

LINCS Canvas Browser v1.0 – User Manual



Written by

Qiaonan Duan and Avi Ma'ayan



Last updated: 4/19/2014

Abstract

For the Library of Integrated Network-based Cellular Signatures (LINCS) project many gene expression signatures using the L1000 technology have been produced. The L1000 technology is a cost-effective method to profile gene expression in large scale. LINCS Canvas Browser (LCB) is an interactive HTML5 web-based software application that facilitates querying, browsing and interrogating many of the currently available LINCS L1000 data. LCB implements two compacted layered canvases, one to visualize clustered L1000 expression data, and the other to display enrichment analysis results using 30 different gene set libraries. Clicking on an experimental condition highlights gene-sets enriched for the differentially expressed genes from the selected experiment. A search interface allows users to input gene lists and query them against over 100,000 conditions to find the top matching experiments. The tool integrates many resources for an unprecedented potential for new discoveries in systems biology and systems pharmacology. The LCB application is available at <http://www.maayanlab.net/LINCS/LCB>.

Installation and Requirements

LCB requires a working internet connection and modern HTML5 compatible web browser capable of displaying and operating dynamic SVG images. Current versions of the browsers FireFox, Opera, Chrome, Safari, and Internet Explorer should work. LCB should also work on mobile devices and respond to touch gestures. LCB is available on various web-sites in different versions. The currently most updated primary version is available at:

<http://www.maayanlab.net/LINCS/LCB>

The other two versions are available at:

<http://apps.lincsccloud.org/LCBL/>

<http://lincs.hms.harvard.edu/explore /canvasbrowser/>

Each version provides almost the same functionality but it is slightly customized. LCB on the HMS site only displays the LJP datasets and has functionality that shows other layers of regulation such cell viability assays. The HMS version does not have a search engine.

The version available on the LINCS apps cloud and the maayanlab.net sites covers the LJP and CPC batches and has the search engine functionality. The version on the maayanlab.net web-site maybe more updated and is used for beta-testing and for creating the tutorials available within this manual.

Instructions

I. Querying for consensus experiments

First, navigate to the **Search** page of LCB at <http://www.maayanlab.net/LINCS/LCB> and shown below:

LINCS Canvas Browser

[Search](#) | [Browse](#)

The screenshot displays the LINCS Canvas Browser interface. On the left, under the heading "Gene Lists", there are two columns: "Up List" and "Down List". The "Up List" contains the following genes: SYNGR3, CDH3, NCK1, TKTL1, NUSAP1, BAG3, NR2F6, NCK1, PCNA, TUBD1, KIF5C, RAB27A, ETV1, APPBP2, TSPAN4, and RAB27A. The "Down List" contains: MTF2, SYNCRIP, SLPI, MRPL19, COX17, KIAA0494, C7ORF42, ANKS1A, PSMB4, AREG, ELOVL6, CSTF2T, CYR61, UGT2B28, and SEC63. Below these lists are "clear" buttons for each. Underneath the lists are checkboxes for "Up" (checked) and "Down", and buttons for "Search", "Example", and "Enrich". At the bottom, there are "Aggravate" and "Reverse" options with checkboxes. On the right, under the heading "Top 50 Consensus Experiments", there are dropdown menus for "Overlap" and "Info (Perturbation, Dose, Time, Cell, Batch)". Below these is a message: "No data available in table". At the bottom of this section, it says "Showing 0 to 0 of 0 entries" and has "Previous" and "Next" navigation buttons.

Step 1: Fill in up/down gene lists in the text boxes by clicking on the example button or providing lists from your own experiments.

Gene Lists

Up List

- SYNGR3
- CDH3
- NCK1
- TKTL1
- NUSAP1
- BAG3
- NR2F6
- NCK1
- PCNA
- TUBD1
- KIF5C
- RAB27A
- ETV1
- APPBP2
- TSPAN4

clear

Down List

- MTF2
- SYNCRIP
- SLPI
- MRPL19
- COX17
- KIAA0494
- C7ORF42
- ANKS1A
- PSMB4
- AREG
- ELOVL6
- CSTF2T
- CYR61
- UGT2B28
- SEC63

clear

Up Down

Search Example Enrichr

Aggravate Reverse

Top 50 Consensus Experiments

Overlap ▾
Info (Perturbation, Dose, Time, Cell, Batch) ▾

No data available in table

Showing 0 to 0 of 0 entries

◀ Previous Next ▶

Step 2: Select the search options using the up and down checkboxes. By checking up only, LCB would only query for matching experiments that have a similar set of differentially expressed up gene lists in the database. By checking the down box, LCB would only match experiments that have a similar down gene list. Checking both boxes would query for experiments that match both a similar up and down gene lists. The Search and Enrichr buttons are only enabled after making the up/down selections.

