Gene regulation network-based drug discovery in plexiform neurofibromas

Team: NF terminators

NFhackathon2020

- *Team name*: NF Terminators

- *Project title:* Gene network based drug discovery in plexifiorm neurofibromas

- We take the "NF Data" challenge and do a datamining on plexiform neurofibromas drug screen data and related cell transcriptomes.

- We also want to compete for the GRANT and INCUBATION prize

- *Team member names*: Daochun Sun (*Researcher, Data Scientist*)

    Rebecca Brown (*Health&Pharma, Researcher, Clinician*)

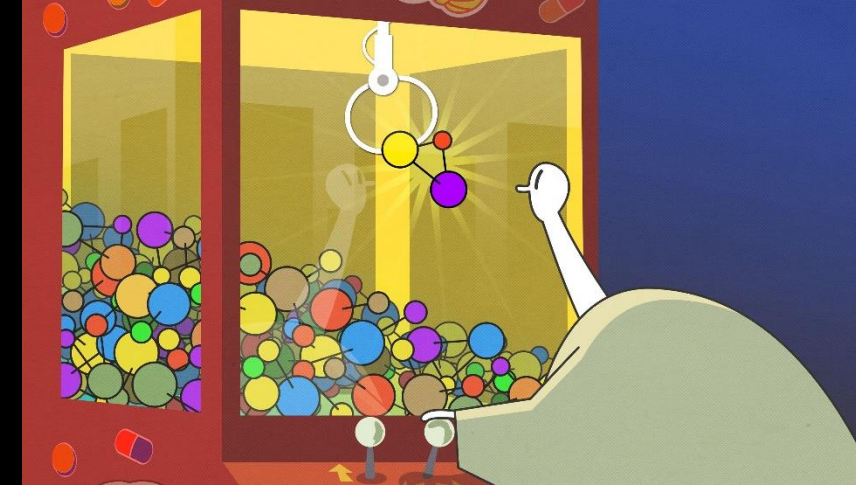   Sameer Farouk Sait (*Clinician, Researcher*)

Before drug screen

After drug screen

What we really want

Challenges for patients, clinicians and researchers in a drug screen:

