

Data Preprocessing

Data normalization, Imputation

Gene Filtering

PCA 85% variation



***k*-Level Ensemble Consensus Clustering**

Partitioning methods

Partitioning around medoids
k-means
Graph partitioning

Agglomerative methods

Average linkage
Complete linkage
Ward metric
Hybrid Bagglo (Karypis)

Probabilistic methods

Expectation maximization
Entropy based clustering
Clustering on subsets of attributes



Multigene-marker (Pattern) Identification

Pathway identification

VALIDATION ON EXTERNAL DATASET