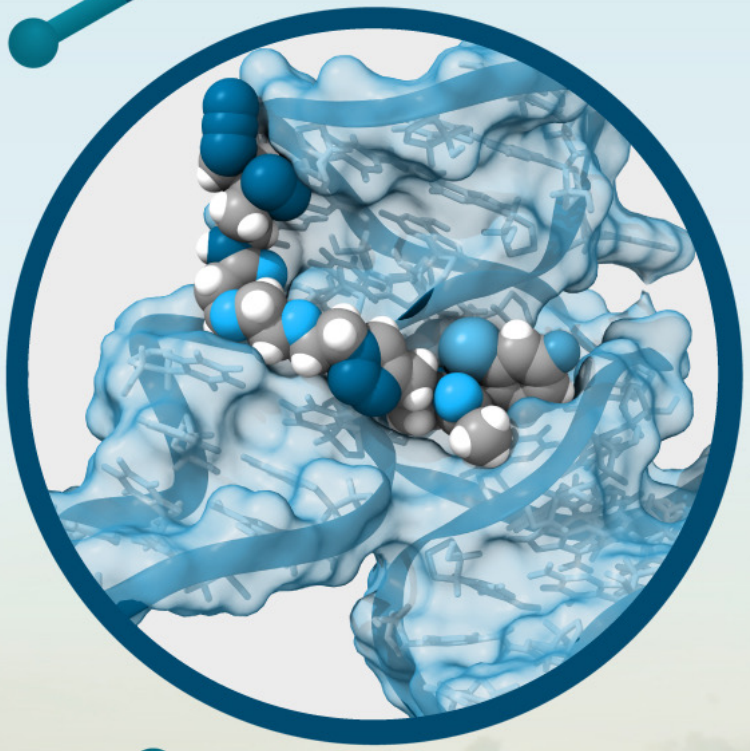


Application of IPA for interpreting RNA-targeted small molecule treatment datasets

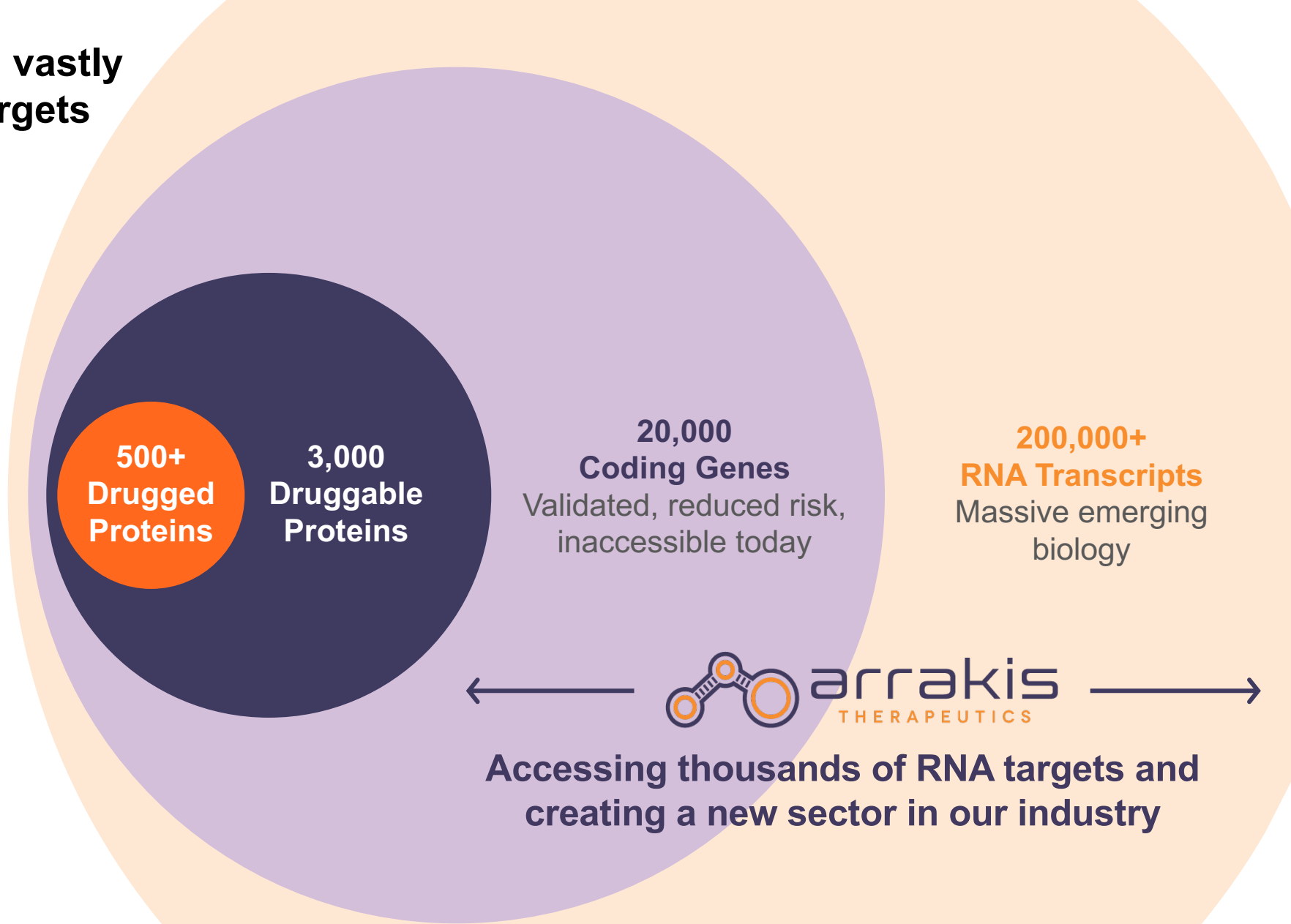
Alex Amlie-Wolf, PhD

IPA User's Group meeting

09/22/22

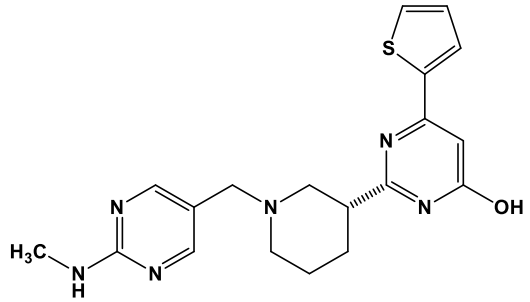
The ability to drug RNA will vastly increase our therapeutic targets

- Targeting RNA extends small-molecule medicines beyond the domain of well-studied protein targets, tapping into **validated but previously inaccessible** biology
- Goal: ligand RNA with drug-like (oral) small molecules
- This requires building a toolkit for identifying and optimizing RNA-targeted small molecules (“rSMs”)