1. We need a ranked candidate list to prioritize.

2. Candidates in a drug screen have low success rate on real tumors.

3. Does the efficacy has a reliable biology (clear targets or mechanisms)?

4. Can we combine different candidates to enhance the response and reduce the toxicity?

A magic strategy using R
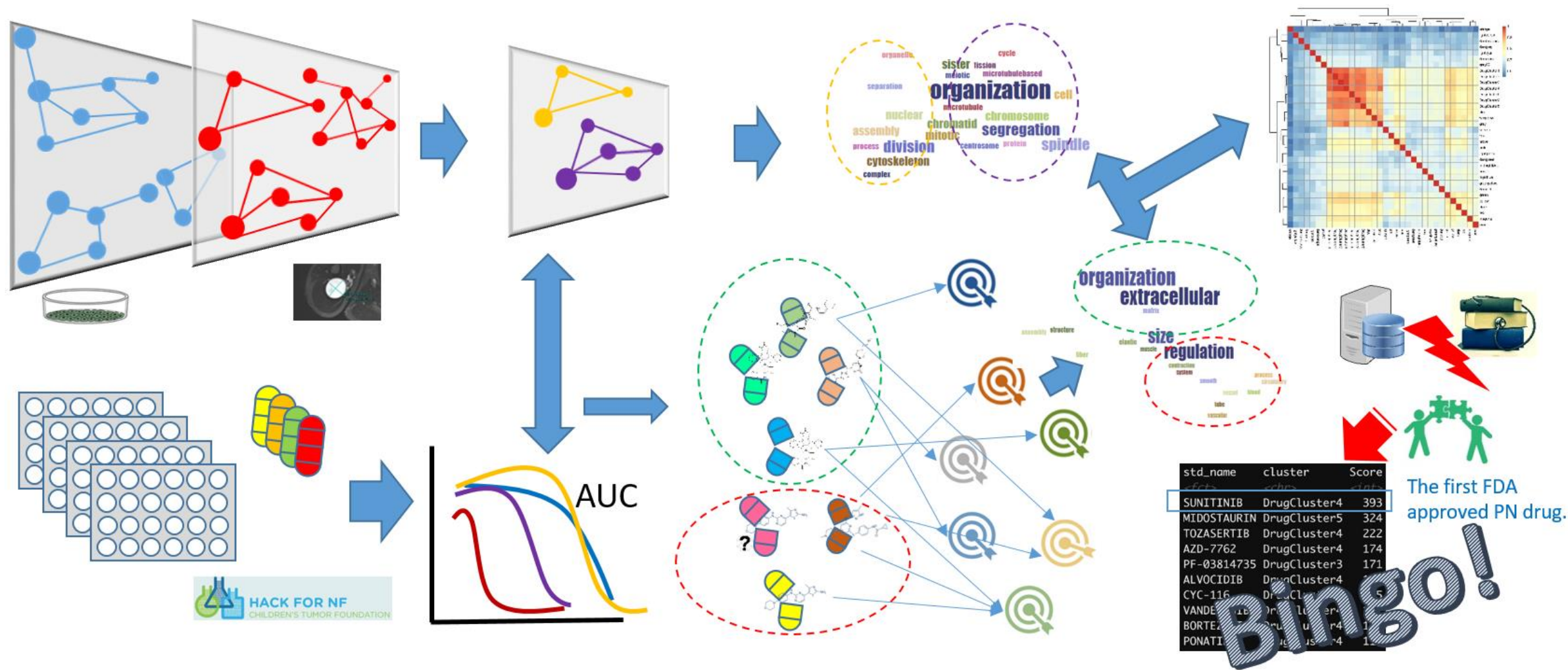
Use preserved gene networks to bridge the gap between models

- Gene expression of PN from GSE41747 study

- Gene expression of immortalized PN cell lines used in drug screen

- WGCNA package in R

PN primary tumors from GEO41747

PN Cells from NTAP

Preserved Eigengene Networks
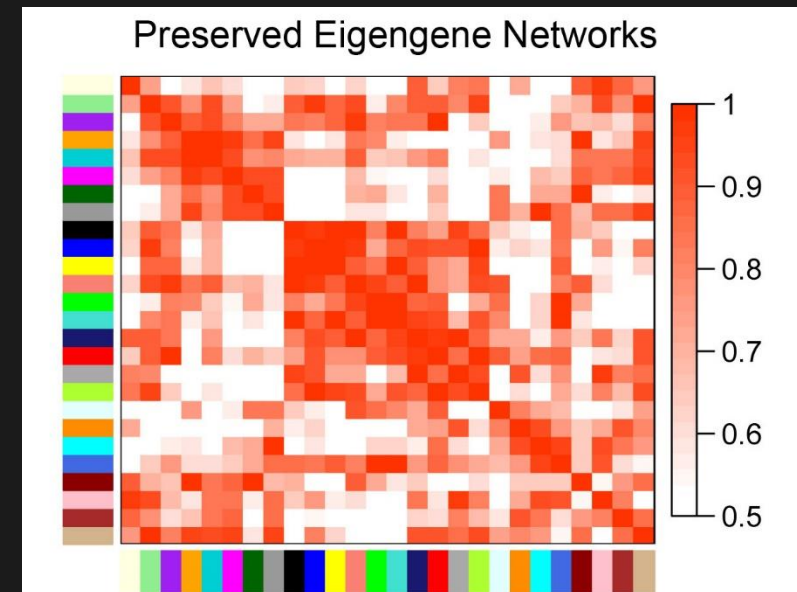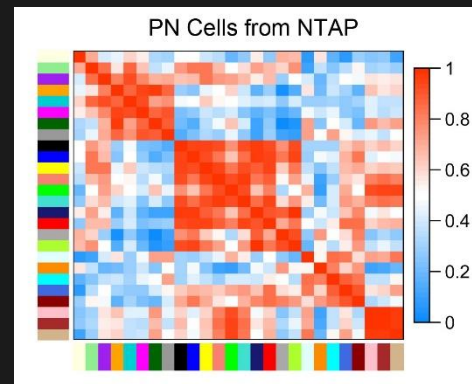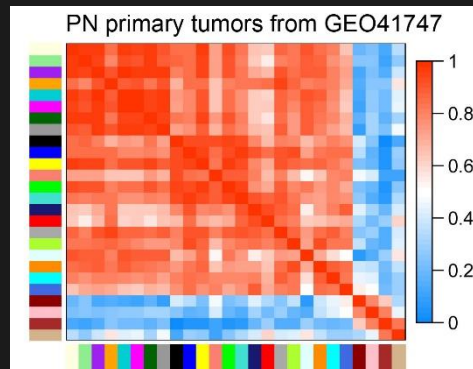
Color blocks represent different gene networks

