

---

# **DEBrowser Documentation**

***Release 1.1.10***

**Alper Kucukural, Nicholas Merowsky**

September 13, 2016



<b>1</b>	<b>Quick-start Guide</b>	<b>3</b>
1.1	Getting Started . . . . .	3
1.2	The Main Plots . . . . .	11
1.3	Quality Control Plots . . . . .	19
1.4	Heat Maps . . . . .	21
1.5	GO Term Plots . . . . .	22
1.6	Data Tables . . . . .	24
<b>2</b>	<b>Local Install Guide</b>	<b>27</b>
2.1	Quick Local Install . . . . .	27
<b>3</b>	<b>DESeq/DEBrowser</b>	<b>29</b>
3.1	Introduction . . . . .	29
3.2	DESeq2 . . . . .	29
3.3	Un-normalized counts . . . . .	29
3.4	Used parameters for DESeq2 . . . . .	30
3.5	EdgeR . . . . .	30
3.6	Used parameters for EdgeR . . . . .	30
3.7	Limma . . . . .	30
3.8	Used parameters for Limma . . . . .	30
3.9	DEBrowser . . . . .	31
3.10	References . . . . .	32
<b>4</b>	<b>DEBrowser Additions</b>	<b>33</b>
<b>5</b>	<b>Indices and tables</b>	<b>35</b>



Contents:



---

## Quick-start Guide

---

This guide is walkthrough for the DESeq Browser from start to finish.

### 1.1 Getting Started

First off, we need to head to the DEBrowser webpage at this url:

<http://debrowser.umassmed.edu/>

Alternatively, if you have the R package installed, you can call these R commands:

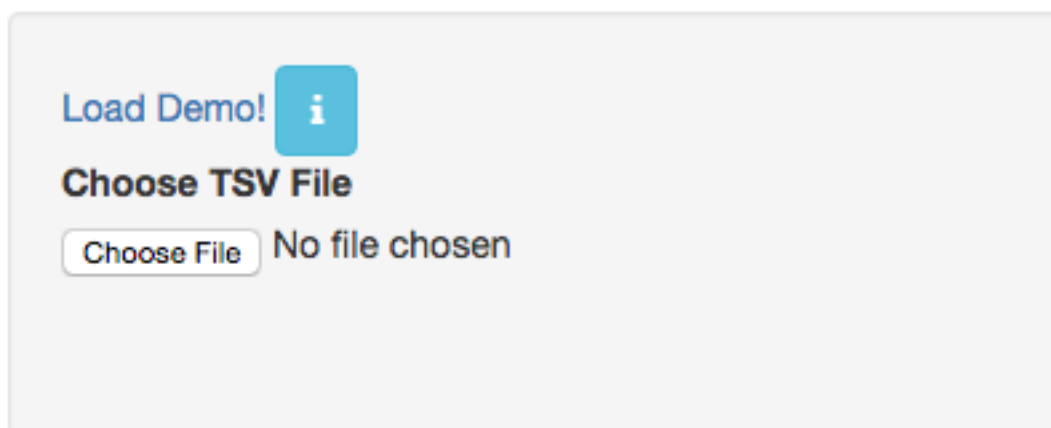
```
library(debrowser)
```

```
startDEBrowser()
```

For more information on installing DEBrowser locally, please consult our Local Install Guide.

Once you've made your way to the website, or you have a local instance of DEBrowser running, you will be greeted with this tab on the left:

## DEBrowser



To begin the DESeq process, you will need to select your Data file (TSV format) to be analyzed using DESeq2. If you do not have a dataset to use, you can select to use the built in demo by clicking on the 'Load Demo!'. To view the entire demo data file, you can download this [demo set](#). For an example case study, try our [advanced demo](#).

The TSV files used to describe the quantification counts are similar to this:

IE:

gene	trans	exp1	exp2	cont1	cont2
DQ714	uc007	0.00	0.00	0.00	0.00
DQ554	uc008	0.00	0.00	0.00	0.00
AK028	uc011	2.00	1.29	0.00	0.00

DEBrowser also accepts TSV's via hyperlink by following a few conversion steps. First, using the API provided by Dolphin, we will convert TSV into an html represented TSV using this website:

<http://dolphin.umassmed.edu/public/api/>

The Two parameters it accepts (and examples) are:

1. source=[http://bioinfo.umassmed.edu/pub/debrowser/advanced\\_demo.tsv](http://bioinfo.umassmed.edu/pub/debrowser/advanced_demo.tsv)
2. format=JSON

Leaving you with a hyperlink for:

[http://dolphin.umassmed.edu/public/api/?source=http://bioinfo.umassmed.edu/pub/debrowser/advanced\\_demo.tsv&format=JSON](http://dolphin.umassmed.edu/public/api/?source=http://bioinfo.umassmed.edu/pub/debrowser/advanced_demo.tsv&format=JSON)

Next you will need to encode the url so you can pass it to the DEBrowser website. You can find multiple url encoders online, such as the one located at this web address: <http://www.url-encode-decode.com/>.

