A Novel Approach for Analyzing and Visualizing High-Dimensional, Multivariate Microarray Data

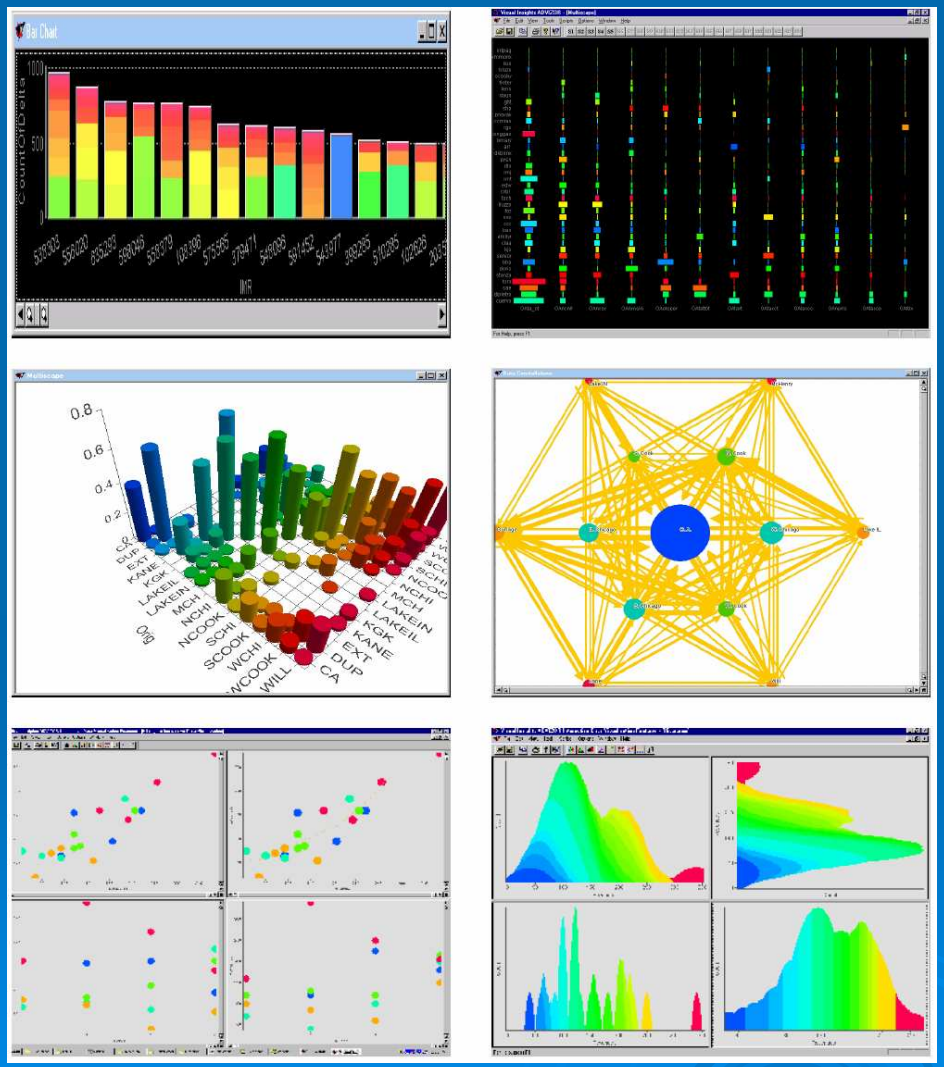
Vetria L. Byrd
University of Alabama At Birmingham

Tarynn M. Witten, MSW, PhD, FGSA, FCSB
Virginia Commonwealth University
BBSI Mentor

