Designing cost-optimal qPCR experiments with PowerNest

PowerNest is a software tool enabling experimenters to explore the effect of sampling on noise propagation throughout qPCR assays. The sampling process is assumed to be comprised of a number of levels; the acquisition of a sample and the preparation of extracted material, reverse-transcription of the mRNA, and the qPCR itself. Given a small set of results, representative of a larger assay, the error at each stage of the experiment is profiled using a nested-ANOVA.

Armed with this information, PowerNest allows the experimenter to explore the effects of modifications to the experimental design on the expected total error of the assay. When given the financial cost of replicates at each level, PowerNest will calculate a cost-optimal sampling-plan, delivering an experiment design that will minimise processing error and maximise the statistical resolution of the assay.


Rob Kitchen[1], Ales Tichopad[2]

[1]University of Edinburgh, National e-Science Centre
[2]Technische Universität München, Physiologie Weihenstephan

qPCR sampling

Typically, samples obtained from studied organisms are processed in several preparative stages before being analysed using qPCR:

1) subjects are randomly selected from the population
2) mRNA samples withdrawn from each subject
3) RT products obtained from each sample
4) qPCR analyses performed on each RT

This hierarchy (see also Figure 1) is susceptible to the introduction of processing errors at each level that will propagate down through successive levels and be reflected in the observed qPCR result. The processing errors present in the final result serve only to confound the interesting biological variation, such as the effect of a treatment or the variation between subjects. It is therefore the prerogative of the experimenter to design the assay in such a way as to minimise the scope for the introduction of this confounding variation.

To create the optimal experimental design, knowledge of the sources of error throughout the sample processing stages is essential.

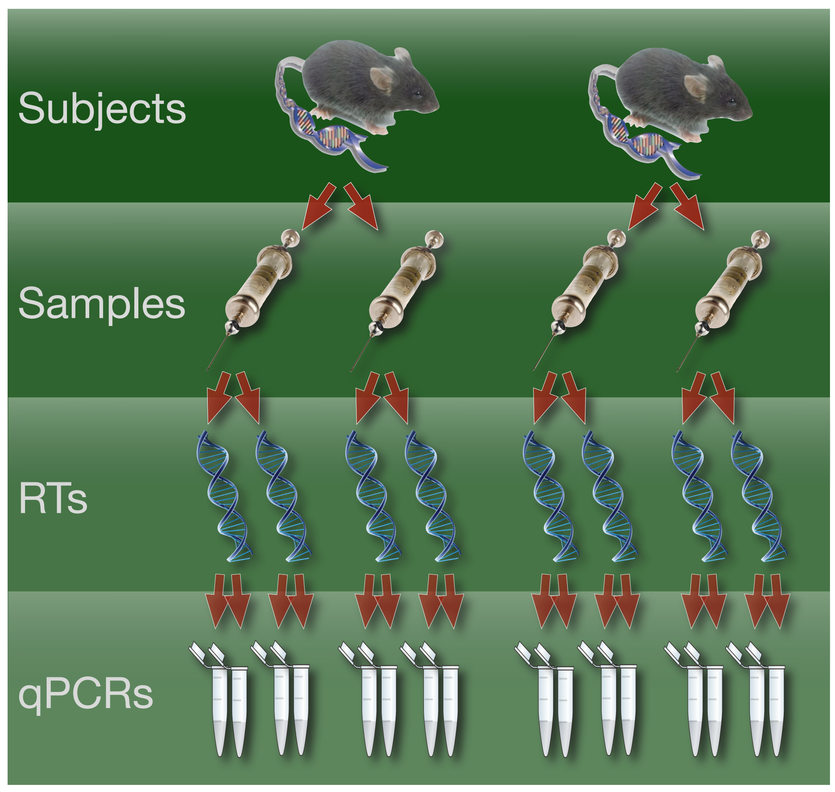

Figure 1: Typical nested experimental design involving two replicates at each level