Encoding our URL will turn it into this:

[http%3A%2F%2Fdolphin.umassmed.edu%2Fpublic%2Fapi%2F%3Fsource%3Dhttp%3A%2F%2Fbioinfo.umassmed.edu%2Fpub%2Fdebrowser%2Fadvanced\\_demo.tsv%26format%3DJSON](http%3A%2F%2Fdolphin.umassmed.edu%2Fpublic%2Fapi%2F%3Fsource%3Dhttp%3A%2F%2Fbioinfo.umassmed.edu%2Fpub%2Fdebrowser%2Fadvanced_demo.tsv%26format%3DJSON)

Now this link can be be used in debrowser as:

<http://debrowser.umassmed.edu:443/debrowser/R/>

It accepts two parameters:

1. jsonobject=[http%3A%2F%2Fdolphin.umassmed.edu%2Fpublic%2Fapi%2F%3Fsource%3Dhttp%3A%2F%2Fbioinfo.umassmed.edu%2Fpub%2Fdebrowser%2Fadvanced\\_demo.tsv%26format%3DJSON](http%3A%2F%2Fdolphin.umassmed.edu%2Fpublic%2Fapi%2F%3Fsource%3Dhttp%3A%2F%2Fbioinfo.umassmed.edu%2Fpub%2Fdebrowser%2Fadvanced_demo.tsv%26format%3DJSON)
2. title=no

The finished product of the link will look like this:

[http://debrowser.umassmed.edu:443/debrowser/R/?jsonobject=http://dolphin.umassmed.edu/public/api/?source=http://bioinfo.umassmed.edu/pub/debrowser/advanced\\_demo.tsv&format=JSON&title=no](http://debrowser.umassmed.edu:443/debrowser/R/?jsonobject=http://dolphin.umassmed.edu/public/api/?source=http://bioinfo.umassmed.edu/pub/debrowser/advanced_demo.tsv&format=JSON&title=no)

Inputting this URL into your browser will automatically load in that tsv to be analyzed by DEBrowser!

For more information about the input file, please visit our DESeq/DEBrowser tab within Readthedocs. Once you've selected your file and the upload has completed, you will then be shown the samples listed within your file uploaded as well as a few options.



**Data Prep**

**Samples**

exper\_rep1 exper\_rep2 exper\_rep3 control\_rep1 control\_rep2 control\_rep3

Go to DE Analysis! Go to QC plots! Reset Samples!

The first option, ‘Go to DE Analysis’, takes you to the next step within the DESeq workflow. In order to run DESeq on your input data you first need to select which samples will go into your conditions. You can run multiple condition comparisons and view the results separately as well. To remove samples from a condition, simply select the sample you wish to remove and hit the delete/backspace key. To add a sample to a condition you simply have to click on one of the condition text boxes to bring up a list of samples you can add to that comparison. Click on the sample you wish to add from the list and it will be added to the textbox for that comparison.

**Data Prep**

**Samples**

exper\_rep1 exper\_rep2 exper\_rep3 control\_rep1 control\_rep2 control\_rep3

Reset Samples!

Please add new comparisons for DE analysis!

**Condition 1**

exper\_rep1 exper\_rep2 exper\_rep3

DE Method: DESeq2 Fit Type: parametric Beta Prior: 0 Test Type: Wald row.sum filter: 10

**Condition 2**

control\_rep1 control\_rep2 control\_rep3

**Condition 3**

exper\_rep1 exper\_rep2 exper\_rep3

DE Method: EdgeR Normalization: TMM Dispersion: 0 Test Type: exactTest row.sum filter: 10

**Condition 4**

control\_rep1 control\_rep2 control\_rep3


**Condition 5**

exper\_rep1 exper\_rep2 exper\_rep3

DE Method: Limma Normalization: TMM Fit Type: ls Norm. Bet. Arrays: none row.sum filter: 10

**Condition 6**

control\_rep1 control\_rep2 control\_rep3

Add New Comparison Remove 

Submit!

