc.).
- 2) Isolate total RNA using a sample prep method appropriate for your sample type(s).
- 3) DNase-treat.
- 4) Quantify.<sup>1</sup>
- 5) Reverse transcribe equal mass amounts of RNA for each sample.<sup>2</sup>
- 6) Test an equal volume of each cDNA in real-time using the candidate normalizer gene(s).
- 7) Assess the variability in Ct among the various sample types.

<sup>1</sup> A traditional A260 reading will measure not only RNA, but also DNA and free nucleotides. If you suspect that samples may contain unequal amounts of these, such that A260 readings may prove inaccurate, consider quantifying using a fluorometer and an RNA-binding dye such as RiboGreen®.

<sup>2</sup> In order to minimize the chance that different samples will reverse transcribe at different efficiencies due to the presence of inhibitors, use *significantly less RNA* than the reverse transcription kit recommends. Even adding one-tenth the suggested amount should still prove sufficient for testing a typical normalizer.

Whether or not the normalizer passes or fails this validation depends on the agreement of Cts among the various samples. That said, rarely if ever will the amplification curves generated during this validation be identical; thus, the researcher must decide if the observed variation is acceptable. The answer really depends on her experimental goals and expectations for the accuracy of final data. One simple guideline is as follows: a normalizer that differs by 1 Ct among the tested

samples can reasonably be expected to introduce a two-fold error into final data calculations if used in an actual experiment on similar samples. So the researcher must ask herself, "Is this an acceptable error?"

As a final note, testing multiple normalizer genes at once can greatly expedite one's search. (We strongly recommend including 18s as one of the genes to be tested.) Besides our large selection of individual pre-designed Assays (including primer-limited, VIC-labeled assays for multiplexing), Applied Biosystems offers several pre-designed panels for doing just this, depending on one's real-time platform. A list of these, along with part numbers, follows.

Individual TaqMan® endogenous control assays: click [here](#).

96-well standard or fast plates . . .

Endogenous Control Plate – 32 potential control genes, each pre-pipetted in triplicate wells:

Human

Part # 4396840 (standard block) or # 4426700 (fast block)

Mouse

Part # 4426701 (standard block) or # 4426697 (fast block)

Rat

Part # 4426698 (standard block) or # 4426699 (fast block)

TaqMan® Array Card (7900 / TaqMan Array thermal cycling block only):

Human Endogenous Control Array – each card contains 16 commonly used control genes in each of 8 loading ports.

Part # 4367563 (2 cards)

Mouse

Part # 4378702

Rat

Part # 4378704

[Return to top](#)

## **Validation #2: Determining dynamic range of reverse transcription**

Choosing an input mass of RNA for reverse transcription reactions is viewed by many researchers as a trivial matter. Most simply use the maximum amount recommended by the RT kit instructions, assuming this guideline is appropriate in all cases. In terms of ensuring that a given reaction does not exhaust the reagents provided by the RT kit, this specification is generally accurate.

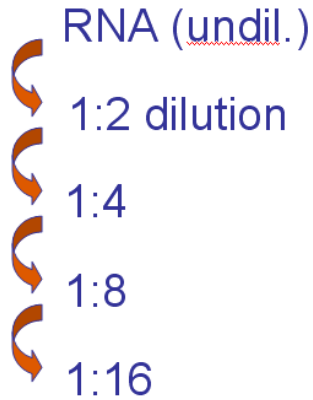
Unfortunately, it is not merely the limitation of RT reagents that influences the acceptable input range. The **customer's samples**, as well as the **genes under study**, also play critical roles in determining how much RNA can safely be added.

It is therefore each user's responsibility to determine the dynamic range of reverse transcription empirically. Failure to do so can lead to erroneous results – errors that, in many cases, will never be detected.

The way in which one tests the dynamic range of RT is by running **RNA dilution curves**. This is a simple and elegant experiment that is carried out as follows.

