

# Proteomics Course

## LECTURE-32 Microarray work-flow: Data analysis



Dr. Sanjeeva Srivastava  
IIT Bombay

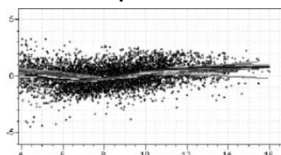


## Lecture outline

- Microarray data analysis
- Discussion on microarray data analysis
  - Normalization
  - Supervised or unsupervised
- Analytical methods
  - Hierarchical clustering
  - Self-organizing maps
  - Principal-components analysis

## Normalization

- Microarray experiments are performed on multiple chips
- To compare multiple microarray measurements, data need to be normalized
- Normalization is performed so that:
  - Data from a single experiment are as accurate as possible (correcting unbalanced PMTs)
  - Data from different experiments can be compared to each other



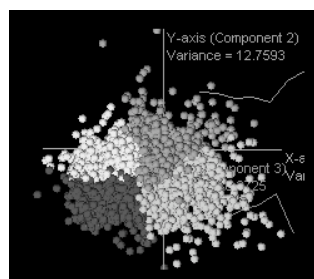
IIT Bombay 3

Proteomics Course

NPTEL

## Principle Component Analysis

- PCA works by finding “supergenes” that: explain the most variance in the sample are orthogonal to each other



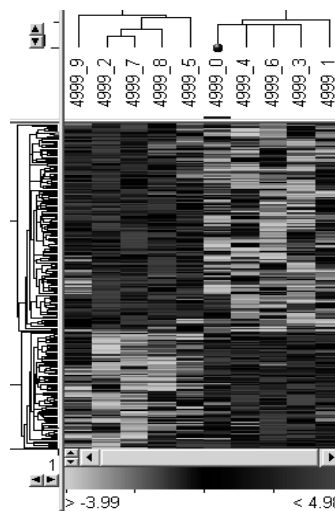
IIT Bombay 4

Proteomics Course

NPTEL

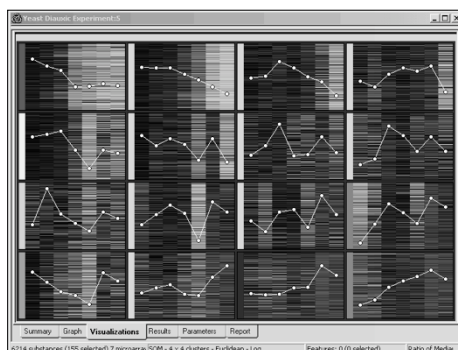
## Clustering

- **Hierarchical:** genes are placed in a hierarchical relationship to each other, as in taxonomy
- **Non-Hierarchical:** genes are placed in clusters that do not necessarily have any relationship to each other



## Self-Organizing Map

- Replicate Dye-Swap microarrays can be quickly inspected for quality using a self-organizing map.



## Summary

- Microarray data analysis
- Discussion and demonstration
- Data analysis requires good software, programming and statistical analysis
- Experimental design is important

## References

- Anders Bengtsson\* and Henrik Bengtsson. Microarray image analysis: background estimation using quantile and morphological filters. BMC Bioinformatics 2006, 7:96 doi:10.1186/1471-2105-7-96
- Latha Ramdas, Wei Zhang. Microarray Image Scanning. Cell Imaging Techniques. Methods in Molecular Biology. Volume 319, 2006, pp 261-273.

## Acknowledgements

- Mr. Pankaj Khanna, Spinco Biotech Private Limited for discussion on microarray scanning and data analysis.