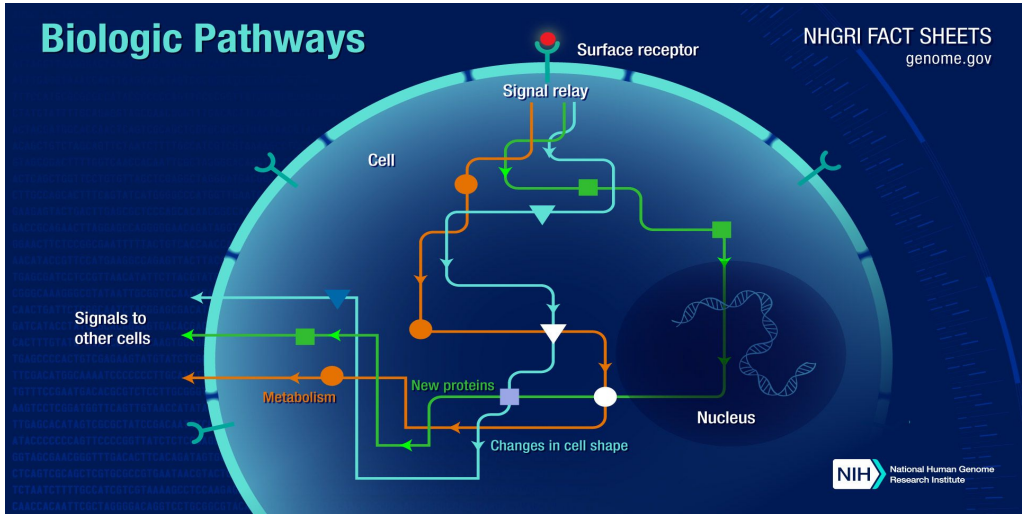
Gene Co-Expression Networks (GCEs): From Clustering to Insights Part 2

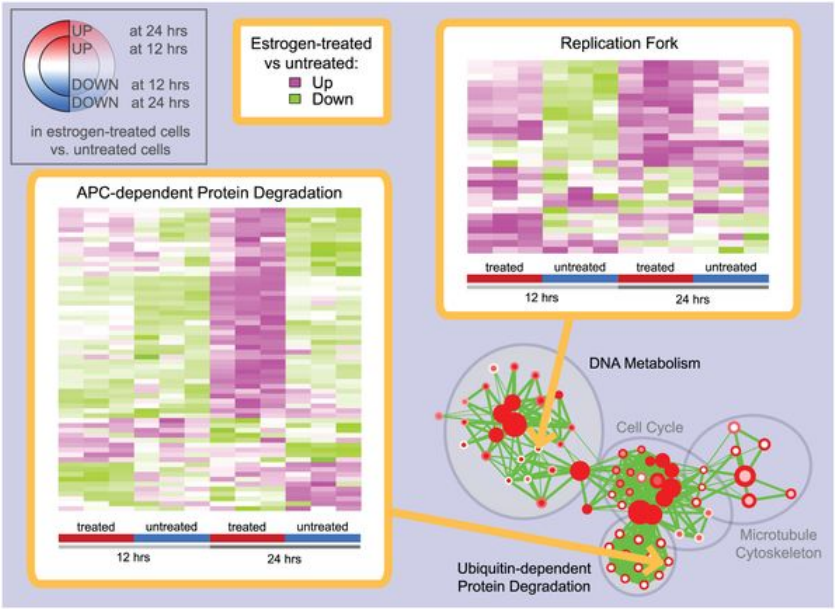
Visualization, Validation & the Diseasome

As. univ. Bozdog Alexandru



Validation & Biological Relevance of Modules

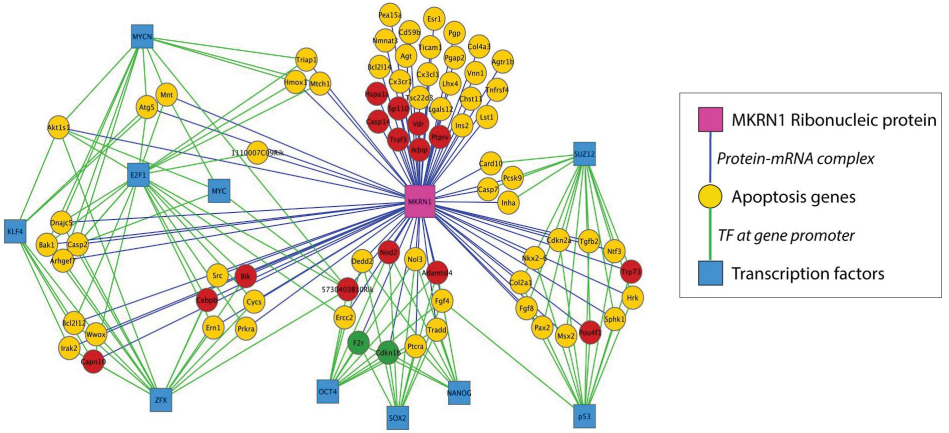
- **Biological validation**: GO/KEGG enrichment analysis.
- **Stability**: bootstrap resampling → do modules persist?
- **Replication**: compare across datasets (e.g. breast vs glioblastoma).
- **Reporting**: correlation metric, threshold/density, community detection algorithm, random seed.



