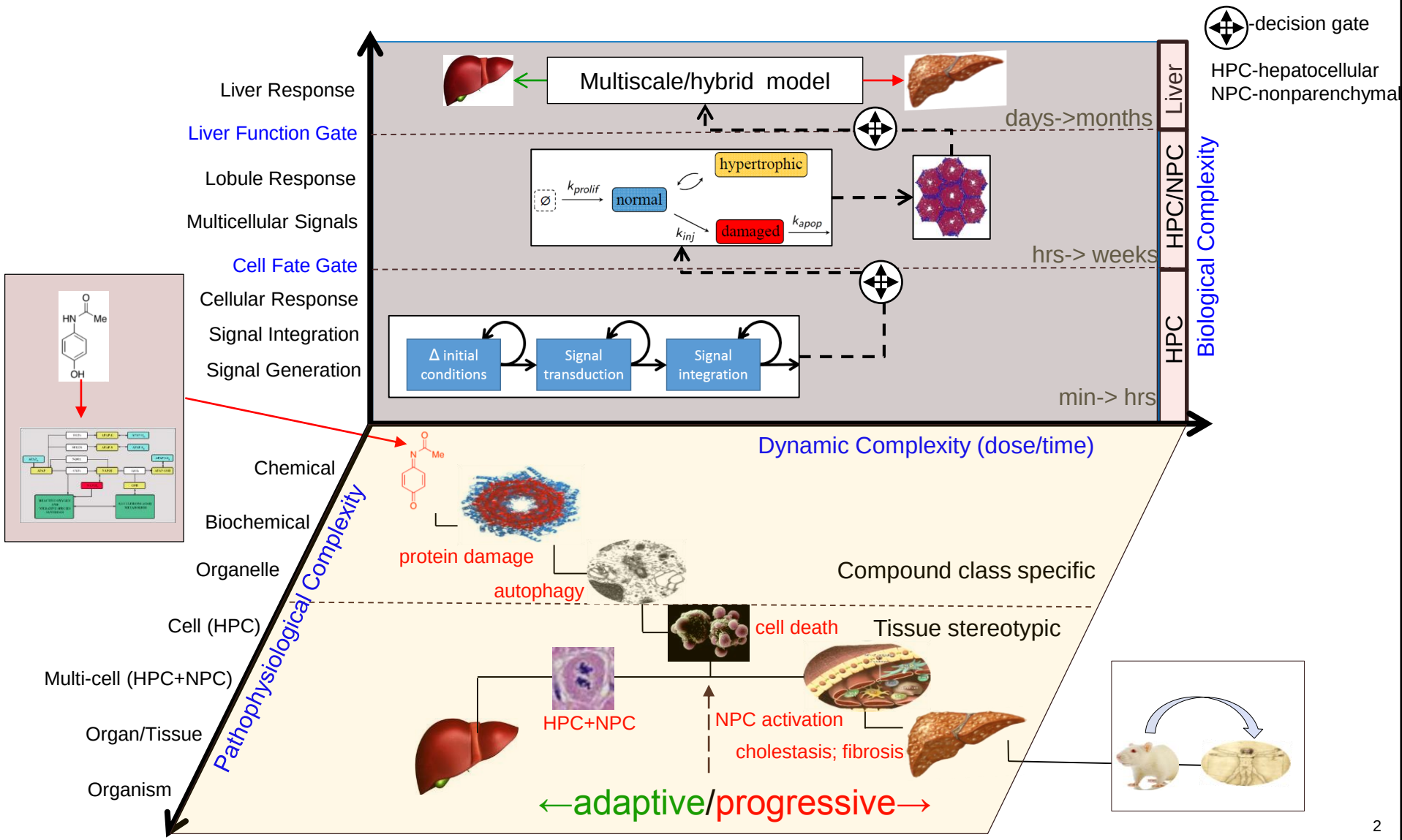
Application of Weighted Gene Co-Expression Network Analysis (WGCNA) to Dose Response Analysis.

Improving interpretation of nonclinical results using modularity to reduce complexity without loss of biological information.

Jim Stevens
Distinguished Research Fellow
Lilly Research Laboratory
stevens_james_l@lilly.com

Lilly

Multi-Scale Modeling of Pathophysiology



Overview

Modeling biological complexity

- The modular nature of complex systems.
- Leveraging modular systems models using gene expression data.
- Translating gene expression data into biological understanding.
- Reducing redundancy in MSigDB.
- Knowing what we don’t know.

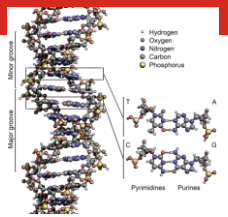
Understanding Molecular Pathogenesis

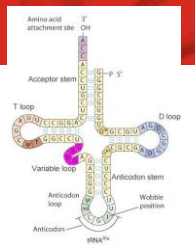
- Correlating expression modules with pathology.
- Closing the loop from transcription factor to pathogenesis.
- Predicting adaptive vs progressive responses.
- Closing the loop on transcriptional control (addendum slides).

Applications of WGCNA to Dose Response Analysis

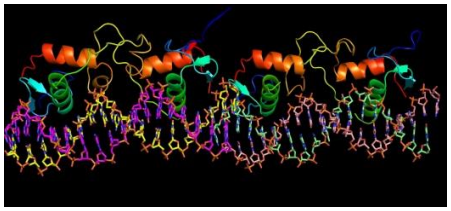
- Separating injury signals from tissue stereotypic response.
- Perturbing network in culture.
- Translation to human.

Biological Systems are Modular Across Scales of Complexity

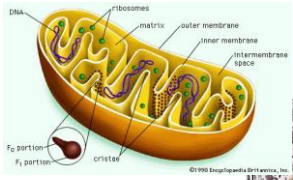


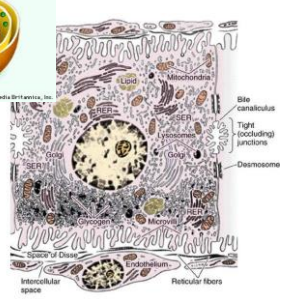


- Modularity refers to *"…pattern[s] of connectedness in which elements of a system (e.g. mRNAs) are grouped into highly connected subsets*." (modified from Wagner et al.[1])
  - Modules can be arranged in hierarchies using looser connections between modules.



