

Reading

The KSQ: Using available scRNA-seq data from cancer cell lines, how would you explore the use of the following FDA-approved antibody therapies in additional cancers?

- **Trastuzumab:** Targets HER2 and is used in the treatment of HER2-positive breast and gastric cancers.
- **Bevacizumab:** Targets VEGF and is used for a variety of cancers, including colorectal, lung, glioblastoma, breast, liver, and kidney cancer.

First, I want to clarify that trastuzumab targets HER2 at the protein level, not just the DNA level. If only cell line data is available, I would begin by analyzing scRNA-seq data from cancer cell lines to assess the correlation between DNA, RNA, and protein expression. For trastuzumab, I would focus on comparing the RNA expression profiles of HER2-positive breast and gastric cancer cell lines to other cancer types to identify potential similarities that could suggest new therapeutic uses.

By understanding the relationship between RNA expression and protein structure, I could confirm whether similar RNA expression in other cancers reflects consistent HER2 protein formation, which would justify further exploration of trastuzumab's efficacy in those cancers.

In addition, I would explore how the expression of HER2 correlates with the tumor microenvironment (proteomic, metabolomics etc. if data is available) and investigate co-expression patterns with other genes to find potential targets for combination therapies.

What are cancer cell lines? Why do we use cancer cell lines?

Cancer cell lines are cultures of cells derived from cancerous tissue samples that are used to represent different types of cancer in the laboratory. They are one of the most commonly used models in cancer research, helping scientists study cancer biology, validate therapeutic targets, and assess drug efficacy.

What is scRNA-seq? Why do we use scRNA-seq data in cancer drug development?

Traditional "bulk" RNA sequencing averages gene expression across potentially millions of cells, masking the differences between individual cells. In contrast, single-cell RNA sequencing (scRNA-seq) analyzes the gene expression of individual cells. After isolating a cell from tissue, its RNA is converted into complementary DNA (cDNA), which is then sequenced to create a gene expression profile for that specific cell. This data allows for a detailed analysis of gene activity in each cell, classification of cell types, identification of rare cell populations, and an understanding of the cellular composition within the tissue.

In a lecture by Dr. Aviv Regev, she highlighted how scRNA-seq revealed that "cold" regions of tumors, which lack T-cell activity, correspond to resistance to immunotherapy in melanoma. This demonstrated how scRNA-seq can be used to map the tumor microenvironment, providing a framework for understanding how tumors grow and respond to treatments, ultimately guiding the development of more targeted cancer therapies.

An analogy would be comparing finding tumor target in a combat situation. We want to target specific group and people without killing civilians or destroying historical sites (aka: killing cancer cells without damaging healthy cells). Most of times, we know the general cell to target (aka: a city). We also know the coordinate (aka the DNA). In order to efficiently target it, we also need to understand the landscape. Is it a skyscraper or multiple houses? Are there defense nearby we should be cautious about? Are there helpers nearby to assist our attack? Knowing RNA expression in addition to DNA expression is like having a 2D map for a better ideas, that is more than the coordinates.

Summary of Cancer Cell Line Encyclopedia

Phase I of CCLE

Goal

1. Characterize genetic and pharmacologic of human cancer models
2. Develop computational method to analyze the link between pharmacologic characteristics and genetic expression and cell lineage patterns
3. Translate cell line genomics by categorizing them according to patient stratification

Material - 1000 cell lines from below cell line repositories

1. ATCC (American Type Culture Collection)
2. DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen)
3. The KCLB (Korean Cell Line Bank)

Method

1. DNA and RNA extraction
2. Generate data
 - a. Affymetrix SNP 6.0 data
 - b. Affymetrix U133 2.0+ expression array data
 - c. Point mutation profiles using a SNP genotyping platform called OncoMap 3.0.
 - d. Hybrid capture exon sequencing of >1600 known
 - e. Putative cancer genes
3. Pharmacologic testing

Results

1. [The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity](#)
 - a. Nature 2012, Mar 28;483(7391):603-7. DOI: 10.1038/nature11003

2. [Pharmacogenomic Agreement Between Two Cancer Cell Line Data Sets](#)
 - a. Nature 2015, Dec 3;528(7580):84-7. DOI:10.1038/nature15736

Phase II of CCLE

Goal

1. Characterize genetic and pharmacologic of human cancer models via NGS
2. Develop computational method to analyze the link between pharmacologic characteristics and genetic expression via NGS and cell lineage patterns
3. Translate cell line genomics via NGS by categorizing them according to patient stratification

Material - 1000 cell lines from below cell line repositories

1. ATCC (American Type Culture Collection)
2. DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen)
3. The KCLB (Korean Cell Line Bank)

Method

1. RNA-seq
2. Exome sequencing
3. miRNA content of all cell lines
4. Quantifying the metabolite abundance of 225 metabolites across CCLE
5. Mass Reaction Monitoring (MRM) mass spec quantification of bulk Histone H3 tail modifications
6. Reverse phase protein array analysis with MD Anderson

Results (Link cancer metabolome to genetic alterations, epigenetic features and gene dependencies)

1. [Global chromatin profiling reveals NSD2 mutations in pediatric acute lymphoblastic leukemia](#)
 - a. Nat Genet. 2013 Nov;45(11):1386-91. DOI: 10.1038/ng.2777
2. [Next-generation characterization of the Cancer Cell Line Encyclopedia](#)
 - a. Nature. 2019 May;569(7757):503-508. DOI: 10.1038/s41586-019-1186-3. Epub 2019 May 8.
3. [The landscape of cancer cell line metabolism](#)
 - a. Nat Med. 2019 May;25(5):850-860 DOI: 10.1038/s41591-019-0404-8. Epub 2019 May 8.

Phase III of CCLE

Goal

1. Characterize genetic and pharmacologic of human cancer models via NGS
2. Develop computational method to analyze the link between pharmacologic characteristics and genetic expression via NGS and cell lineage patterns
3. Translate cell line genomics via NGS by categorizing them according to patient stratification
4. Understand the protein content of cell lines

Material - 1000 cell lines from below cell line repositories

1. ATCC (American Type Culture Collection)
2. DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen)
3. The KCLB (Korean Cell Line Bank)

Method

1. Interrupt or alternating protein function with
 - a. Antibody-drug conjugates
 - b. Antibody mediated cellular cytotoxicity (ADCC)
 - c. CART-T cells
2. Quantify the abundance of proteins by performing Tandem-mass tagging mass spectrometry
 - a. serine/threonine phosphorylation events were quantified by cxxxxx
 - b. tyrosine phosphorylation was quantified in a small set of cell lines under conditions, by Carr Mass Spectrometry platform at the Broad Institute

Results

1. [Quantitative Proteomics of the Cancer Cell Line Encyclopedia](#)
 - a. Cell. 2020;180(2):387-402.e16. DOI:10.1016/j.cell.2019.12.023
2. [Video Lecture: Cell atlases as roadmaps to understand and treat cancer by Dr. Aviv Regev](#)

Bonus questions

How is HER2 positive protein being targeted with tumor heterogeneity?

Could you clarify how the mechanisms of action of these two antibodies are different?