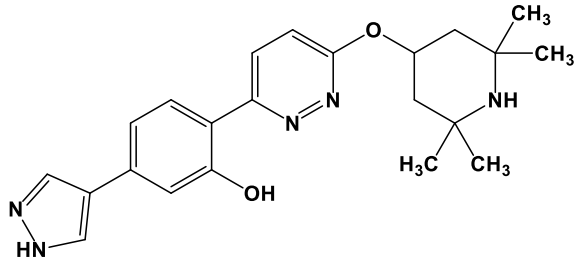
Bioactive, RNA-targeting, drug-like SMs already exist

MERCK



Ribocil-A FMN riboswitch
Nature 2015

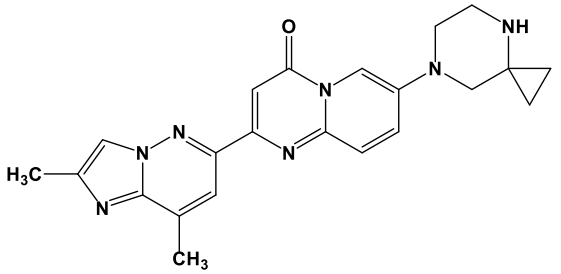
NOVARTIS



Branaplam (LMI070) SMN2
Nature Chem Bio 2015

PTC THERAPEUTICS

Roche



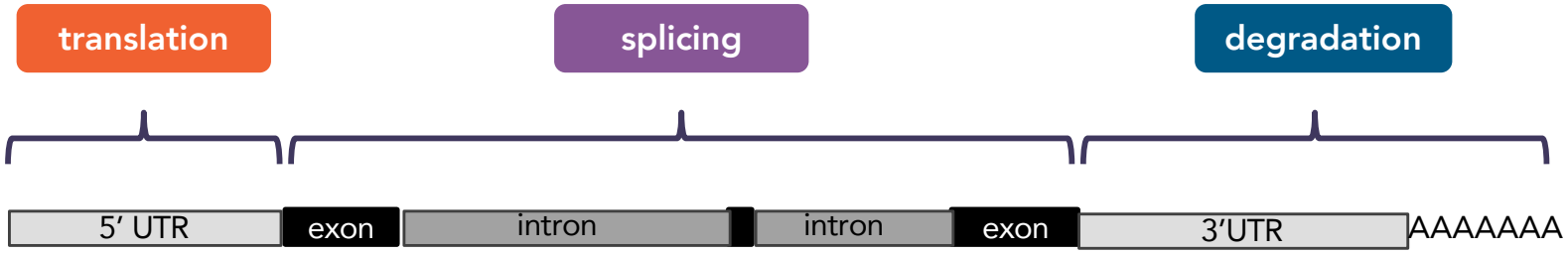
Risdiplam (RG7916) SMN2
J Med Chem 2018

- These compounds were identified in phenotypic screens and **later** discovered to act on RNA targets
- Branaplam is in Phase 1/2 clinical testing and risdiplam was recently approved by the FDA for the treatment of spinal muscular atrophy (SMA) and is marketed under the name Evrysdi™
- Our approach is to start with an RNA structure and screen small-molecule libraries to identify selective binders followed by assessment of their activity in cells

arrakis THERAPEUTICS

Our approach to drugging mRNA

- Focus on mRNA to modulate the expression of undruggable proteins
- Compounds could affect expression by acting at any stage of the mRNA lifecycle – splicing, translation, mRNA decay



Use bioinformatics data from public and proprietary sources to select targets:

- mRNA isoform complexity
- Conservation
- Human SNPs

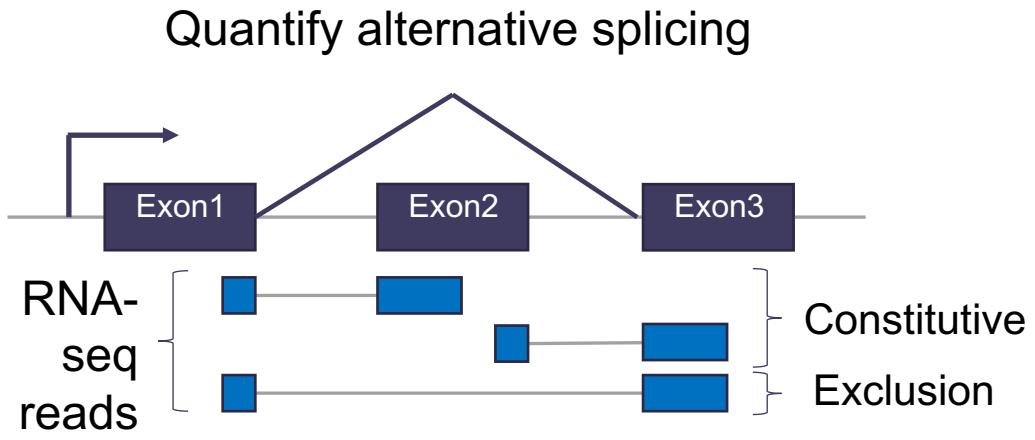
- Transcriptional start sites
- Translation start sites
- Translational efficiency (e.g., ribosomal profiling)

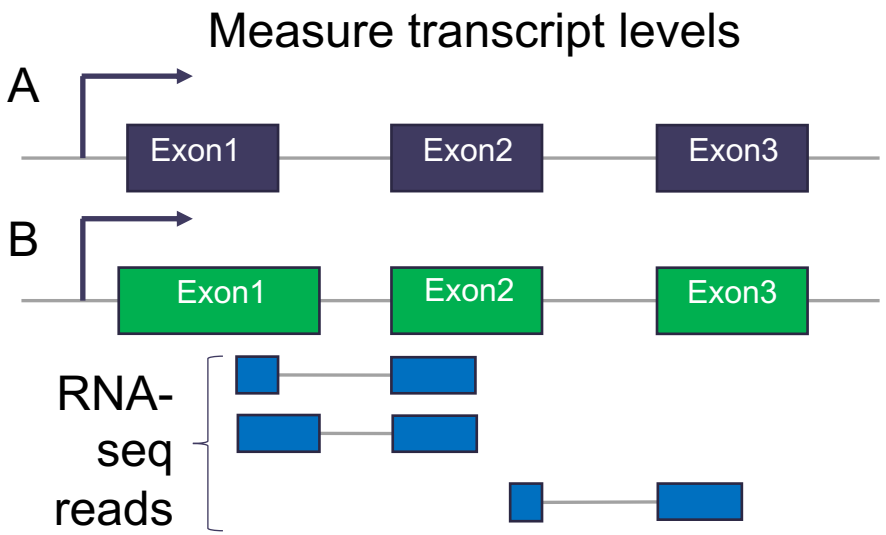
- Splicing efficiencies at each intron, assessing basal and induced mis-splicing
- RBPs (binding and functional impact)