# Correlate the drug responses to the preserved networks
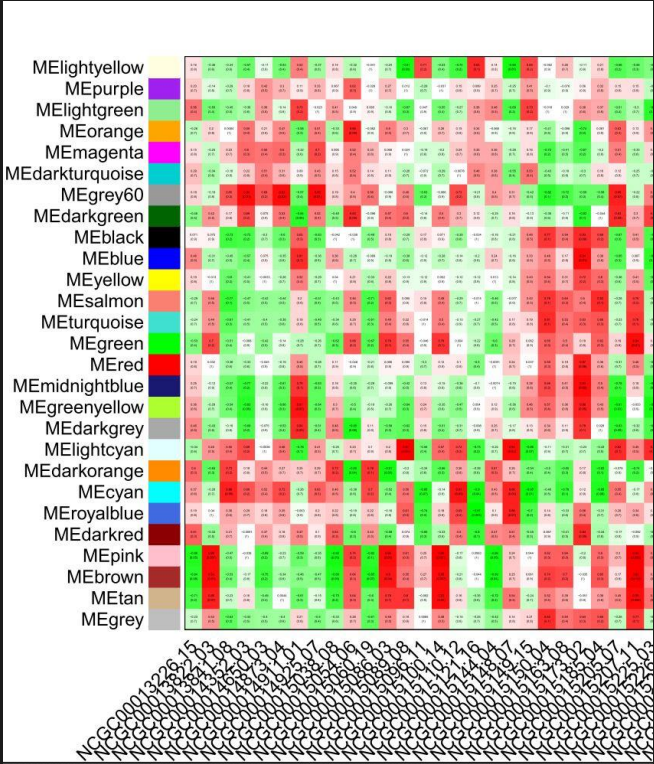
- In the heatmap, red indicates a positive correlation between drug efficacy and networks and green indicates a negative correlation.
- Color blocks on each row represent preserved networks
- Each column is a drug candidate
- This is a partial data without grouping the drugs

- Drugs can be clustered into unique groups based on correlation between their effects and various subcellular networks identified

DrugCluster1   DrugCluster3   DrugCluster5   DrugCluster7

DrugCluster2   DrugCluster4   DrugCluster6

Gene ontology (GO) semantic analysis

- GO terms enriched in each gene networks were calculated (GOSemSim R package)

- According to the drug annotations, the known target genes in any drug cluster can be determined, and GO terms enrichment were calculated

- GO term similarity were determined by semantic analysis respectively in three GO branches: molecular function (MF), biological process (BP) and cell component (CC)

