Double Trouble: Understanding Sex Differences in Synthetic Lethal interactions in Human Cancers

Project Updates

Alexander Turco

July 18, 2023

UHN Princess Margaret Cancer Centre

Recall: The Combined Inactivation of Two Genes Can Lead to Synthetic Lethal Interactions

- Synthetic lethal interactions describe the relationship between two genes whose coupled inactivation, but not their individual inactivation, causes cell death or reduces cell viability

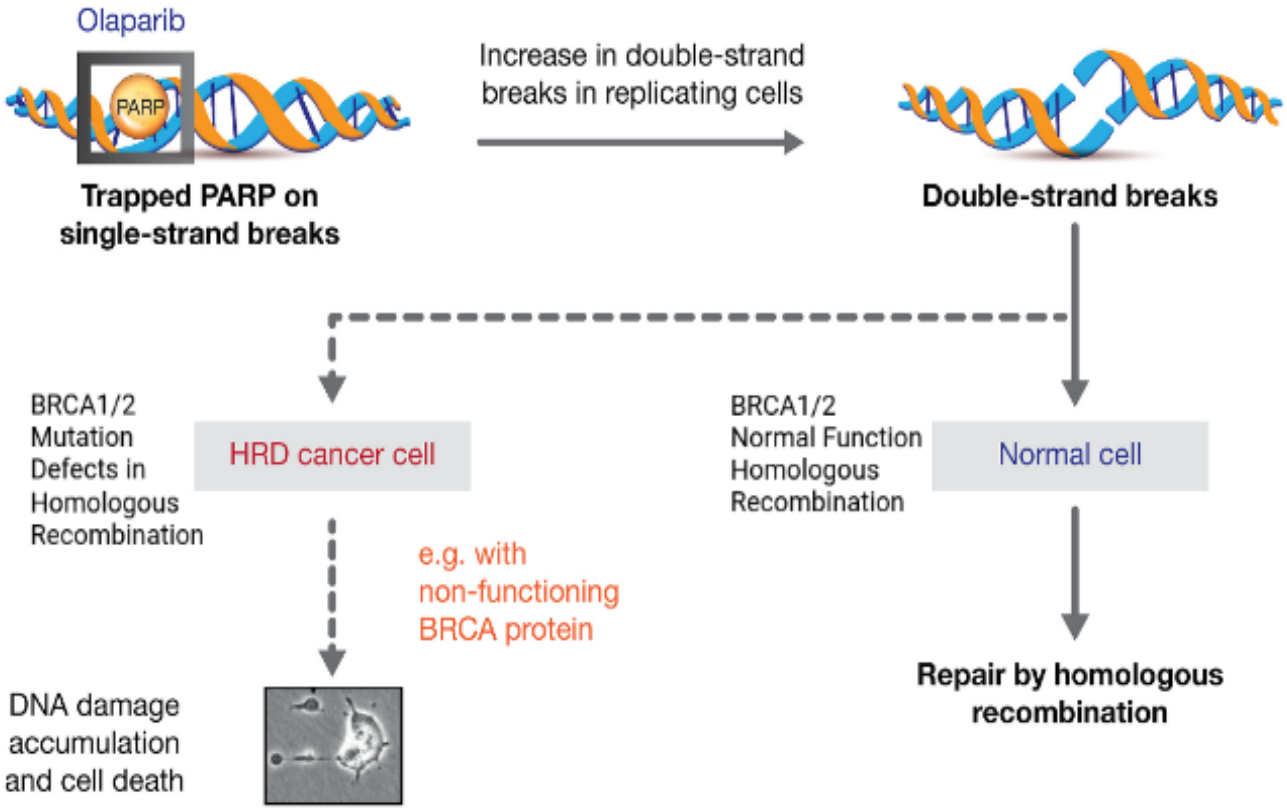
- Inactivation: Preventing or disabling normal function of a gene (e.g mutation)