Enrichment map for early-onset colon cancer and overlap with known disease genes.
From:Merico, Daniele, et al. "Enrichment map: a network-based method for gene-set enrichment visualization and interpretation." *PloS one* 5.11 (2010): e13984.

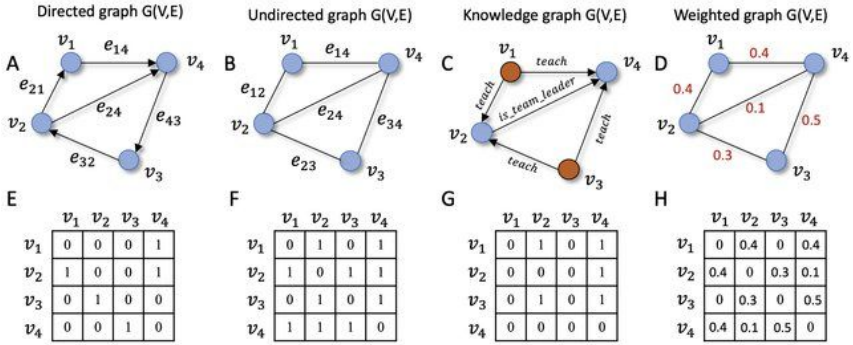
Visualization and Biological Context

- Visualization Tools: Cytoscape, Gephi, NetworkX.
- Identify clusters, hub genes, and patterns.
- Example: Liver tissue—modules enriched fo linked to detoxification.



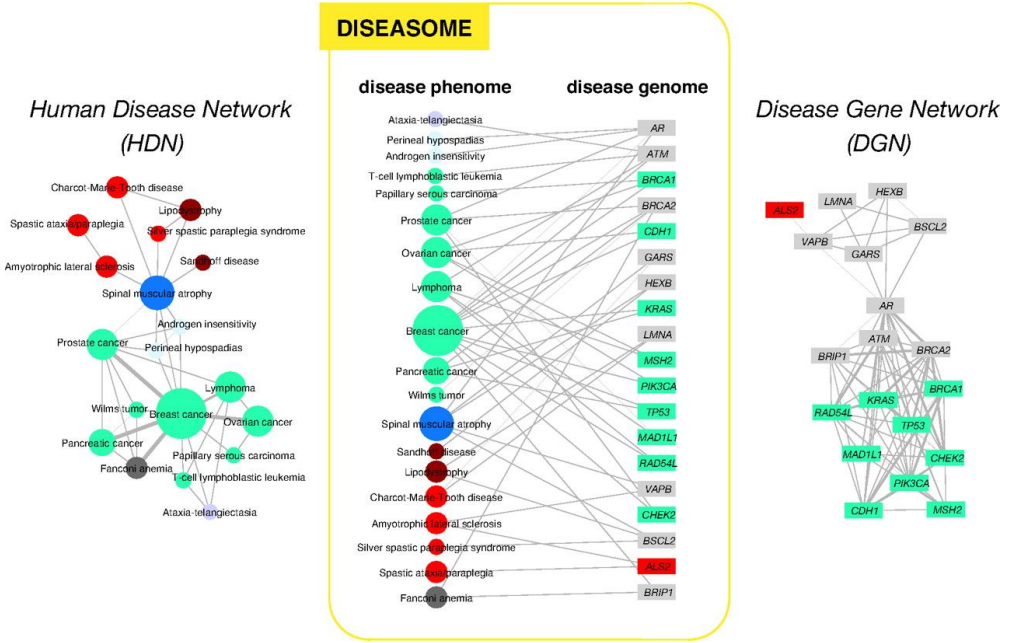
Challenges in GCE Analysis

- Noisy data: Requires careful preprocessing and filtering.
- Threshold choice: Balancing noise reduction and meaningful connections.
- Scaling methods:
  - Memory efficiency: Adjacency list vs. matrix.
  - Parallelizing calculations for large datasets.

