

# GenePattern

## **Lu.Getz.Miska.Nature.June.2005.clustering.ALL**

**Module name:** Lu.Getz.Miska.Nature.June.2005.clustering.ALL  
**Description:** Hierarchical clustering of 73 ALL samples with various types of genetic alterations  
**Author:** Gad Getz (Broad Institute), gp-help@broad.mit.edu

### **Summary**

Hierarchical clustering of 73 ALL samples with various types of genetic alterations can identify sub-populations which are associated with the genetic alterations.

### **Hierarchical clustering of 73 ALL samples with various types of genetic alterations**

We analyzed 73 bone marrow samples obtained from patients with acute lymphoblastic leukaemia (ALL). Hierarchical clustering revealed non-random partitioning of the samples into three major branches: one containing all five t(9;22) BCR/ABL-positive samples and 10 out of 11 t(12;21) TEL/AML1 samples; a second branch containing 15 out of 19 T-cell ALL samples; and a third branch containing all but one of the samples with an MLL gene rearrangement.

Hierarchical clustering was performed with average linkage, unweighted average distance (UPGMA), using (1-Pearson correlation) as the distance measure. Before clustering, data were filtered to eliminate genes with expression lower than 7.25 (on a log<sub>2</sub> scale) in all samples. Next, all features (miRNAs) were centered and normalized to a mean of 0 and a standard deviation of 1. Clustering was performed on both the samples and genes axes.

### **References:**

- Lu, Getz, Miska, et al. "MicroRNA Expression Profiles Classify Human Cancers," Nature 435, 834-838 (9 June 2005)
- Jain, A. K. & Dubes, R. C. Algorithms for clustering data. Prentice-Hall Inc., Upper Saddle River (1988).