Gene Lists

Up List

- SYNGR3
- CDH3
- NCK1
- TKTL1
- NUSAP1
- BAG3
- NR2F6
- NCK1
- PCNA
- TUBD1
- KIF5C
- RAB27A
- ETV1
- APPBP2
- TSPAN4

clear

Down List

- MTF2
- SYNCRIP
- SLPI
- MRPL19
- COX17
- KIAA0494
- C7ORF42
- ANKS1A
- PSMB4
- AREG
- ELOVL6
- CSTF2T
- CYR61
- UGT2B28
- SEC63

clear

Up Down

Search Example Enrichr

Aggravate Reverse

Top 50 Consensus Experiments

Overlap
Info (Perturbation, Dose, Time, Cell, Batch)

No data available in table

Showing 0 to 0 of 0 entries

◀ Previous Next ▶

Step 3: Choosing either aggravate or reverse search. If reverse search is chosen, LCB would query for experiments with up gene-list similar to the input down gene-list and the down gene-list is similar to input up gene-list. This option may be used to search perturbations that would reverse a disease condition. The aggravate option will simply do the opposite, matching the up with the up and the down with the down.

Gene Lists

Up List

- SYNGR3
- CDH3
- NCK1
- TKTL1
- NUSAP1
- BAG3
- NR2F6
- NCK1
- PCNA
- TUBD1
- KIF5C
- RAB27A
- ETV1
- APPBP2
- TSPAN4

Down List

- MTF2
- SYNCRIP
- SLPI
- MRPL19
- COX17
- KIAA0494
- C7ORF42
- ANKS1A
- PSMB4
- AREG
- ELOVL6
- CSTF2T
- CYR61
- UGT2B28
- SEC63

Up Down

Aggravate Reverse

Top 50 Consensus Experiments

Overlap ▾ Info (Perturbation, Dose, Time, Cell, Batch) ▾

No data available in table

Showing 0 to 0 of 0 entries

Step 4: Clicking the search button will initiate querying the database. It will usually take less than 20 seconds to complete a single search.

Gene Lists

Up List

- SYNGR3
- CDH3
- NCK1
- TKTL1
- NUSAP1
- BAG3
- NR2F6
- NCK1
- PCNA
- TUBD1
- KIF5C
- RAB27A
- ETV1
- APPBP2
- TSPAN4

clear

Down List

- MTF2
- SYNCRIP
- SLPI
- MRPL19
- COX17
- KIAA0494
- C7ORF42
- ANKS1A
- PSMB4
- AREG
- ELOVL6
- CSTF2T
- CYR61
- UGT2B28
- SEC63

clear

Up Down

Aggravate Reverse

Top 50 Consensus Experiments (Up/Aggravate)

Overlap ▼	Info (Perturbation, Dose, Time, Cell, Batch) ⌵
0.9762	SCF.300ng/ul.1h.BT20.LJP004
0.2200	AG14361.25um.6h.A673.CPC006
0.1600	temsirolimus [CI779].10um.6h.SW480.CPC006
0.1600	Picotamide monohydrate.10um.6h.HA1E.CPC001
0.1600	LY-231514.88.8um.6h.NCIH508.CPC006
0.1600	HKI-272.10um.6h.SNUC4.CPC006
0.1600	DIHYDROERGOTAMINE MESYLATE.10um.6h.HA1E.CPC004
0.1600	C3393M50.10um.6h.HEPG2.CPC013
0.1400	tadalafil.10um.6h.A375.CPC011
0.1400	synephrine.10um.24h.MCF7.CPC010

Showing 1 to 10 of 49 entries ◀ Previous Next ▶

Step 5: Table of matching experiments. After the search is complete, the top 50 matching experiments/conditions are listed in a table as shown above. The table has 5 pages with each page displaying 10 entries. Use the next and previous buttons at bottom to navigate through these pages. The first column of the table shows a similarity score with 1 representing a perfect match and 0 representing no match. The second column of the table shows descriptive information about each experimental condition. The information provided includes perturbation, dose, time point, cell-line and batch ID.

Gene Lists

Up List

- SYNGR3
- CDH3
- NCK1
- TKTL1
- NUSAP1
- BAG3
- NR2F6
- NCK1
- PCNA
- TUBD1
- KIF5C
- RAB27A
- ETV1
- APPBP2
- TSPAN4

clear

Down List

- MTF2
- SYNCRIP
- SLPI
- MRPL19
- COX17
- KIAA0494
- C7ORF42
- ANKS1A
- PSMB4
- AREG
- ELOVL6
- CSTF2T
- CYR61
- UGT2B28
- SEC63

clear

Up Down

Search Example Enrichr

Aggravate Reverse

Top 50 Consensus Experiments (Up/Aggravate)