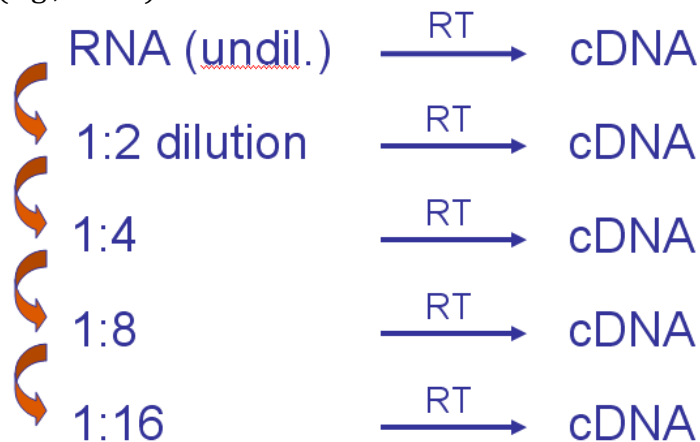
The user begins by selecting an RNA sample from among his or her unknowns. This sample must be one that was isolated in the same manner as all the other samples: same sample prep approach, same approach to removing residual EtOH, same final diluent, and preferably the same user. If the quality among one's samples is thought to be variable, then testing multiple samples might be in order. Alternatively, some researchers combine multiple samples and use the mixture as a starting point for this validation, on the thought that a pool will be more or less representative of all samples.

A good concentration to begin with is the one recommended by the RT kit: e.g., 1 ug in a 20 uL RT. But before doing the actual conversion to cDNA, the customer should dilute this sample in nuclease-free water, as follows.

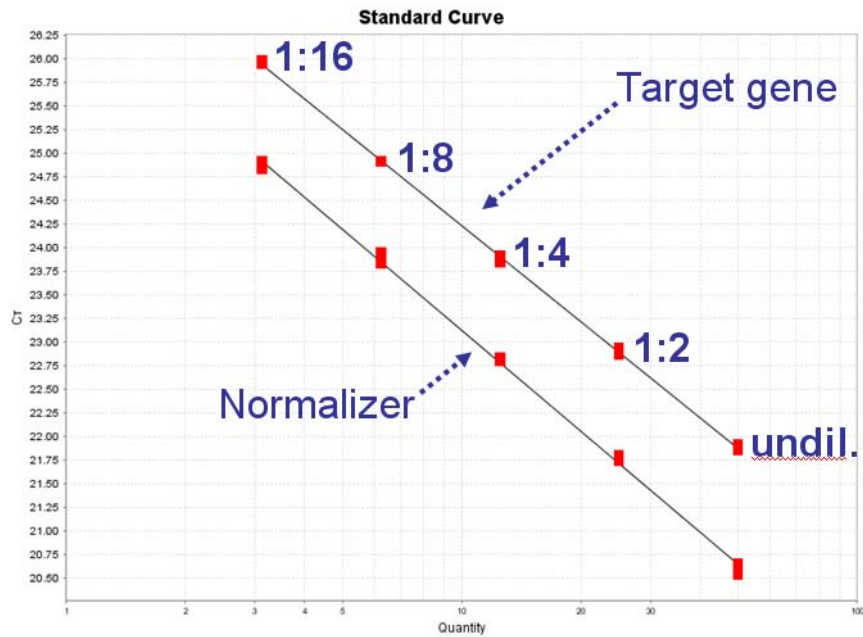


(Note: Depending on several factors, including sample quality, a larger starting amount and / or dilution factor may be desirable or even necessary. As a result, some researchers will start with more RNA than is recommended by the RT kit and perform 1:5 or even 1:10 dilutions. In either case, a minimum five-point curve is suggested.)

Next, each of these dilutions should be converted via RT in the same reaction volume (e.g., 20 uL).



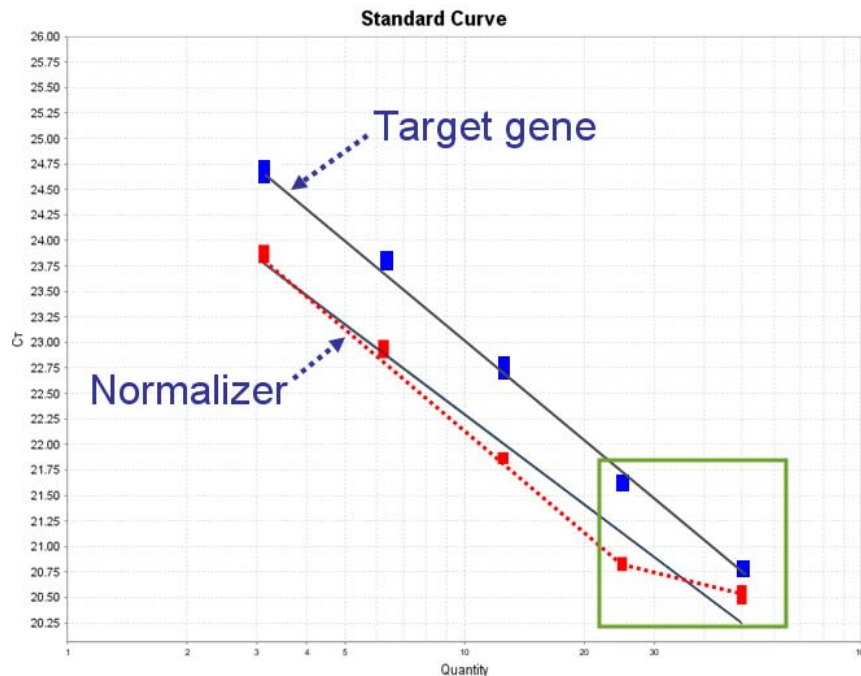
Once the reaction is complete, one simply amplifies in real-time an equal volume of each resulting cDNA with both the normalizer Assay and the target Assay(s). If the user employs the Standard Curve template in AB's real-time instrument software, curves similar to the following will be generated.