Synthetic = Combining

[2] Lee et al., 2018

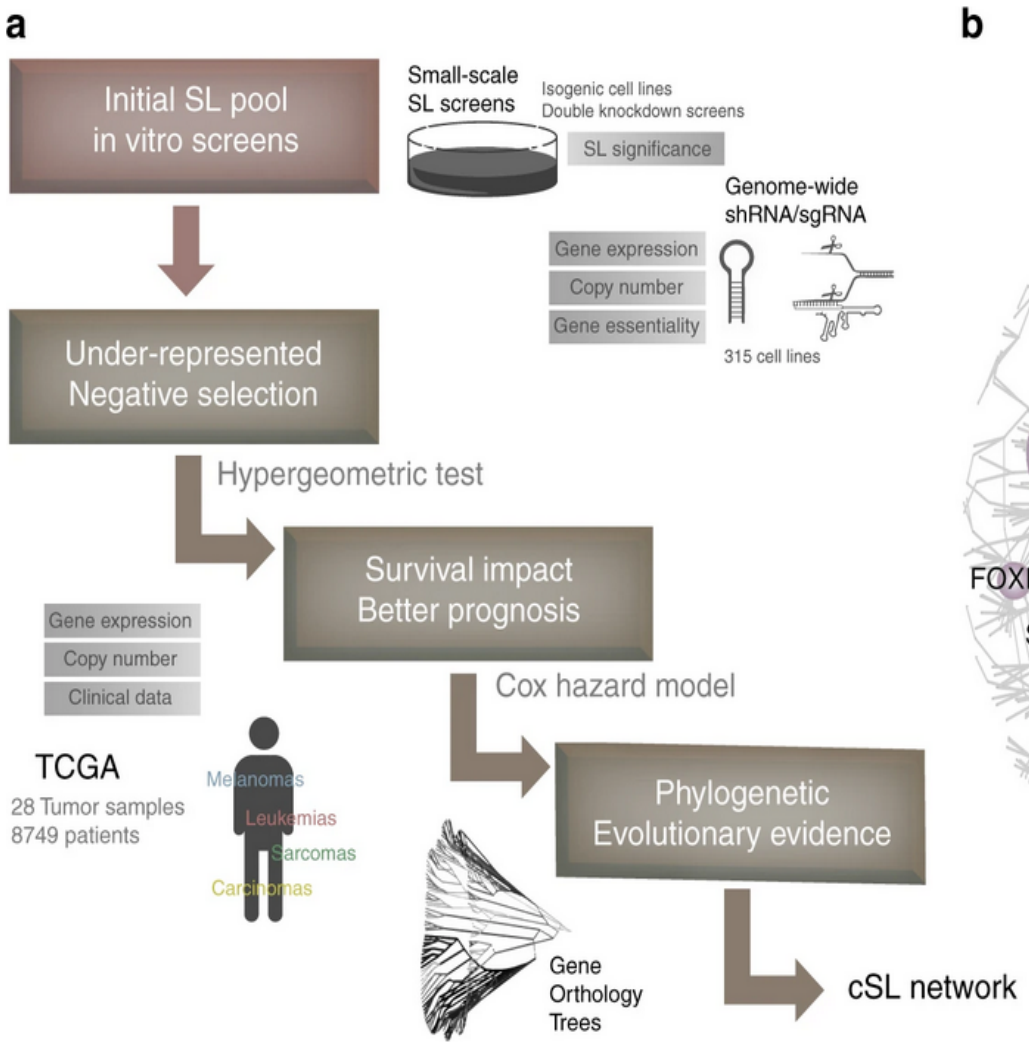
Recall: Synthetic Lethal Interactions are Harnessed for Precision Oncology

