

Host: This is the podcast from *Clinical Chemistry*. I am Bob Barrett. qPCR, or quantitative PCR, is a valuable technique for accurately and reliably profiling and quantifying gene expression. Samples obtained from the organism of study typically have to be processed through several preparative steps before qPCR.

A report published in the October issue of *Clinical Chemistry* recommended the use of sample replicates preferentially to any other replicates when working with solid tissue, cell cultures, and single cells, and recommended the use of RT replicates when working with blood. The study also showed how an optimal sampling plan can be calculated for a limited budget.

Dr. Ales Tichopad is a co-author of the paper and is a reader at The Technical University, Munich, where he specializes in real-time PCR technology and its applications. He has developed several methods of data analysis and has been a freelance consultant to the biotechnology industry since 2006. He is our guest in this podcast.

Dr. Tichopad, the focus on error in the quantitative real-time PCR experiments moves from the instrument precision to sample preprocessing noise. So how important is the preprocessing error?

Dr. Ales Tichopad: In light of our results, the preprocessing error actually dominates over the instrument imprecision in most cases. Yet, if you look at most papers today, reporting data from quantitative real-time PCR experiments, you will actually see that technical replicates are used rather than any other replicates. Only seldom replicates are done for RT or sampling.

This is not an adequate approach, however, we believe. The error borne in the real-time instrument is minor, as compared to the noise due to sample processing, for example. And if quantification is performance in solid tissue sample or even in a single cell, this preprocessing noise may account for more than 99% of the entire noise.

Host: So what other error sources are involved in the real-time qPCR assay?

Dr. Ales Tichopad: In a typical experiment, there are various operations required before we obtained the C_q value. One should think of an experiment as of a hierarchical process, where one processing step is followed by the other.

Typically, samples are obtained from experimental animals, we call them subjects here. These samples may be stored, sometimes even for years, and then they are taken for

mRNA extraction. The mRNA is then reversed transcribed, and the cDNA is then quantified. Therefore, we have at least some three levels where a noise may enter the experiment: the sampling, it's the RT, and it's the qPCR.

Of course, this number can even be larger, if you, for example, consider the extraction as a separate step, or you may even take particular steps from the extractions and create replicates here. So, therefore, all the chain is actually an essential entry for an error to the experiment.

Host: Tell us, how exactly do various sample types compare in terms of error structure?

Dr. Ales Tichopad: It seems that there are two main types of biological matrices: suspension sample, such as blood or cell culture in a medium, and solid samples, such as tissue specimen and single cells.

The solid tissue is fairly common material to use, and it is the major challenge actually, because of the large processing noise.

We found out that there is a substantial variation between tissue samples obtained from the same subject, at the same time even. Three, four differences were not seldom in our experiment.

This imprecision is larger than any common difference between two experimental animals; we use cows in our study. One actually cannot reliably distinguish between two individuals having only one tissue sample from each of them. That means the error that is there from the sampling process is much larger than the common inter-subject variation.

The sampling noise accounts for more than 90% of the total noise. The RT-related noise was approximately about 5% of the overall noise. And finally the PCR noise was only around 1%.

Surprisingly, low-copy transcripts behave differently. We could see much larger effect of the RT and the PCR. It is not that obvious to us why, but perhaps it has something to do with the pipetting of low concentrated analyte. This is something that is often referred to as PSO error or PSO effect.

(00:04:53)

The single cell samples are much like the solid tissue, each in the sample variability, more than 90% of the total

variance. Just accounts for the difference between different cells obtained from the same organism.

In blood, we get various results. Most important, the sample, that is actually the tube with the blood, is finally precise and consistent. The noise at this level is rather motive, the RT noise seems to dominate in terms of percentage of the total noise.

The inter-subject differences are rather visible, and we can actually reliably work with the single sample obtained from any subject.

If you look at the comparison, blood has a solid tissue, at times it's pretty obvious that the reason for the large imprecision is the sample heterogeneity, rather than the extraction. Blood is simply more homogeneous and delivers more consistent measurements.

Last, the cell culture data showed that the various samples, meaning here the different flasks or the different cell culture wells, show approximately the same noise of the RT. Here again the PCR noise is minimal. Note, however, that there is typically nothing like subject in the cell culture experiment. Usually people work with the polygonal population of cells, this is little bit unique for these types of experiments.

