Network Clustering

MCL clustering based on -log of CONNECTIVITY

Ranking Differential Expression

• Collapsing probes to genes
• SAM based ranking of the chronologically closest time-points

GS Enrichment Analysis

• 7871 GS retained (>10 & <1000 genes); 1000 GS permutations

Normalization for Age-related SG Development

• Exclusion of GS showing the same trend in both strains during the same time period (FDRq < 0.05; nom. p-value < 0.005; TAGS ≥ 50%)

Requirements for Nodes
For EM-related GS (ancestor either GO_0071944, GO_0031012, GO_0005911):
• (FDRq < 0.05)
For EM-associated GS:
• 1st degree neighbor of a EM-related GS (≥ 8% overlap in the LEs of the EM-related GS and the EM-associated GS)
• FDRq < 0.05; nom. p-value < 0.005; TAGS ≥ 50%

Requirements for Edges
• CONNECTIVITY score ≥ 0.08 (corresponding to ≥ 8% overlap in the LEs the two GSs)

Network Clustering
• MCL clustering based on -log of CONNECTIVITY

Data acquisition & resources

GS Resources

n = 21855
• Gene Ontology (GO) 12467
• NCI (NC) 219
• PFAM (PF) 4147
• KEGG pathways (KE) 225
• Biocarta pathways (BI) 249
• Reactome pathways (RE) 943
• Transcription factor targets (TF) 615
• miRNA targets (MI) 793
• Genebands (GB) 393

Data analysis

Leading Edge (LE) Analysis

• Identification of genes contributing to each GS’s significance and computation of LE metrics (TAGS, LIST, SIGNAL)
• Flagging LE genes located in Aec1 and Aec2
• Computation of a CONNECTIVITY coefficient matrix representing the overlap between the LE genes of all GSs paired one on one

Network building

Leading Edge Analysis

VVLJQLÀFDQFHDQG computation of LE metrics (TAGS, LIST, SIGNAL)