MF

BP

CC
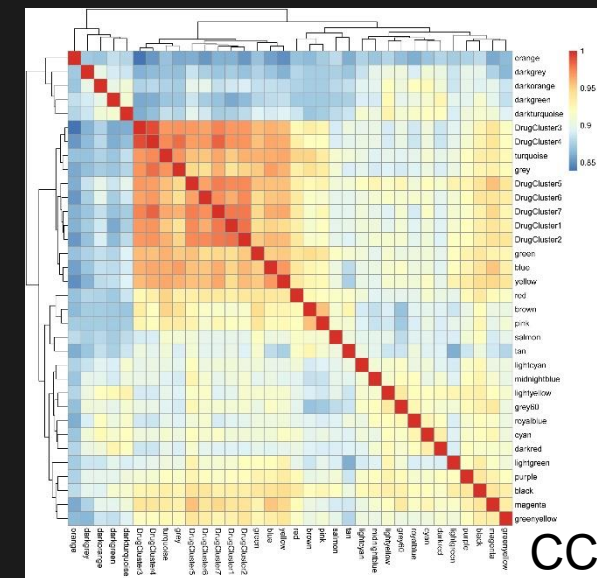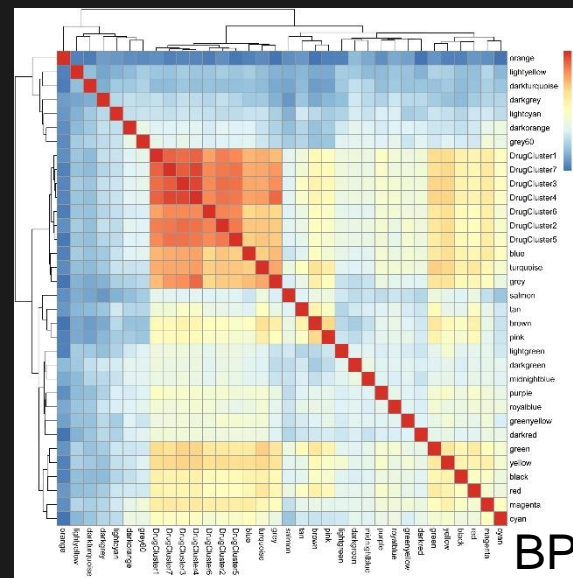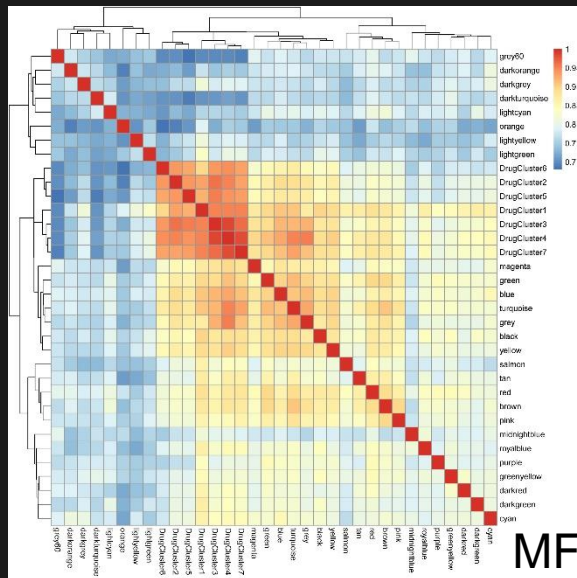
Assumption: High similarities indicate strong biology relevance, suggesting the drugs potentially target the networks

Why our strategy is great?

- Using the preserved networks, we enrich candidates targeting both cells and tumors

- The algorithm weights higher on the "well-annotated" drug candidates

- The gene networks can help researchers to nail down the mechanisms

- The drug clusters can serve as a start point to explore the targets of a novel candidate

- Drugs can be combined according to their clusters

- ……

More we can do given more time…

- Weighted toxicity scores (PMID: 29739789).

- The FDA drug label database to design a combination of candidates

- MPNST drug screen can be integrated for candidates to inhibit both

- ……

References:

- MEK inhibition exhibits efficacy in human and mouse neurofibromatosis tumors (J Clin Invest. 2013 Jan;123(1):340-7)

- Immortalization of Human Normal and NF1 Neurofibroma Schwann Cells (Lab Invest 96(10):1105-15, 2016)

- WGCNA package: https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/Tutorials/

- GOSemSim: https://bioc.ism.ac.jp/packages/release/bioc/html/GOSemSim.html

- R Bioconductor project: https://www.bioconductor.org/

- Synapse NFhackathon2020: https://www.synapse.org/#!Synapse:syn22336443/wiki/605694

- CTF NFhackthon2020: https://www.ctf.org/news/hack-for-nf-2020

- The slack discussions in the general and scientific channels

- The R codes provided by the NFhackathon organizers

Acknowledgements

- The teamwork among the NFhackathon community

- My 5-year-old daughter's strategy to end NF.