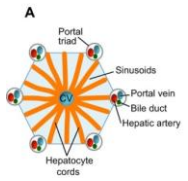
- Modular behavior can be captured in unsupervised network models using coalescent properties of the system.
  - Physical interactions – protein interaction networks
  - Dynamic interactions – gene regulatory networks
  - Statistical interactions – individual elements connected to phenotype

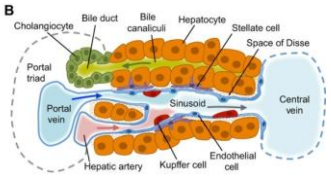




- Co-regulation in transcriptional networks is a coalescent property of biological systems – networks self-assemble.
  - Connected at level of transcriptional control, e.g. Hox gene networks, Nrf2, etc.
  - Defined/modeled statistically to yield co-expression modules.

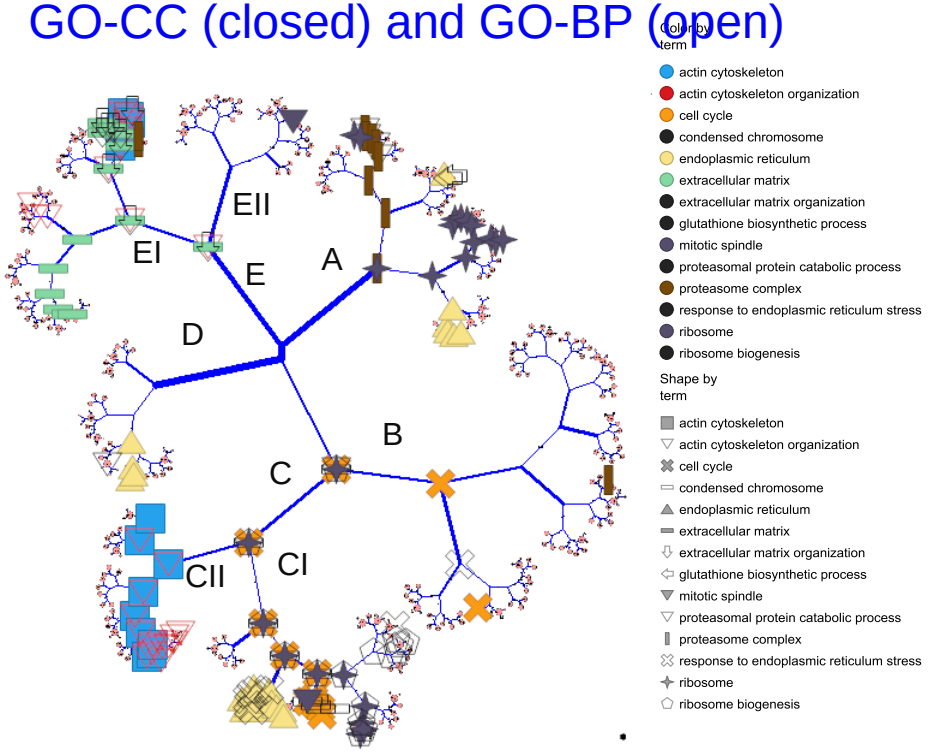
- Modeling complex systems as networks/modules has advantages:
  - Avoids the 'curse of dimensionality.'
  - If $2X10^4$ genes form $2X10^2$ modules complexity is reduced by 99%.
  - Biological content is retained.
  - Network visualization applied to modular systems improves data interpretation.

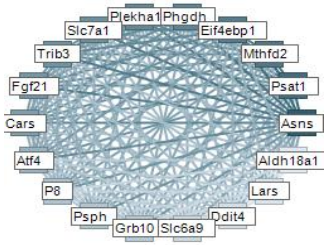




WGCNA^2 - Form Follows Function

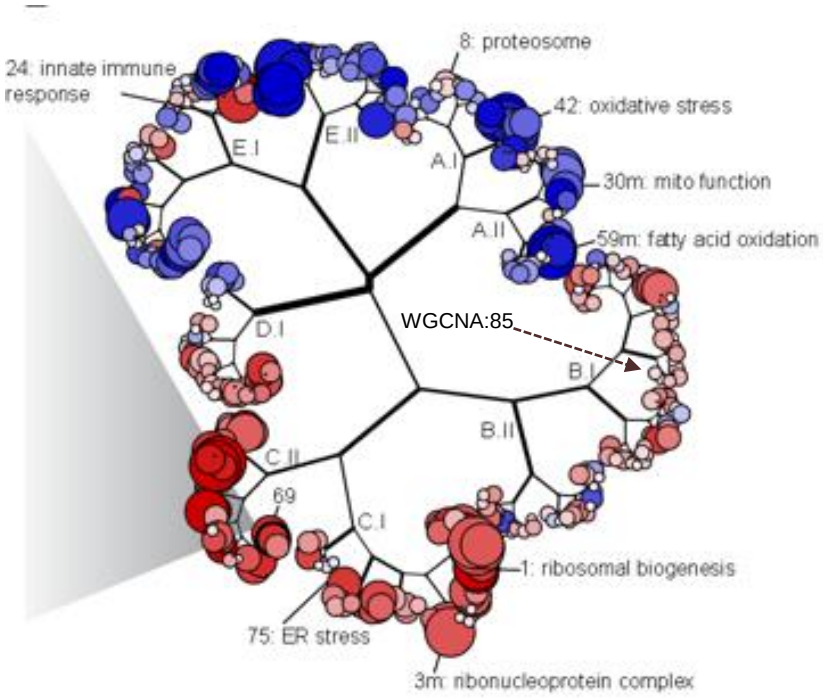
GO-CC (closed) and GO-BP (open)

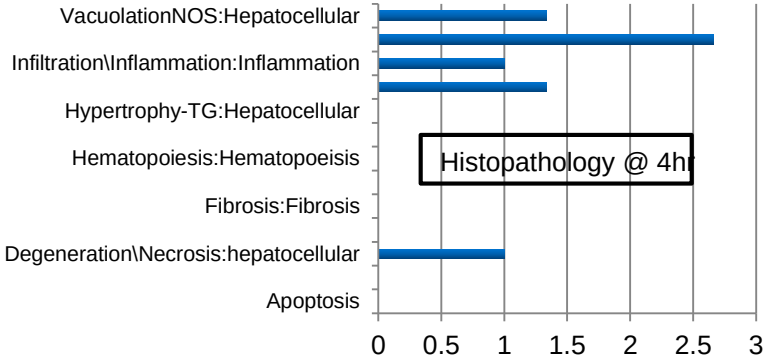




WGCNA:85 - Atf4 nutrient depletion

Co-expression modules (genes that respond similarly to drugs): 1 readout per module; the Eigengene (EiG).



