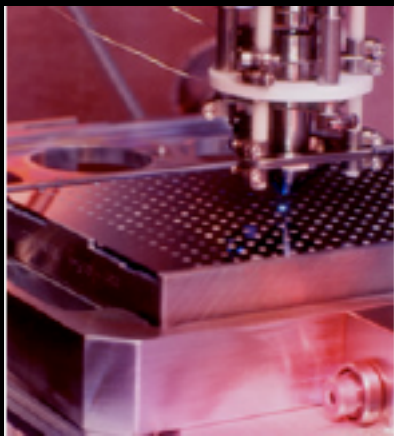


# Integrative Analysis of multiple large-scale molecular biological data



SRI PRIYA PONNAPALLI

GENOMIC SIGNAL PROCESSING LABORATORY

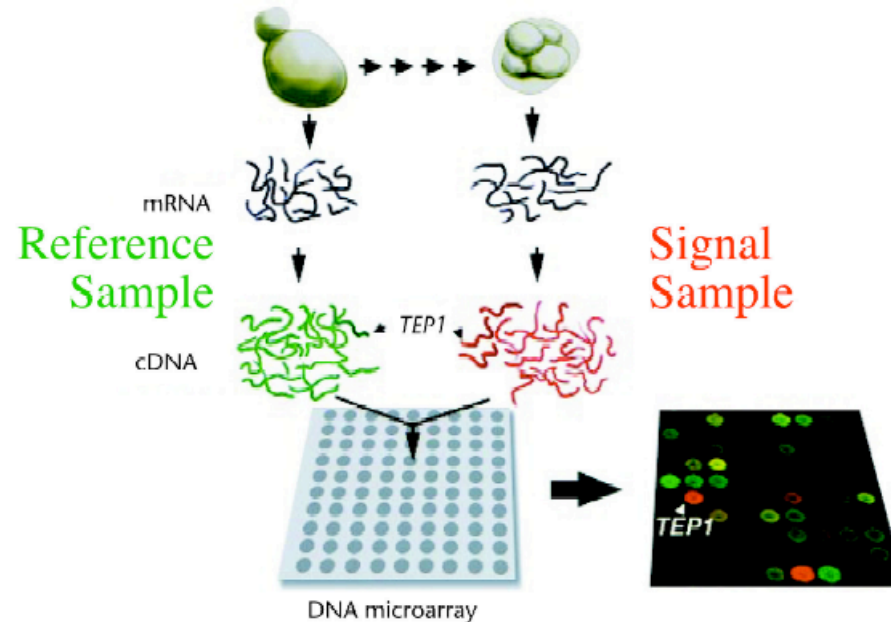
THE UNIVERSITY OF TEXAS AT AUSTIN

# New High Throughput Technologies

Sequencing of the human genome, and the genomes of other model organisms, such as yeast, is completed or well underway.

DNA microarray hybridization technology allows probing fluorescently tagged cDNA copies of mRNA from a single sample with thousands of DNA targets or synthetic oligonucleotides simultaneously.

Gene expression levels can now be monitored on a genomic scale.

