

Experimental Design Considerations

Because it is not practical to develop one experimental design that is appropriate for all microarray experiments, we will discuss several factors to take into consideration when designing microarray experiments for measuring gene expression changes in biological systems.

The arrays should represent replication at the highest level by incorporating biological variability because the primary goal of most microarray experiments is to understand the gene expression differences between 2 or more classes of biological (RNA) samples. Replicate arrays derived from one RNA sample will provide information on variability introduced during the microarray analysis process itself rather than variability found in the biological system.

For most experiments, we recommend including a minimum of 3 independent biological replicates for each sample class; however, there are instances where more than 3 and less than 3 replicates are appropriate. Please see the suggestions below for experimental designs that include 3 replicates, more than 3 replicates, and less than 3 replicates.

3 Replicates

- Each biological replicate should consist of RNA isolated from pooled tissue collected from 3 or more animals when possible. (This may not be appropriate if the experimental goal is to specifically measure animal-to-animal variability.) Pooling samples is not a substitute for experimental replication but will normalize animal-to-animal variability and variability introduced by sample collection procedures.

More than 3 Replicates

Below are some examples of studies that may require more than 3 replicate samples.

- Studies where higher levels of variability might be expected (i.e., tissue or cells that are difficult to isolate and may be more vulnerable to variability introduced during the collection process).
- An experiment that involves the analysis of tissues that are heterogeneous (i.e., tumor cells or tissue and organ biopsies).
- Studies to develop class prediction algorithms based on gene expression data typically require larger data sets for analysis.
- Studies where RNA samples need to be pre-amplified before processing by standard labeling protocols are known to exhibit greater variability and may require additional replicates.

Less than 3 Replicates

The following example illustrates a study where fewer than 3 replicates may be suitable.

- A preliminary study involving a broad time course variable where the goal is to define a narrower timeframe for follow up with additional replicates. Although statistical confidence in specific gene changes may be lower due to fewer replicates, significant global gene expression changes can usually be detected confidently and used to identify key time points for further study.