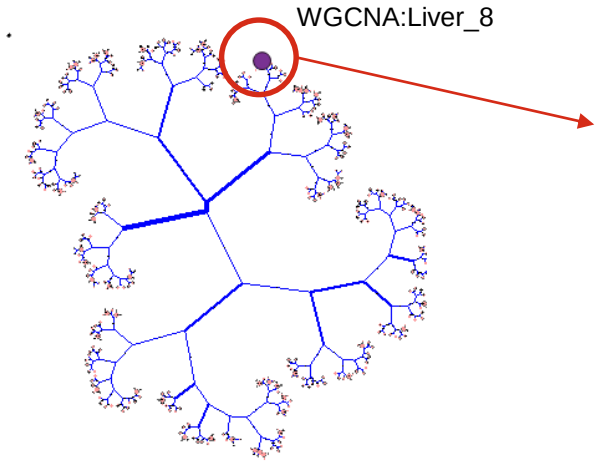


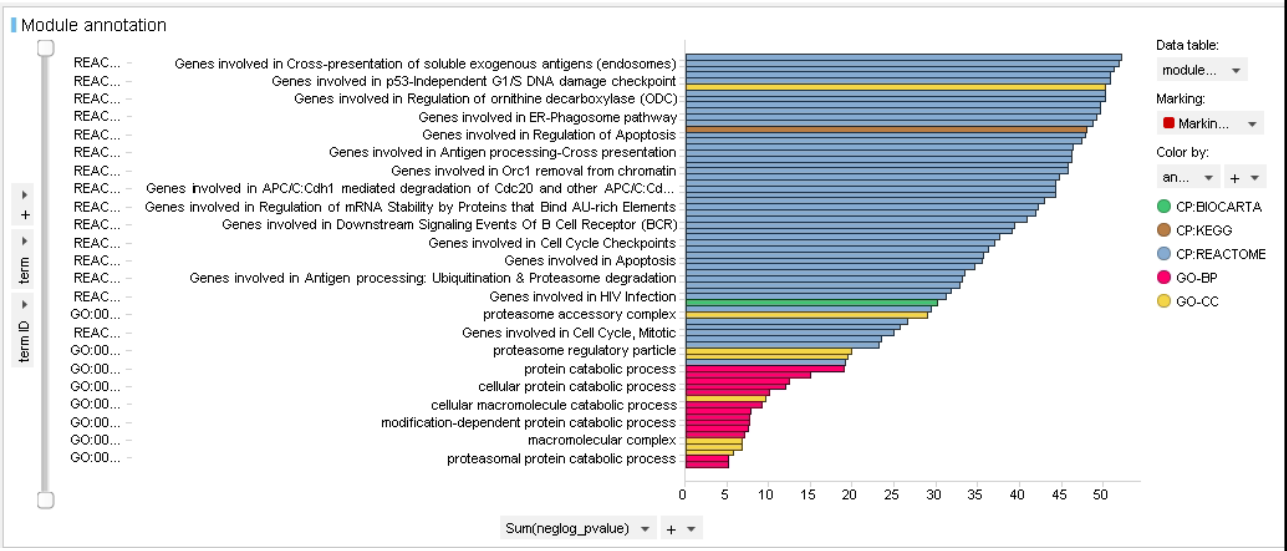
Reducing Redundancy of MSigDB Information.

GO and TF terms

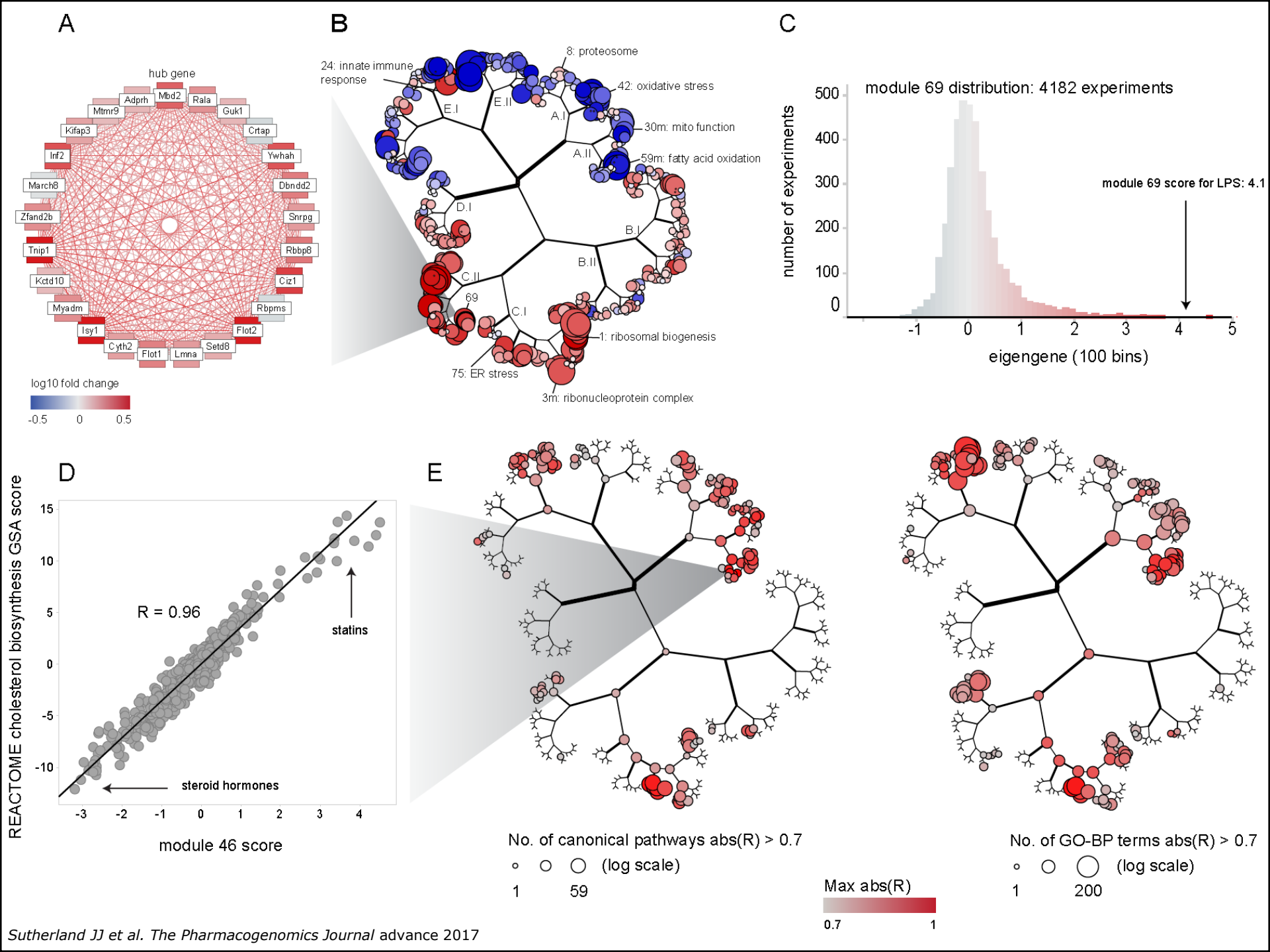
module	annotation ty...	term ID	term	neglog_pv...	number of g...	URL	Filtered to at...	genes in mo...	Fraction of g...
WGCNA|Liver:8	CP:REACTOME	REACTOME_VI...	Genes involved in Vif-mediated degradation of APOBEC3G	52.21	34	http://www.re...	Untagged	147	0.23
WGCNA|Liver:8	CP:REACTOME	REACTOME_C...	Genes involved in Cross-presentation of soluble exogenous anti...	51.99	33	http://www.re...	Untagged	147	0.22
WGCNA|Liver:8	CP:REACTOME	REACTOME_C...	Genes involved in CDK-mediated phosphorylation and removal o...	51.39	33	http://www.re...	Untagged	147	0.22
WGCNA|Liver:8	CP:REACTOME	REACTOME_P...	Genes involved in p53-Independent G1/S DNA damage checkpoint	50.82	33	http://www.re...	Untagged	147	0.22
WGCNA|Liver:8	CP:REACTOME	REACTOME_A...	Genes involved in Autodegradation of the E3 ubiquitin ligase COP1	50.82	33	http://www.re...	Untagged	147	0.22
WGCNA|Liver:8	GO-CC	GO:0000502	proteasome complex	50.31	37		Untagged	147	0.25
WGCNA|Liver:8	CP:REACTOME	REACTOME_C...	Genes involved in CDT1 association with the CDC6:ORC:origin c...	50.27	33	http://www.re...	Untagged	147	0.22
WGCNA|Liver:8	CP:REACTOME	REACTOME_R...	Genes involved in Regulation of ornithine decarboxylase (ODC)	50.27	33	http://www.re...	Untagged	147	0.22
WGCNA|Liver:8	CP:REACTOME	REACTOME_D...	Genes involved in Destabilization of mRNA by AUF1 (hnRNP D0)	49.75	33	http://www.re...	Untagged	147	0.22
WGCNA|Liver:8	CP:REACTOME	REACTOME_S...	Genes involved in SCF-beta-TrCP mediated degradation of Emi1	49.75	33	http://www.re...	Untagged	147	0.22
WGCNA|Liver:8	CP:REACTOME	REACTOME_E...	Genes involved in ER-Phagosome pathway	49.30	34	http://www.re...	Untagged	147	0.23
WGCNA|Liver:8	CP:REACTOME	REACTOME_P...	Genes involved in p53-Dependent G1 DNA Damage Response	48.77	33	http://www.re...	Untagged	147	0.22
WGCNA|Liver:8	CP:KEGG	KEGG_PROTE...	Proteasome	48.01	31	http://www.ge...	Untagged	147	0.21
WGCNA|Liver:8	CP:REACTOME	REACTOME_R...	Genes involved in Regulation of Apoptosis	47.85	33	http://www.re...	Untagged	147	0.22
WGCNA|Liver:8	CP:REACTOME	REACTOME_A...	Genes involved in Assembly of the pre-replicative complex	47.42	33	http://www.re...	Untagged	147	0.22
WGCNA|Liver:8	CP:REACTOME	REACTOME_S...	Genes involved in SCF(Skp2)-mediated degradation of p27/p21	46.42	32	http://www.re...	Untagged	147	0.22

