Analysis of time-course microarray data using the ANCOVA framework

Tzu L. Phang and Katherina Kechris
University of Colorado Health Sciences Center
Microarray Time Course Experimental Design and Analysis

- ANCOVA
- Significant Genes
- Stepwise Regression
Downsides of the previous approach

1. ANCOVA violates the homogeneity of regression slope assumption
2. Contrast comparisons compared to a baseline factor, making non-baseline comparisons non-trivial
3. Only one categorical factor is allowed
4. No convenient tool for biologists to select comparison of interest
Proposed Analysis Flow

\[ y_{i,j,k,l} = \alpha + \beta_j^T + \beta_k^S + \beta_{jk}^{TS} + \gamma t_i + \delta_j^T t_i + \delta_k^S t_i + \delta_{jk}^{TS} t_i + \theta t_i^2 + \phi_j^T t_i^2 + \phi_k^S t_i^2 + \phi_{jk}^{TS} t_i^2 + \epsilon_{ijkl} \]

F-stat + FDR = Significant Genes

Categorical Effects
(Strain, Trmt, Strain x Trmt)

Continuous Effects
(Time)

Interaction Effects
(Strain x Time, etc)

Johnson-Neyman Procedure
Results

Example Plot 1
✗ Categorical Effect
✓ Time Effect
✗ Interaction Effect

Example Plot 2
✗ Categorical Effect
Time Effect
Interaction Effect

Example Plot 3
✗ Categorical Effect
✗ Time Effect
✓ Interaction Effect
Conclusion

• We have established a comprehensive statistical framework to study the effect of different variables for multiple-factors microarray time course data.
• Each of these effects; categorical, continuous, and interaction, should be handled with appropriate methods (Tukey HSD, Johnson-Neyman, etc)
• Future challenges:
  – Build a user friendly tool to select genes of interest amid the sea of interacting effects.
  – Establish the relationship between these effects across genes and experiments using network graphing technology.