CELLX analysis

Mikhail Dozmorov

2018-02-24

Data from CELLX

- Go to http://54.149.52.246/cgi-bin/RPPA/cellx.cgi
- Select "Expression" tab out of "CNV"/"Expression"/"Mutation"/"Other" tabs
- Select "RSEM-barplot" option on the sidebar. Read more about RSEM gene expression measures at https://deweylab.github.io/RSEM/
- Select any cancer-associated gene from http://cancer.sanger.ac.uk/census, e.g., "ERBB2". Alternatively, use any gene name you may find biologically interesting
- Enter lower-case gene name into the "HUGO" textbox at the bottom of the page
- Click "Submit" - the page will refresh in ~20 sec
- Save the tab-separated data using "Download table" link into `data/RSEM_expression_GENESYMBOL.data.tsv`
- Import the downloaded data into R

Gene MIA analysis

There are 11188 unique samples from 59 tissue sources ($N_{cancer} = 35$, $N_{normal} = 24$) indexing expression of the gene *MIA*. The graph below shows the expression in of *MIA* in normal versus cancerous tissues.

Cancer vs. normal expression boxplots, all cancers

RSEM Expression MIA
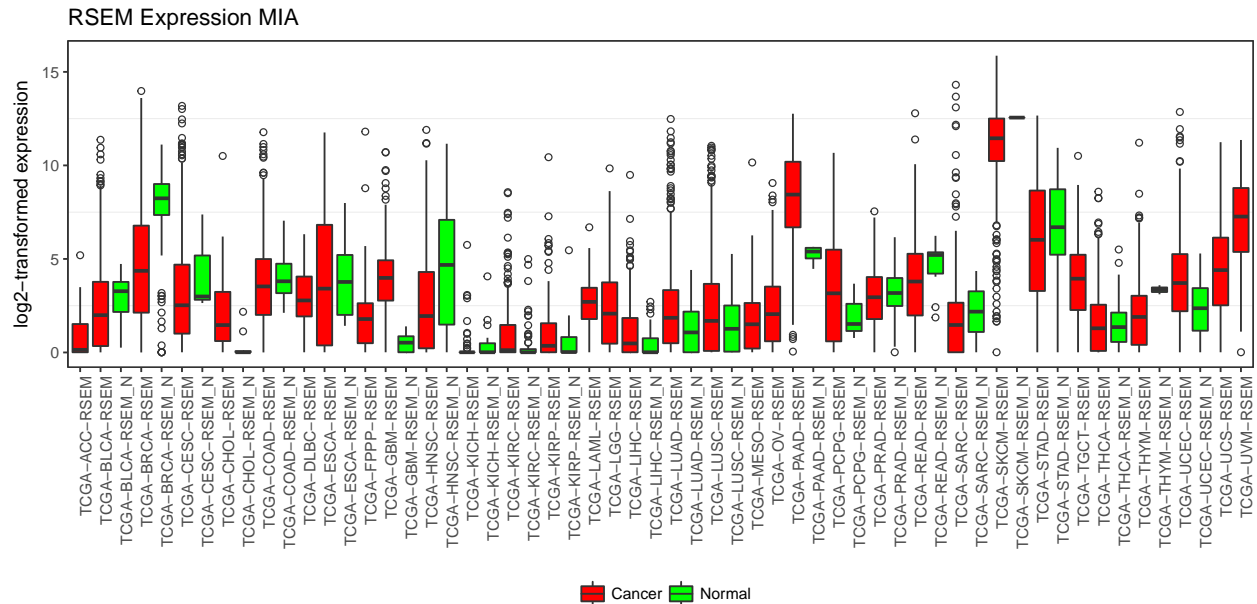
Differential expression of gene MIA between tumor and normal tissues

The table below displays differential expression statistics comparing the expression levels of *MIA* in normal versus cancer tissues (where normal comparison tissues are available) by Welch two-sample t-test.

Column names: "Cancer.Name", "Acronym" - cancer type, "Fold_change" - fold change (positive/negative - upregulated/downregulated in cancer), "p_value", "t_statistic" - results of Welch two-sample t-test, "Number_of_normal", "Number_of_cancer" - number of samples, "Mean_expression_normal", "Mean_expression_cancer" - log2-transformed RSEM expression.

The table is sorted by "p_value" column, from most to least significant.

Cancer.Name	Acronym	Fold_change	p_value	Number_of_normal	Number_of_cance
Breast invasive carcinoma	BRCA	-2.8436	0.000000	114	111
Lung adenocarcinoma	LUAD	1.0103	0.000001	59	53
Uterine Corpus Endometrial Carcinoma	UCEC	1.5245	0.000001	35	55
Liver hepatocellular carcinoma	LIHC	0.6395	0.000004	50	37
Glioblastoma multiforme	GBM	3.4999	0.000007	5	16
Pancreatic adenocarcinoma	PAAD	2.9830	0.000054	4	17
Lung squamous cell carcinoma	LUSC	0.9630	0.000063	51	50
Head and Neck squamous cell carcinoma	HNSC	-2.1107	0.000154	44	52
Kidney renal clear cell carcinoma	KIRC	0.4686	0.000595	72	53
Cholangiocarcinoma	CHOL	1.6090	0.000876	9	3
Thymoma	THYM	-1.2162	0.037682	2	12
Colon adenocarcinoma	COAD	-0.4421	0.049839	41	47
Bladder urothelial carcinoma	BLCA	-0.5429	0.088539	19	41
Kidney renal papillary cell carcinoma	KIRP	0.3421	0.106697	32	29
Rectum adenocarcinoma	READ	-0.7149	0.179512	10	16
Pheochromocytoma and Paraganglioma	PCPG	1.3781	0.250577	3	18
Prostate adenocarcinoma	PRAD	-0.2264	0.253370	52	50
Thyroid carcinoma	THCA	0.1435	0.422987	59	51
Stomach adenocarcinoma	STAD	-0.4351	0.435034	35	41
Cervical and endocervical cancers	CESC	-0.9946	0.581499	3	30
Kidney Chromophobe	KICH	-0.1154	0.590879	25	6
Esophageal carcinoma	ESCA	0.0795	0.910630	11	18
Sarcoma	SARC	-0.1782	0.947917	2	26

Cancer vs. normal expression boxplots, individual cancers

The figure below is a plot that highlights the differences in *MIA* expression distribution in normal and cancer tissues for each cancer type.

RSEM Expression MIA