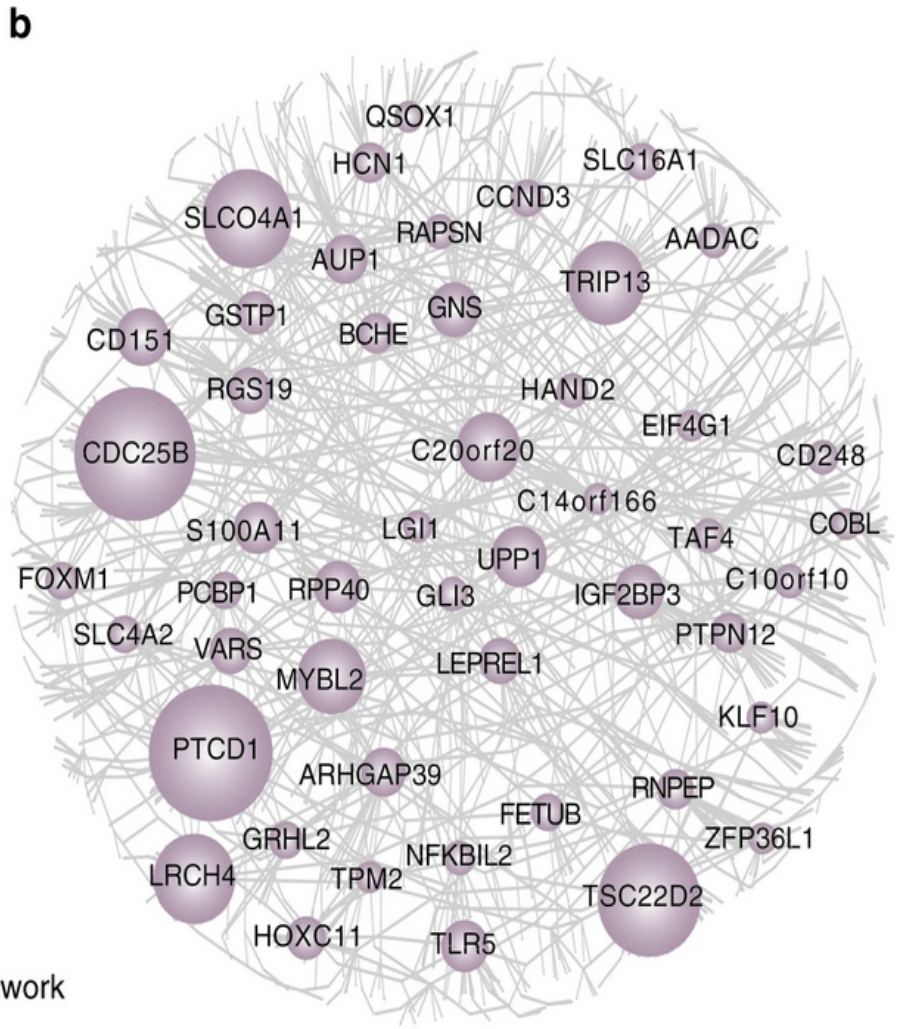
- Four FDA approved anti-cancer drugs are Poly [ADP-ribose] polymerase 1/2 (PARP1/2) inhibitors that work via a synthetic lethal mechanism



[4] Figure From O'Connor, 2015

Recall: Building Pan-Cancer Synthetic Lethality Networks





[2] Lee et al., 2018

Recall: Sex Differences Add an Additional Layer of Complexity

Human sex differences are mainly caused by;

1. Gonadal hormone secretions
2. Genes located on the sex chromosomes (X and Y)

This leads to differences in the frequency of certain cancer types and the efficacy of treatments in males and females

The Objective

Can we build sex-specific synthetic lethality networks for various cancer types?

More specifically, we are trying to elucidate the differences in synthetic lethal interactions between males and females using a network based approach.

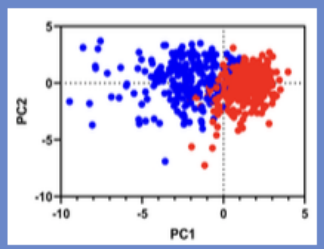
Overall Project Workflow

DATA COLLECTION

	Sample 1	Sample 2
Gene 1	6	3
Gene 2	1438	739
Gene 3	2361	1852
Gene 4	400	951
Gene 5	299	142

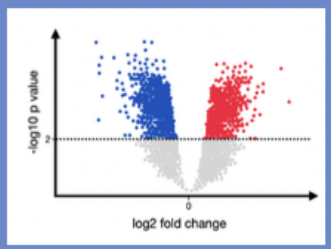
Obtain RNA-seq data from TCGA in the form of raw counts. (Genes are features, samples are data points/variables)

DATA PROCESSING



Normalize and transform raw RNA-seq data, identify sources of variation, batch effects. Do samples cluster according to biological conditions?