Overlap	Info (Perturbation, Dose, Time, Cell, Batch)
0.9762	SCF_300ng/ul_1h_BT20_LJP004
0.2200	AG14361_25um.6h.A673.CPC006
0.1600	temsirolimus [CI779]_10um.6h.SW480.CPC006
0.1600	Picotamide monohydrate_10um.6h.HA1E.CPC001
0.1600	LY-231514_88.8um.6h.NCIH508.CPC006
0.1600	HKI-272_10um.6h.SNUC4.CPC006
0.1600	DIHYDROERGOTAMINE MESYLATE_10um.6h.HA1E.CPC004
0.1600	C3393M50_10um.6h.HEPG2.CPC013
0.1400	tadalafil_10um.6h.A375.CPC011
0.1400	synephrine_10um.24h.MCF7.CPC010

Showing 1 to 10 of 49 entries

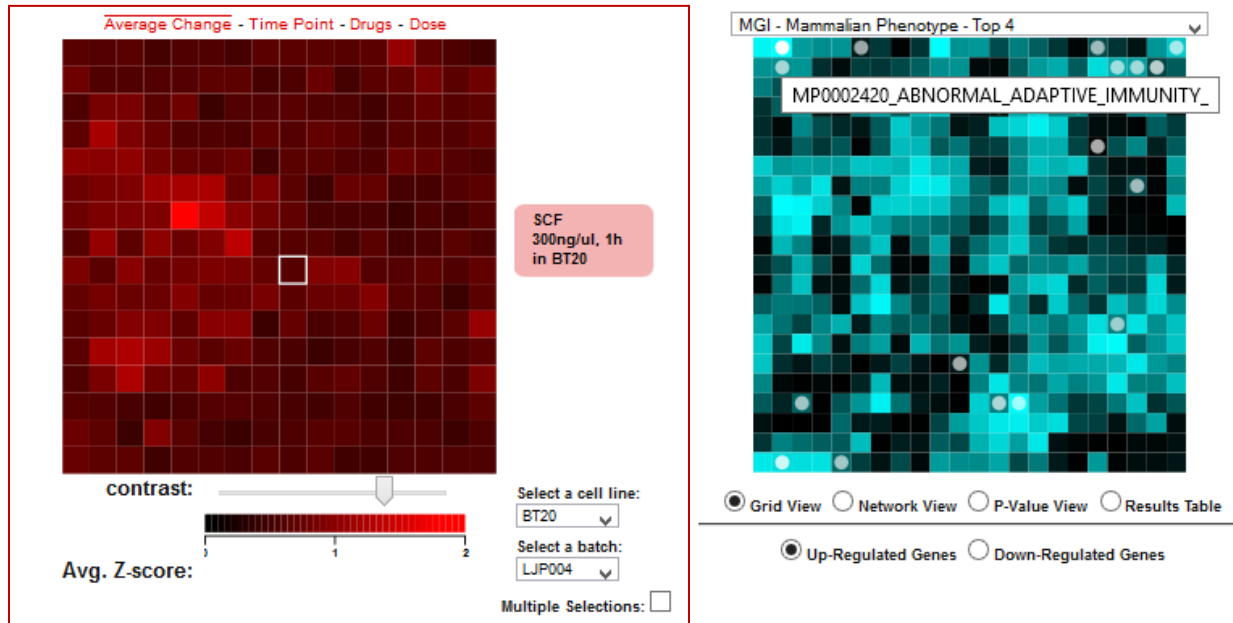
◀ Previous Next ▶

Step 6: Clicking on an entry in the table will visualize the experiment and its enrichment analysis results on canvases in a new tab.

LINCS Canvas Browser

[Search](#) | [Browse](#) | [Help](#)

Search:

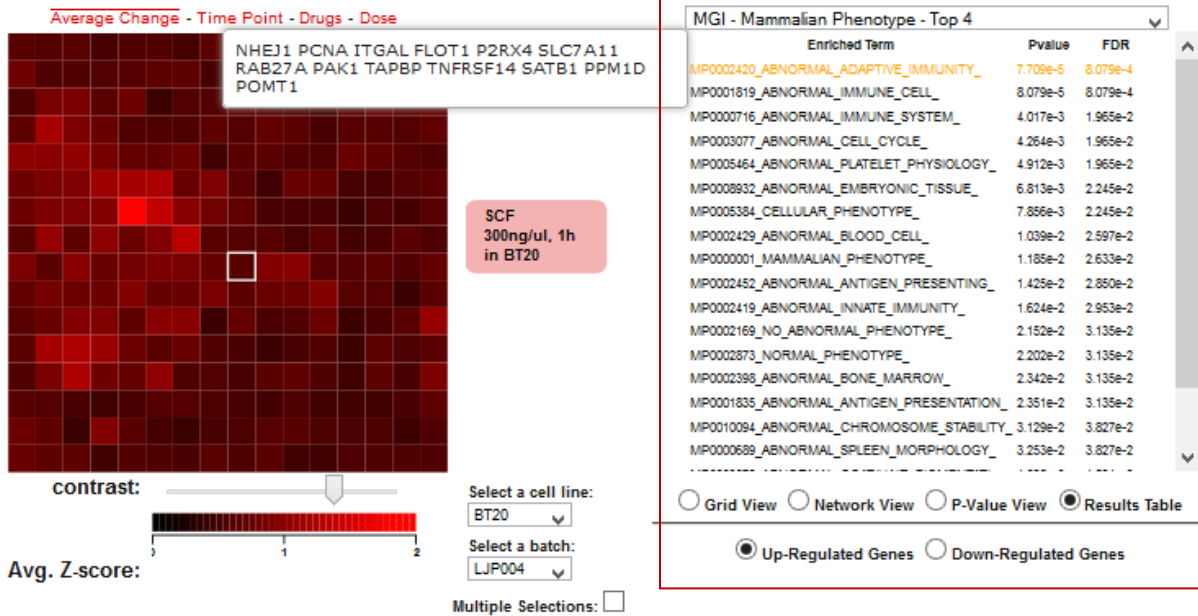


Step 7: Experiments displayed on a canvas. The red canvas on the left visualizes all the experiments applied on one cell-line in one batch. Each experiment is represented by a tile on the canvas. The canvas in the figure shows all the experiments applied on the BT20 cell-line in the LJP004 batch. The tiles are arranged by gene expression similarity so that neighboring tiles have similar gene expression profiles. The coloring on the canvas indicates signal strength. A bright red tile represents a strong perturbation while a darker tile a weak perturbation. Different coloring schemes that represent other aspects of the data could be chosen by click on the navigation bar above the canvas. The experiment that was clicked on in the search results page is highlighted by a white square.

LINCS Canvas Browser

[Search](#) | [Browse](#) | [Help](#)

Search:

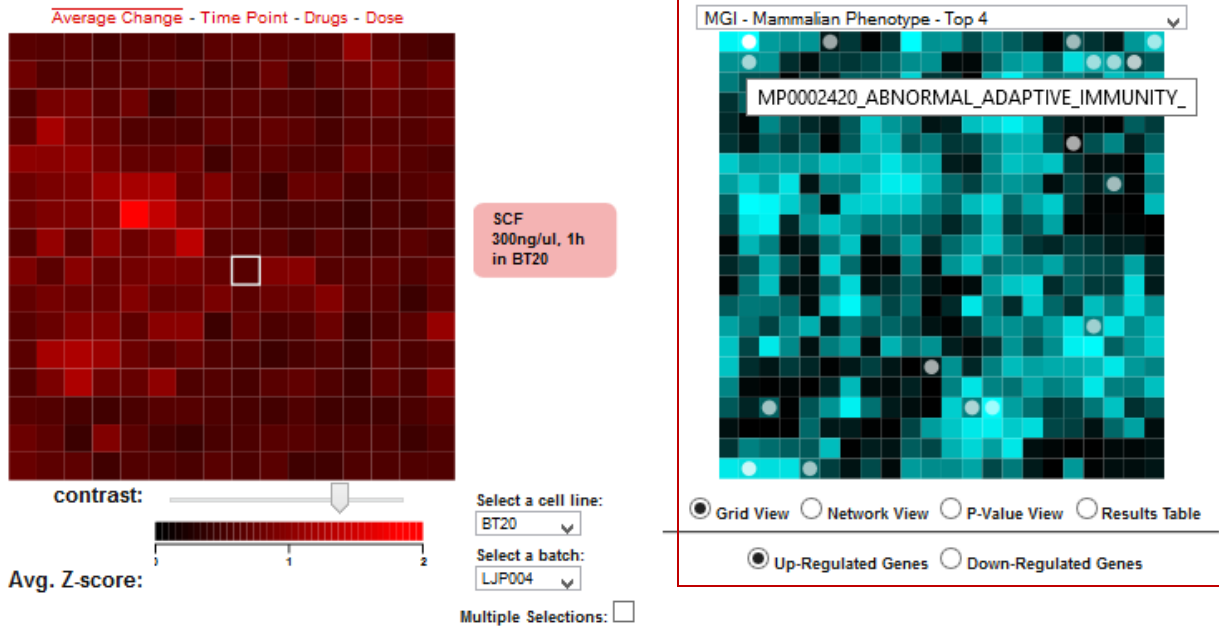


Step 8: Enrichment results table. The enrichment results table lists all the significantly enriched terms for the up-regulated genes of the consensus experiment using the Mammalian Phenotype gene-set library created from the MGI mouse phenotypes ontology browser. Enrichment is computed using the Fisher exact test with BH correction. Selection of the down-regulated genes radio button would display significantly enriched terms for the down-regulated genes for that experiments. A tooltip showing the overlapping genes between the up-regulated genes and the genes of a gene-set entry would pop-up after mouse hovering on that entry. Different gene-set libraries could be selected using the dropdown menu above the enrichment results table.