August 02, 2005

From the Desk of Vetria L. Byrd *

Graphical Representation of Data



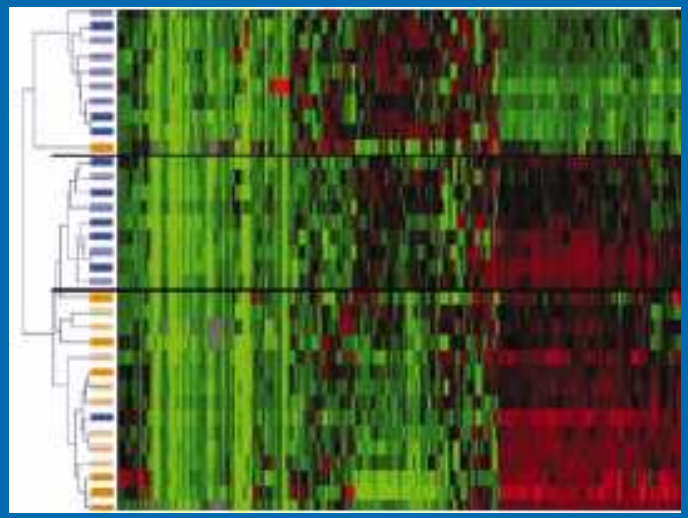
- Charts
  - Bar, Line, Pie
- Matrix View
- Multiscape
- Network view
- Scatter Plots
- Histograms

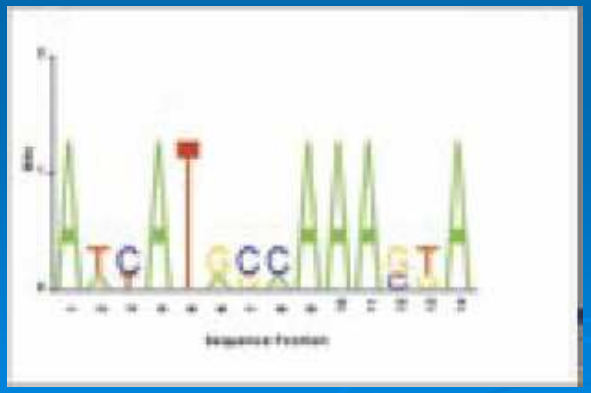
Eick, S. G. and Karr, A. F. Visual Scalability. Technical Report Number 106. National Institute of Statistical Sciences, 2000.

From the Desk of Velia L. Byrd

Common Visualization for Microarray Data

- Cluster grams
- Pairwise and multiple sequence alignments
- Parallel coordinates
- Heat maps





```
WYFGKITR---RESERLLLNPENPRGTFLVRESET-TKGAYCLSVSDFDNAK-----GLN
-YHGKISR---ETGEKLLLA-TGLDGSYLLRDSES-VPGVYCLCVLYHG----------Y
WNDGCIMGFISKERERALLK-DQQPGTFLLRFSESSREGAITFTWVERSQNGGE-PDFHA
YIMGFISK----ERERAILS-TKPPGTFLLRFSESSKEGGVTFTWVEKDISGK--TQIQS

VKHYKIRRLDSG-------GFYITSRTQ--FSSLQQLVAYY
IYTYRVSQTETG-----------------------------
VEPYTKKEL--------------------------------
VEPYTKQQLNNMSFAEIIMGYKIMD-AT--NILVSPLVYLY
```

From the Desk of Vetta L. Byrd

Problem Statement

- Microarray data contains answers to biological questions that haven't been asked or formulated yet
- No way to intuitively discern or extract biologically relevant data without visualization tools

Tilstone, C. (2003). Vital Statistics. Nature 424:610-612.

From the Desk of Velma L. Byrd

Research Goal

- To enable researchers to assess and visualize high dimensional, multivariate data in a way that is equivalent to (or improved upon) current state of the art techniques for visualizing such data

Presentation Overview

- Goals/Tasks
- Approach to achieving goals/tasks
- Questions
- Suggestions & Comments

From the Desk of Velma L. Byrd

Goals and Tasks

- Assess existing visualization tools
- Determine what, if any, unmet visualization needs exist
- Develop a tool/method to aid in the analysis and visualization of multivariate, high dimensional data

