

Unsupervised
Analysis

Training Set - 54,675 probesets
5 Normal Fetal Cartilage and
41 Normal Non-cartilage Arrays

Supervised
Analysis

Data Transformation and Filtering

1. RMA normalize and $\log_2$ transform
2. Adjust median of each cartilage probe signal to 0
3. Adjust all other probes relative to cartilage median
4. Filter for genes with signals 2 SD away from the mean in at least 2 arrays

Unsupervised Results

9483 probesets representing the most differentially expressed genes among all tissues.

2 way Hierarchical Clustering

Pearson Correlation
9483 probesets and 46 Arrays

Unsupervised hierarchical clustering results

1. Fetal cartilage expression is consistent among samples
2. >64 genes consisting of many known cartilage-specific genes drive cartilage array clustering.

Data Transformation and Filtering:

1. RMA normalize and $\log_2$ transformation
2. SAM (100 permutations)
3. Probes 3 fold greater expression in cartilage
4. Probes with FDR=0

Supervised Results

2446 probesets representing
1713 genes with higher expression in cartilage

Validation Set (2446 Probesets)

2-5 normal fetal cartilage samples
34 normal non-cartilage tissues
224 arrays

124 U133A Arrays

1. RMA normalize
2. Median normalize 56 GNF and 68 Celsius Arrays
3. 1363/2446 probesets

74 U133B Arrays

1. RMA normalize
2. 882/2446 probesets

26 U133A Plus 2.0 Arrays

1. RMA normalize
2. 201/2446 probesets

Cartilage-selective algorithm

1. 5 fold higher expression than the median expression level of all non-cartilage tissues
2. Analog of CV <50% of non-cartilage samples

Cartilage Selective Results

161 (185 probesets) cartilage-selective genes