



4340 Duncan Ave | Suite 220
Saint Louis, MO | 63110
+1 (314) 833-9764

Jimmy Buffett
Margarita Pharmaceuticals
March 3, 2022

Interactive Data Analysis With Rosalind™ NanoString

Canopy Biosciences offers an end-to-end NanoString workflow which includes immune cell-type profiling analysis for specific panels. As part of this service, we provide the following report which includes sample quality details and introduces Rosalind™, our interactive and collaborative data analysis software. Once registered with Rosalind™ you will have immediate access to all your differential expression data, allowing collaboration across the globe.

NanoString Sample Quality Control

An initial quality control check was performed on all samples. The results can be found in the table below. Briefly, the concentration was determined using a fluorimetric-based measurement (Qubit) and the quality was assessed using a Bioanalyzer. The RNA Integrity Number (RIN) is derived from an algorithm evaluating the 28S to 18S rRNA ratio. RIN values can be from 1 to 10, with a score of 10 representing perfect quality. DV₂₀₀ is reported for all FFPE-derived samples and is a measurement of the % of molecules >200 nucleotides.

Sample ID	NanoString File Name	Group	Qubit (ng/ul)	RIN	DV 200
C1	20190910_C1_01.RCC	1	6500	10	97
C2	20190910_C2_02.RCC	1	7000	10	97
C3	20190910_C3_03.RCC	1	6300	10	96
C4	20190910_C4_04.RCC	2A	2370	6	79
C5	20190910_C5_05.RCC	2A	6300	9	92
C6	20190910_C6_06.RCC	2A	3950	8	85
C7	20190910_C7_07.RCC	2B	6600	10	96
C8	20190910_T1_08.RCC	2B	4010	10	94
C9	20190910_T2_09.RCC	2B	4800	10	96
C10	20190910_T3_10.RCC	3A	1410	8	90
C11	20190910_T4_11.RCC	3A	720	8	87
C12	20190910_T5_12.RCC	3A	8200	9	94

Sample Quality Control (continued)

Quality Flags

The samples listed in the table below have been flagged for quality control issues. If necessary, samples can be excluded from analysis if it is determined that their inclusion could have detrimental effects to the experimental comparisons.

Sample ID	Flag	Explanation	Keep or Drop
C1	mRNA Content Normalization Flag	Low binding density. Counts significantly lower relative to the other samples and mRNA content normalization would not be appropriate in this case.	Keep
C12	Binding Density Flag	High binding density. Positive control values are still linear, and percent of probes above background (set at >50 counts) are similar to other samples.	Keep



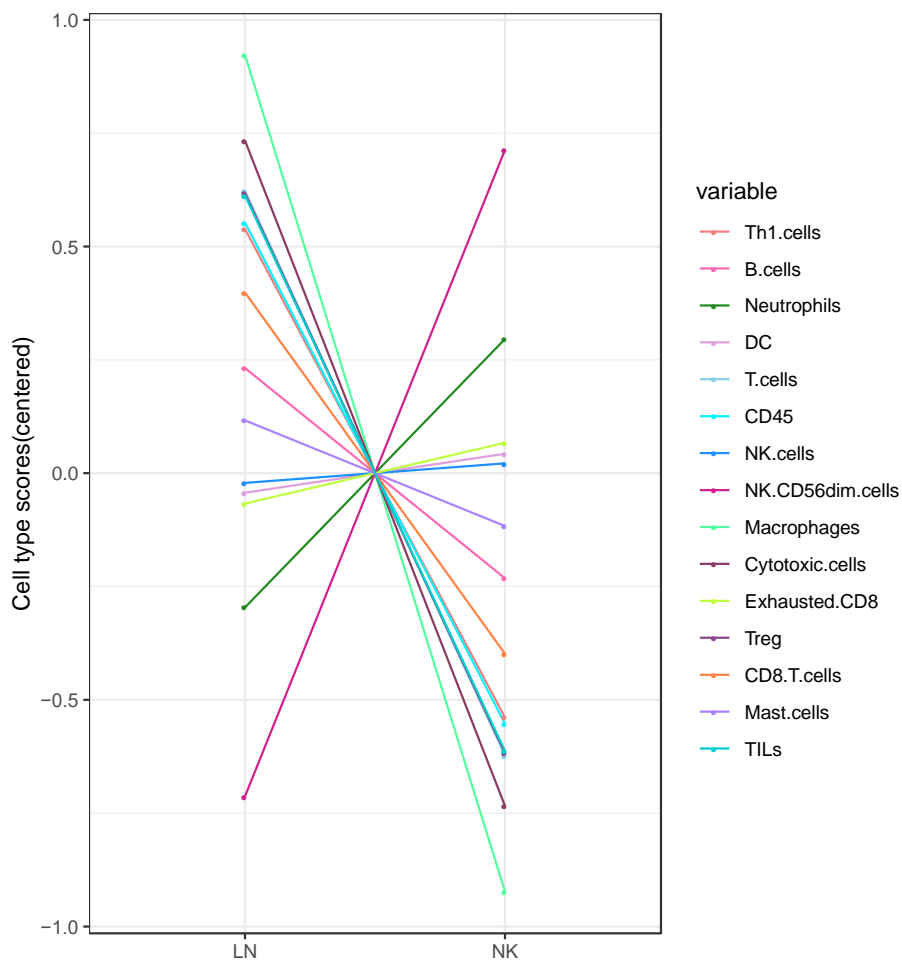
Immune Cell Profiling Results 1 vs 2A



1 vs. 2A Immune Cell Profiling

Cell Type Score Summary

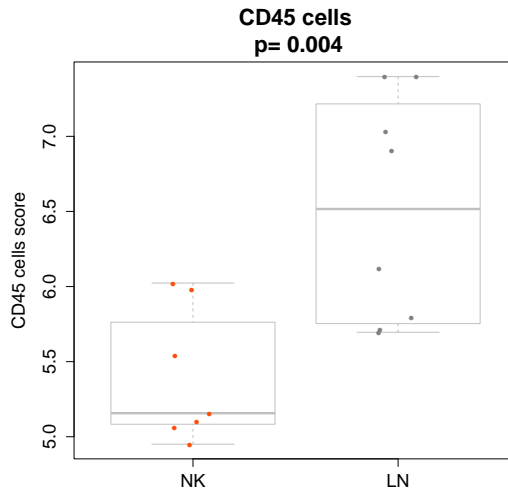
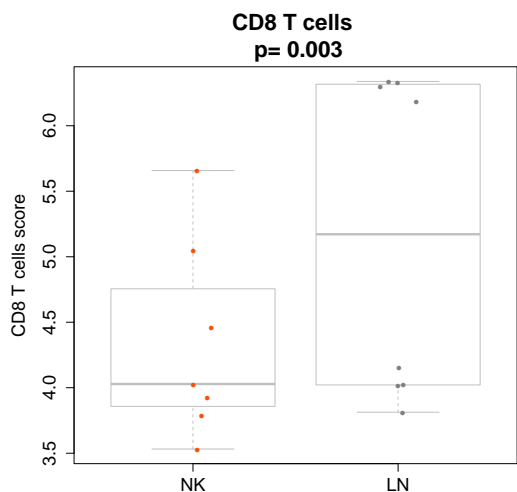
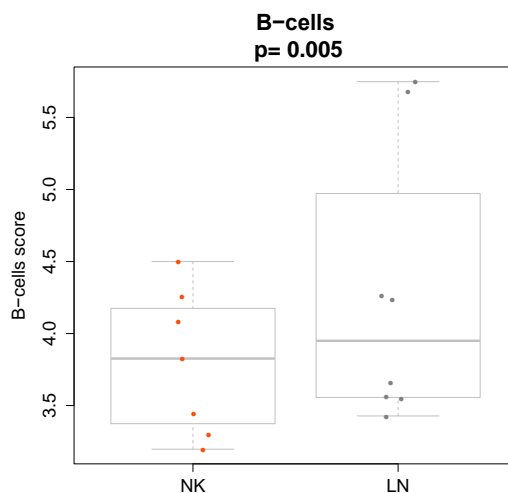
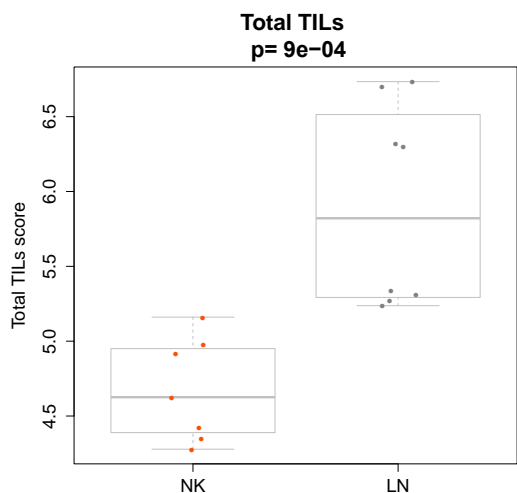
Cell type profiling was performed by the method described in Danaher et al. (2017) using NanoString nCounter Advanced Analysis 2.0 to measure the abundance of various cell populations. This method derived 60 marker genes whose expression levels measure 14 immune cell populations. Below is a summary of the cell type scores. Change in expression between one experimental group and the other can be visualize for each cell type. Individual cell abundance scores are available on subsequent pages.