- 2D RNA folding
- 3D structure analysis to identify potential ligandable pockets

Goal: characterize and understand off-target rSM effects

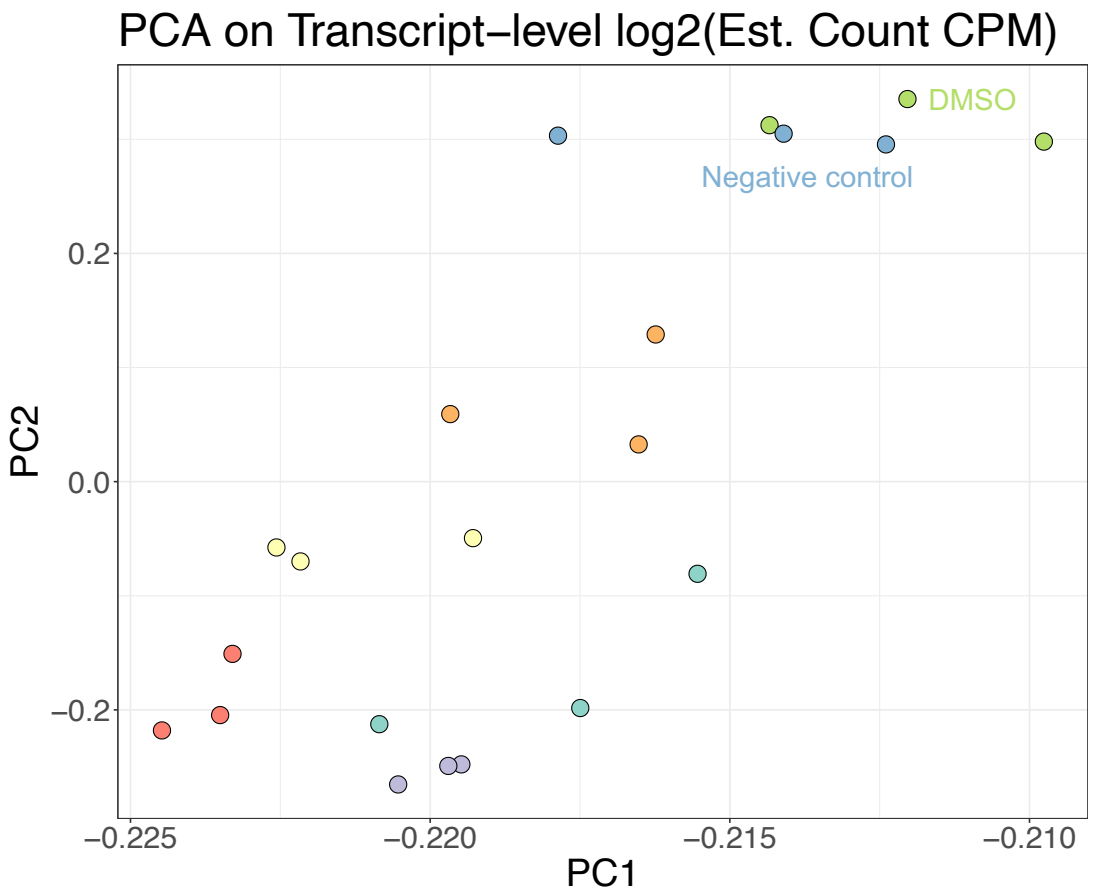
- Whatever the mechanism of action, rSMs interact with RNA molecules
- Once we screen and identify rSMs with quantifiable RNA binding, we need to understand their function
- Specific hypotheses about how our molecules act on RNA can be sensitively tested with bespoke assays
- However, ultimately we need to know what happens when rSMs get into cells!
- RNA-seq and proteomics are essential tools for understanding global effects on other RNAs and on proteins, respectively