DIFFERENTIAL EXPRESSION



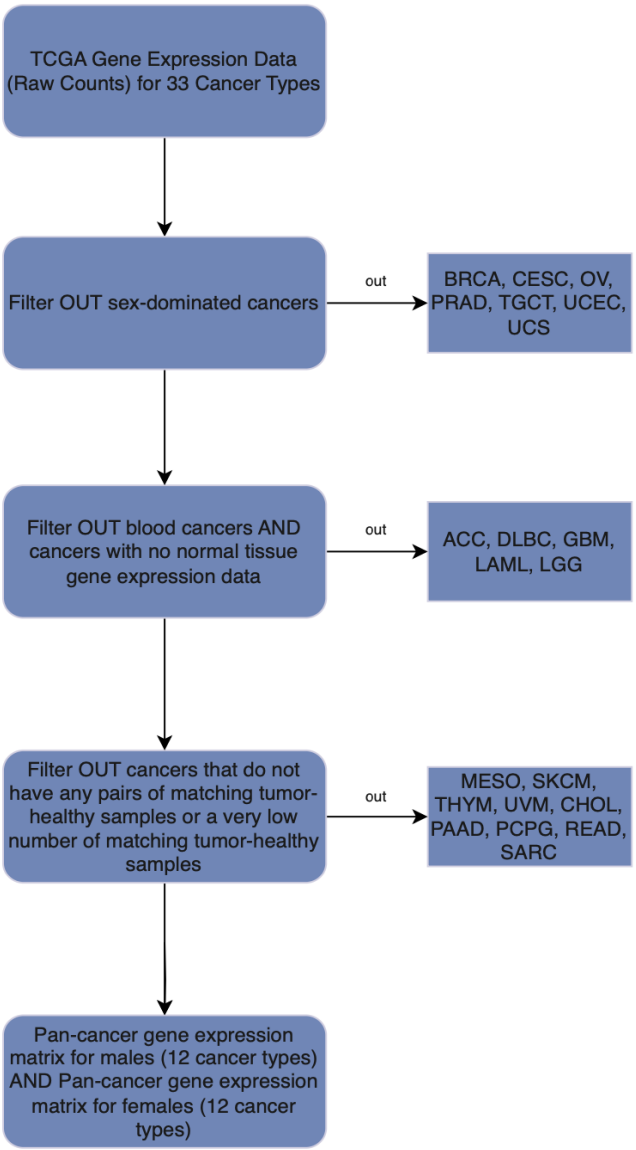
Perform differential gene expression analysis to identify genes that are differentially expressed in tumor tissue.

SYNTHETIC LETHALITY

depmap | Cancer Dependency Map

Find potential candidate SL pairs for differentially expressed genes using CRISPr gene essentiality data from DepMap.

Selection of TCGA Cancer Types



- There is a lack of healthy tissue RNA-seq samples in the TCGA database (12 cancers with more than 10 pairs in M and F)
- Raw count pan-cancer matrices created with 12 remaining TCGA cancer types (BLCA, COAD, ESCA, HNSC, KICH, KIRC, KIRP, LIHC, LUAD, LUSC, STAD, THCA)

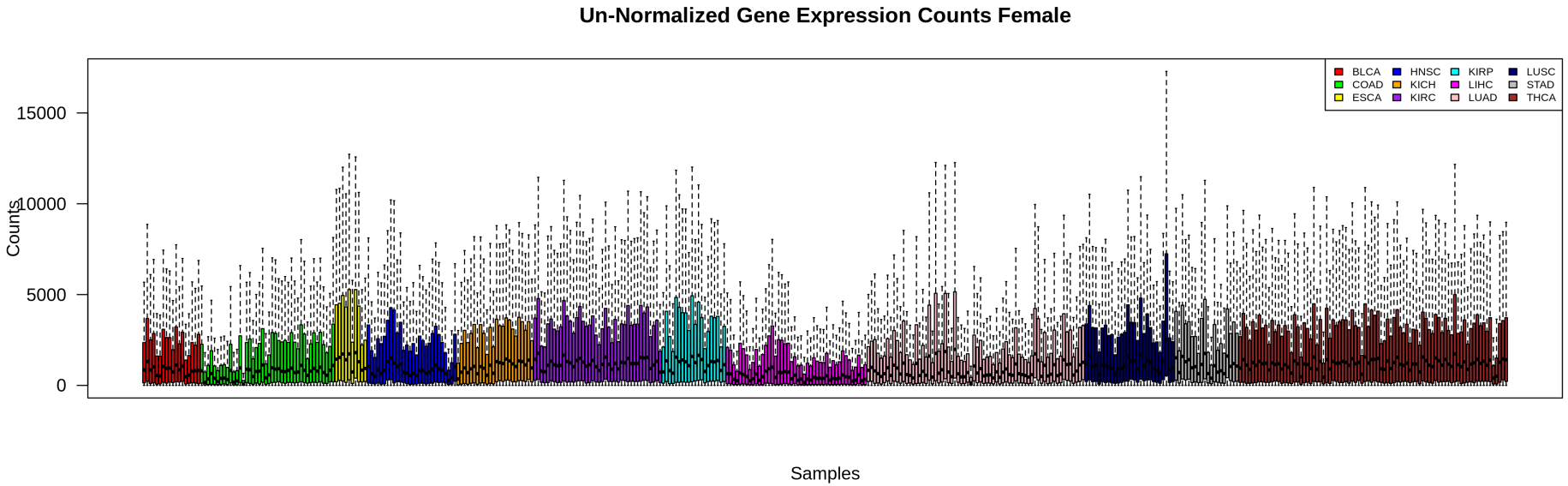
Creating and Pre-Filtering Pan-Cancer Expression Matrices