LINCS Canvas Browser

[Search](#) | [Browse](#) | [Help](#)

Search:

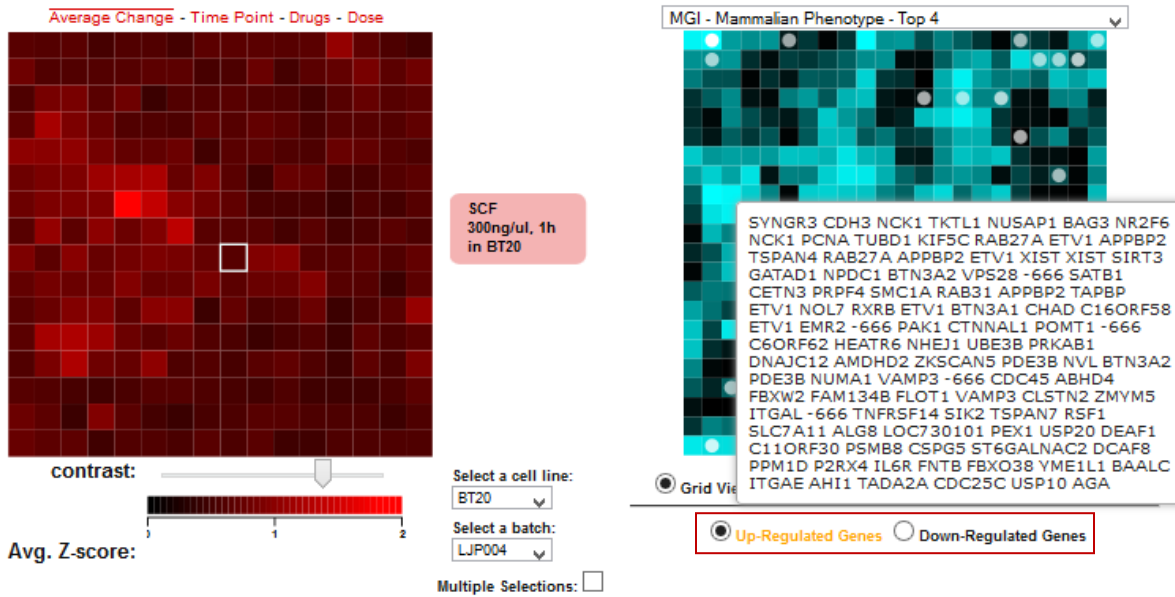


Step 9: Alternative views. Enrichment analysis results could also be visualized on gene-set library canvases by selecting the grid view radio button. These gene-set library canvases show all gene-sets in a specific gene-set library organized by their gene content similarity. Significantly enriched gene-sets that match the selected L1000 experiment are highlighted by white circles where brighter circles denote stronger enrichment scores.

LINCS Canvas Browser

[Search](#) | [Browse](#) | [Help](#)

Search:

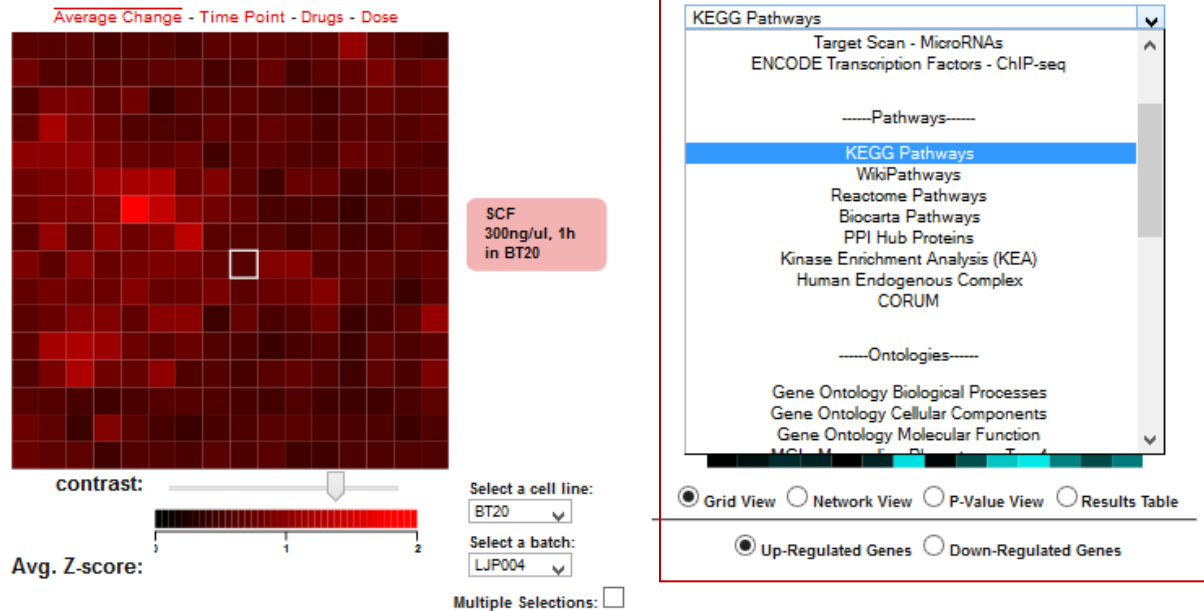


Step 10: Displaying the up- or down-regulated genes. Mouse-hovering the “Up-Regulated Genes” text or the “Down-Regulated Genes” text would display the respective gene lists.

LINCS Canvas Browser

[Search](#) | [Browse](#) | [Help](#)

Search:

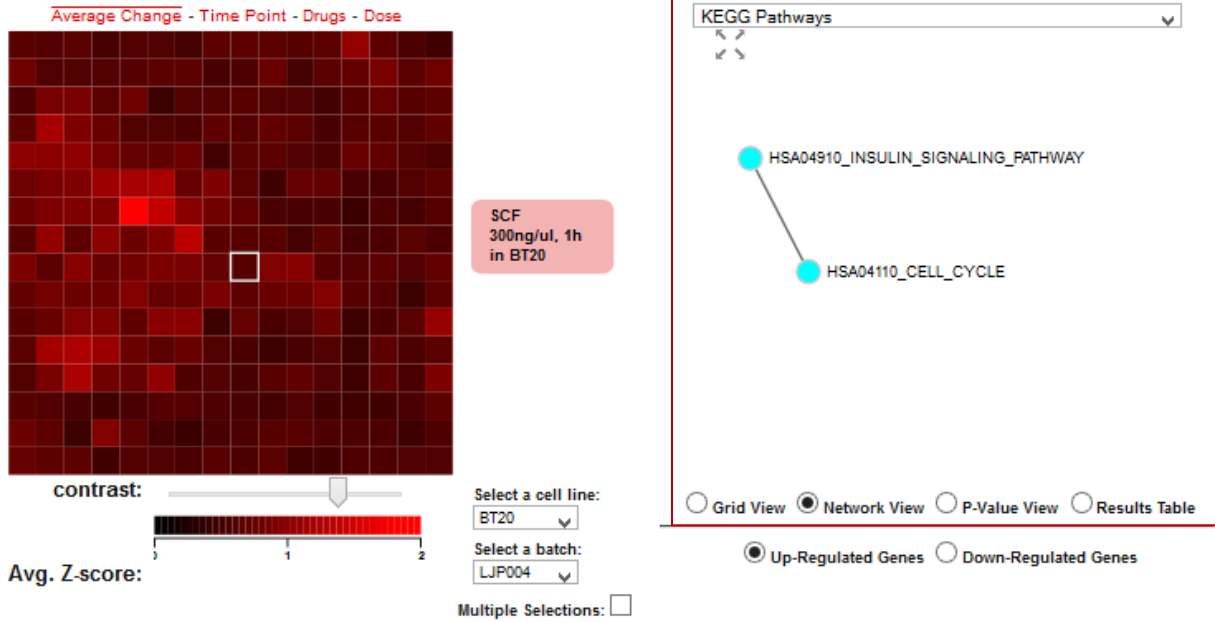


Step 11: Changing the gene-set library enrichment file. There are 30 different gene-set libraries that can be used to perform enrichment. These can be selected from the drop down list above the enrichment analysis panel.

