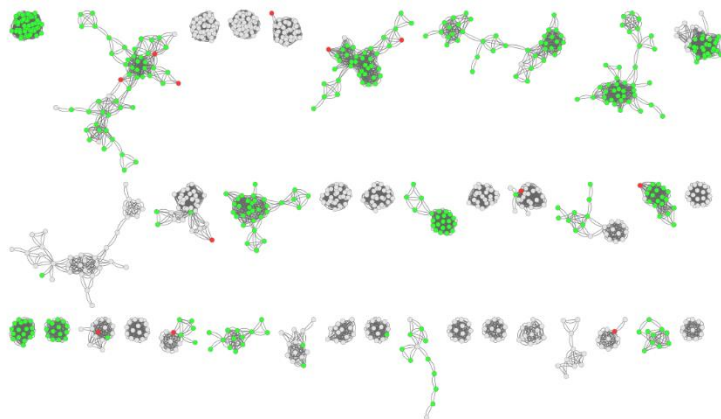# NTP's Proposed Approach to Biological Interpretation of Genomic Dose-Response Results

Scott S. Auerbach Ph.D., DABT
Biomolecular Screening Branch
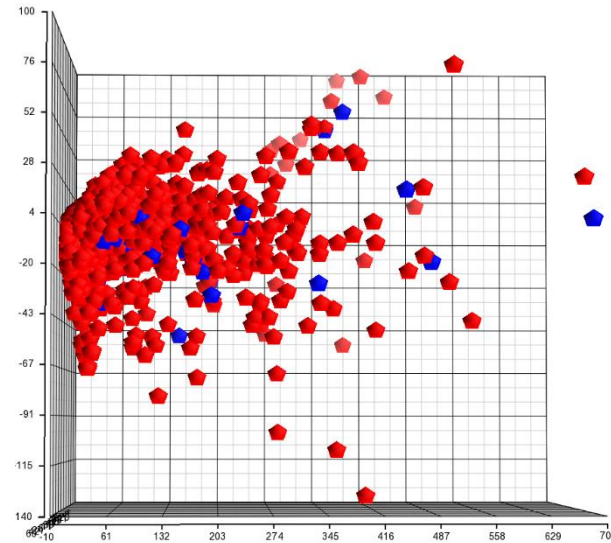National Institute of Environmental Health Science

Expert Panel Meeting on the Peer Review of Draft NTP
Approach to Genomic Dose-Response Modeling
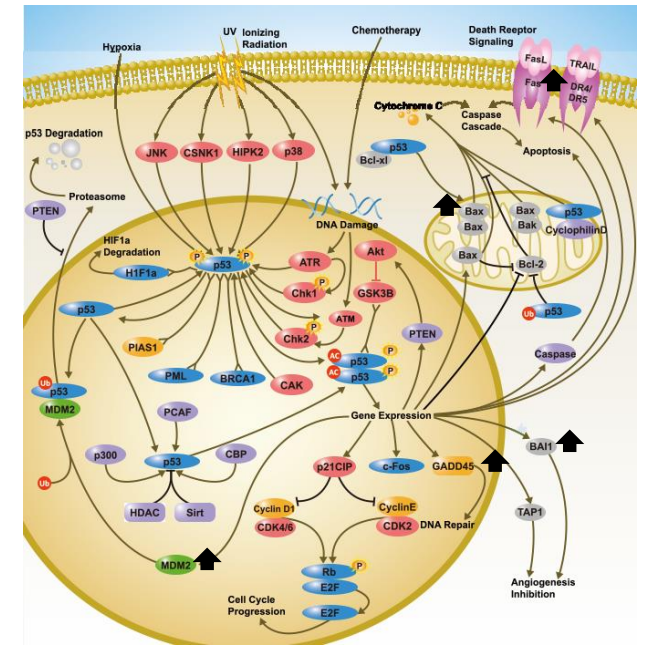October 25, 2017

- Historically

  – Gene Ontology, KEGG, Ingenuity, GeneGo, MSigDB C2, Toxicity Signatures, Co-expressed gene sets