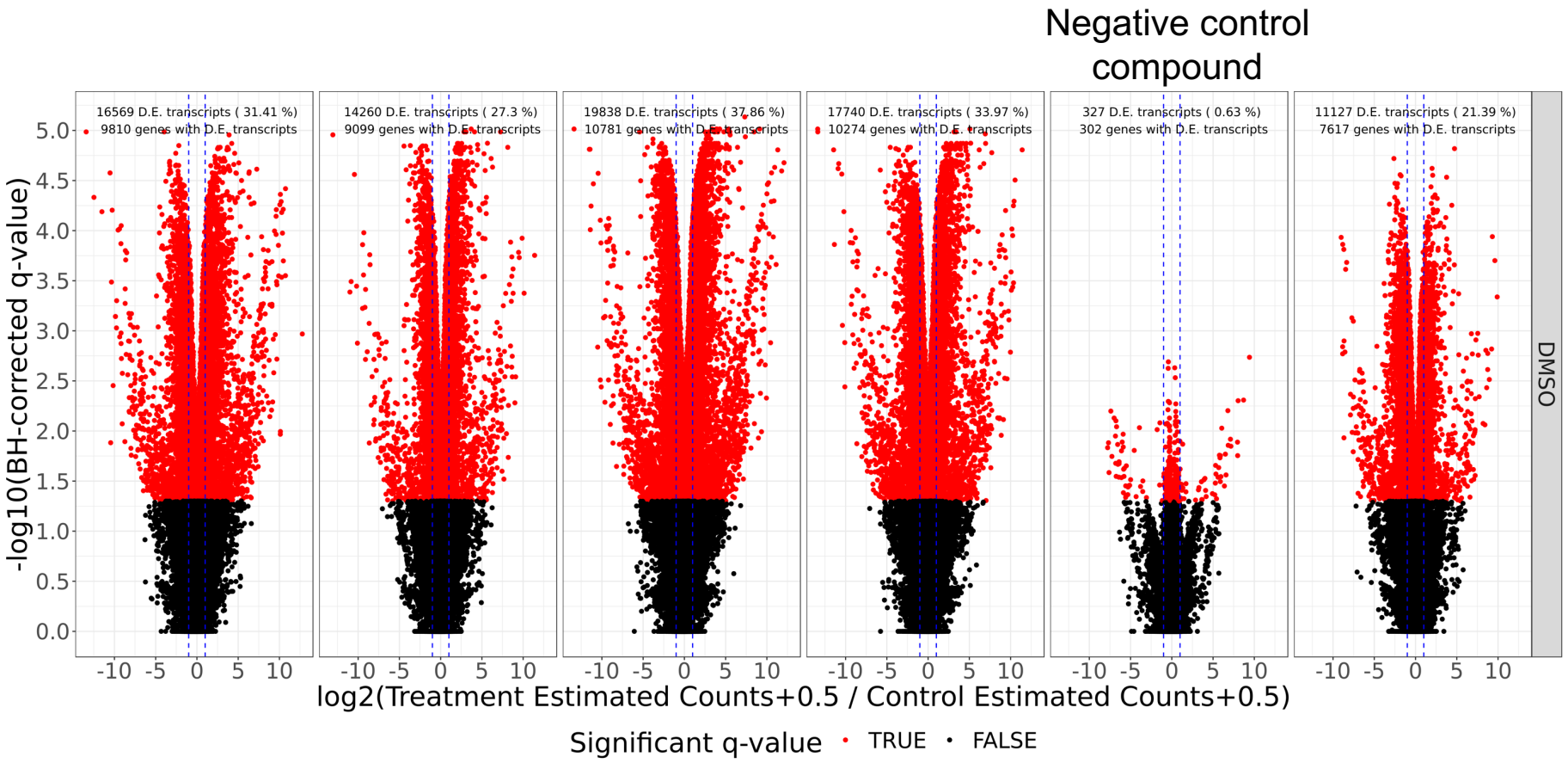


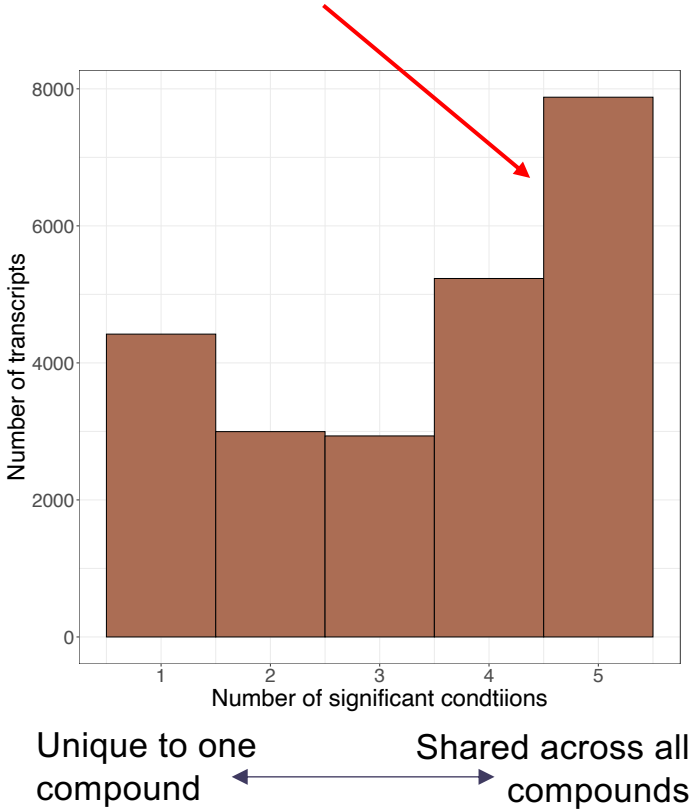
Case study: early compound for internal program

- Testing 5 analogs of cell active compound, along with structurally-related negative control compound against DMSO
- Three replicates per condition, 85-140M reads per sample
- Differential expression analysis used Kallisto estimates followed by Sleuth, with p-value correction across multiple condition comparisons
  - Found similar results with htseq-counts + DESeq2



Differential expression analysis shows widespread and shared compound effects



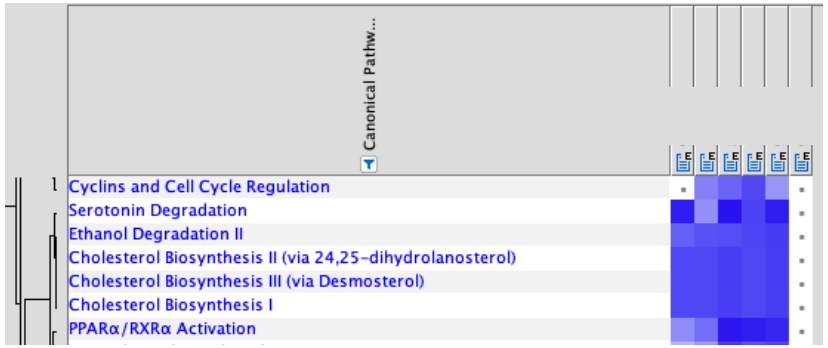


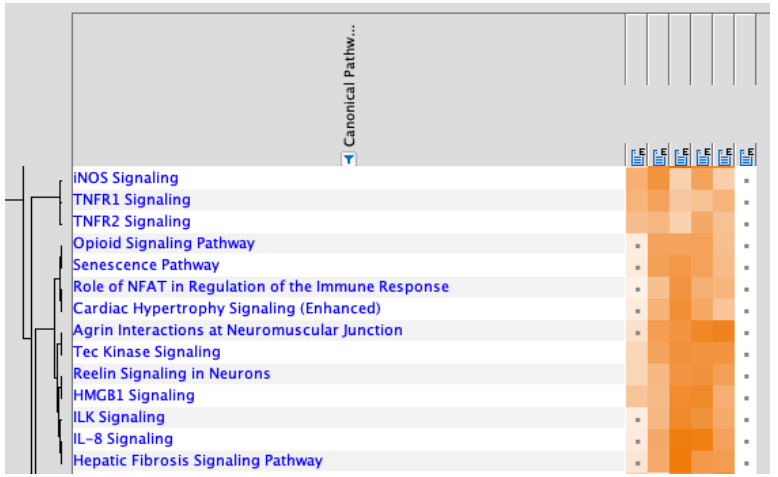
Widespread transcriptomic dysregulation is apparent, can we use IPA to pick out specific processes?

Uses Kallisto estimated counts with 100 bootstraps for transcript estimation to perform response error linear modeling and quantify differential expression with Sleuth tool, statistics based on likelihood ratio test between model with and without treatment variable

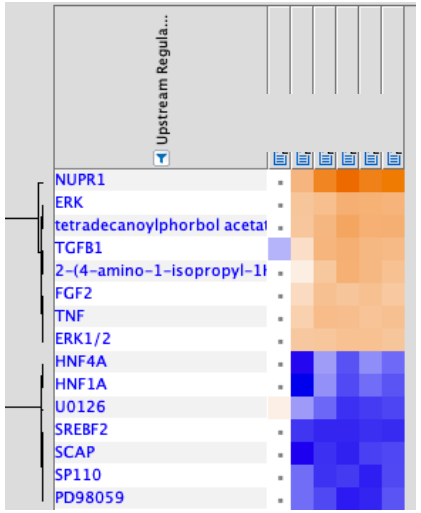
IPA characterizes common and compound-specific effects

