

PowerArray

Jack Liu, CIIT => GSK
Stan Young, NISS

1

PowerArray Features

1. Data inport and cleaning
2. rSVD-based signal extraction
3. rSVD for data block
4. Powerful visualization functions (7 ways)
5. Volcano plots and error-bar plots
6. K-means clustering
7. ANOVA based testing and many multiple testing adjustments
8. Least angle regression
9. Gene ontology classification
10. Automatic annotations
11. Pathway analysis!!!

Essentially "full featured" micro array analysis package.

2

Software Things

1. Over 30k lines of interface/subject specific code.
2. Over 300k lines of algorithms.
3. Uses Windows ".NET", 5Mb of code.
4. Part of the "Power" family,
PowerMarker, PowerArray, PowerMV.

3

Key Statistical Things

1. $N \ll p$, $N=10-50$, p 1k to 20k.
2. Micro array has missing data and noise.
3. rSVD is used to advantage.
4. Multiple visualization methods.

4

Robust Singular Value Decomposition

Data = Fit + Error

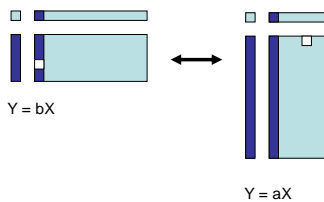


$$X_{ij} = a_i b_j + e_{ij}$$

5

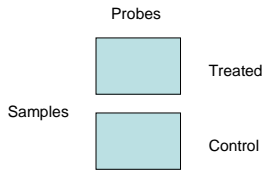
rSVD Algorithm

Alternating Robust Regression



6

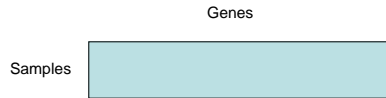
Data Cleaning



Use rSVD to sort the rows and columns.
Remove probe and sample outliers.
Do this for each of the 20k genes.

7

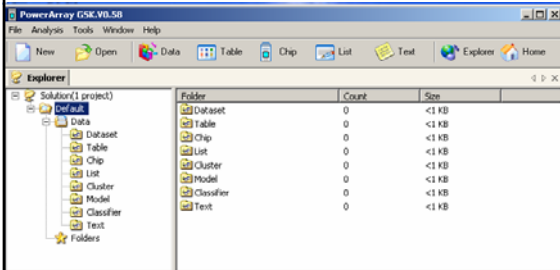
Statistical Analysis of Genes



Analysis : 20k t-test, ANOVAs, linear contrasts.
Bonferroni adjusted p-values.
rSVD, etc.

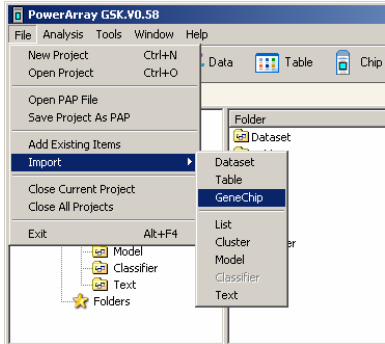
8

Screen Shot 1



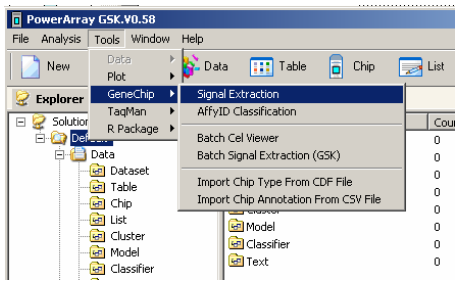
9

Screen Shot 2



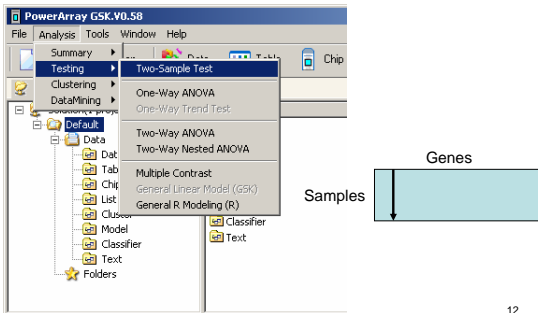
10

Screen Shot 3



11

Screen Shot 4



12

Download Software – Wait !

Google(PowerArray)

http://www.powermarker.net/PowerArray/PowerArray_Installer.msi

Windows .NET

User manual planned for this summer, GSK and SanofiAventis.