- Challenges

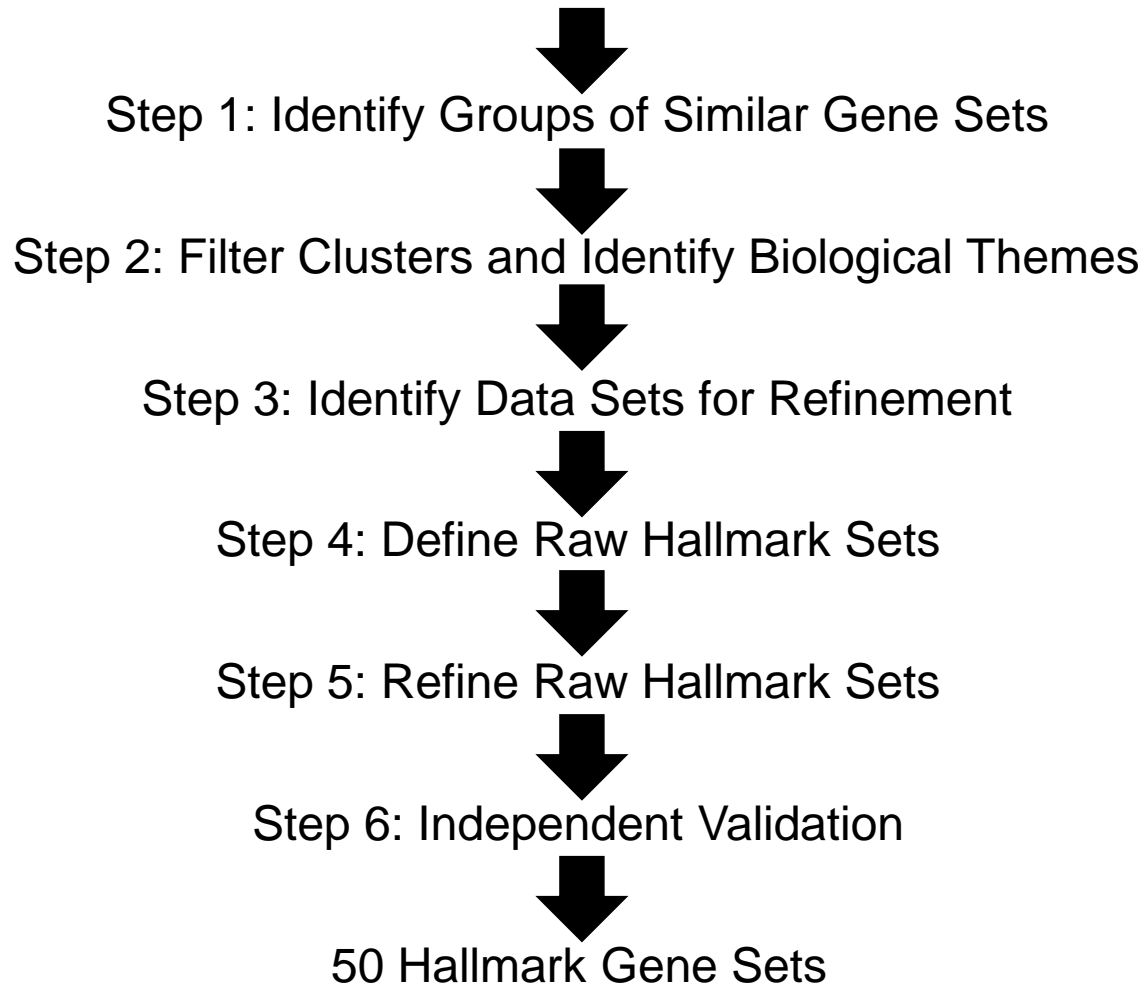 – High redundancy, Incomplete coverage of gene space, Partial congruency with gene expression data

CPDB Pathway Clusters

KEGG Pathway Genes

⬆ Expression Change

MSigDB
Molecular Signatures Database

Step 1: Identify Groups of Similar Gene Sets

Step 2: Filter Clusters and Identify Biological Themes

Step 3: Identify Data Sets for Refinement

Step 4: Define Raw Hallmark Sets

Step 5: Refine Raw Hallmark Sets

Step 6: Independent Validation

50 Hallmark Gene Sets

Liberzon et al., Cell, 2015

- Limited redundancy

- High percentage of each gene sets is regulated at the level of transcript abundance