Host: Is it common to report gene expression data for individual samples? You introduced this subject as the experimental unit. What's the difference?

Dr. Ales Tichopad: Yes, subject is actually the experimental organism, say a mouse or human. This is what we study and interpret in most cases. qPCR is used by biologists, and biologists study biological organisms, subjects. We do not really interpret samples of — there may be actually several samples taken from a single animal, and they may be very different one from the other, even if taken from the same individual.

Think of an experiment where you compare an effect of the tract with placebo on a hamster liver, for example. Typically, you compose two experimental groups, one treated with the tract, the other treated with the placebo. A reasonable experiment would be to take several preferably, tens or hundreds of hamsters, where each treatment group would be created and consist of several animals.

Then the statistics is calculated comparing the means of the groups. The sample is only the middleman here actually. You always have to average out the samples to get one value for each of the subjects.

Of course, no one would like to do an experiment with tens or hundreds of animals, but sometimes it may just be needed to observe minor effects.

We should know that the number of subjects can be decreased however, if a proper design with the sufficient number of sample replicates is used. Why? Because this increases the accuracy of each single subject's measurement. If the accuracy is high, we more easily see the effect. This is something one should consider before too many animals are sacrificed in experiments actually.

The variation between individual subject should however not be considered a noise. It is natural that individuals worry for various reasons. Therefore, as long as we intent to conclude that say a drug has a general effect on any considered biological population, we should not try to minimize this natural variability by say taking only clonal subjects or taking siblings, which is relevant when working with cell cultures, only seldom people consider nature or genetic variability by taking several cell lines. Usually only very homogeneous cell culture samples are studied.

It is fine, and it generates consistent results, but nevertheless, we have to be careful with any general conclusion in such experiments. Easily the result we got may be valid only for cell line, which was used in our experiment.

Host: Using reference genes is a method to correct for noise. Isn't this enough to clean the noise?

Dr. Ales Tichopad: The reference gene is a gene that is assumed to have a constant level of its transcript in the cells. Provided this assumption holds, the normalization with reference gene or several genes is justified. However, we also assume that this is not really stressed enough, that the sample processing noise that affects the target is the same as the noise that affects the reference, mainly in terms of the error structure that propagates throughout the sample processing.

(00:09:52)

For example, the extraction efficiency should be the same for both genes. We actually say that these two genes should have this very high covariance, that is the same trend in variation throughout the sample processing should be given.

We could, however, clearly see in our experiment that low-copy genes and high-copy genes have completely different processing characteristics. Therefore, one should never

normalize low-copy target with the high-copy reference or vice versa.

Host: I believe that experimental design is no longer only a matter of precision and accuracy but also a financial issue. Is there any way to optimize costs for experiment?

Dr. Ales Tichopad: Unfortunately, when designing an experiment, we have also other objectives in mind than just getting the strongest experimental evidence by including as many subjects and replicates as we can get. Money is always an issue.

Nowadays, I see an increasing discussion about pilot experiments, a small experiment done before real, full-scale experiment is conducted. I consider this extremely useful.

I believe that this kind of pilot studies will establish routinely the time. Such a small-scale pilot is useful to set up the food scale design mainly, in terms for replicates and reference genes. One can relatively easily calculate the error components in all preprocessing steps. This is actually what we did in our paper.

We also provide and reference in the paper a software called PowerNest, that is a free wire. This software can slice the overall noise into components using something that's called "nested Anova" analysis, a particular statistical procedure, and then one just has to assign cost to each type of sample preprocessing operation.

Obviously, the error or noise gets more with increasing number of replicates at every level. Ideally, one just creates as many replicates at the highest level that would be taking as many biological subjects as possible, but this is exactly why we are limited to the top budget.

Therefore, it may actually be more cost-effective to increase number of replicates at the sampling level, for example, rather than to increase the number of RT and PCR replicates that is frequently done today.

People stress a lot the importance of the technical replicates, however, there is no rationale behind that.

And then knowing the cost for each processing step, one may calculate optimal design for a given budget, or limited number of subject, and this is actually what PowerNest software is doing.

Host: Dr. Ales Tichopad is a reader at The Technical University, Munich, and a freelance consultant to the biotechnology industry. He has been our guest in this podcast from *Clinical Chemistry*. I am Bob Barrett. Thanks for listening.

Total Duration: 13 Minutes