1 vs. 2A Immune Cell Profiling

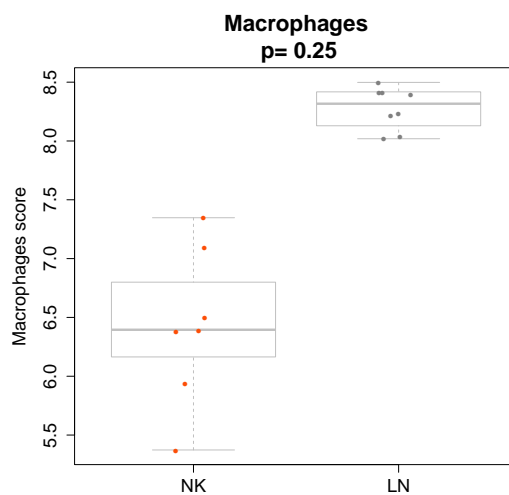
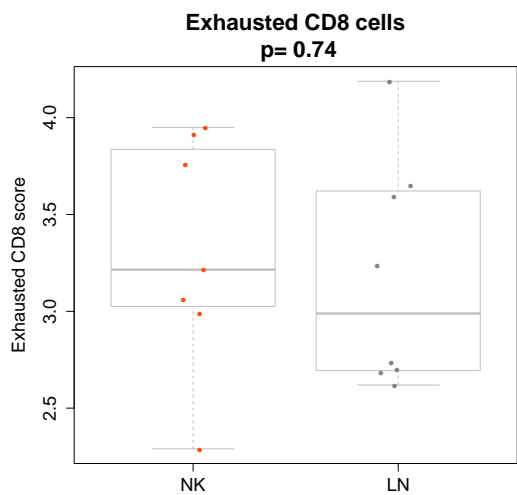
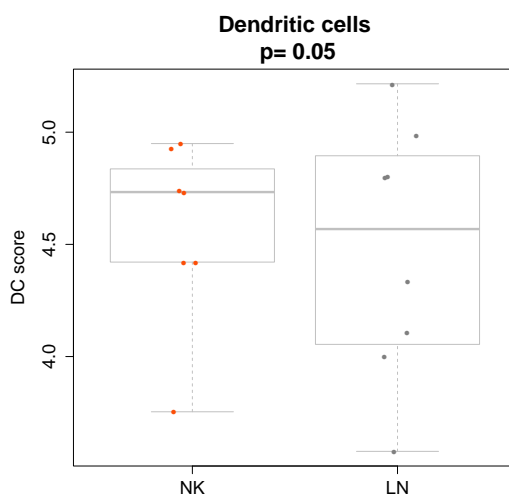
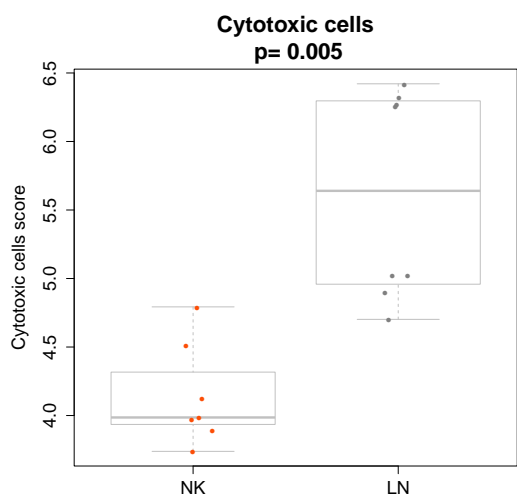
Total TILs and Individual Cell Type Scores

Total TILs (top left) and raw cell type abundance graphs are shown below and on the following slides. Abundance estimates are given on the log2 scale, so a unit increase in score corresponds to a doubling of a cell type's abundance. These scores do not support claims about whether one cell type is more abundant than another. Rather, they permit claims that a cell type is more abundant in one group than in another.



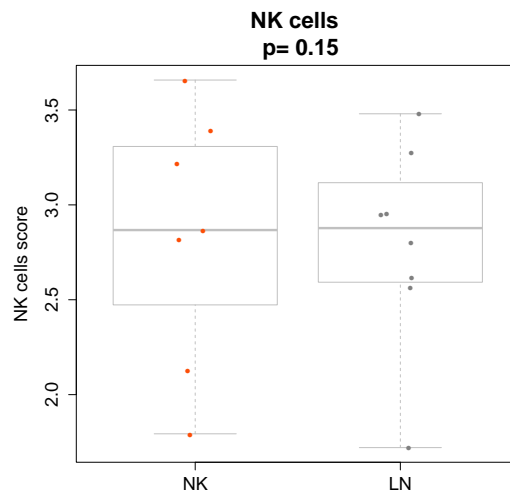
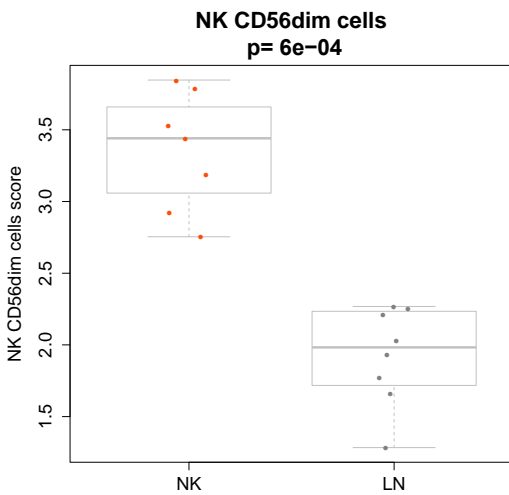
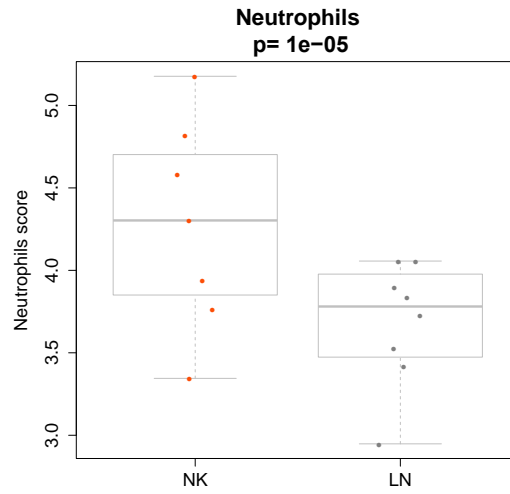
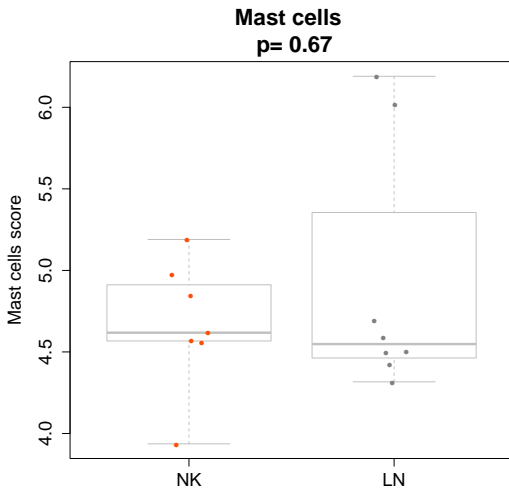
1 vs. 2A Immune Cell Profiling