Data table: module annotat...
Marking: Marking (5)





Screen shot from Spotfire TXG_MAP tool for WGCNA:Liver_8, a module highly enriched in genes associated with the proteasome.



Sutherland JJ et al. The Pharmacogenomics Journal advance 2017

Overview

Modeling biological complexity

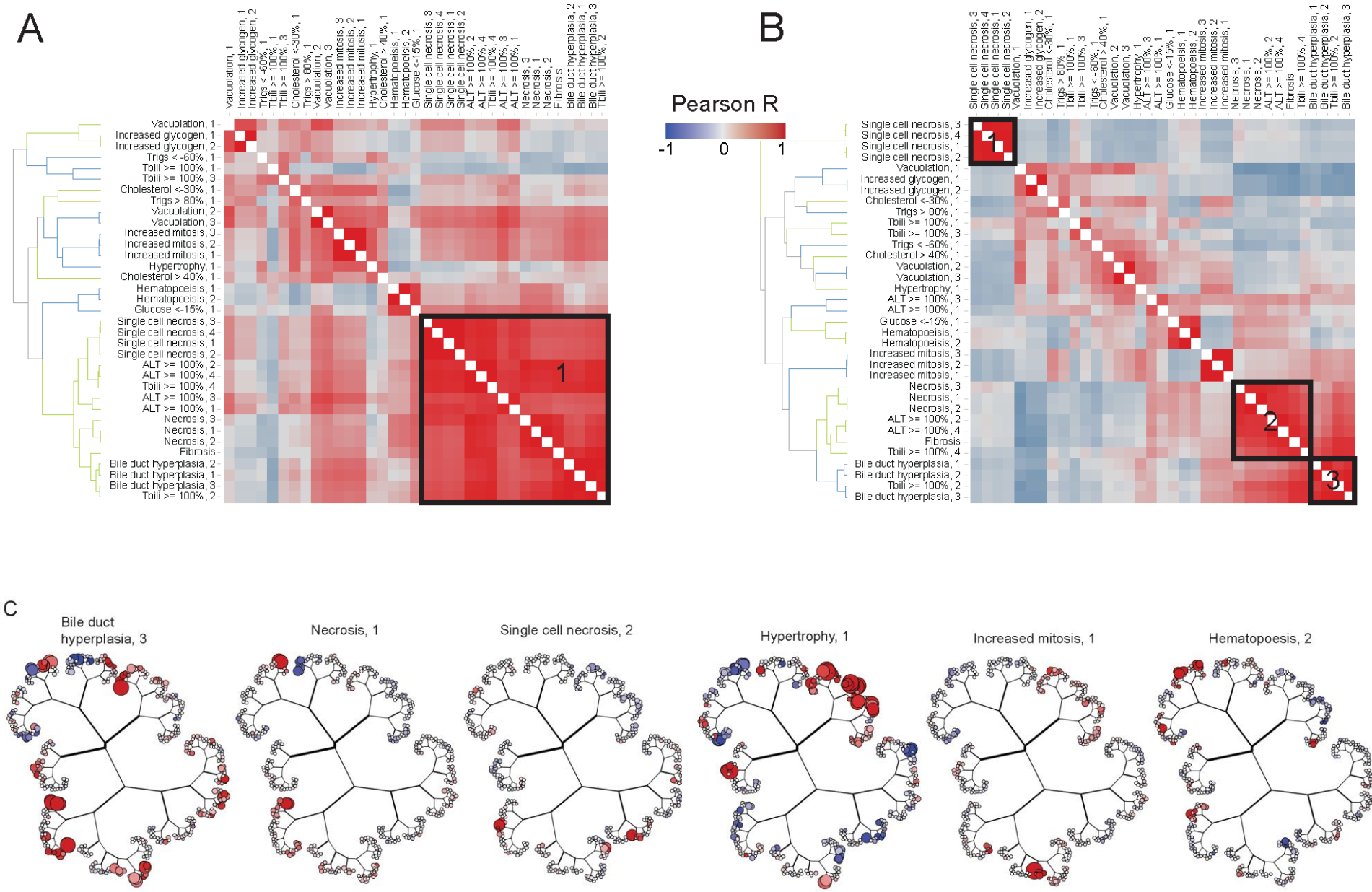
- The modular nature of complex systems.
- Leveraging modular systems models using gene expression data.
- Translating gene expression data into biological understanding.
- Reducing redundancy in MSigDB.
- Knowing what we don’t know.