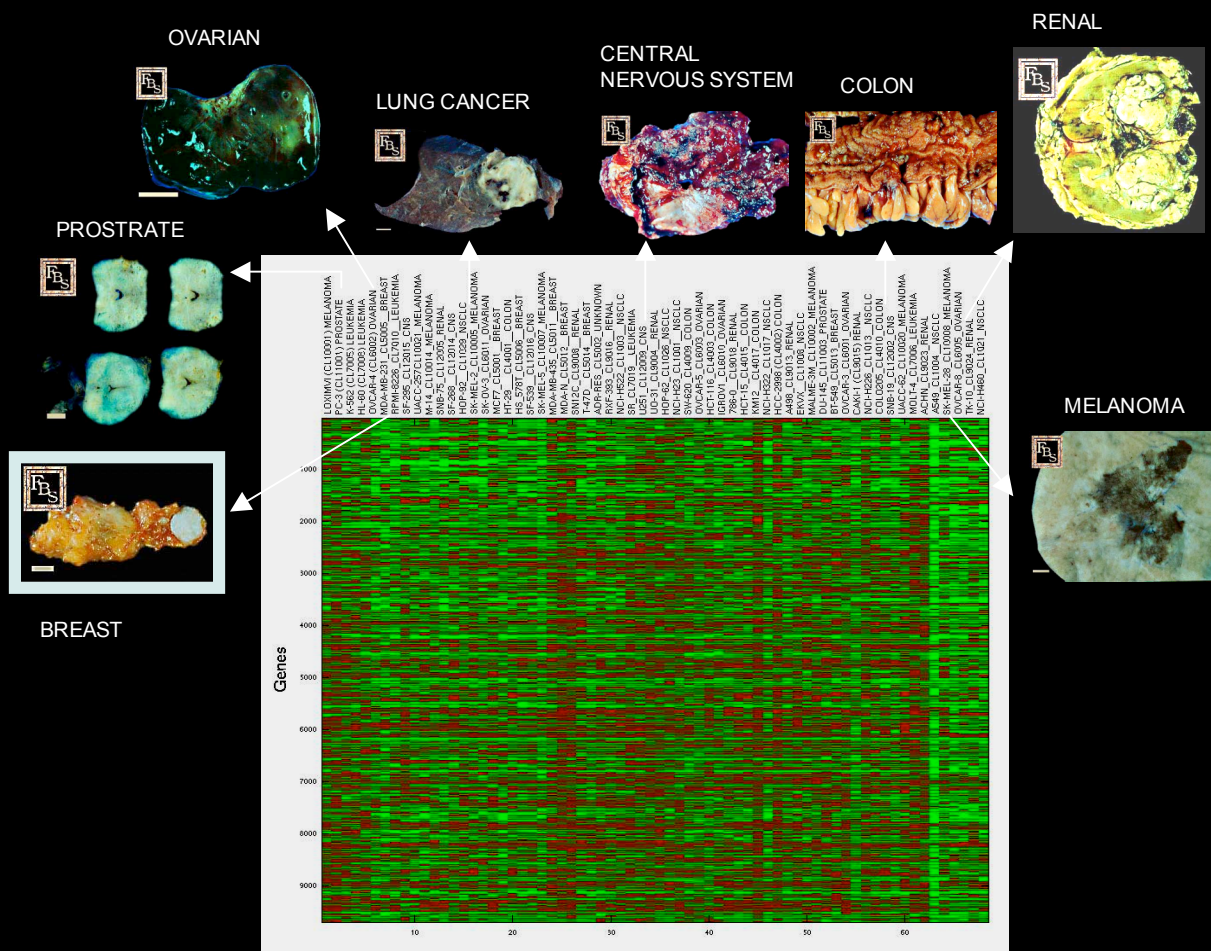


Brown & Botstein, *Nature Genetics* 21, 33 (1999);

Lipshutz, Fodor, Gingeras & Lockhart, *Nature Genetics* 21, 20 (1999).

**Motivation: The NCI60 Human cancer cell lines are extensively used as experimental models.**

**Goal: Perceive relationships between three datasets containing three different types of attributes of the NCI60 – genome-scale expression, sensitivities to more than 70,000 chemical compounds and chemotherapeutics, and proteomic profiles.**