Individual Cell Type Scores



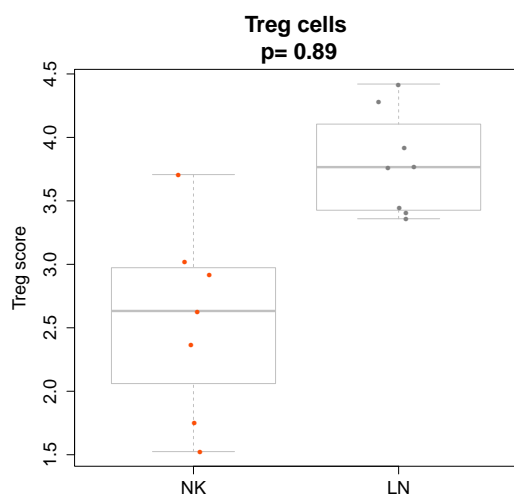
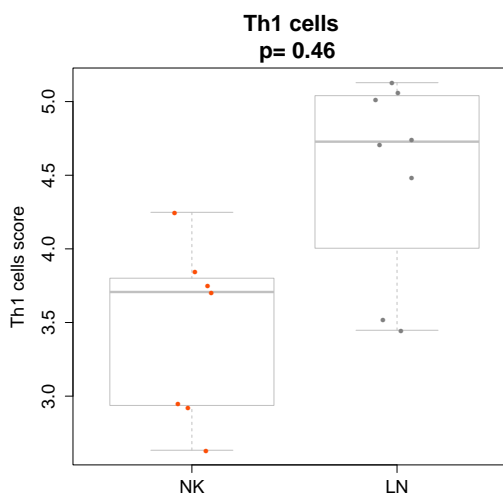
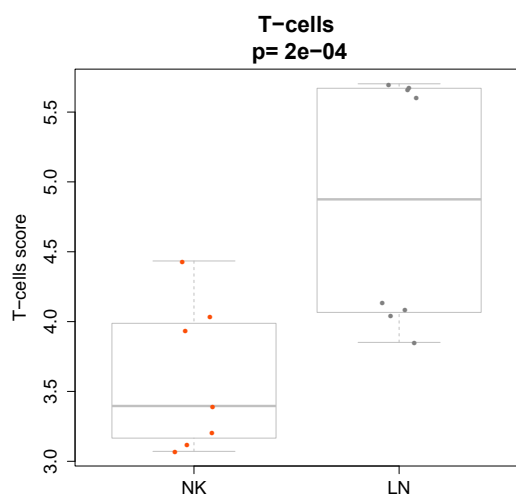
1 vs. 2A Immune Cell Profiling

Individual Cell Type Scores



1 vs. 2A Immune Cell Profiling

Individual Cell Type Scores



Introduction to Rosalind™ Analysis Software



Intuitive Interface & Easy Collaboration

The Rosalind™ platform powered by ONRAMP™ is a very user-friendly interface. With icons and tool tips, tutorials and videos, and statistical overviews, this software helps even a novice gain powerful insights into complex experiments and data sets.

You can easily share your project space with others in your team to collaborate on the results and next steps. Data images, spreadsheets and other files are available for immediate download. The next page explains how to get started with Rosalind™.

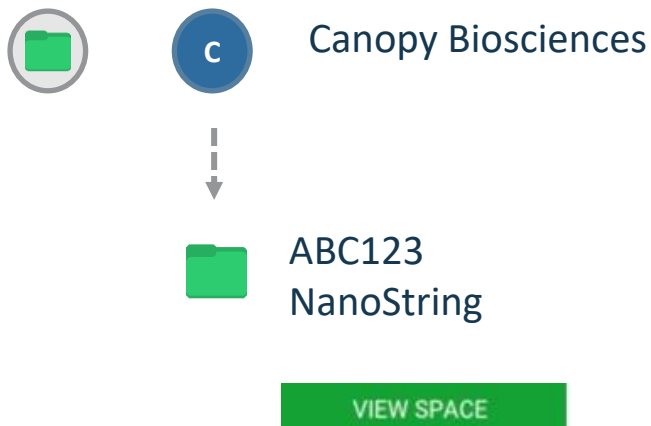
Getting Started with Rosalind™

1. Check your inbox for an e-mail for Rosalind and follow the instructions to create an account.
2. Once an account has been created, you will receive an e-mail notification that an analysis space has been shared (see image below)
3. Log-in to your account, accept the analysis space invitation and begin exploring your data!

ROSALIND™



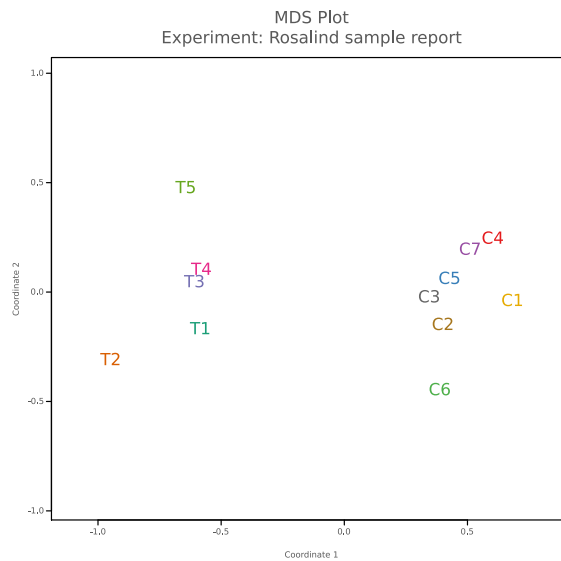
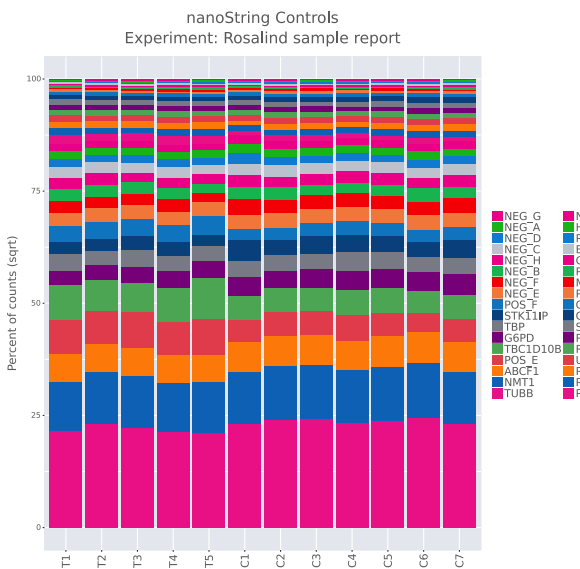
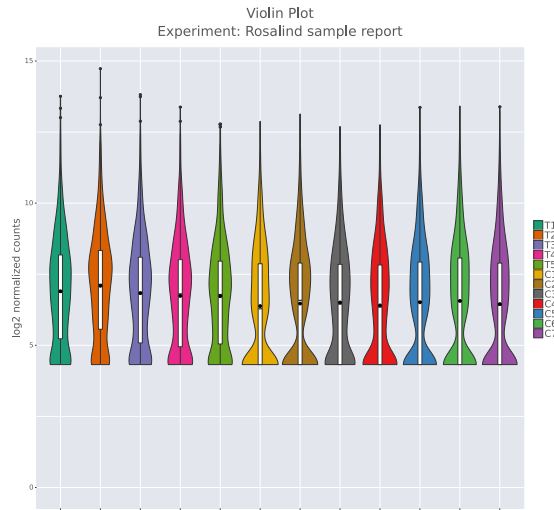
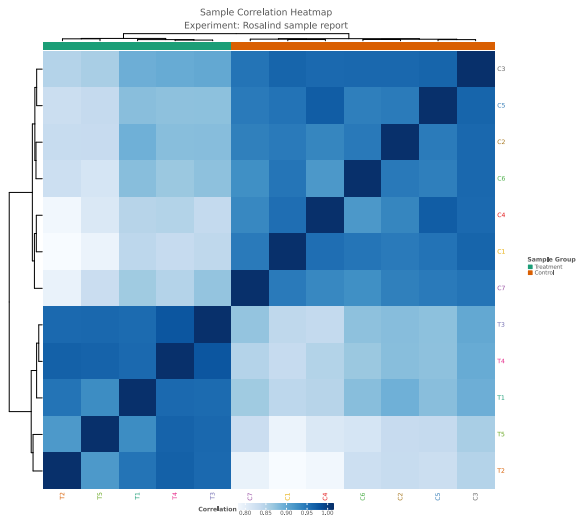
Canopy Biosciences has invited you to join a Space and collaborate!