Understanding Molecular Pathogenesis

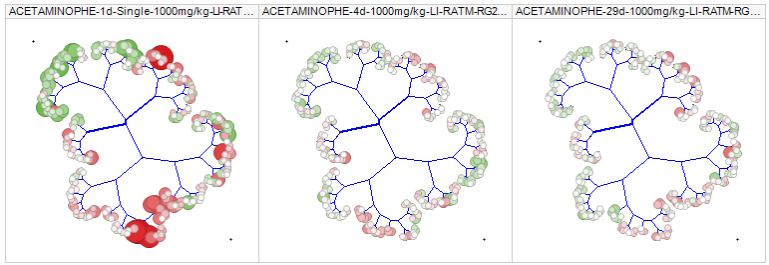
- Correlating expression modules with pathology.
- Closing the loop from transcription factor to pathogenesis.
- Predicting adaptive vs progressive responses.
- Closing the loop on transcriptional control (addendum slides).

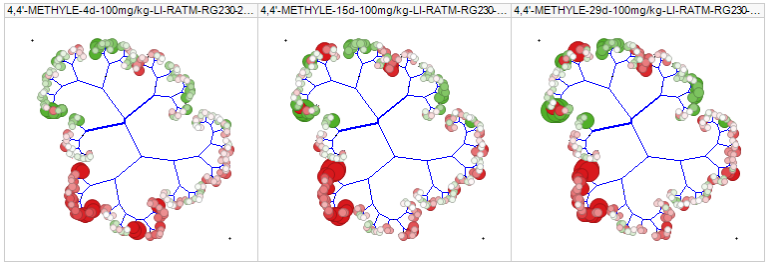
Applications of WGCNA to Dose Response Analysis

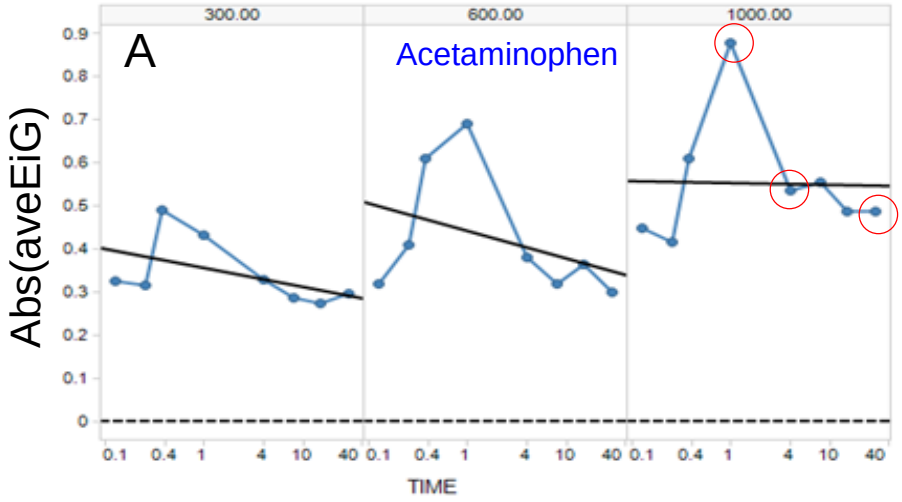
- Separating injury signals from tissue stereotypic response.
- Perturbing network in culture.
- Translation to human.

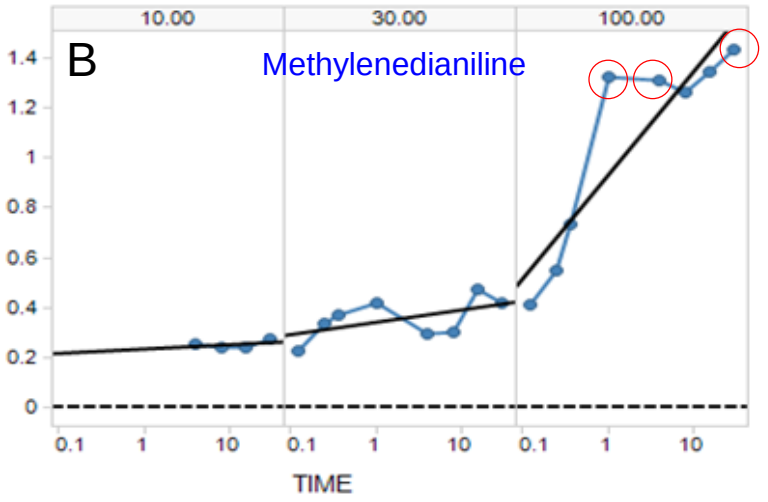

A
B
Pearson R
-1
0
1
C
Bile duct hyperplasia, 3
Necrosis, 1
Single cell necrosis, 2
Hypertrophy, 1
Increased mitosis, 1
Hematopoesis, 2

Predicting Adverse Responses for Adaptive vs. Progressive Tox-Phenotypes



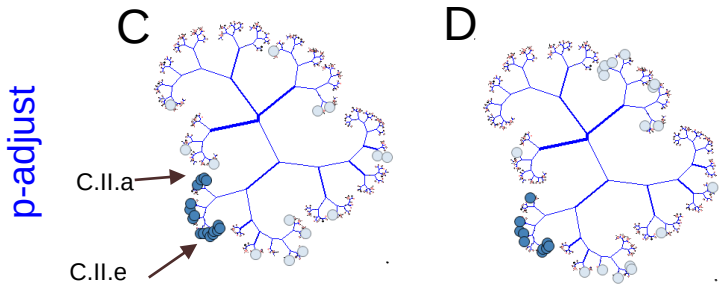






Adverse Concurrent

Adverse@29d



Panels A and B: Average absolute eigengene score (Abs(aveEiG) for all module changes for three dose levels of acetaminophen (A) and methylenedianaline (B) from TG-GATES 28d repeat dose rat liver studies. Note the scales for A and B differ. The TXG_MAP outputs shown correspond to the highest dose in the upper panel at different time points corresponding to the red circled time points.