The second option, ‘Go to QC plots!’, takes you to a page where you can view quality control metrics on your data input. The page opens with an all-to-all plot displaying the correlation between each sample. Left of this plot is a panel which contains various parameters to alter the look of your plot such as width and height. You can change the type of dataset being viewed within these QC plots by selecting the dataset you want at the top of the left panel. Each dataset can have its own unique parameters to change and alter your QC plots.

In addition to the all-to-all plot, you can also view a heatmap representation of your data as well as a Principal Component Analysis (PCA) plot by selecting the specific plot option on the left panel under ‘QC Plots’. You can also select the type of clustering and distance method for the heatmap produced to further customize your quality control measures. Additionally, you can also view a density plot for your sample data for your raw data and the data after normalization (Figure 6). IQR and Density plots are another great visualization too to help you spot outliers within your sample data incase you want to remove or look into any possible discrepancies.

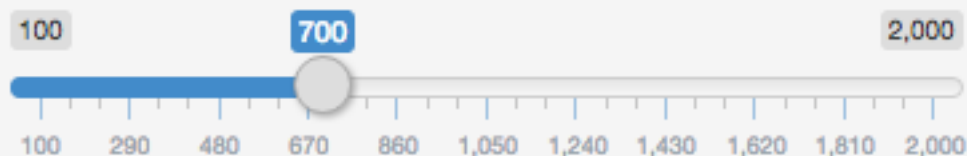
# DEBrowser

## QC Plots:

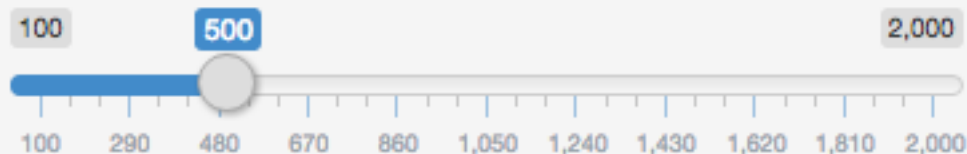
- ☒ All2All  
☐ Heatmap  
☐ PCA

Submit!