What we learn from the above result is that RT efficiency is consistent across all starting concentrations of RNA (1 ug down to 62.5 ng). We know this because all dilution points fall in a straight line, as indicated by the nice correlation coefficient ( $R^2 = 0.99$  in this case).

**Conclusion:** with this sample set and these particular gene Assays, the researcher can presumably use any starting amount of RNA in his or her RT from within the tested range (assuming consistent sample prep).

Unfortunately, the results of this test are not always so favorable. Occasionally, they will look more like the following plot:



In this instance, the normalizer and target genes are behaving differently during reverse transcription for the most concentrated amount of starting RNA. While the target gene is converting in a linear manner, the normalizer is noticeably suppressed, as indicated by the higher-than-expected Ct. This effect is likely caused by one or more **RT inhibitors** – compounds that slow down the enzymatic reaction when present at substantial concentrations. Diluting the RNA beyond a certain point (1:2, in our case) alleviates the problem, such that all subsequent dilutions indicate linear reverse transcription for both Assays.

(Note: it is not at all unusual to see such an effect, whereby the higher concentrated gene Assays are inhibited to a much greater degree during RT compared to lower-expressed targets.)

Why do we care that this is happening? Simple: if this user had not identified this issue up front, and had instead converted all of his unknown samples into cDNA using the highest concentration of RNA, he would have risked a variable RT efficiency among his different samples. This could in turn have led to incorrect measurements of his normalizer during real-time amplification – errors that would then have adversely affected the final numbers.

**Conclusion: where inhibition of RT is observed, the user should consider either (a) using an amount of starting RNA that does fall within the linear dynamic range of RT, or (b) rethinking his or her approach to sample prep, such that the inhibitors are substantially removed prior to RT.**

What are these inhibitors? There are two categories to consider.

First, there are inhibitory compounds native to the sample itself. These include melanin, hemoglobin, humic acids, polysaccharides, and so on. Basically, anything that can inhibit PCR can also be a problem for reverse transcription. One should always choose a sample prep method that eliminates these compounds in so far as that is possible. (See Ambion's [Real-Time PCR Decision Tree](#) for assistance in choosing an appropriate kit.)

The more common – and insidious – class of inhibitors are sample prep reagents themselves. Compounds such as phenol, chloroform, guanidinium, proteinase K, and ethanol are potent inhibitors of RT. The last in this list deserves special mention, as ethanol is employed by almost every sample prep approach during the step prior to final elution / resuspension. Removing as much EtOH as possible, while not overdrying the RNA pellet (such that resuspension becomes problematic), is critical.

One common question is as follows: “Do I have to test every Assay I intend to use for quantification?” The answer is, “Generally not.” This is because limits on the dynamic range of reverse transcription tend to be mass-dependent. So simply testing the normalizer (generally the highest expressed target in one’s study) and one or two targets of varying expression levels is probably adequate.

What about testing every sample? Again, this should be unnecessary, if two precautions are followed: (1) care is taken to be very consistent in one’s sample prep across all samples (e.g., in the removal of residual EtOH); and (2) one consistently uses an amount of starting RNA that is considerably less than the highest amount tested. So if the researcher tests 1 ug / 20 uL RT as the highest concentration for this validation experiment, using 100 ng of each experimental sample should eliminate concern of sample-dependent inhibition.

A final note: many researchers assume that a 260/280 spec reading (with attention to 230, where organic compounds can often be detected) is sufficient for assessing sample purity. However, laboratory tests indicate otherwise. Even 260/280 ratios for RNA considered by most to be “excellent” (1.9-2.0) have been associated with significant levels of protein contamination. For this reason, a more certain way to test the dynamic range for reverse transcription is through the validation described above.