PANELS C and D: Top-ranked modules selected for effect size >1.0 and p-adj $<10^{-3}$. are shown for any tox-phenotypes considered adverse and either concurrent (C – present at the same time as gene expression) or predictive of adverse tox-phenotypes occurring at any time later than 1 day (D). Discussed in detail in Fig 2 of Sutherland et al. (2017).

Overview

Modeling biological complexity

- The modular nature of complex systems.
- Leveraging modular systems models using gene expression data.
- Translating gene expression data into biological understanding.
- Reducing redundancy in MSigDB.
- Knowing what we don’t know.

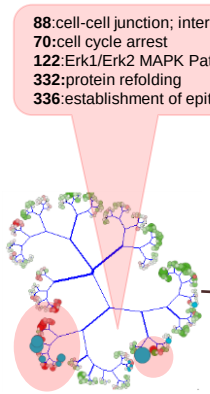
Understanding Molecular Pathogenesis

- Correlating expression modules with pathology.
- Closing the loop from transcription factor to pathogenesis.
- Predicting adaptive vs progressive responses.
- Closing the loop on transcriptional control (addendum slides).

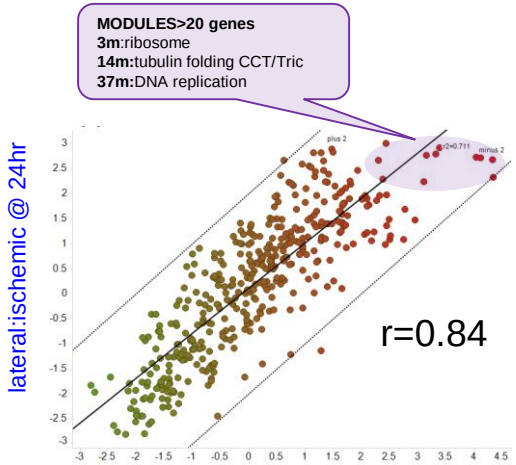
Applications of WGCNA to Dose Response Analysis

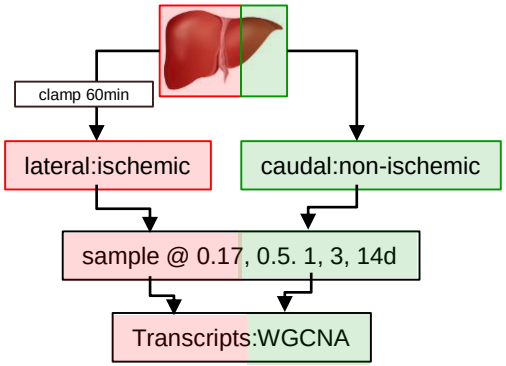
- Separating injury signals from tissue stereotypic response.
- Perturbing network in culture.
- Translation to human.

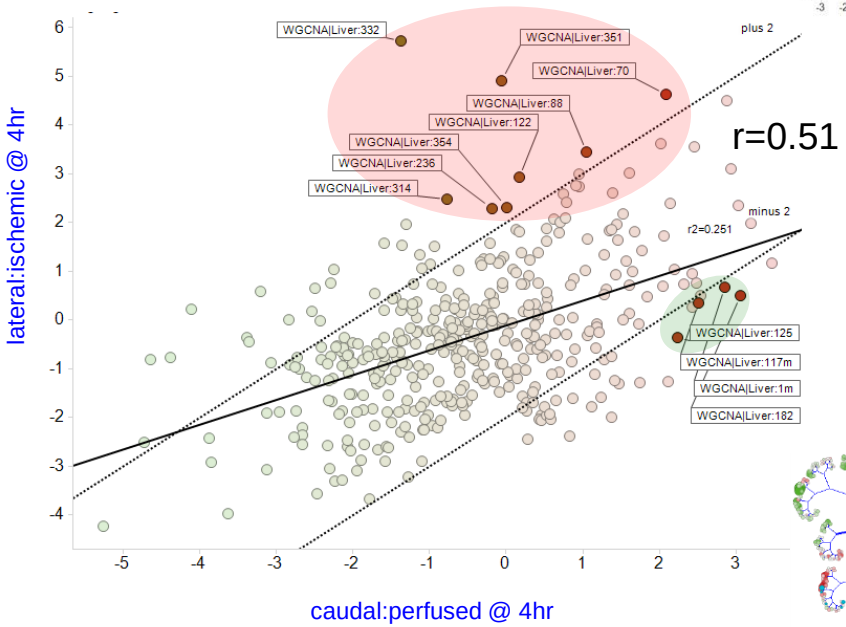
The Importance of Time Series



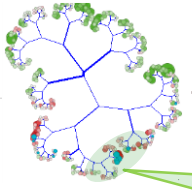
BEST COMPARITORS:
-bortezomib r=0.7
-cycloheximide r=0.69
-N,N diemethylnitrosamine r=0.68
-phorone r=0.68







BEST COMPARITORS:
Partial hepatectomy 6 hr r=0.92
Partial hepatectomy 3 hr r=0.87



117m: RNA processing
125:nucleospeck
1m:KEGG:Spliceosome
198:micro-ribonucleoprotein complex