- For all 6 compounds, identify transcripts with significant differential expression and at least 2-fold change in either direction, then collapse to genes
- Maximum fold change of a given transcript is used for collapsing (under advanced analysis settings)

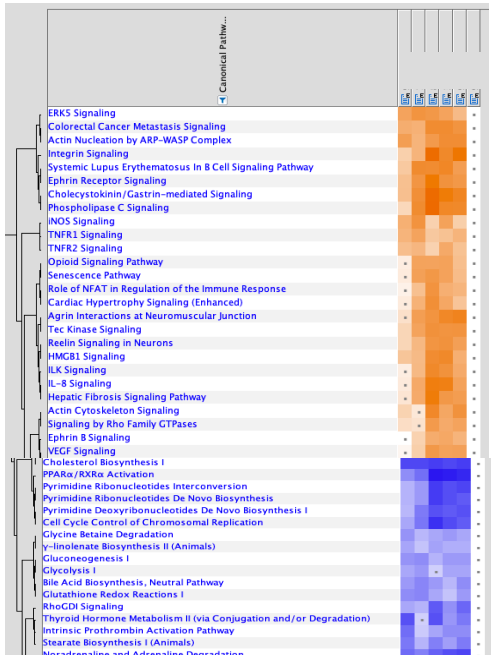




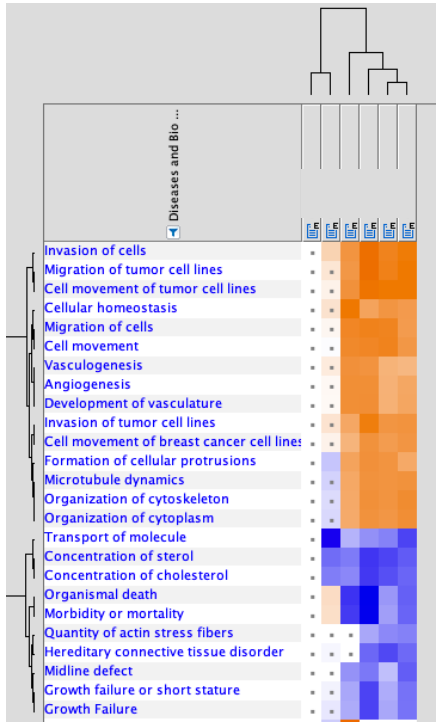
Pathways



Regulators



Tox functions

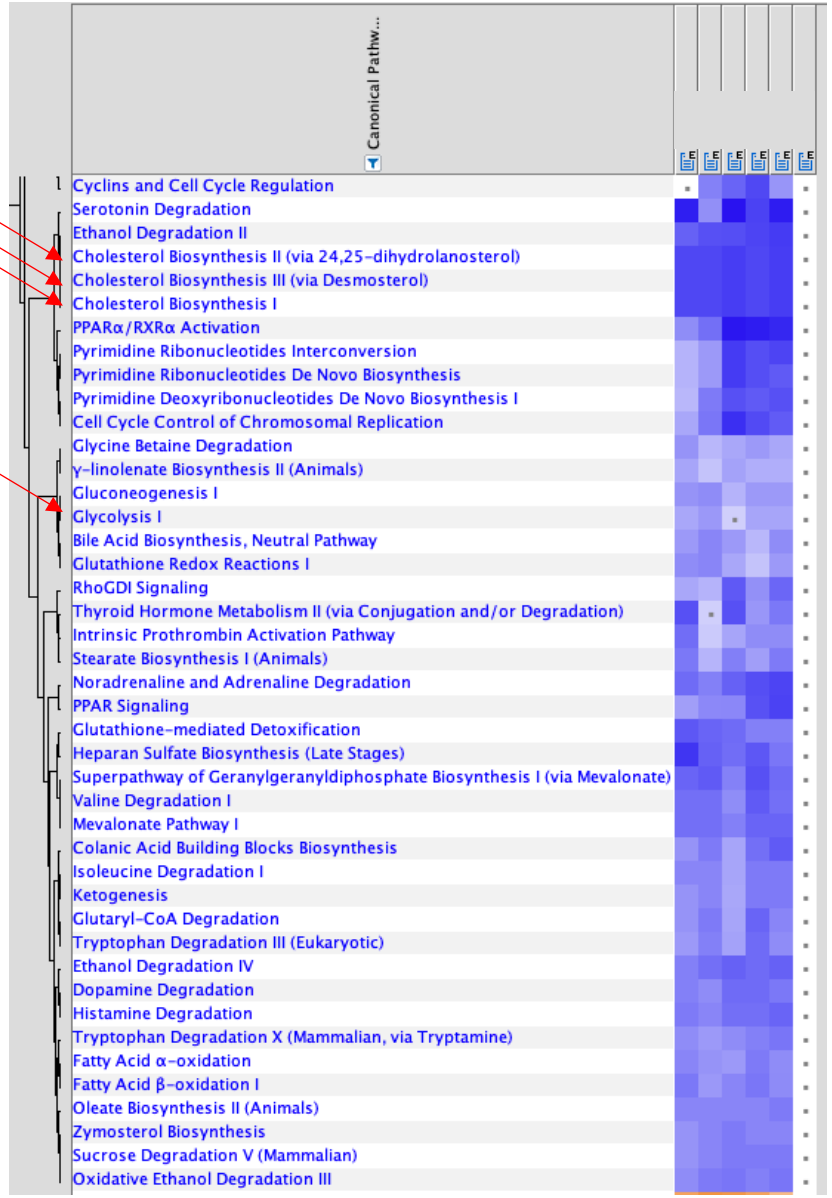


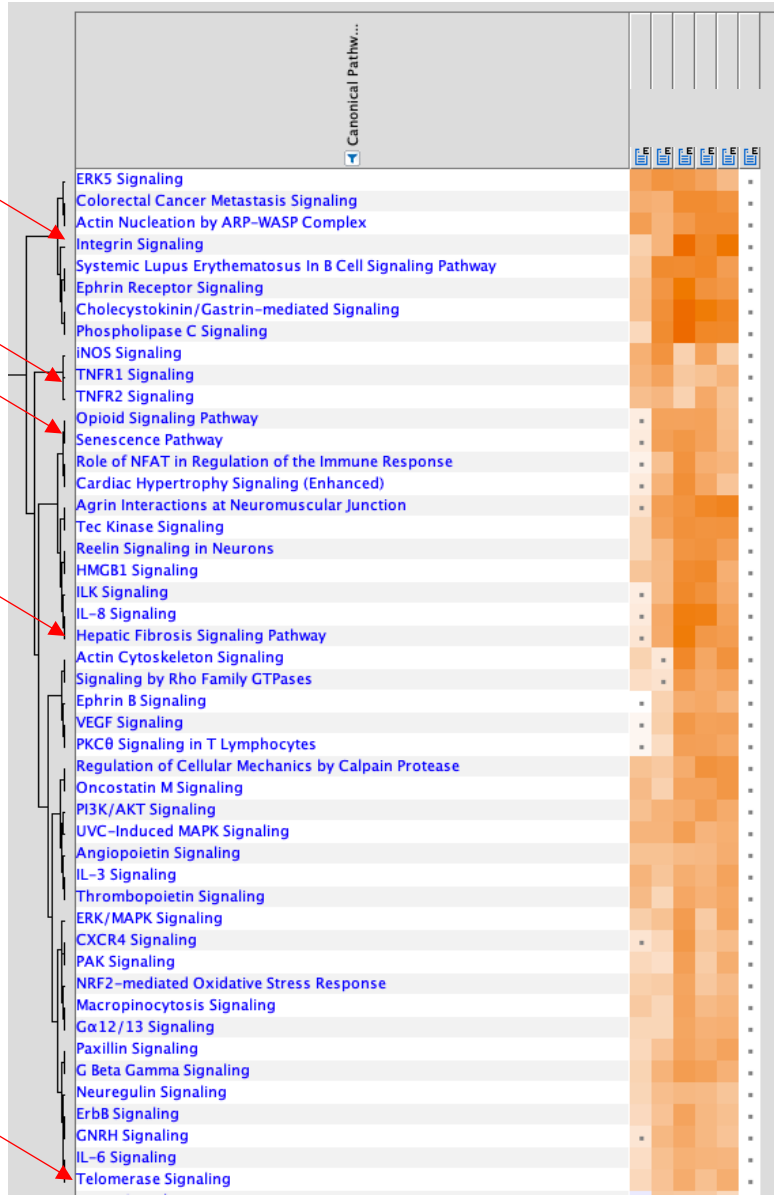
Bio functions