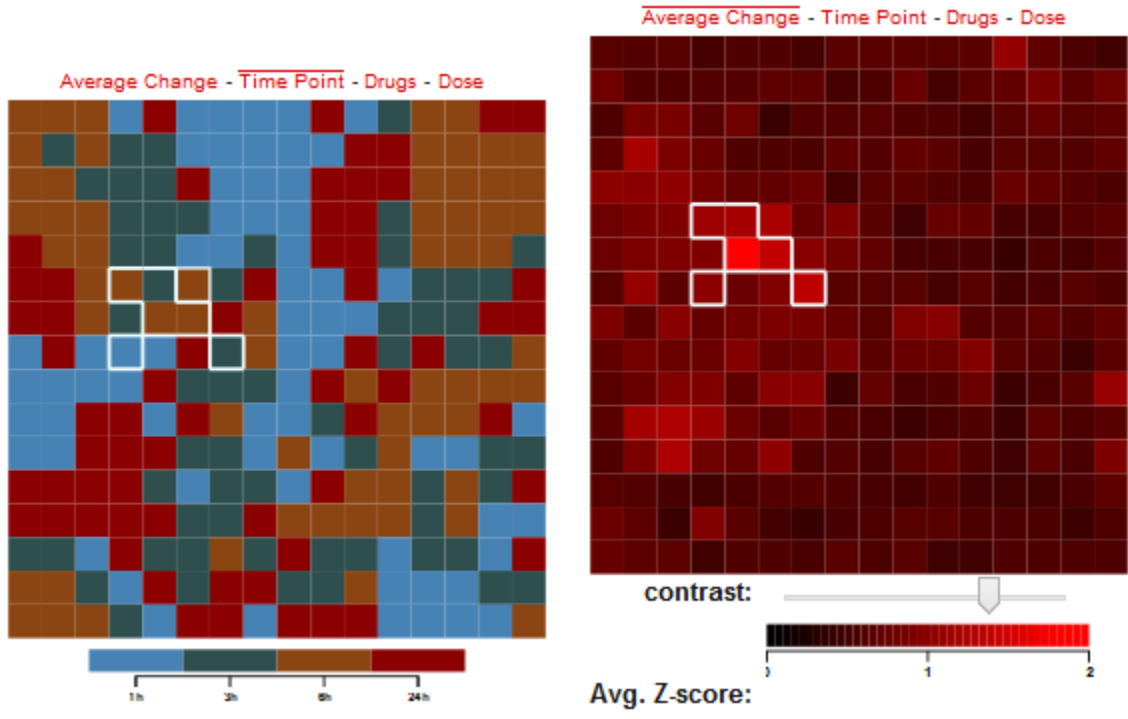
LINCS Canvas Browser

[Search](#) | [Browse](#) | [Help](#)

Search:



Step 12: Viewing the results as a network. The network view visualizes the enriched terms as a network that connect enriched terms based on shared genes. The network can be viewed in a pop-up window.



Step 13: Using the multiple selection option. When toggling between views on the experimental canvas multiple tiles can be highlighted using the multiple selection option checkbox.

II. Enrichment analysis on the input gene lists with Enrichr

LINCS Canvas Browser

[Search](#) | [Browse](#)

Gene Lists

Up List

- SYNGR3
- CDH3
- NCK1
- TKTL1
- NUSAP1
- BAG3
- NR2F6
- NCK1
- PCNA
- TUBD1
- KIF5C
- RAB27A
- ETV1
- APPBP2
- TSPAN4

clear

Down List

- MTF2
- SYNCRIP
- SLPI
- MRPL19
- COX17
- KIAA0494
- C7ORF42
- ANKS1A
- PSMB4
- AREG
- ELOVL6
- CSTF2T
- CYR61
- UGT2B28
- SEC63

clear

Top 50 Consensus Experiments

Overlap ▾ Info (Perturbation, Dose, Time, Cell, Batch) ▾

No data available in table

Showing 0 to 0 of 0 entries

◀ Previous Next ▶

Up Down

Search Example **Enrichr**

Aggravate Reverse

Enrichment analysis can be performed on input gene lists by clicking on the Enrichr button. The Enrichr button would send your input gene lists to the Enrichr web service. The up/down checkboxes determine which gene list to send. If both are checked, the up and down gene lists will be merged into one gene list and sent to Enrichr. You can learn more about how to use Enrichr at <http://amp.pharm.mssm.edu/Enrichr/index.html#help>.



[Login](#) | [Register](#)

Transcription Pathways Ontologies Disease/Drugs Cell Types Misc

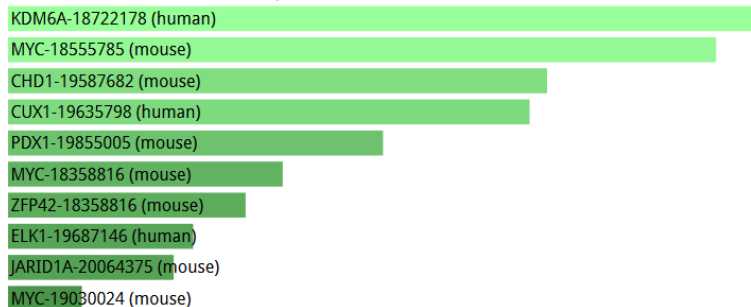
Description Up List (Aggravate) (95 genes)