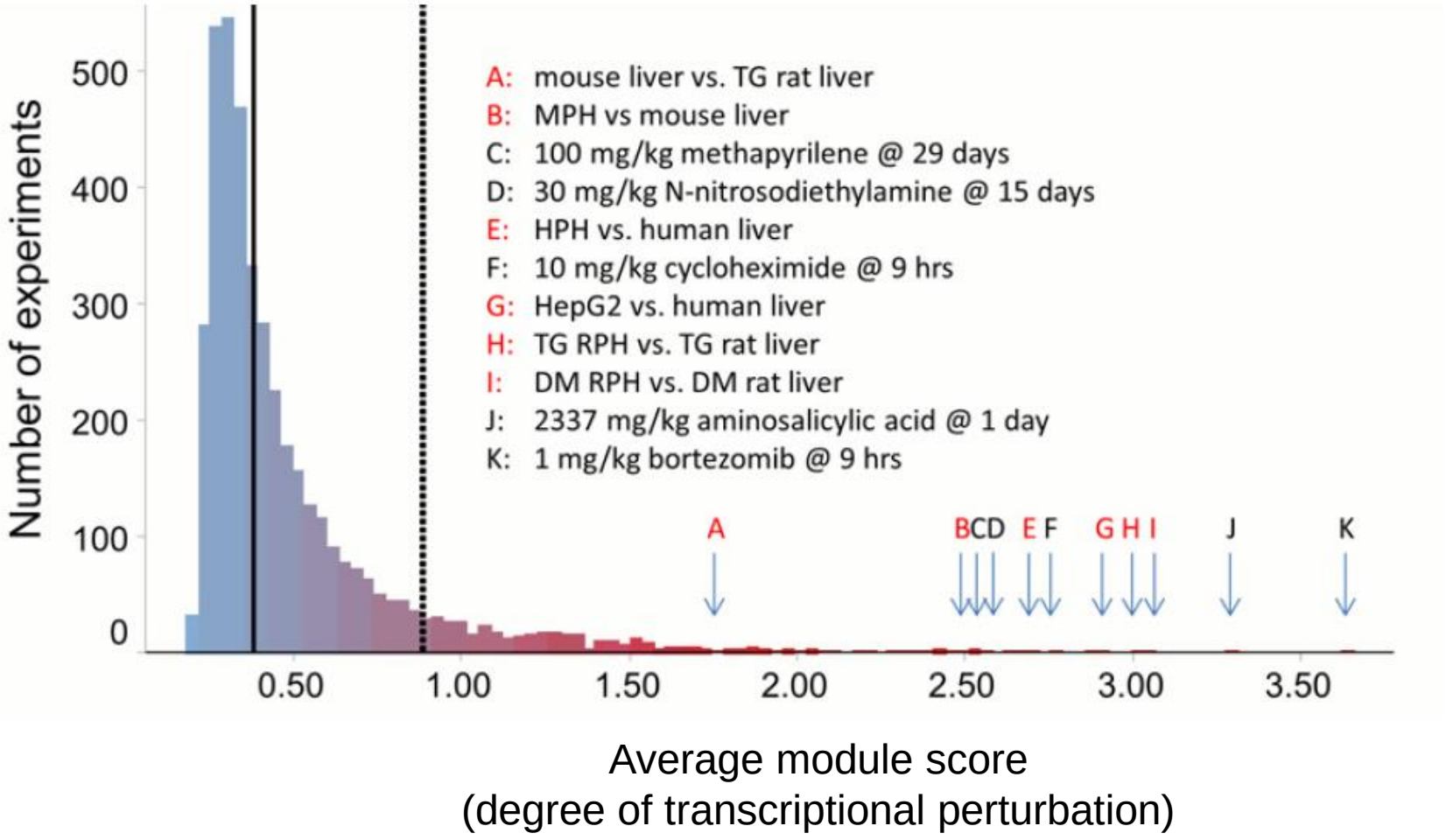
ISCHM:nonISCHM @ 4 hr r=0.51

- Lateral difference drivers relate to heat shock, cell cycle arrest, cell-cell junction changes
- Caudal difference drives relate to ribosomal RNA processing
- Evidence of distinct responses

ISCHM:nonISCHM @ 1d r=0.84

- Similarity driven by DNA replication, ribosomal biogenesis and tubulin formation
- Evidence of convergent pathobiology

Perturbations Caused by Placing Hepatocytes in Culture.



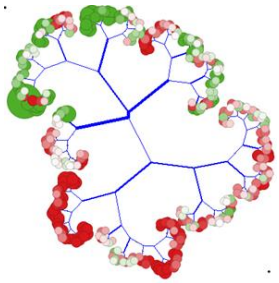
Sutherland, JJ et al. PLoS Comp Biol.

HYPOTHESIS: Probability of Human Liver Toxicity Given Nonclinical Toxicity is a Function of Network EiG, Preservation and Effect Size.

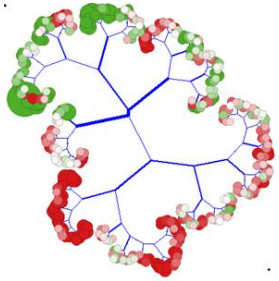
rLIVER
(RL)

p(hLT|rLT)
p(rLT|hLT)

hLIVER
(hL)



p(hLT|rLT)
p(rLT|hLT)



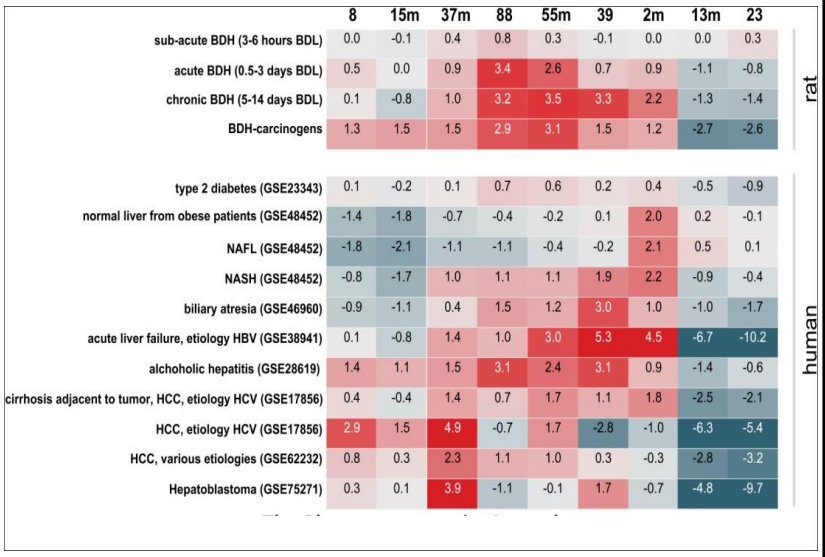


Fig. 6. (Sutherland et al. 2017) Module scores are averaged across treatments in each BDH subtype (rats) from Figure 5 (top heatmap) and human samples (bottom heatmap) available in each Gene Expression Omnibus (GEO) series, identified via their accession number.

p(hLT|rLT) = ƒ (networks)