Downregulated; Upregulated

Pathway analysis of compound treatment effects

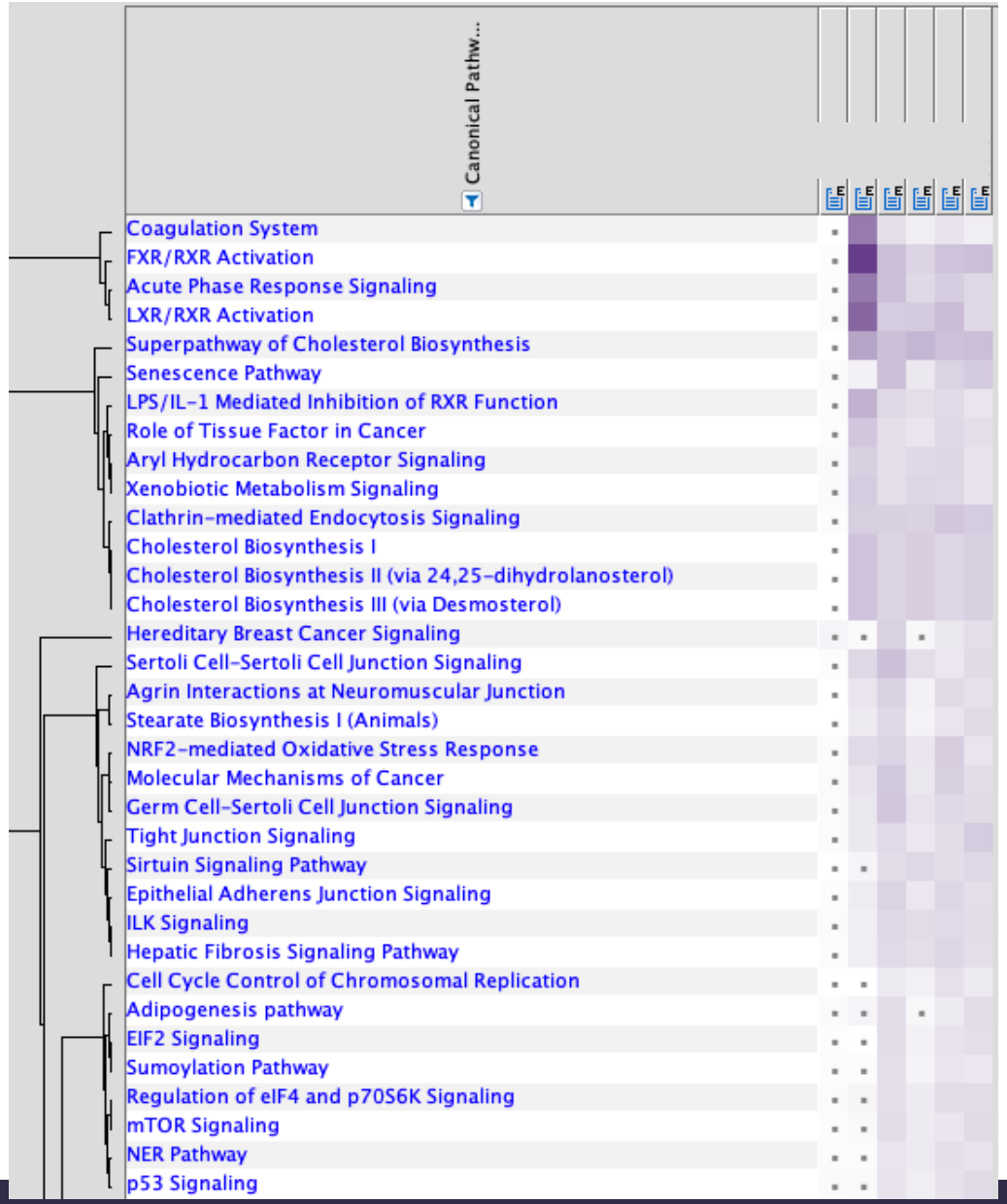
- Hierarchical clustering identifies sets of consistently downregulated (blue) and upregulated (orange) pathways
- Most strongly enriched pathways were found across all compound treatments
- Broadly seems to suggest cell stress / death pathways (upregulation in signaling and senescence pathways, downregulation of biosynthesis especially cholesterol and cell cycle)