- Empirically validated/curated

- Challenges
  - Limited gene coverage
  - No toxicological interpretation

- We want to develop *Hallmarks+*

- Mine the GEO database to identify co-regulated gene sets not currently captured in the Hallmark Gene sets

  - Due to our effort with the development of the S1500+ gene set most of the required data has been organized and curated

- Mine existing phenotypic-anchored signatures such as those that contained in the DrugMatrix database and those from the published literature

- Remine MSigDB and CPDB in manner similar to the what was done to create the Hallmark gene sets to identify additional sets that may have been overlooked

Deepak Mav and Logan Everett

- NextBio (Illumina Correlation Engine)

    – Identify toxicological treatments and phenotypes that are associated with changes in the Hallmark sets

    – Identify curated gene sets (e.g., KEGG/DrugMatrix pathways) that exhibit significant overlap with the gene sets

    – Identify tissues and cell types when the gene sets are over-expressed

    – Provide representative citations

- Using the NextBio curation we can provide a toxicological and mechanistic interpretation that is species and organ/tissue specific

HALLMARK_XENOBIOTIC_METABOLISM_M5934

- **Tissues/Cell types**: exhibits high level of expression in liver, in particular hepatocyte

- **Associated Hallmark sets:** HALLMARK_FATTY_ACID_METABOLISM_M5935

- **Toxicological Interpretation**

  - **Down-regulation:** Indicates de-differentiation of tissues or cells (particularly liver and hepatocytes) that occurs following significant tissue damage and associated regeneration.

  - **Up-regulation:** Indicative of general chemical stress in the liver and other tissues to a lesser extent. The changes in gene expression can be mediated through a number of transcription factors including factors AhR, CAR, PXR,, PPAR alpha or Nrf2. Activation of these signaling pathways have been associated with number of pathological outcomes in rat liver including hypertrophy, hyperplasia, and liver neoplasia. In addition effects on these signaling pathways is often associated with disturbance of lipid homeostasis both in the liver and the blood.

- **Mechanistic Interpretation:** Xenobiotics are foreign small molecules which are introduced into the body from the environment (e.g., inhalation of air pollutants), through the ingestion of food (e.g., fruits, vegetables, food preservatives), and the use of medication. Xenobiotics are eventually excreted via either the biliary or the renal route but may require extensive metabolism in order to be converted to a form that the excretion system can handle. All major routes of exposure (skin, nasal mucosa, lung, liver, kidney, & gastrointestinal tract) have substantial xenobiotic metabolism capabilities. Highly water soluble molecules often enter Phase III metabolism directly and are excreted unchanged in the urine. Xenobiotics that are less water soluble but with hydroxyl, amino, carboxyl, or sulfhydryl functional groups can enter Phase II where they are acylated, sulfated, conjugated with amino acids, or glucuronidated in order to be excreted via either the biliary or the renal route. The majority of the marketed drugs is highly lipophilic and requires activation prior Phase II metabolism. This Activation process is Phase-I xenobiotic metabolism and involves oxidation, reduction, and/or hydrolysis of the a drug molecule in order to introduce hydroxyl and carboxyl groups into the molecule for Phase II metabolism. Many of the enzymes and transporters involved in xenobiotic metabolism are inducible through drug interaction with AhR, PXR, or CAR. The flavin-containing monooxygenases (FMOs) are constitutively expressed and their expression levels appear to be regulated by endogenous steroids. (DrugMatrix)

Supporting Citations (PMID): 15307212; 27760163; 25661872; 26145830; 18187584

Lets Be Provocative!

Gene Set	Percentage	Genes Passing All Filters	P-value	BMDL Median (mg/kg/day)	BMD Median (mg/kg/day)	BMDU Median (mg/kg/day)	Genes Up	Genes Down	Overall Direction
Response to DNA Damage Stimulus	14	26	0.00001	9	17	34	22	4	Up

GO Definition: Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a stimulus indicating damage to its DNA from environmental insults or errors during metabolism.

What does NTP say, if anything?

- Use of the Hallmark gene sets

- Expansion of the Hallmarks

- Curation of the Hallmarks

- Other variables to consider

  - Use existing curated gene sets?

  - Don't attempt toxicological interpretation because of uncertainty?