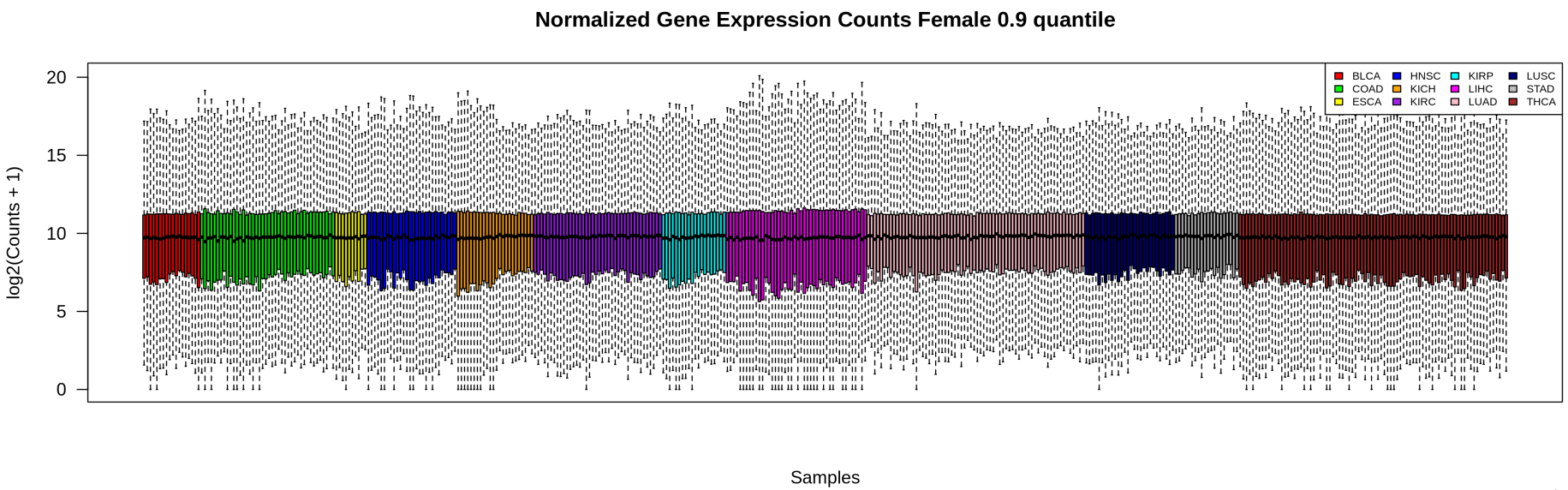
- For each gene, if expression is $>$ 90th quantile of overall expression in AT LEAST 1 sample, we keep it, otherwise we filter it out
- We reduce the amount of genes with extremely low expression, thereby reducing noise and improving sensitivity to detect differentially expressed genes (genes with weak expression are more susceptible to technical noise arising from library size, library composition, etc)

Normalization for RNA-seq Data Analysis

Required to identify genes that are differentially expressed due to some biological phenomena and not due to technical variation

Normalizing Pan-Cancer Expression Matrices





content...