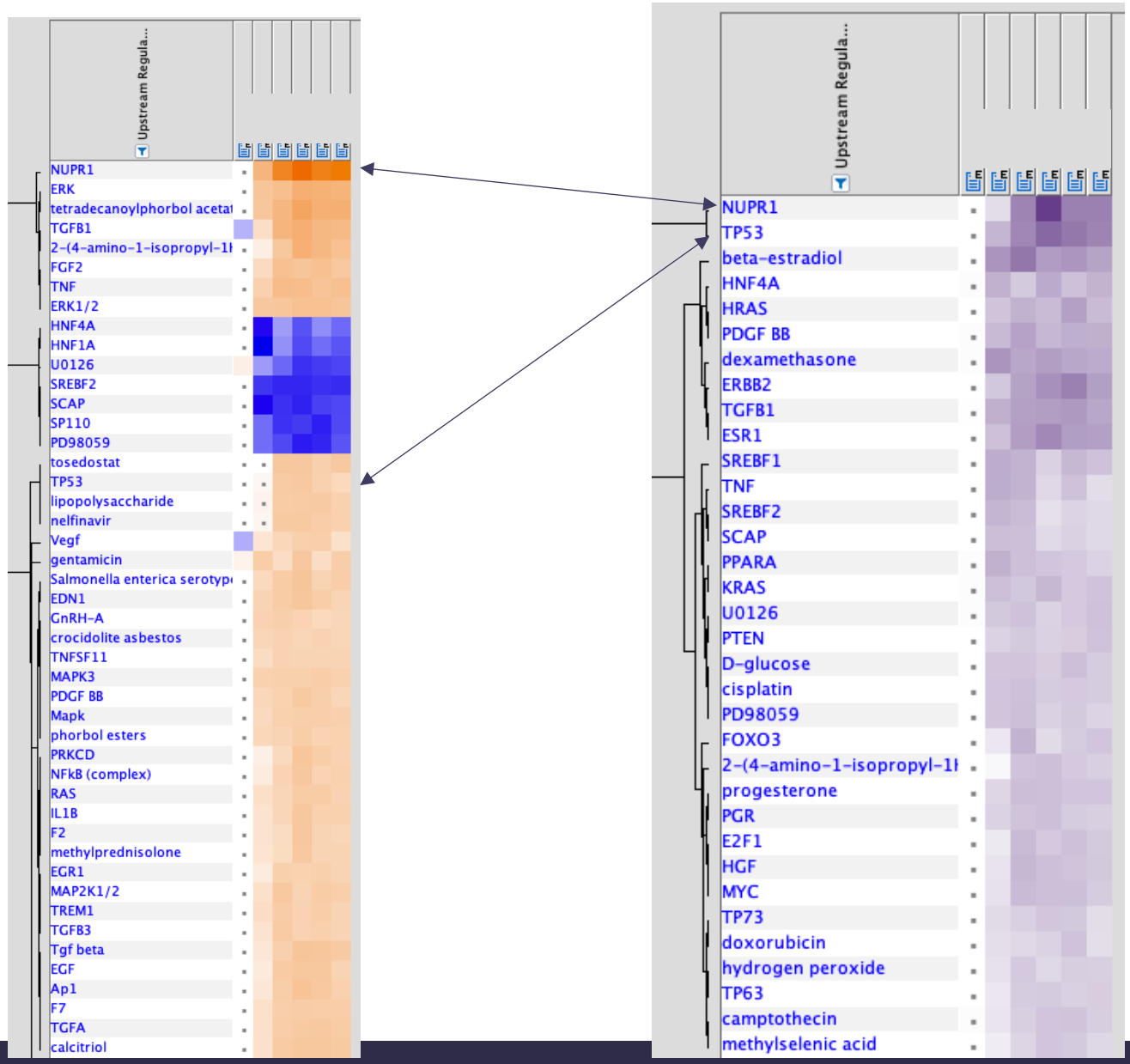
Strongest pathways by p-value

- Enrichment p-value as opposed to fold change-based Z-score encodes enrichment of affected genes in a pathway
- Picks out coagulation, RXR and phase response signaling as specific effect of one compound, senesecence and cholesterol biosynthesis more broadly



Upstream regulator analysis supports common effects mediated by specific factors

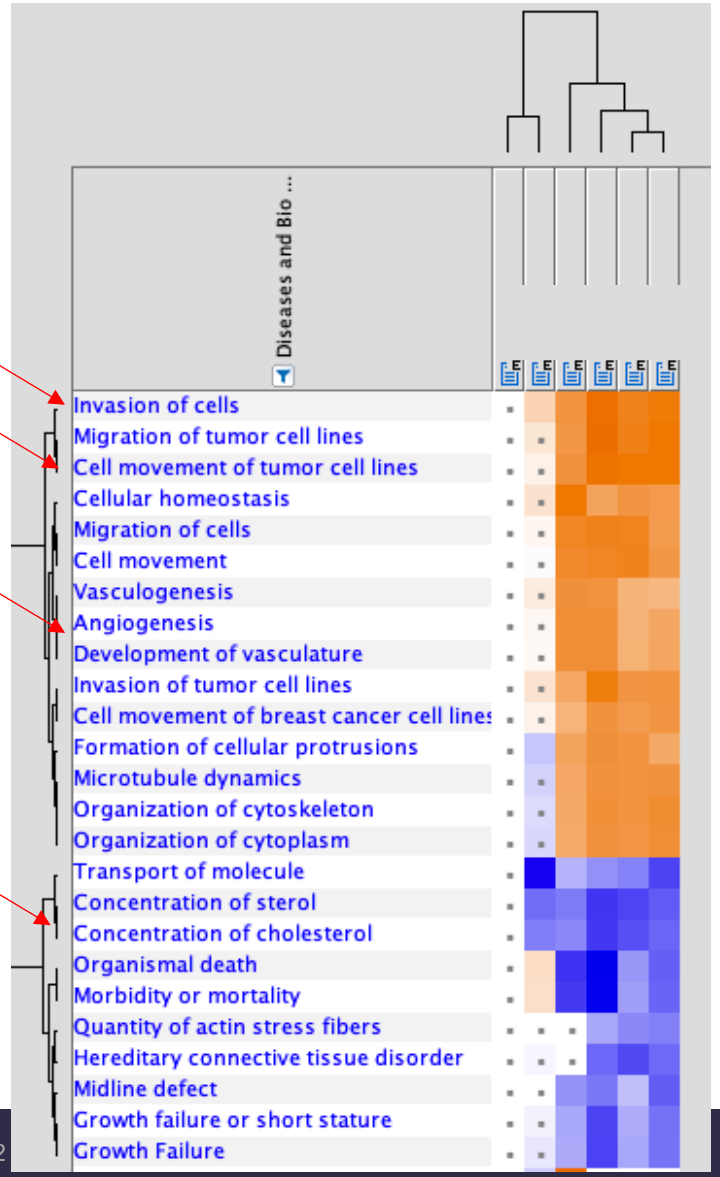
Visualizing Z score (blue is downregulated, orange is up) shows clusters of both up-regulated and down-regulated regulatory pathways



