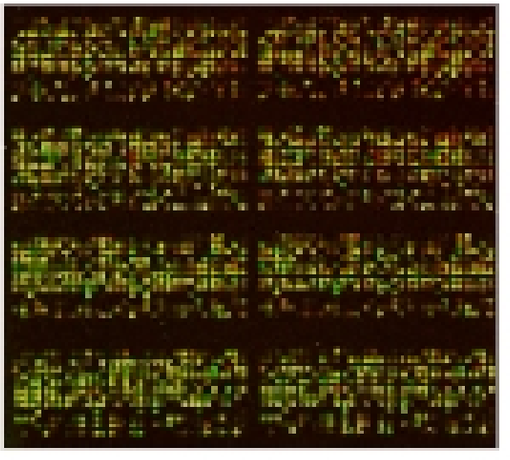
Margin Trees for High-dimensional Classification

Tibshirani and Hastie

Errata (confirmed by Tibshirani)

- Section 2 (a) about the property of 'single link age'. M should be M_0
- Section 2.1 close to the last line of second pa ragraph. "at least" should be "at most"
- The statements about complete/single linkage are misleading. In fact, they use standard defi nition of complete/single linkage except the di stance metric is replaced with margin betwee n pairwise classes.
  (I traced their code to confirm this).

Targeted Problem



- Multi-class
  - #class >> 2
- High-dimensional, few samples
  - #features >> #data →linear separable
  - already good accuracy, need interpretable model
- Ex. micro-array data
  - feature : gene expression measurement
  - class: type of cancer
  - Instances: patients