GCB/CIS 535 Microarray Topics

John Tobias
November 3rd, 2004

Overview

Experimental Design
Technology
Replicates
Experimental Execution
Data Processing

Overview

Experimental Design
Technology
Replicates
Experimental Execution
Data Processing

Experimental Design

Identify Critical Conditions
Identify Critical Comparisons
Minimize Conditions
Maximize Replicates
Consider conditional changes in cell type representation - Do you need laser capture?

Pooling

Benefits
More RNA - avoid double amplification
Caveats
Learn information about the pool - may not apply to individuals
A bad sample can have a broad effect

Overview

Experimental Design
Technology
Replicates
Experimental Execution
Data Processing

Technology Choice

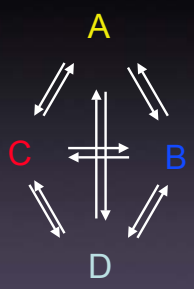
One Color - Affymetrix
Simplifies experimental design
Robust manufacturing process
Two Color - cDNA or long oligo spotted
Lower per-array cost
Custom arrays - for organism or application

Design Examples

Conditions	One Color	Two Color
2		
3		 x3 or  x3

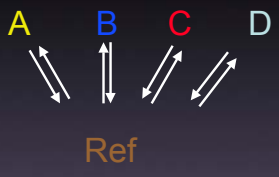
Two Color Examples

Loop



Two Color Examples

Common Reference



Two Color Design Choices

Loop
All pairwise comparisons are direct
Expensive for several conditions
Hard to add conditions

Common Reference
Design remains simple for multiple conditions
All pairwise comparisons are indirect
Choice of appropriate reference is tricky

Overview

Experimental Design
Technology
Replicates
Experimental Execution
Data Processing

Replicates - you can't have too many

What is a replicate?
Replicates add statistical power to your data
How many you need depends on the variation inherent in your system
You probably don't need technical replicates

Overview

Experimental Design
Technology
Replicates
Experimental Execution
Data Processing

Pilot Experiments

When to do them
Known dysregulated genes can be confirmed
Plan for how results will drive experimental design
Verification of sample preparation protocol
Pitfalls
No reliable statistics to do meaningful discovery
Often cannot be added to a dataset produced later

Controlling Variation

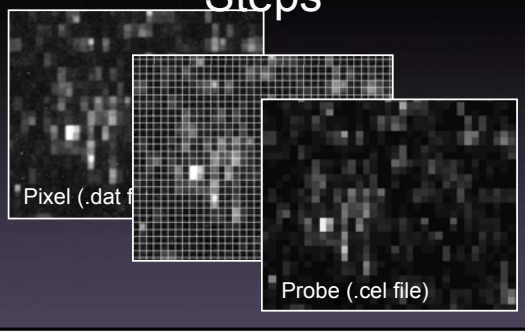
Avoiding Batch Effects
Reagent Lots
Incubators
Animal Handling
Time
Affects all of the above
Hard to "add replicates later if I need them"

Overview

Experimental Design
Technology
Replicates
Experimental Execution
Data Processing

Affymetrix Terminology

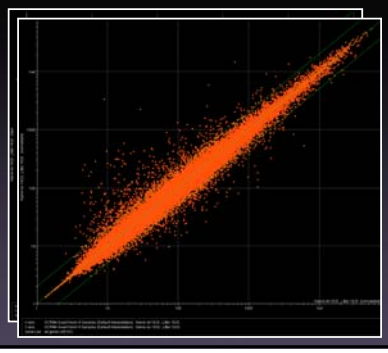
Affy Summarization Steps



From Probes to Genes

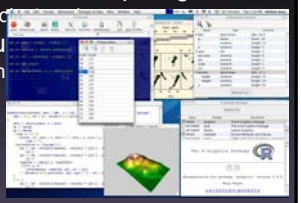
Affymetrix Microarray Suite v5 - "statistical"
Model Based Choices:
RMA -
http://www.stat.berkeley.edu/~bolstad/RMAExpress/RMAExpress.html
GC-RMA - http://www.bepress.com/jhubiostat/paper1/
dChip (MBEI) - http://www.dchip.org/
PLIER – Affymetrix (open source release)

MAS5 *vs.* GCRMA



R

The R Project for Statistical Computing
http://www.r-pro
Biocondu
http://www.biocon



Contact Information

Penn Bioinformatics Core - 13th Floor Blockley Hall
John Tobias - 1314 - jtobias@pcbi.upenn.edu
Reserve Computers
http://core.pcbi.upenn.edu/

PBiC
Penn Bioinformatics Core