To join the space and begin collaborating with others you will need to login to your account and accept this invitation

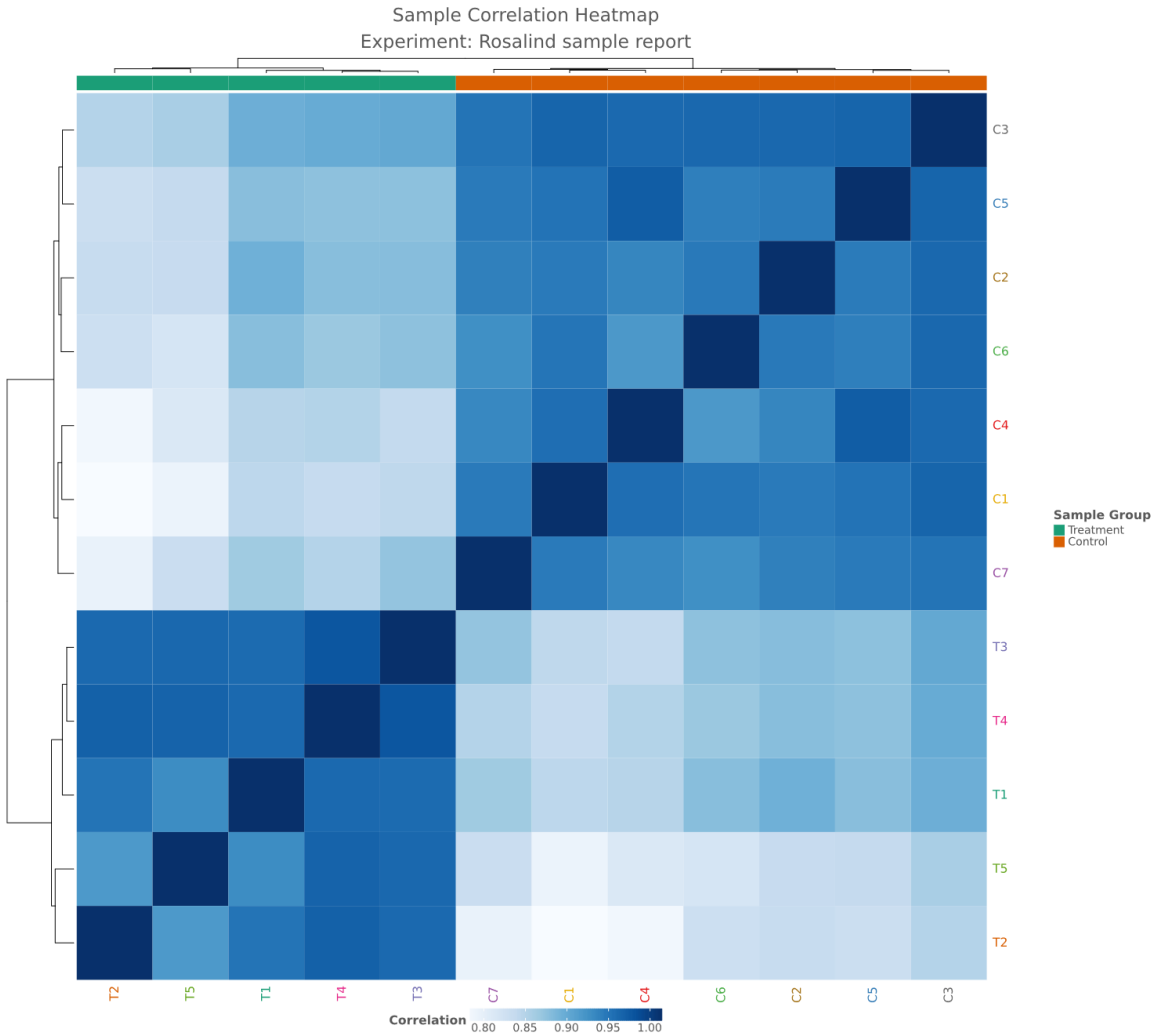
Quality Control

For each project space that has been shared, there will be several quality control plots to help visualize the performance of samples in the experiment. Careful assessment of the quality control information can be helpful in interpreting the differential expression data. The next 4 pages of this report will explain the four quality control plots that Rosalind™ displays for each NanoString analysis project. Please note that these are example images and do not represent your actual data.



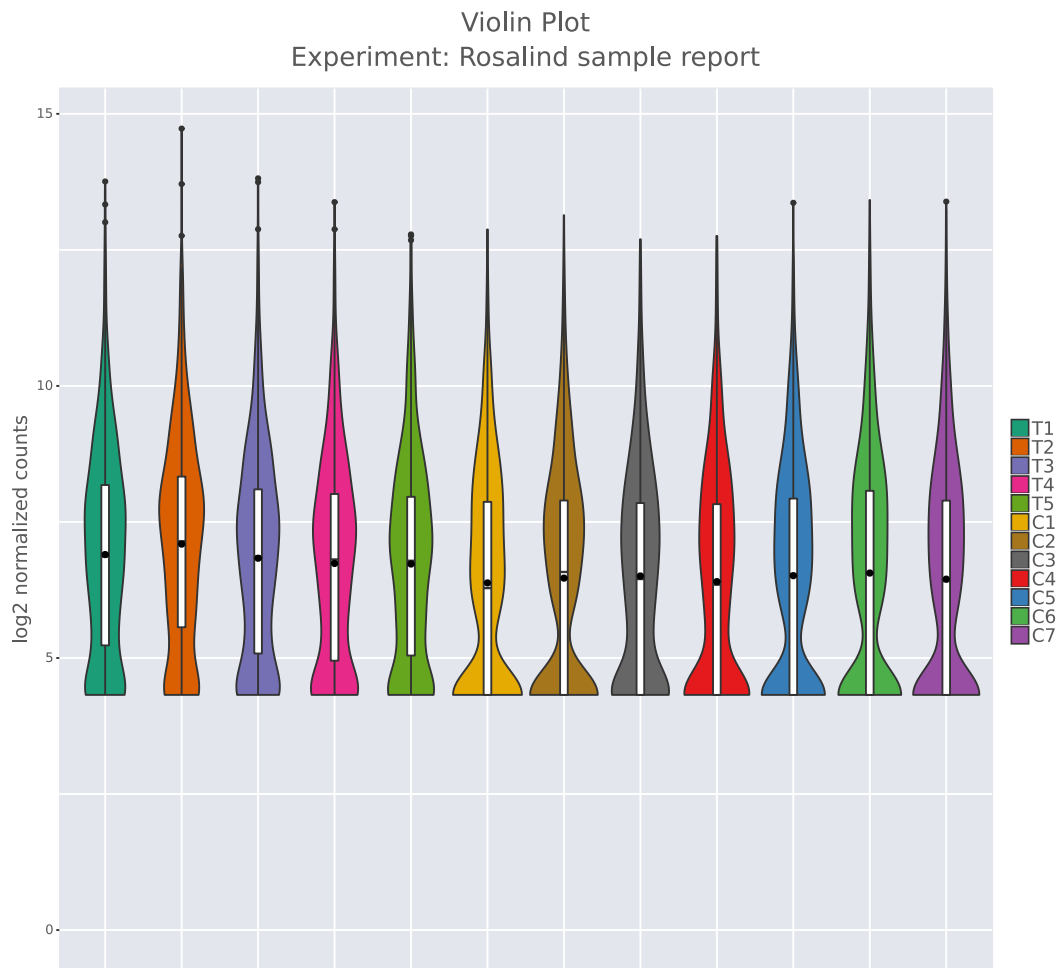
Sample Correlation Heatmap

A sample correlation heatmap provides a graphical representation of data, in which the individual values contained in a matrix are represented as colors. In this case, the data matrix contains correlation values between samples, with the darkest blue representing the strongest correlation. The dendrogram annotation on the top axis provides valuable information regarding the clustering of samples. You should expect to see samples that are closely related (i.e. those in the same replicate group) to be strongly correlated together in the plot, and to be closest branches of the dendrogram.