[Return to top](#)

### **Validation #3: Validating assay amplification efficiencies**

Knowing the efficiencies of one’s target and normalizer assays is important in a gene expression experiment, as this information is used to decide which

experimental approach one should employ: the **comparative Ct (ddCt) method**, or the **relative standard curve method**. In short, only when the efficiencies of the target and normalizer assays are identical or very similar can the former method be used. (For more information on these two methods, including how to make an appropriate choice between them, please see the [Guide to Performing Relative Quantitation](#).)

The most common way to test an assay's efficiency is by amplifying a dilution series in real-time and examining the resulting slope of the regression line. While doing so may seem like a fairly simple endeavor, this kind of experiment is in fact fraught with pitfalls. Indeed, one of the most common errors among real-time users is generating inaccurate dilution curves, and therefore misinterpreting the tested assay's efficiency. We will here make several specific recommendations to help minimize the most frequent errors. (For an excellent reference on both pitfalls and solutions to accurate efficiency determination, see [Amplification Efficiency of TaqMan\(TM\) Assays](#).)

First, one should generate the curve from a dilution of an inhibitor-free nucleic acid. Most users choose a concentrated cDNA containing their assay-specific target in significant quantities, while others prefer artificial template such as plasmid containing their target of interest. Indeed, the choice should not matter, as long as that nucleic acid is both clean and sufficiently concentrated to allow for a series of dilutions.

What dilution should one use, and how many points? **We recommend 1:10 dilutions over at least five points.** Such a curve covers a full four logs of dynamic range – large enough to minimize the effect of minor pipetting errors. One of the most common mistakes is generating, say, a three-point curve of two-fold dilutions. A curve of such small magnitude will almost inevitably produce inaccurate results, as even small pipetting errors become magnified.

Assessing the confidence with which the software establishes the regression line following the real-time run is also critical. The linearity of the dilution points, combined with the precision of pipetting replicates, affect the correlation coefficient (denoted as the  $R^2$  value that accompanies each curve). A value of 0.99 or higher is generally considered acceptable. A value lower than 0.99 indicates that the software has drawn the regression line with some reservations, due to one or more points not falling on the line. Assuming outliers are few (i.e., one or perhaps two points per curve), one can generally omit these wells and greatly improve his results. However, multiple outliers likely indicate a more widespread problem, such as an improperly diluted curve. In these cases, we suggest repeating the experiment.

A few additional suggestions for generating accurate dilution curves follow.

- 1) Always dilute the template in a serial manner, rather than preparing each dilution from the stock tube.

- 2) Mix each dilution point very thoroughly by pipetting or vortexing before making the subsequent dilution. (Simple tapping or inverting the tube is not generally adequate.)
- 3) Always amplify at least three replicate wells for each dilution point.
- 4) Following the run, if the standard point representing the highest concentration does not fall in line with subsequent dilutions, the standard may contain inhibitors. Remove this point entirely from consideration.
- 5) For this validation experiment, have more than one person in the lab generate a dilution curve for each new assay, preferably using different pipettors. If the two slopes disagree, search out the source of error. (User? Pipettors?)

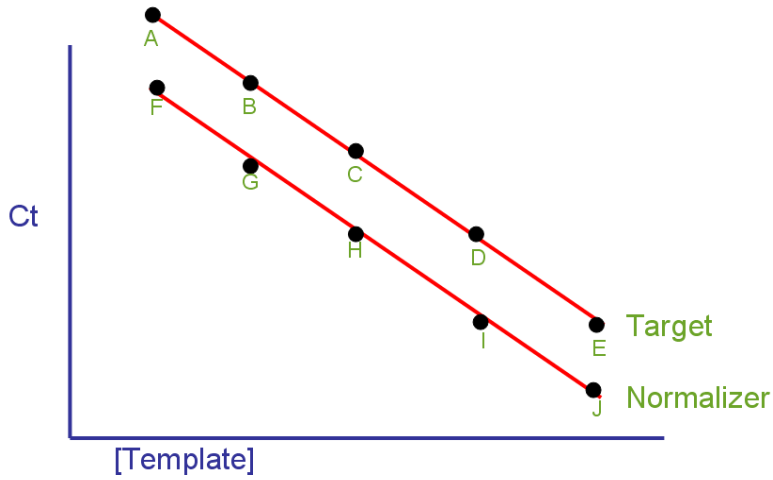
A common question concerns calculated efficiencies greater than 100%. Clearly, PCR can do no more than double the target with each round of PCR, so slopes that are less negative than -3.3 (say, -3.0) should, in theory, be impossible. A miscalibrated pipetman is often the culprit; for this reason, well-maintained pipettors are essential for trustworthy results. That said, even using accurately calibrated pipetmen and exercising great care in performing the experiment can occasionally lead to “impossible” slopes – a fact that highlights the difficulty of achieving perfectly accurate dilution curves.