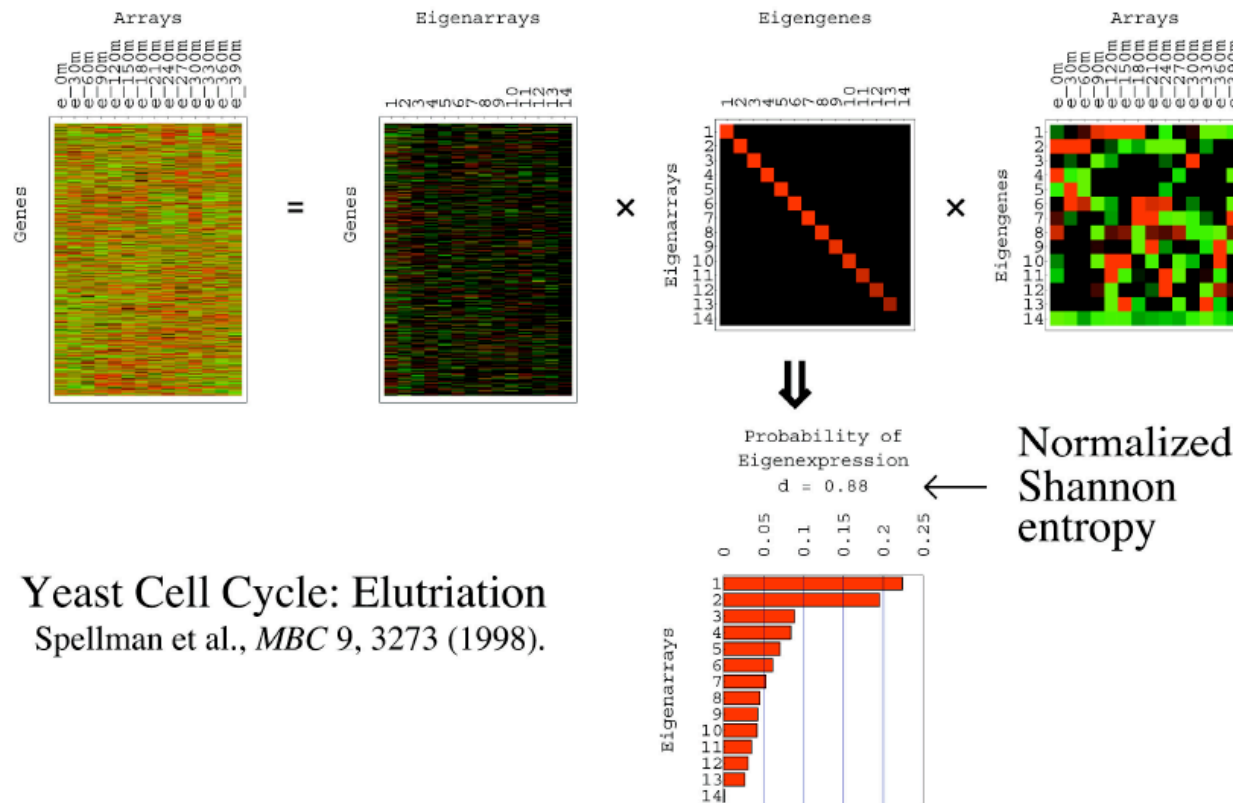


Proposed data set for analysis :  
NCI60  
HUMAN  
CANCER  
CELL  
LINES

# KEY PAPER 1 : [ Alter et al, 2000]

## Singular Value Decomposition (I)

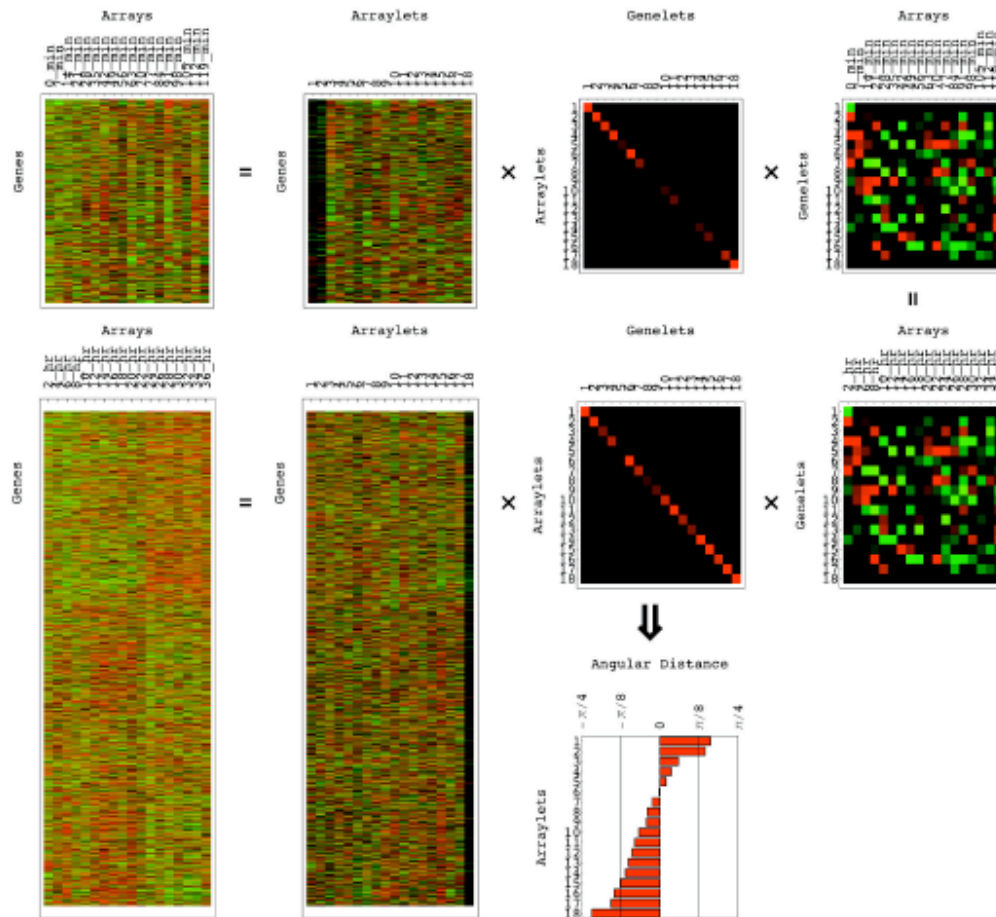
Linear transformation of gene expression data from **genes**  $\times$  **arrays** space to **reduced diagonalized** “**eigengenes**”  $\times$  “**eigenarrays**” space.



# Generalized SVD (I)

Linear transformation of two datasets from **two genes**  
**× arrays spaces** to **two reduced diagonalized**  
**“genelets” × “arraylets” spaces.**

Yeast Cell Cycle: Alpha Factor  
 Spellman et al., *MBC* 9, 3273 (1998).



KEY PAPER 2 :  
 [ Alter et al, 2003 ]

## KEY PAPERS 3,4 & 5: **Data Sets**

- Proteomic Profiling of the NCI-60 Cancer Cell Lines Using New High-density Reverse-phase Lysate Microarrays

[Nishizuka et al ,2003]

- A Gene Expression Database for the Molecular Pharmacology of Cancer

[Scherf et al, 2000]

- Systematic variation in gene expression patterns in human cancer cell lines

[Ross et al, 2000]