20.109 RNAseq Ex4: CCLE data

Amanda Kedaigle and Ernest Fraenkel

1/6/2019

Review your work so far

In our original dataset for this project, we looked at the gene expression of the DLD1 colorectal cancer cell line with and without a mutation in the BRCA2, with and without DNA damage due to etoposide treatment. In the last lab, we looked at gene expression in a different cell line with and without etoposide treatment. What have we learned so far about etoposide? How did the BRCA2 mutation affect our results?

Explore the Cancer Cell Line Encyclopedia webtool

In this lab, we'll be using data from the Cancer Cell Line Encyclopedia Consortium (CCLE). You can read more about this huge effort to make public cell line data available to reseachers in these papers (link1, link2). In short, researchers have made available gene expression data from hundreds of cell lines, and then data about the survival of those same cell lines after treatment with several drugs or after CRISPR knockdowns of various genes. This trove of data allows researchers like us to look at drug sensitivity and its relation to gene expression over hundreds of cell lines without having to do hundreds of experiments.

We've downloaded some gene expression data from the CCLE into the rstudio.cloud Exercise #4 space for you. But first, check out the CCLE data at https://depmap.org/portal/. You can look up our cell line, DLD1, in the search bar on that page. You can see that they have several datasets available for that cell line, including several CRISPR screens and omics data such as gene expression and protein data.

To compare our cell line to others, click on Tools > Data Explorer on the top right of the website. This will bring up a chart. Using the drop down menus on the left, you can chart whatever relationship you're interested in with the CCLE data. For example, we might want to plot the Drug Sensitivity of the cell lines to etoposide versus the Gene Dependency of the cell lines on the BRCA2 genes. Create this chart on the Data Explorer and expand the Statistics tab on the left to show the regression lines and see the p value of the regression between these variables. You can try different datasets and see how this changes the data (for example, there are several different datasets for drug sensitivity, the CTD dataset or the Sanger dataset, which may contain different subsets of cell lines and different data).

Make sure you understand the axes on this graph. The AUC score for Drug Sensitivity is the area under the curve of "drug concentration" vs "cell viability". In other words, a low AUC indicates a high response (cell death) of the cell line to the drug. Meanwhile, another name for Dependency score is "gene knockdown viability effect." A high Dependency score for Gene Dependency indicates cell growth after knockdown of the gene by CRISPR or by RNAi, whereas a negative score indicates cell death after knockdown. What are these charts telling us about the relationship between BRCA2 knockdown and etoposide dependency in different cell lines?

For the rest of this lab, you'll do some data exploration on CCLE data and decide on a different gene that you think might have more of an impact on etoposide sensitivity. You can use this information to propose future experiments for this class!

Use your RNA-seq analysis skills on CCLE data

In rstudio.cloud, we have uploaded gene expression data from the CCLE for you. Using the script.R file in that space, write code to analyze differential gene expression between cell lines that are sensitive to etoposide

versus those that aren't. Genes with large changes in expression between these cell lines are candidates for genes that affect etoposide sensitivity.

Using the top most differential genes from that analysis, make more charts on the CCLE Data Explorer website. Find a gene that you think is a good candidate for future experiments. The gene(s) you pick might be different from your classmates' - that's fine! As long as you have a good justification for your proposal to run experiments on that gene. In your final report, you'll summarize what you've learned in these labs about the mechanism and affect of etoposide treatment, and propose future experiments using the gene(s) you identified to further explore this important cancer drug.