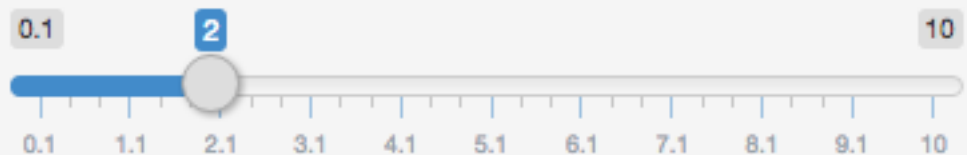
### width



### height




### corr font size

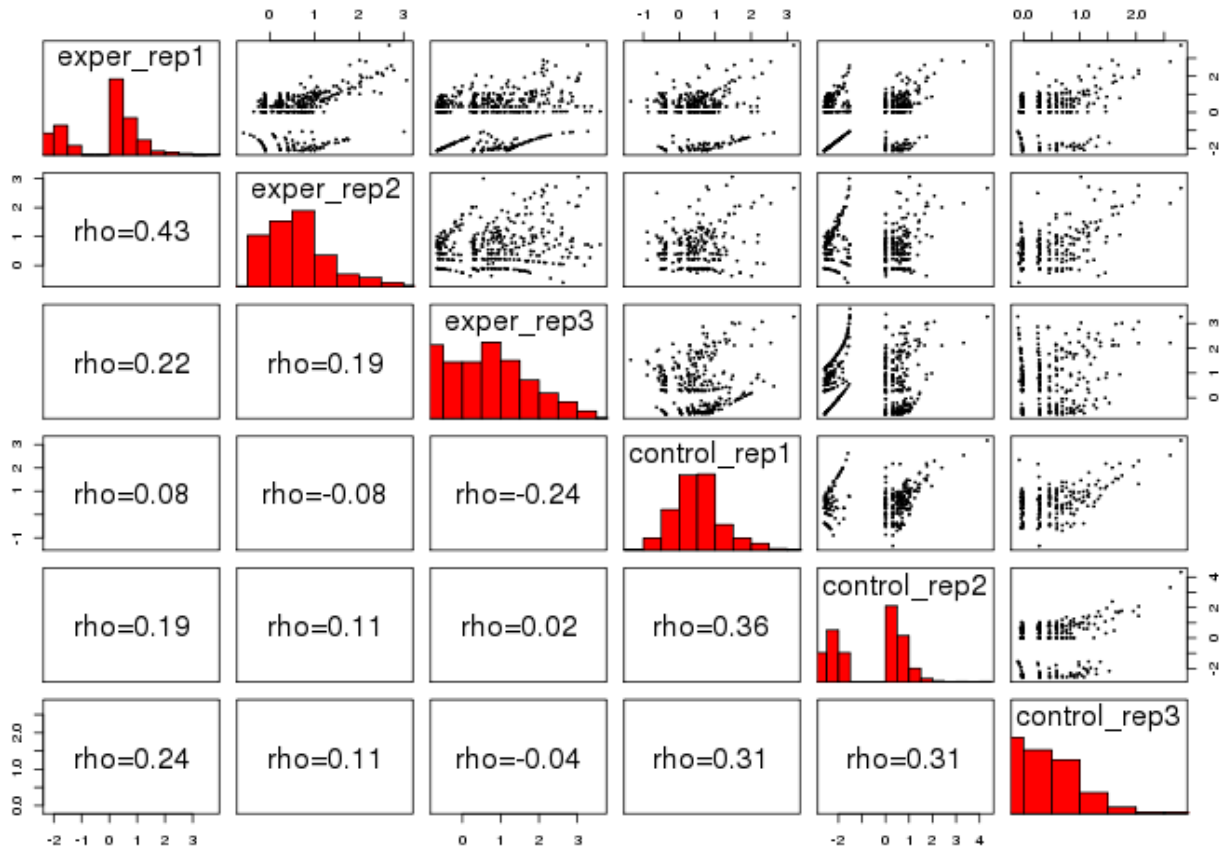