Once the slope has been determined for your Assays, the following calculation can be used to determine their respective numerical efficiencies:

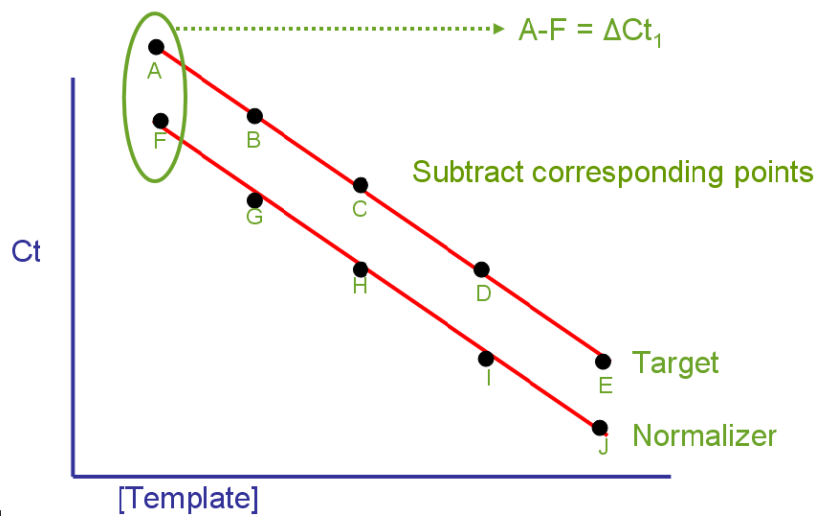
$$E = 10^{(-1/\text{slope})} - 1$$

The user’s next task is to determine if the ddCt method for relative quantification is appropriate for her given each pair of assays. Generally speaking, the slopes of the curves (and hence their efficiencies) must either be identical or very similar to one another in order to subsequently make use of the ddCt method. (We must allow for some error, as rarely will the calculated values be exactly the same.) The trick, then, is to mathematically determine the difference, then to set a cutoff for the acceptable difference. The most commonly accepted approach for doing so is as follows.

First, imagine that each point on the curve has a unique letter, as shown below:



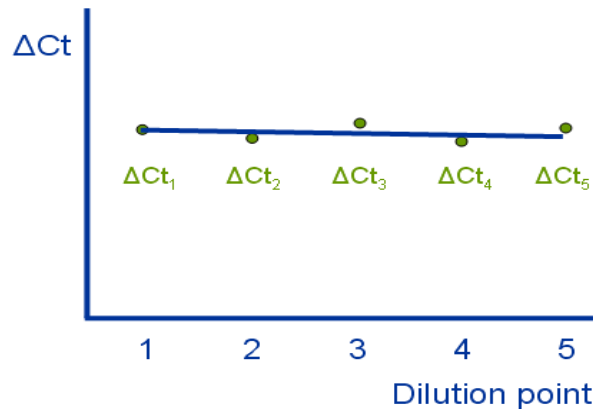
Now, subtract F from A in order to determine the distance (in Ct) between the two lines at this point.



Now do the same for the other paired points (B-G, C-H, etc.), as follows.

$A-F = \Delta Ct_1$ $B-G = \Delta Ct_2$ $C-H = \Delta Ct_3$ $D-I = \Delta Ct_4$ $E-J = \Delta Ct_5$
---

Now graph these points in Excel.



If the two dilution curves are parallel, clearly the regression line pictured above will have a slope of 0. However, as long as the slope is greater than -0.1 and less than 0.1, the difference in efficiency between the two assays is not significant, and the ddCt method is appropriate with this assay pair. If not, one should probably consider using the relative standard curve method.

While perhaps overly subjective to some, another useful test of efficiency equality between two Assays is a simple visual inspection of the amplification curves. When two Assays produce amp curves whose geometric phases are parallel, regardless of the slopes of their respective dilution curves, the efficiencies are either identical or very close. This is a crude but useful test to carry out for confirming whether or not one's dilution curves are accurate. (Example: if dilution curves for two Assays are producing slopes that differ greatly, but the amplification plots are clearly parallel, there is almost certainly a problem with the dilution curves.)

One final note: Applied Biosystems has well over one million pre-developed Assays, thousands of which have been tested and validated for the ddCt method. Please see the following white paper for details.

[Amplification Efficiency of TaqMan\(TM\) Assays.](#)

[Return to top](#)