Statistical aspects of validating microarray results with qRT-PCR

Claus-Dieter Mayer

Biomathematics & Statistics Scotland, BioSS Office, Rowett Research Institute, Aberdeen, UK

For a complex high throughput technology like microarrays it is very difficult to guarantee the correctness of the results for each and every gene present on the array. For this reason it has become a standard practice to validate results for a small list of genes, which showed the most significant changes in the microarray experiments. A commonly used validation method is quantitative real-time RT-PCR (qRT-PCR). Typically the validation uses the same samples that have been analysed in the microarray experiment. Doing this might remove general doubts about the microarray technology, but there is also a widespread misconception in the life science community that statistical uncertainty (usually reflected in a p-value or a false discovery rate (FDR)) will be eliminated at the same time.
By introducing a simple model for the joint distribution of the two types of measurement, we can derive tests for differential expression that are based on both technologies and demonstrate that the gain in power by adding the validation data is very small in typical situations. This is quite obvious from a statistical point of view as this type of validation is a form of pseudo-replication that can only reduce technical but not biological error. Nevertheless we think it is an important point to make and to communicate in order to ensure good scientific practice.
In a second part we will study this two-stage procedure (stage 1: microarray experiment, stage 2: qRT-PCR validation) for the case that biologically independent samples are being used in the validation step. As the genes selected for the validation depend on the outcome at stage 1, we can regard this as a special case of a (group) sequential testing procedure. On the other side the combination of two different experiments resembles problems in meta-analysis and we will discuss how methods derived for these statistical areas can be applied in this situation. We will particularly study the problem of controlling the overall FDR in this two-stage testing procedure.