Statistical model

This experimental **hierarchy is modelled using a nested-ANOVA**; similar to a crossed design except factors meet in unique combinations. The nested design is applied as any given qPCR replicate cannot be dispensed from two different RT tubes, just as an individual sample cannot be obtained from two different subjects. The model is as follows:

$$CT_{gijkl} = \mu_g + a_{i(g)} + b_{j(gi)} + c_{k(gij)} + d_{l(gijk)}$$

This states that the observed CT value is a linear combination of the mean treatment effect, μ_g, for each group, g, and the random contributions of replicates at each of the sample processing steps; $a_{i(g)}$ is the contribution of the i^{th} subject, $b_{j(gi)}$ that of the j^{th} sample from the i^{th} subject, $c_{k(gij)}$ that of the k^{th} RT from the j^{th} sample from the i^{th} subject, and $d_{l(gijk)}$ that of the l^{th} qPCR obtained from the k^{th} RT from the j^{th} sample from the i^{th} subject.

- this **model accurately reflects the linear propagation of processing errors** through successive levels, and enables the estimation of the relative and absolute processing errors introduced at each level from the raw CT data
- it is assumed **that increasing the number of replicates**, n, at a given level will result in a **reduced contribution to the overall error** propagation from that level:

$$\sigma^2_{\overline{CT}} = \hat{\sigma}^2_i / n_i + \hat{\sigma}^2_j / n_i n_j + \hat{\sigma}^2_k / n_i n_j n_k + \hat{\sigma}^2_l / n_i n_j n_k n_l$$

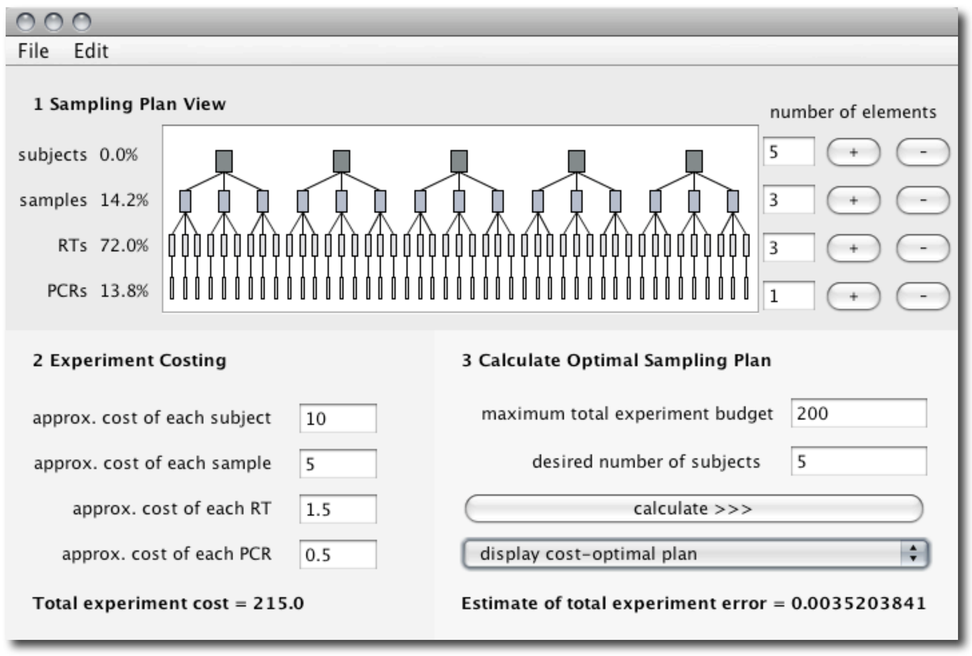

Figure 2: PowerNest allows the user to interact with the design of the experiment and observe the effect of changing its structure on the expected total error.

PowerNest

Illustrated here in Figure 2, the features of PowerNest provide an intuitive means by which to optimise the qPCR sampling process. In summary, the software:

- **profiles the variance structure** of a qPCR assay using the nested ANOVA model; the data are input by either:
  - providing an Excel spreadsheet containing Ct data
  - manually specifying the variance component at each level
- exploits these measured or expected variances to **extrapolate the experiment design** to greater numbers of replicates
  - the number of replicates at each level can be manually adjusted and the software provides an updated estimate of the total error of the assay
- provides the opportunity to **calculate the total cost of an assay** based on the individual cost of replicates at each level
- offers **automatic and exhaustive calculation of the optimal experiment design**, given a maximum budget and/or number of subjects, based on the variance structure of the assay.