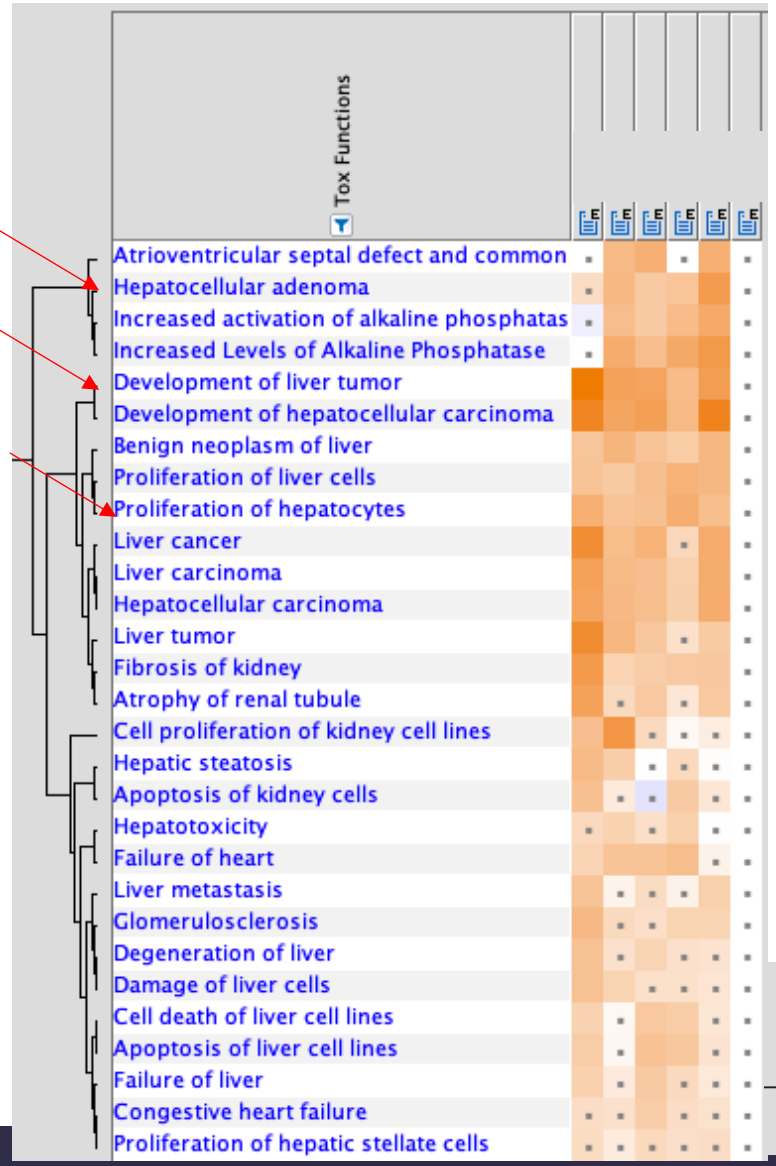
Visualizing corrected p-value with hierarchical clustering shows strong NUPR1 (high z-score) and TP53 (moderate z-score) signals across compounds

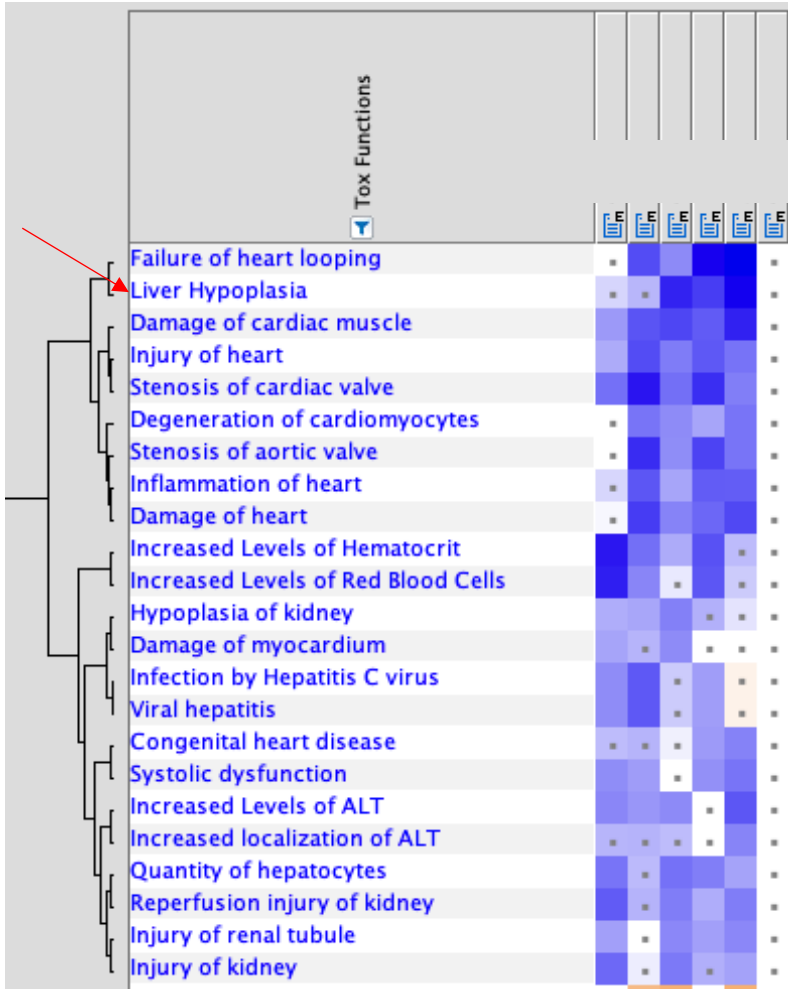
Toxicology and disease functions suggest broad carcinogenic effects

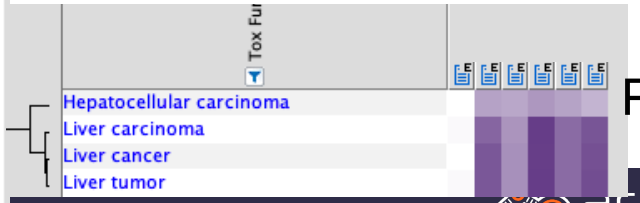
Disease / bio function



Tox function







P-values

Challenges in interpreting pathway analysis results

- **With such widespread transcriptomic compound effects, it can be overwhelming to look at pathway enrichments**
  - Hard to avoid the pitfall of looking for your favorite/expected enrichments, but on the other hand can help to give hints about MoA
  - Any other strategies and tips to get closer to interpretable results?
- **Combination of p-value and z-score results can help to tease out 'real' enrichments**
- **Upstream regulator analysis is powerful approach to try to explain observed changes, even with widespread cellular effects**

aamliewolf@arrakistx.com