From the Desk of Velvia L. Byrd

Goal #1: Assess Existing Visualization Software

Commercial

GeneSpring
GeNet
Array-Pro Analyzer
CloneTracker
ImaGene
GeneSight
ArrayPack
GenePort
HDFR
Array-Pro Analyzer
arraySCOUT
arrayTAG
GenePix
Z3 2-D PAGE Image Analysis

Free

ANOVA Programs
CLUSFAVOR
SCAN-ALYZE
CLUSTER/TREEVIEW
D-CHIP
GENE-CLUSTER
PaGE
VERA and SAM
R-package in R and S-PLUS subroutines
SVDMAN
TREE-ARRANGE and TREEPS
BRB Array Tools
TIGR MultiExperiment Viewer
TIGR ArrayViewer
TIGR SpotFinder
F-SCAN
P-SCAN
GeneX
MAPS (Mircroarray Project System)

This is NOT an exhaustive list

From the Desk of Velvia L. Byrd

Goal #2: Determining Visualization Needs

- Survey Research Instrument
  - Survey Design & Implementation
  - VCU Institutional Review Board
  - Identify Survey Targets
- Survey Targets
  - Research Universities
  - National Laboratories
  - Research Institutions
  - Biotech Companies

From the Desk of Velma L. Byrd

Survey Goal

- To ascertain, from researchers who do microarray analysis, what features and functions the next visualization tool should offer to allow them more biological insight into the analysis of multivariate, high-dimensional microarray data.

From the Desk of Velvia L. Byrd

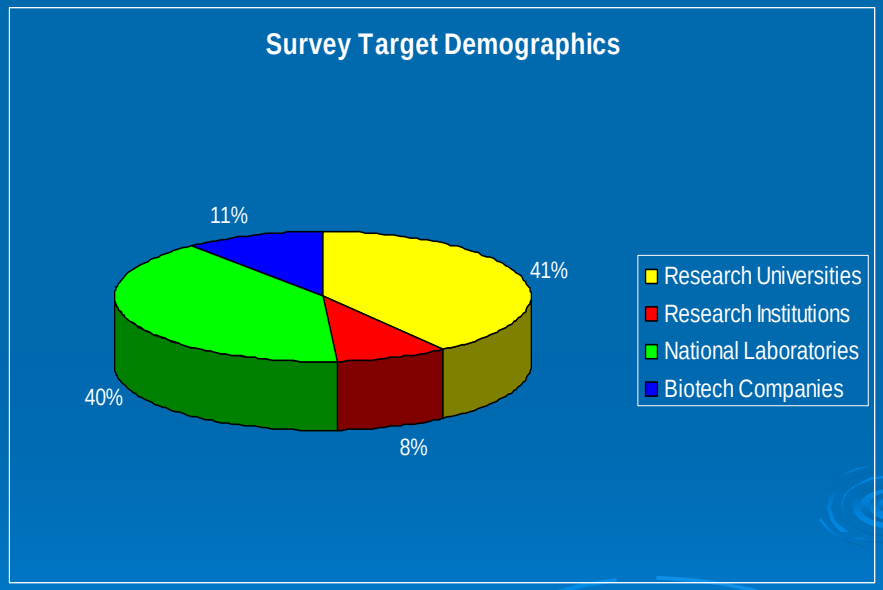
Participation in the Study

- Voluntary
- No Identifiable data collected
- Will provide researchers with insight
  - What visualization tools are currently used in the analysis of microarray data,
  - How the tools are used and
  - If the current tools meet the needs of the analyst using these tools.

From the Desk of Vetia L. Byrd

Survey Instrument

- Needs Assessment for Scientific Visualization of Microarray Data
- VCU IRB Approved: IRB# 5065
- 600+ Survey Targets
- 500 Survey Targets randomly selected
- Implemented and distributed this summer from VCU
- Results pending



Acknowledgements

Dr. Tarynn M. Witten

BCCL Colleagues

BBSI Colleagues

BBSI Faculty & Staff

VCU Postal Service

From the Desk of Vetia L. Byrd

Questions?

Comments?

Suggestions?

From the Desk of Velma L. Byrd