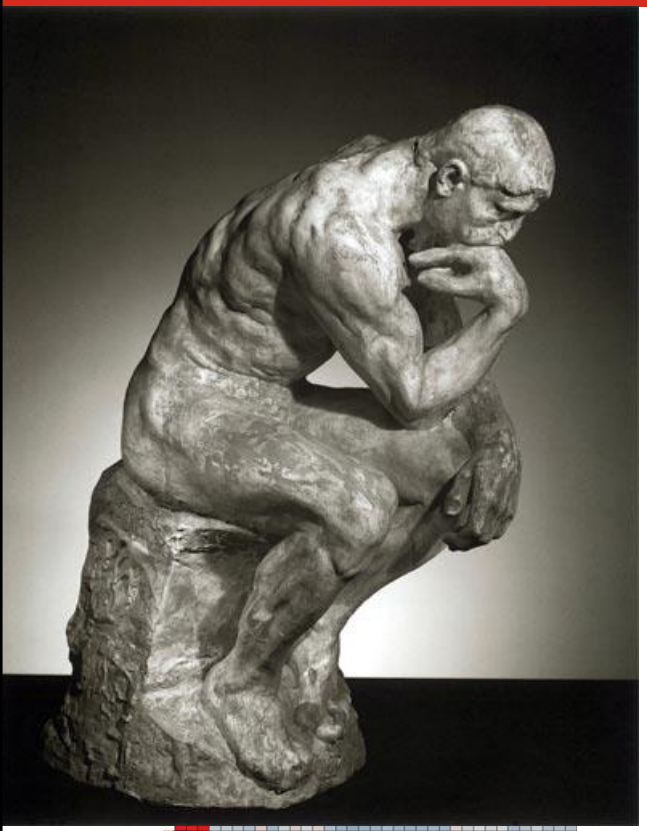
p(hLT|rLT)= ƒ (EiG, eff size, preservation)

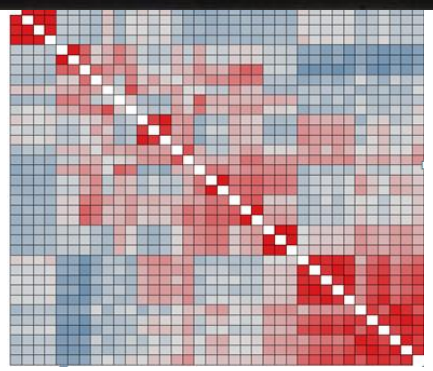
Where:

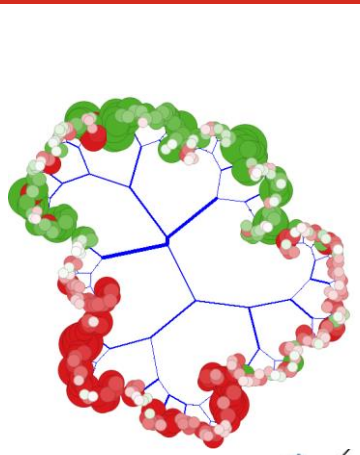
- *EiG – eigengene score (How much did it change?)*
- *Eff Size – Is the network associated with adverse outcomes?*
- *Preservation - Z-score (Is the network preserved in human?)*

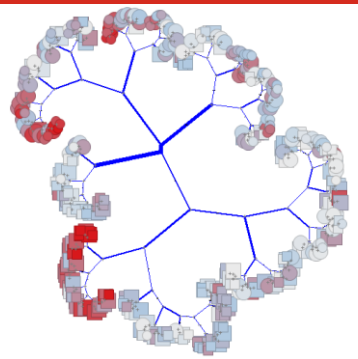
transQST

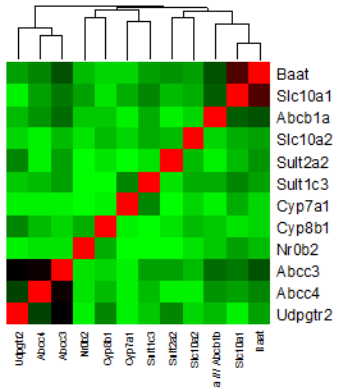
QUESTIONS?

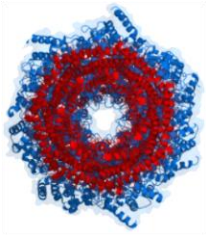


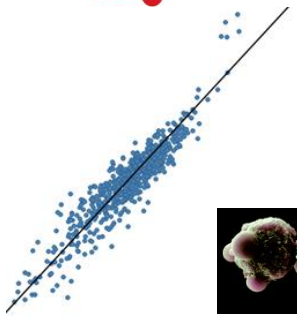


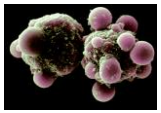


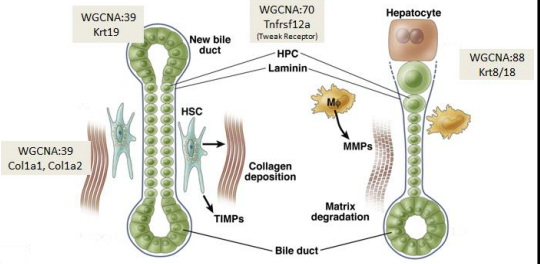


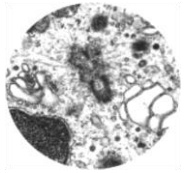

Baat
Slc10a1
Abcb1a
Slc10a2
Sult2a2
Sult1c3
Cyp7a1
Cyp8b1
Nr0b2
Abcc3
Abcc4
Udpgtr2

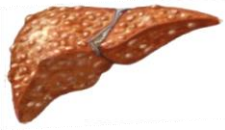


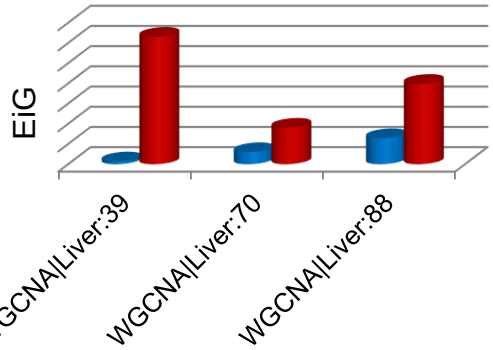





WGCNA:39
Krt19
New bile duct
WGCNA:70
Tnfrsf12a
(Tweak Receptor)
HPC
Laminin
Hepatocyte
WGCNA:88
Krt8/18
HSC
Mφ
MMPs
WGCNA:39
Col1a1, Col1a2
Collagen deposition
TIMPs
Matrix degradation
Bile duct






EiG
WGCNA|Liver:39
WGCNA|Liver:70
WGCNA|Liver:88

REFERENCES

1. *Wagner, G. et al., The road to modularity. Nat. Rev. Genetics 8:921-931 (2007)*

2. Horvath, S. et al. Weighted Gene Co-expression Network Analysis (WGCNA) http://labs.genetics.ucla.edu/horvath/CoexpressionNetwork/

3. Sutherland, J.J. et al. Toxicogenomic module associations with pathogenesis: a network-based approach to understanding drug toxicity. Pharmacogenomics Journal, in press (2017).