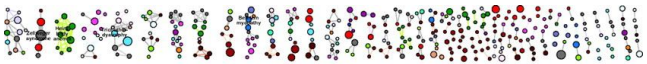


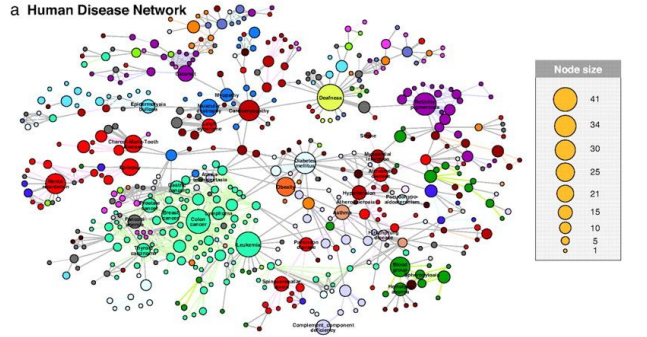
The DISEASOME: Linking Genes, Modules, and Diseases

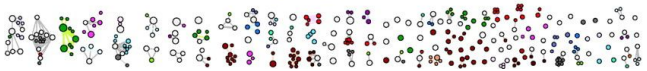
- The **DISEASOME** = network of human diseases connected by shared genes, pathways, or molecular mechanisms (Barabási et al., 2007).
- **Co-expression modules** map to **functional gene groups** implicated in specific diseases.
- **Linking modules** → disease phenotypes highlights **comorbidity patterns** and **shared risk factors**.
- Integrates **genomic, transcriptomic, and clinical data** -> systems-level disease map.

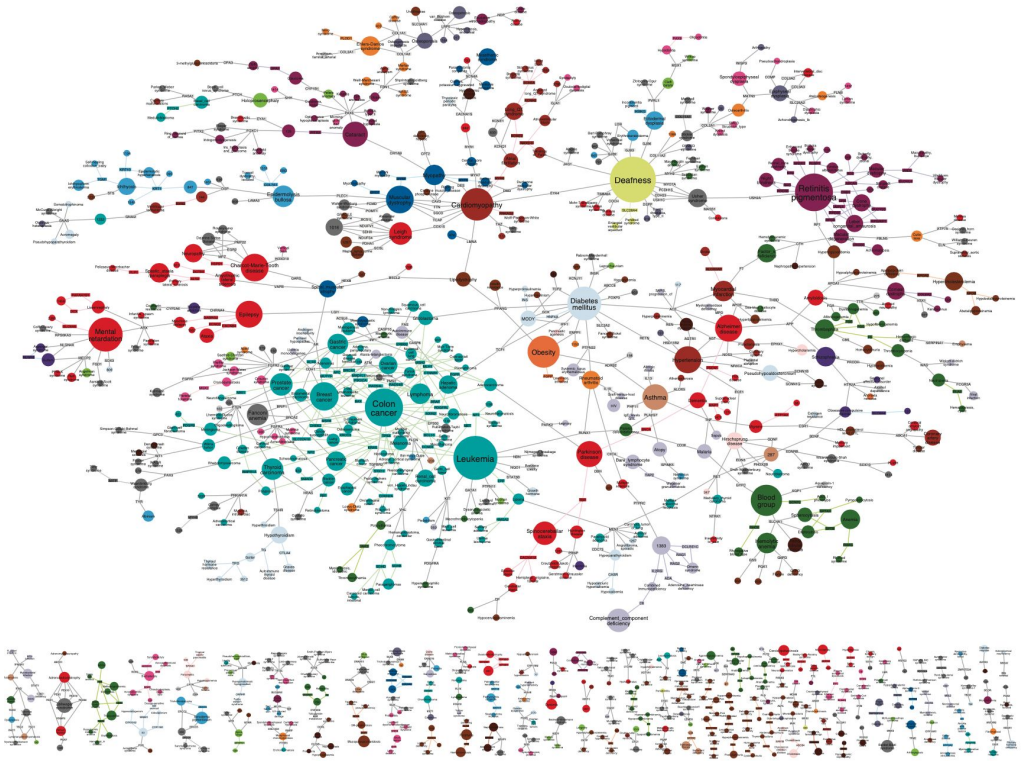


The human diseasome









From: https://barabasi.com/art/work/diseasome

Key Takeaways and What's Next?

- Gene co-expression networks can be visualized and interpreted using tools like Cytoscape, Gephi, or NetworkX.
- Modules represent groups of genes with correlated expression → often linked to biological functions or pathways.
- Validation is essential: enrichment analysis, stability checks, and replication across datasets.
- The diseasome integrates these modules into a global disease network, connecting genes to phenotypes and comorbidities.
- Applications: cancer research, drug repurposing, systems medicine.
- Next Lab: Applying machine learning for disease classification.