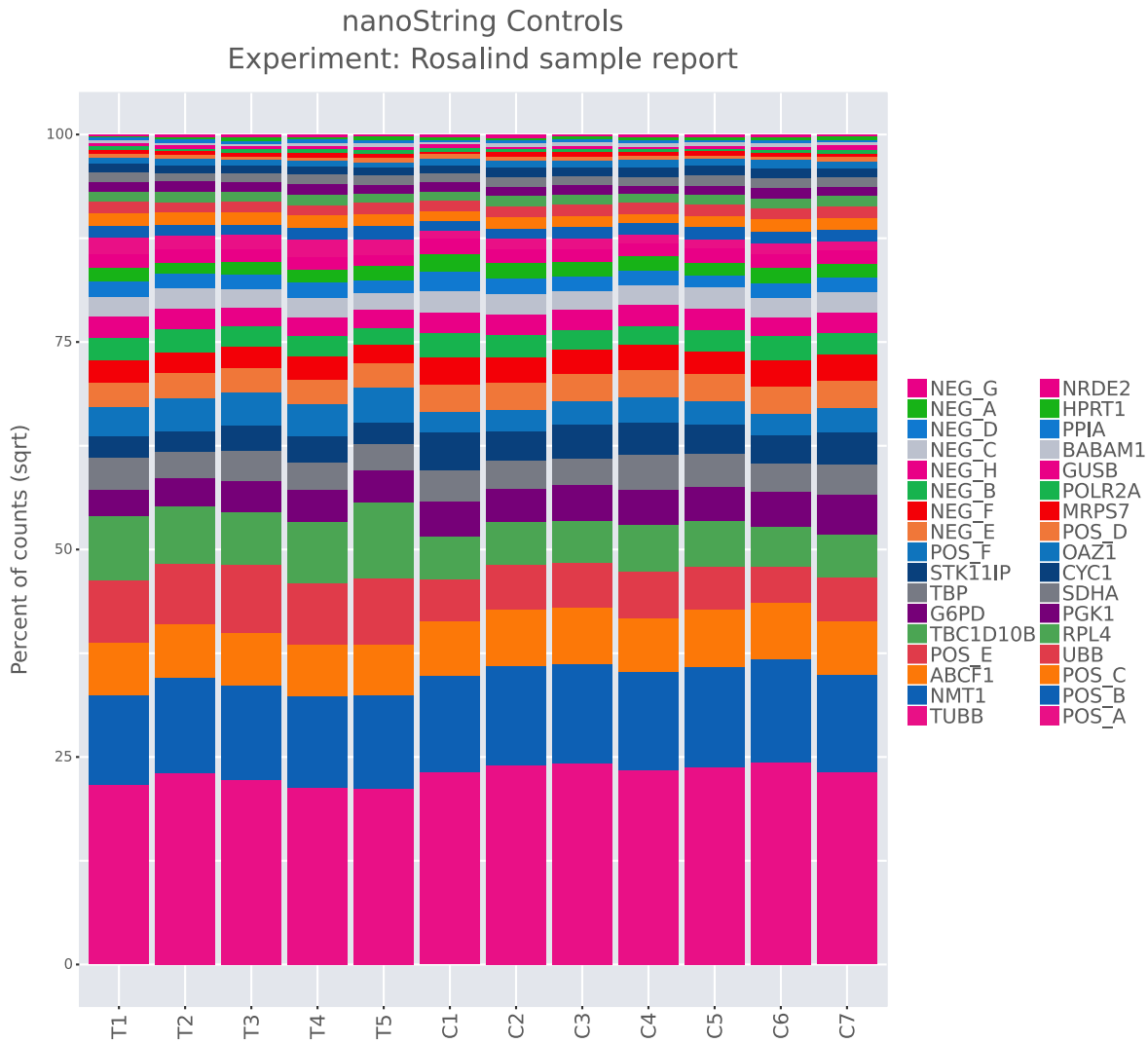
Combined Box and Violin Plot

An example of a combined box/violin plot is depicted below. The graph displays the distribution of the log-transformed and un-normalized gene counts for each sample in your experiment. Box plots are a standardized way to display the distribution of data based 5 metrics: minimum, first quartile, median, 3rd quartile and maximum. Sometimes the mean is also displayed as a dot in the rectangle delimited by the 25th and 75th percentiles. While box plots only graph summary statistics, violin plots show the actual distribution of the data. They are rotated density plots on each side of the box plot, which give you the ability to clearly observe the variation in the data.



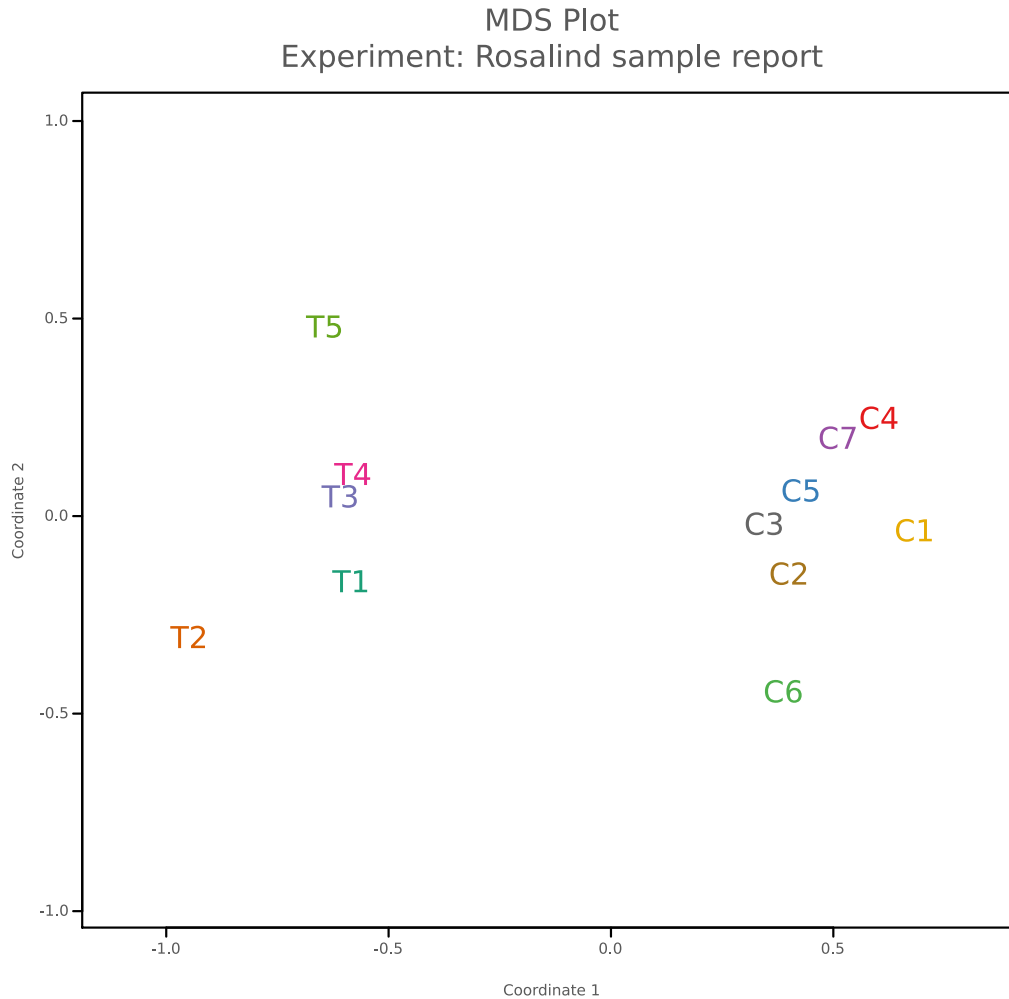
NanoString Controls

This plot shows the expression level of spike-ins and negative controls. The square root of the expression level is used to show the lower expression values. Each column sums to 100%. Again, please note that this is only a sample plot and does not represent your actual data.



