

Statistical Bioinformatics: Tools for Microarray Analysis

Cooperative Design Meeting

August 19, 2003

Danh Nguyen, Ph.D.

Division of Biostatistics



Challenges in Analyzing Microarray Data

■ Challenges:

- Biomedical research: Massive array data
- Multiple platforms: gene expression
chemical/protein interaction arrays
- How extract useful info./analyze data

■ NCI caBIG response: caARRAY foundation

➤ Data
Management

Data
Analysis

Knowledge
Integration

- Data analysis: *Pre-processing, classification, clustering, visualization, etc.*

Responding to Challenges at UCD/UCDCC

- UCD Microarray Bioinformatics Resources:
Unique network addressing microarray analysis
- Example:
Product—Data **pre-processing** methods and tools

UCD Microarray Bioinformatics Resources

- Key people:
 - David Rocke—Director, CIPIC; Head, Statistical Bioinformatics; Leader pre-processing methods
 - Laurel Beckett—Head, Division of Biostatistics, Director, Biostatistics Shared Resource for CC
 - Division of Biostatistics: Nguyen, Tsodikov
 - Active collaboration: Rowe Genetics, Statistics,...
- Key Partnerships:
 - LLNL (David Nelson)
 - Existing & expanding partnership: Sun Microsystems (David Rocke)

Example: Pre-processing Methods

Pre-processing methods for:

- Measurement errors in gene expression arrays – tested on cDNA, Affy, and CodeLink platform
- Data transformation: variance-stabilizing
 - Across full range of expression
- Software: LMGene 1.0
 - R based —is open source, interoperable (Perl, Java); Bioconductor

Downstream Analyses Benefit from Pre-processing

- More precise comparisons of gene expression
 - Identifying down/up-regulated genes in treatment/control groups, cancer/normal cells
- Common univariate & multivariate analysis
 - Cancer classification/prediction
- More comparable & reproducible results across sites, studies

caBIG and UCD Cancer Center: Benefits and Contributions

- Further develop/extend our software & its implementation/distribution
- Improved local availability & access to analytical tools
- Integrating (e.g.) pre-processing tools on caARRAY
 - Wider access
 - Pilot & demonstrate use of our tools across sites & CCs
 - Encourage further contributions/sharing

Measuring Improvement and Success

- Product: Enhanced analytical tools/software
- Provide access to interoperable analytic tools
 - Support cancer researchers:
 - Accelerate methodological development & knowledge
- Pilot & demonstrate the use of our tools across sites → improvements: refinement of methods
- Promote wider use & contributions to caBIG