ChEA

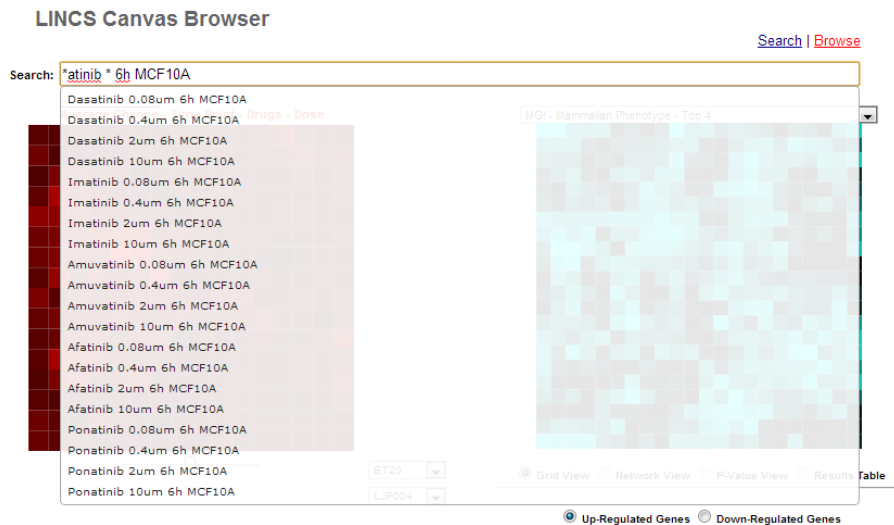
Bar Graph Table Grid Network ⚙

Click the bars to sort. Now sorted by **combined score**.

SVG PNG JPG



III. Searching for and visualizing L1000 perturbations



To search for a specific perturbation in the dataset users can utilize the search bar above the canvases. The search bar supports "*" as a wildcard. The example in the figure above searches for all experiments that use MCF10A cell line, have a perturbation name ended with atinib and a time-point of 6 hours. Selection of a matched experiment on the dropdown menu would visualize the experiment in the appropriate canvas and show its enrichment analysis results in a table as in show in step 6 of part I.

IV. Implementation details

The experiment canvases

The similarity between experiments is calculated from the z-score signatures of the 1000 measured genes using Pearson's correlation. A z-score signature is a vector of moderated z-scores calculated by the CMAP group at the Broad Institute from the replicates of an experimental condition as it compares with the population background. 'Moderated' means that the z-scores take into account the quality of each replicate and give low weight to bad quality replicates when averaging them. Z-score signatures were downloaded using LINCS L1000 application programming interface (API). The Network2Canvas algorithm is applied to compute canvas layouts. This algorithm clusters experimental conditions based on their similarity using simulated annealing where each experiment is represented by a canvas tile. The canvases are then rendered by HTML SVG elements that are created and manipulated using the D3 JavaScript library.

The gene-set library (enrichment) canvases

LCB includes 30 gene-set libraries that were borrowed from the Network2Canvas and Enrichr applications. Canvases were created from gene-set libraries using the Sets2Networks algorithm and then the Network2Canvas algorithm. The enrichment analysis in LCB implements the Fisher exact test to compute enrichment. To integrate the enrichment analysis canvases in LCB,

a listener function listens to click events. These events are registered to each tile on the experiment canvases. When a user clicks on a tile, the listener function is invoked, and then the function passes the differentially expressed genes of that tile to the enrichment canvas. The enrichment canvas then takes the genes it just received, performs the enrichment analysis, and displays the result.

Web communication protocols

LCB requests data for canvases on-the-fly using AJAX calls. LCB sends an AJAX request when a user selects a new batch. A server-side PHP script reads this request and sends back the requested JSON file that describes the canvas as plain text file. LCB interprets the file into JavaScript arrays using the eval function. Similarly, a second server-side PHP script handles the gene-set library requests and sends back the gene-set canvases data for visualization.

The search interface

Three text files are created from the entire L1000 dataset for the search purpose: 1) up genes file; 2) down genes file; and 3) up/down gene file. Each line in these files stores up or/and down genes for one experimental condition and all the experiments are included in each of these files. The organization within this file is maximized for query efficiency by preloading the file into to server side code at compile time. Once a user clicks the search button, LCB posts the user input list/s with the description tags to a server-side PHP script. The PHP script parses the gene lists and determines which text file to use based on the description tags. The parsed lists are passed to a Python script for the actual query process. The Python script compares the overlap between the input list/s and the gene lists of each experiment, and returns the top 50 matched experiments with the highest overlap. Since the up and down gene lists in the L1000 background signature dataset are equal in length, a simple overlap computation is fast and accurate. The overlap score is calculated as follows: $|D_{up} \cap L_{up}| / |L_{up}|$ for up gene set input query D_{up} ; $|D_{down} \cap L_{down}| / |L_{down}|$ for down gene set input query D_{down} ; and $(|D_{up} \cap L_{up}| + |D_{down} \cap L_{down}|) / (|L_{up}| + |L_{down}|)$ for up and down gene-set query where L_{up} and L_{down} are the up and down L1000 lists. D is used to denote user input lists and L , L1000 lists. The PHP script encodes the result into a JSON format and sends it back to LCB. An "Access-Control-Allow-Origin: *" header was set in all PHP scripts to enable universal embedding of LCB.

V. Contact, help and support

Please send suggestions, report bug and ask for help by contacting Avi Ma'ayan at avi.maayan@mssm.edu.