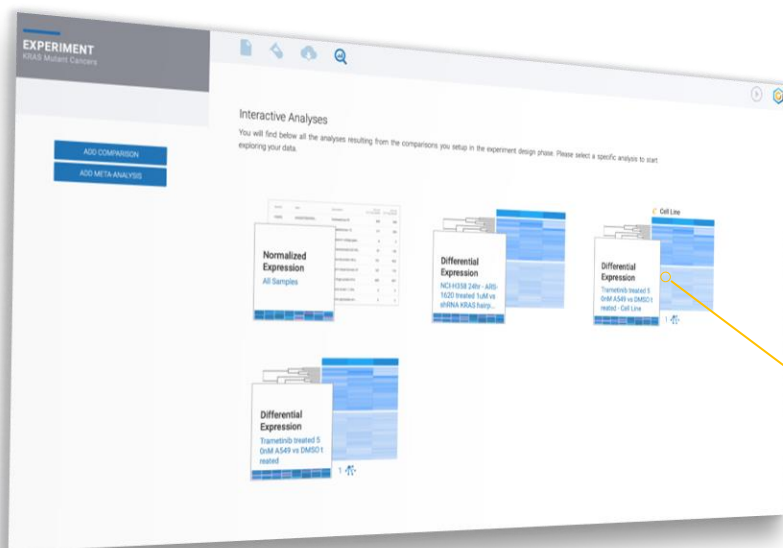
MDS Plot

An example of a multidimensional scaling (MDS) plot is provided below. The MDS plot provides visual representation of the patterns of proximities (i.e., similarities or distances) among a set of objects. In this case, we plot the expression differences between samples within your experiment. Samples that are clustered together generally have similar patterns of expression. If one or more samples are highly dissimilar from one another, one would expect these samples to segregate in the MDS plot.



Comparisons

Group comparison analysis is performed as requested. In your shared analysis space, you will be able to view multiple figures and tables for each comparison. The top figure below is a snapshot of what you will see in the Discovery & Analysis section. Selecting a specific comparison will take you to a page that allows you to explore all the differential expression data and figures. Please note that all following data are only examples. The data from your experiment may be viewed in Rosalind™ and downloaded as needed.



Click on the comparison you would like to explore

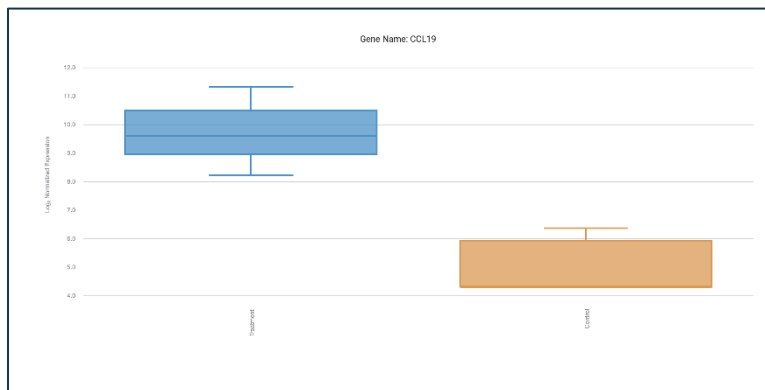
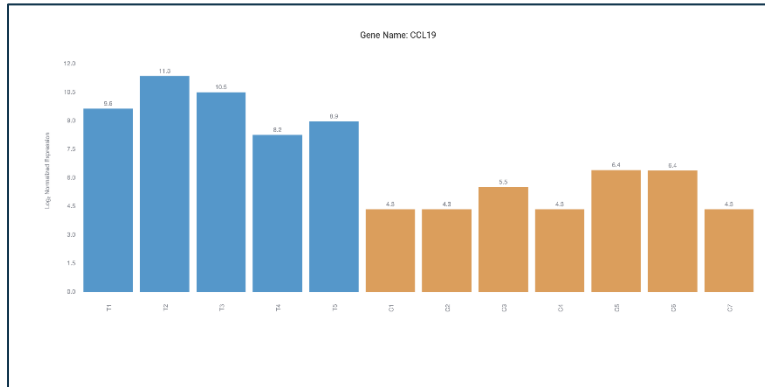


Differentially Expressed Gene List

You can visualize gene lists that pass the parameters of the filter (please see note below on filters). You can further explore the individual genes identified and their expression levels using bar graphs and box and whisker plots. Bar graphs show gene expression across all the samples while box and whisker plots graphically depict groups of gene expression data through their quartiles.

DIFFERENTIAL EXPRESSION		
Treatment vs Control		
CCL19	TPSAB1/B2	HLA-DPA1
CCL21	HLA-DPB1	TYROBP
ATP6V1C2	EGF	CD209
LTF	CD52	IL7R
CXCL1	LTB	CYBB
ITGAX	CCL5	NKG7
CCL2	C1QA	C1QB
FCN1	IL10RA	ITGAL
IRF4	CD45R0	IL33
IRF4	IKZF1	PTPRC
FCGR3A/B	CD84	CD45RB
CPA3	FCGR2B	MS4A4A
HLA-DRA	VWF	KLRB1
CD8A	GZMA	LILRB1
CD3D	IKZF3	SERPINF2
CD68	HIST1H4E	F13A1
HLA-DMB	LYN	CSF3R
IL2RG	ITGB2	PRF1
IFITM1	ITK	CSF2RA
FAS	SIGLEC1	CXCL8
SELL	CD14	CD4
JAK3	HAVCR2	ICAM1
CD3E	PLCG2	CASP1
MS4A1	CTSC	RASSF5
ALOX5	TNFRSF1B	TNF
ZAP70	VCAM1	ANXA1
FCGR1A	ISG20	SPI1
PANX1	ITGAM	CD44
GMPR	CXCL12	DLL1
FGR	ITGA4	UBA52
CD74	KIT	RHOA
S100A8	BIRC3	TNFSF10
ID1	CD27	GBP2
PTEN	ARRB2	NFKB1
VAV1	CSF1	IL27RA
RELB	HLA-DOA	CFH
INFO	LISTS	NP2
		GATA3

