Practical aspects of assessing two-color spotted array quality

Jean Yee Hwa Yang
University of California, San Francisco
http://www.biostat.ucsf.edu/jean/

Statistics for Microarray Data
IBC Short Course, 11 July 2004

Quality assessment: practice

Array Hybridization quality
- Qualitative: Diagnostic plots
- Quantitative: Composite score

Spot quality

Print quality
- 9mers
- QC Hybs

mRNA quality
- Bio-analyzer
- OD readings

Generating quality diagnostic plots

A directory of gpr files from the same print-run

Download R and Bioconductor packages

library(arrayQuality)
ppQuality(organism="Ma")

Diagnostic plots (png)
Quality reports

arrayQuality package

- Graphical tools to look at print-run and array hybridization quality
- Print-run quality:
  - 9mers analysis
  - Quality Control hybridization (QCHyb)
- General hybridization quality:
  - Diagnostic plots
  - Composite quality score
  - Bioconductor package, written in R
  - All code is open source and can be adapted to suit individual core facilities.

Print quality I

Aim: verify that each slide has been printed correctly
Random 9mers hybridization:
- Spot morphology
- Presence or absence of spotted long-ols or cDNA.
- Output includes list of probes with insufficient amount of spotted DNA.

Print quality II

Mouse:
- Liver vs Pooled

QC hybridization:
- Signal reproducibility comparison to previous QC hybridizations.
- Verify accuracy of GAL files.
Lecture 1: Spotted Microarrays

General hybridization quality

Spatial plots: Which is better?

Dot Plots: Which is better?

Histograms: Which is better?

Example

Signal to Noise

\[
\text{Signal/Noise} = \log_{10} \left( \frac{\text{spot intensity}}{\text{background intensity}} \right)
\]

Things to look for:
- Values of the mean and variance.
- Position of the density curve for "empty" and "negative" controls spots.

Positive controls
Potential for spike

Negative controls

Precision of control spots

17 replicates / probe

Empty

Data provided by Christina Lewis, UCSF

Mean 0.28, var 1.28

Mean 2, var 3.28

Mean 0.87

Var = 1.28

Mean = 1.35

Var = 2.17
gpQuality: Comparative boxplot

- Create a collection of "good" arrays as reference.
- Generating a series of quantitatively
  quality measures.
  - Variation of Green background.
  - Signal to noise ratio.
  - Median to mean pixels intensity.
  - Variance of replicated control spots.
- For every new spot arrays:
  - Check if the new slide value is
    within this range by plot the scaled
    results for visual comparison.
  - Calculate the corresponding
    percentile value for each measures.
  - Calculate the number of measure
    below acceptable range.

Summary & future directions

- arrayQuality provides graphical tools for quality assessment at
  the print-run and array hybridization levels.
- All code is open source and can be adapted to suit individual core
  facilities.
- Future: In collaboration with Ash Alizadeh and others, we are
  designing a new MEE80 70mers oligo set which contains a lot more
  QC control probes in the array design.
- Information from the QC control probes will be incorporated in our
  future quality assessment.

Acknowledgment

- Core Facility members:
  - David Erle
  - Andrea Barczak
  - Jay Sandler
  - Michael Salazar
  - Agnes Paquet
  - Christina Lewis