SELECTED LIST		
237	≥ 1.5	0.05000
Genes	Fold Change	p-Adj

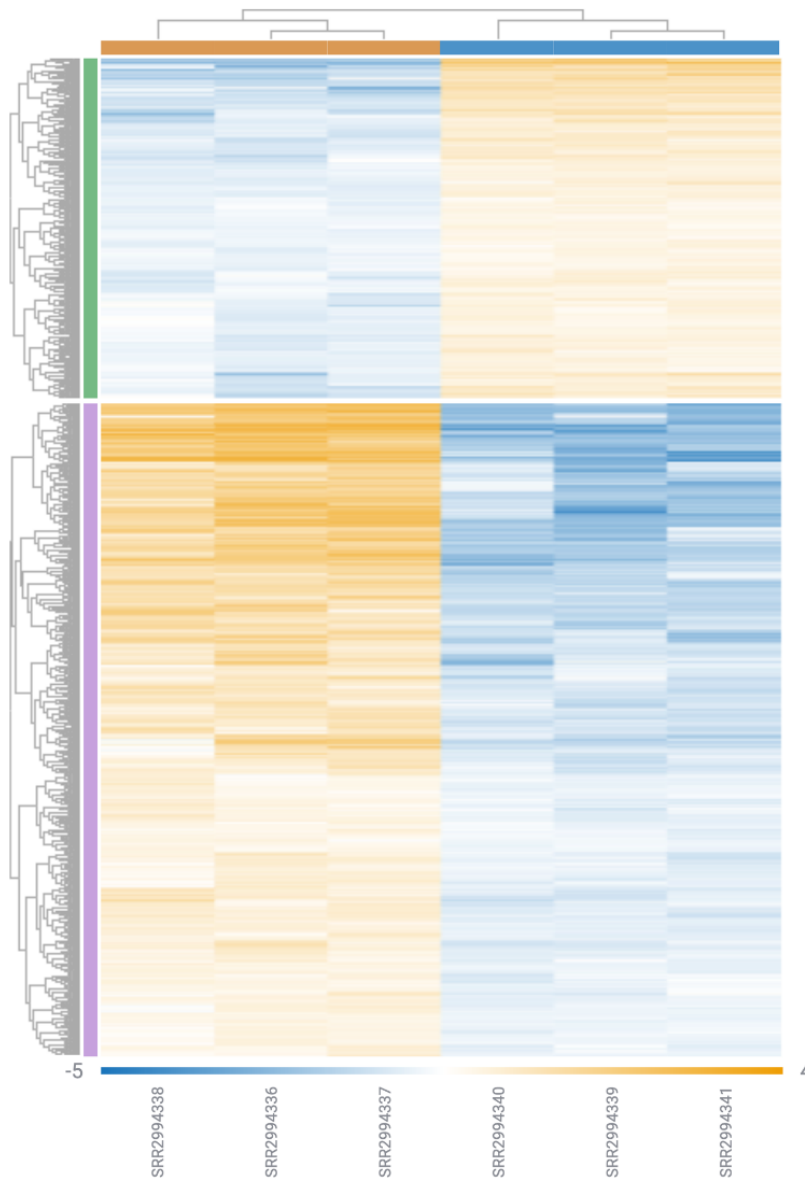


Data will be filtered using Rosalind's Dynamic Default Filter. This filter uses a default of +/- 1.5 for fold change and 0.05 FDR (adjusted p-value). The filter then dynamically adjusts to allow for a reasonable number of genes through the filter, which is why some data sets will show a final filter that is different from +/- 1.5, 0.05 FDR.

Heat Maps

A heatmap provides a graphical representation of data, in which the individual values contained in a matrix are represented as colors. In this case, the heatmap displays a grid in which each row represents a gene and each column represents a sample. The color and intensity of the boxes are used to represent log normalized counts for each of the genes that pass the filter selected: darker blue indicates higher expression.

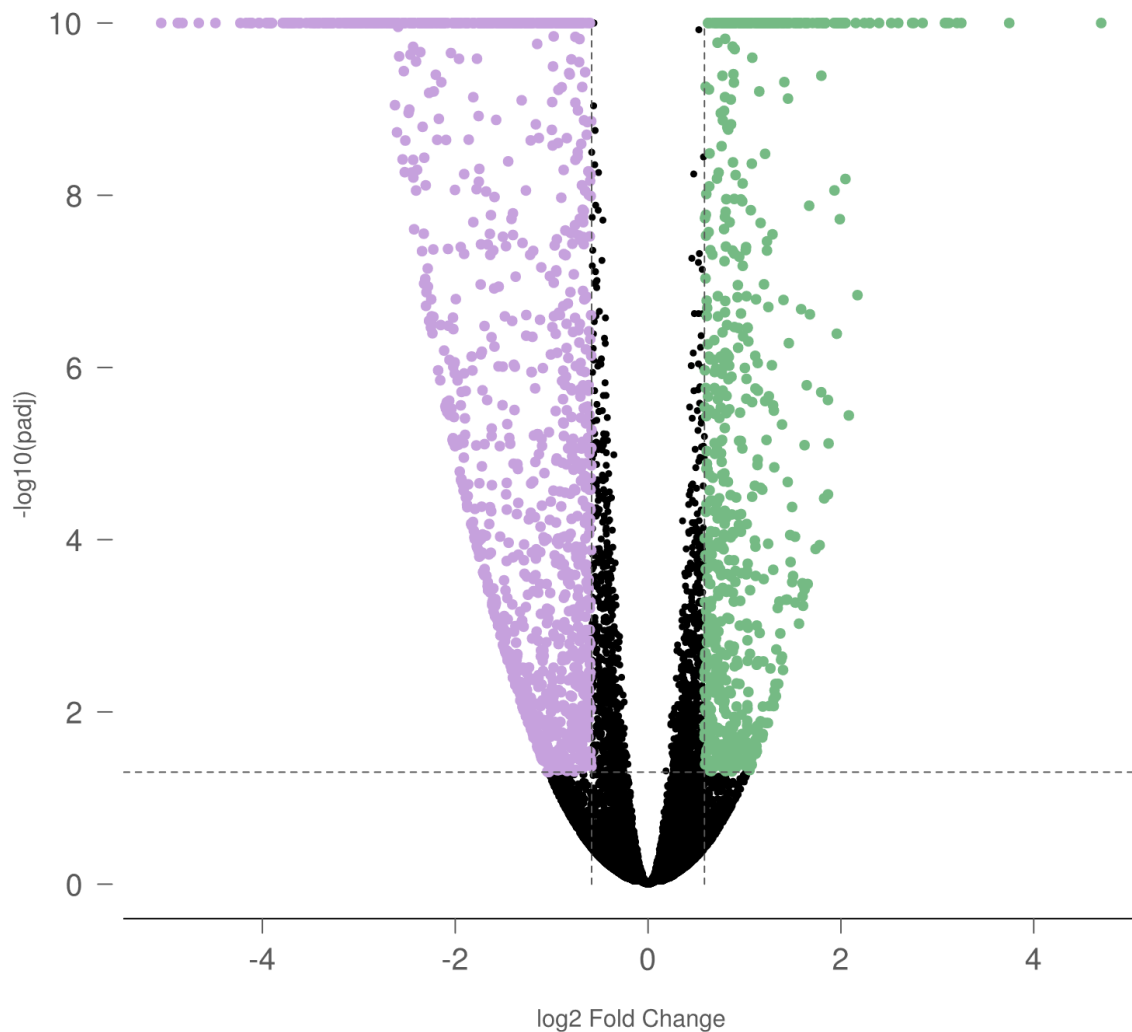
Clustering of genes for the final heatmap of differentially expressed genes was done using the PAM (Partitioning Around Medoids) method using the fpc R library.



Volcano Plot

The volcano plot shows differences in gene expression between the two groups selected for comparison. The log ratio of the fold change is on the X axis, and the negative log of p-adj/p-value is on the Y axis. Each dot represents a gene within the comparison performed.

The coloring on the dots reflects the clustering information for each gene, and those in black are genes that do not pass the parameters of the filter selected.



Pathway Mapping

The Rosalind™ platform provides access to best-in-class data visualization and pathway interpretation, including enrichment with over 20 leading knowledge bases:

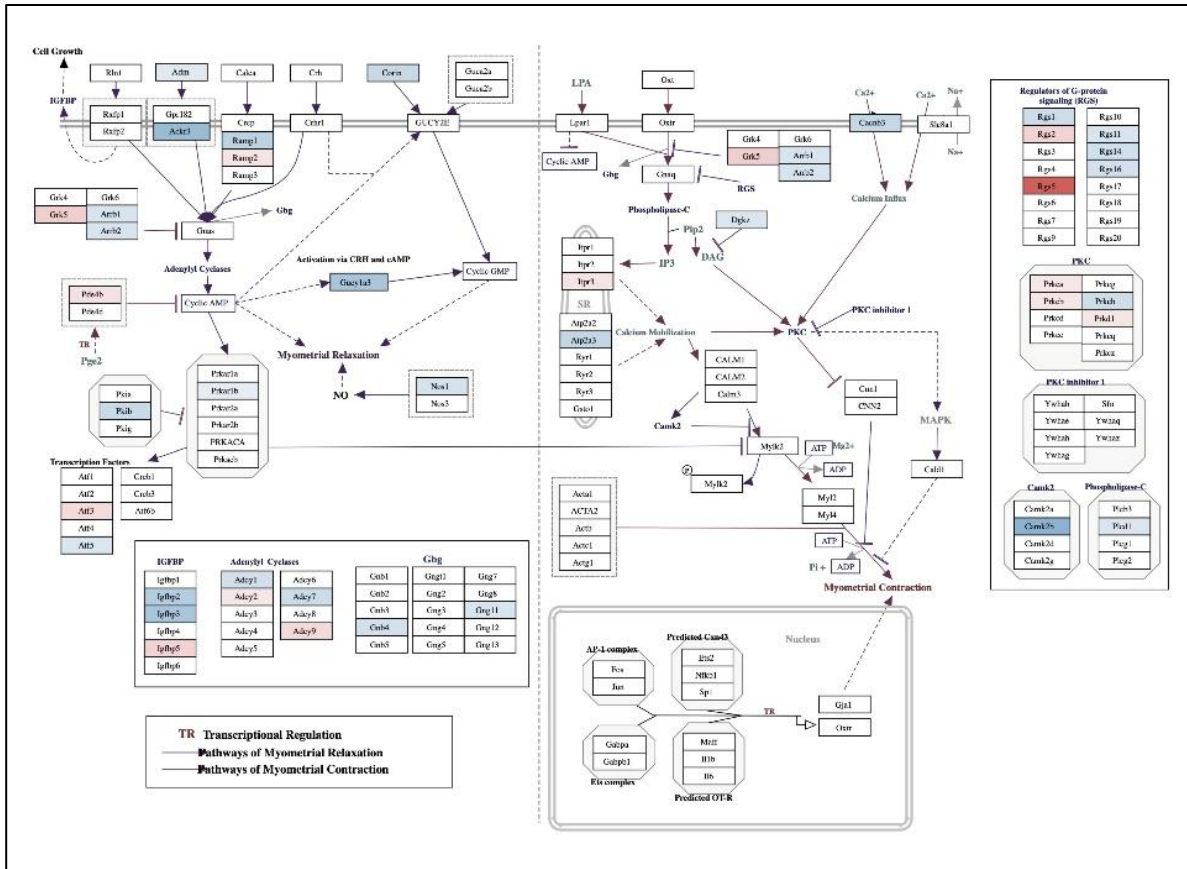
Pathways: WikiPathways, Reactome, BIOCYC, Pathway Interaction Database, and the Small Molecule Pathway Database

Gene Ontology: Biological Processes, Molecular Functions and Cellular Component

Proteins: Protein-Protein Interactions, Interpro, Pfam, SMART, GENE3D, Prosite and PRINTS

Other Knowledge Bases: MSigDB, COSMIC, GWAS catalog, Lipid Maps and Chromosome location

You can perform dynamic and interactive analyses on your data using pathway explorer.



of Genes in Target: 7 7↑ 0↓

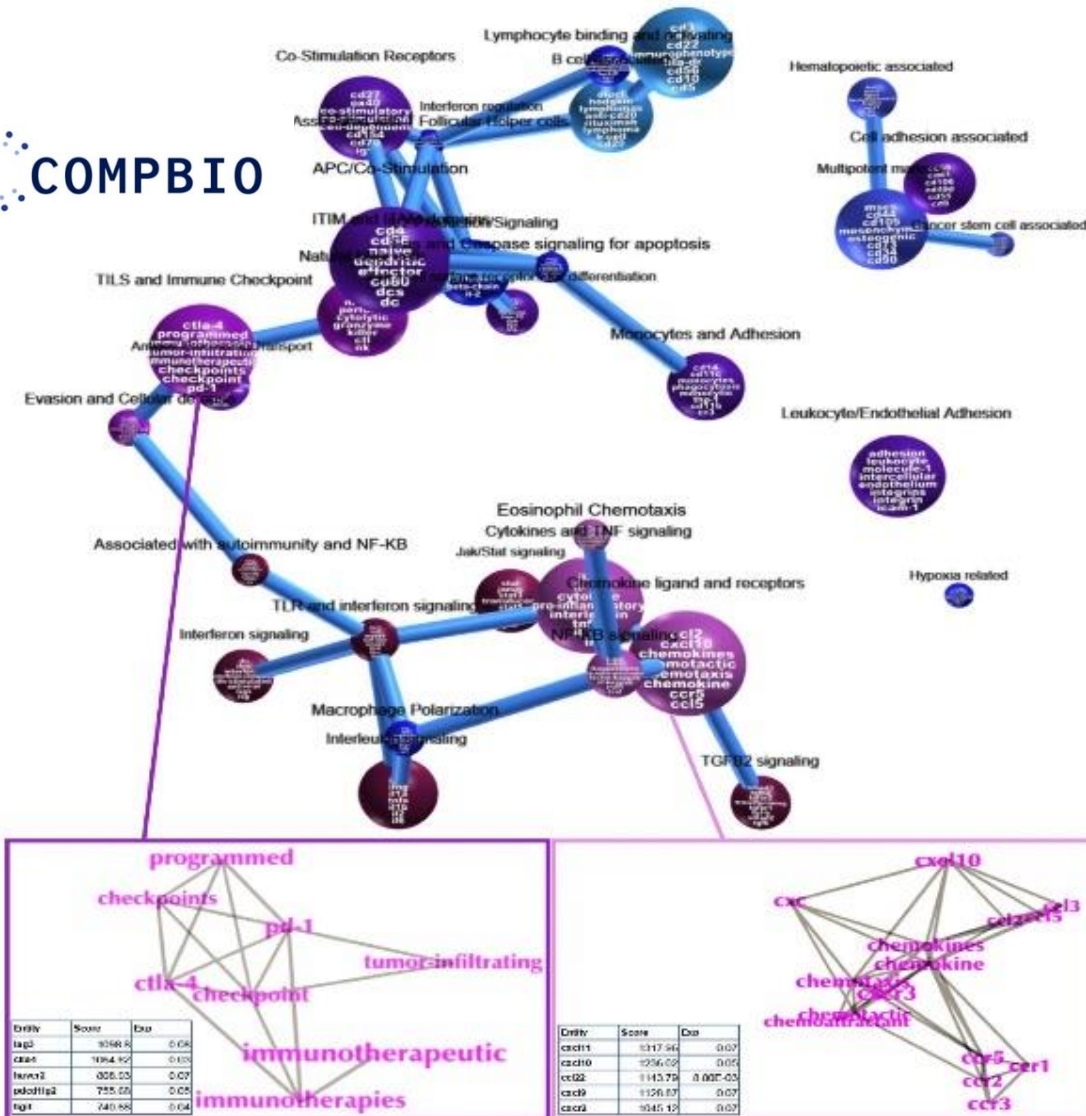
Gene	-4	+4	Fold Change	p-Value
Gbp2b			1.13438	0.011663
Il1b			1.01812	0.018924
Cxcl10			0.927602	0.021508
Irf8			0.900242	0.000392
Cxcl9			0.836603	0.020949
Irf1			0.746798	0.007001
Icam1			0.706458	0.005597



In the interactive view, you can see genes present in your selected pathway.

Recommendation & Contact Information

Canopy also offers a deeper understanding of your comparisons. In a partnership with PercayAI, we offer the CompBio analysis tool. This powerful tool uses your list of most important genes (based on fold changes from your comparisons) to create novel pathways based on the current literature. This is very different than current pathway analysis tools in that it is not looking for a best-fit pathway. Instead, it uses augmented intelligence software to connect genes found to be important in your comparisons to other genes, transcripts, and proteins in the literature. It does the work that you would manually do in order to make connections between your findings and published data. The results are presented in the form of multiple figures (see example below) and tables. This tool has been demonstrated to save an enormous amount of time in drawing conclusions about data. For example, one can use this for mechanism of action validation or to identify novel targets for possible combination studies. Please contact us to learn more.



Contact Information

We are here to help! If you need additional support or analysis, please contact us.

- For questions on the report, please contact info@canopybiosciences.com.
- For additional analysis, please contact your sales representative or sales@canopybiosciences.com for a quote.