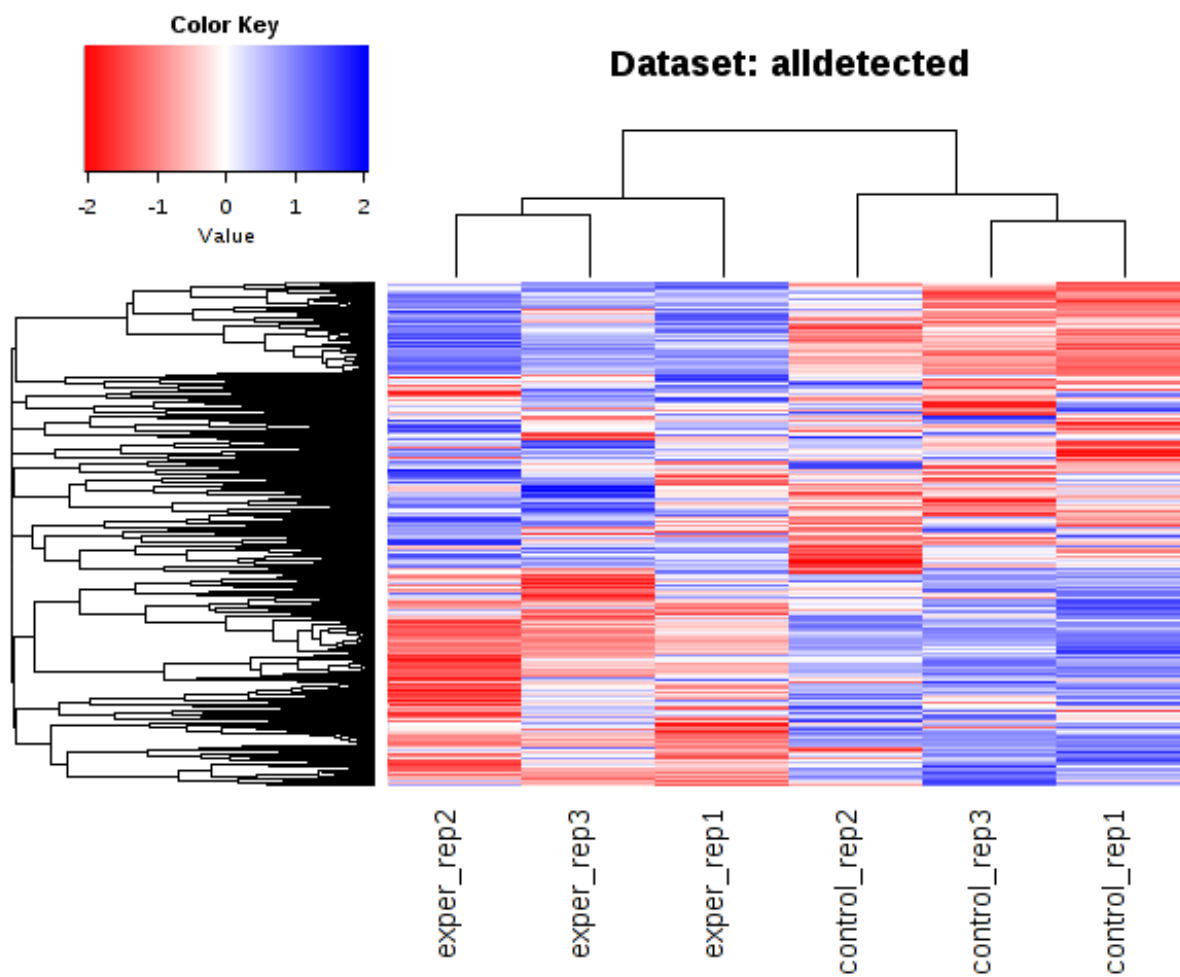
 Download Plot

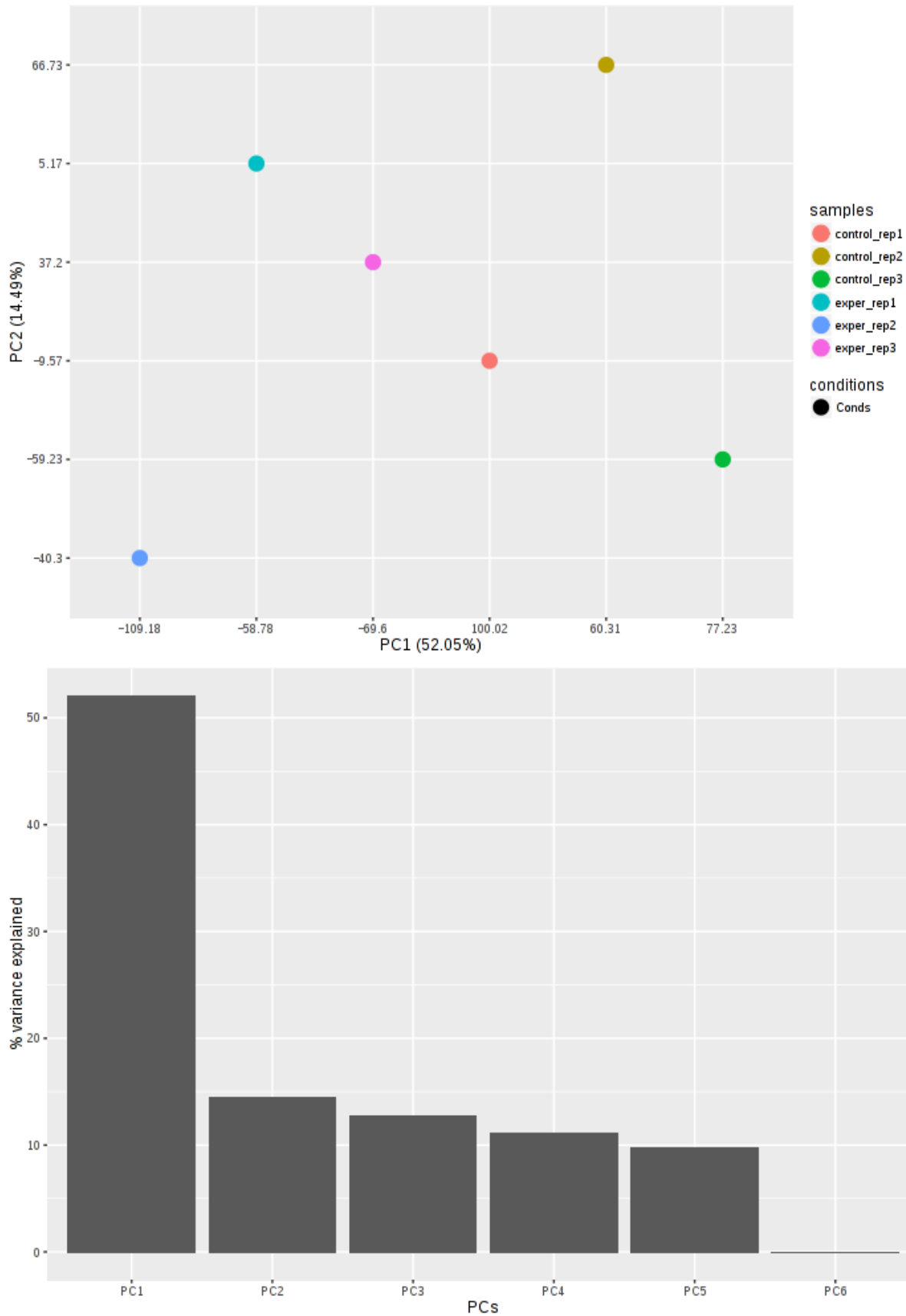
### Choose a dataset:

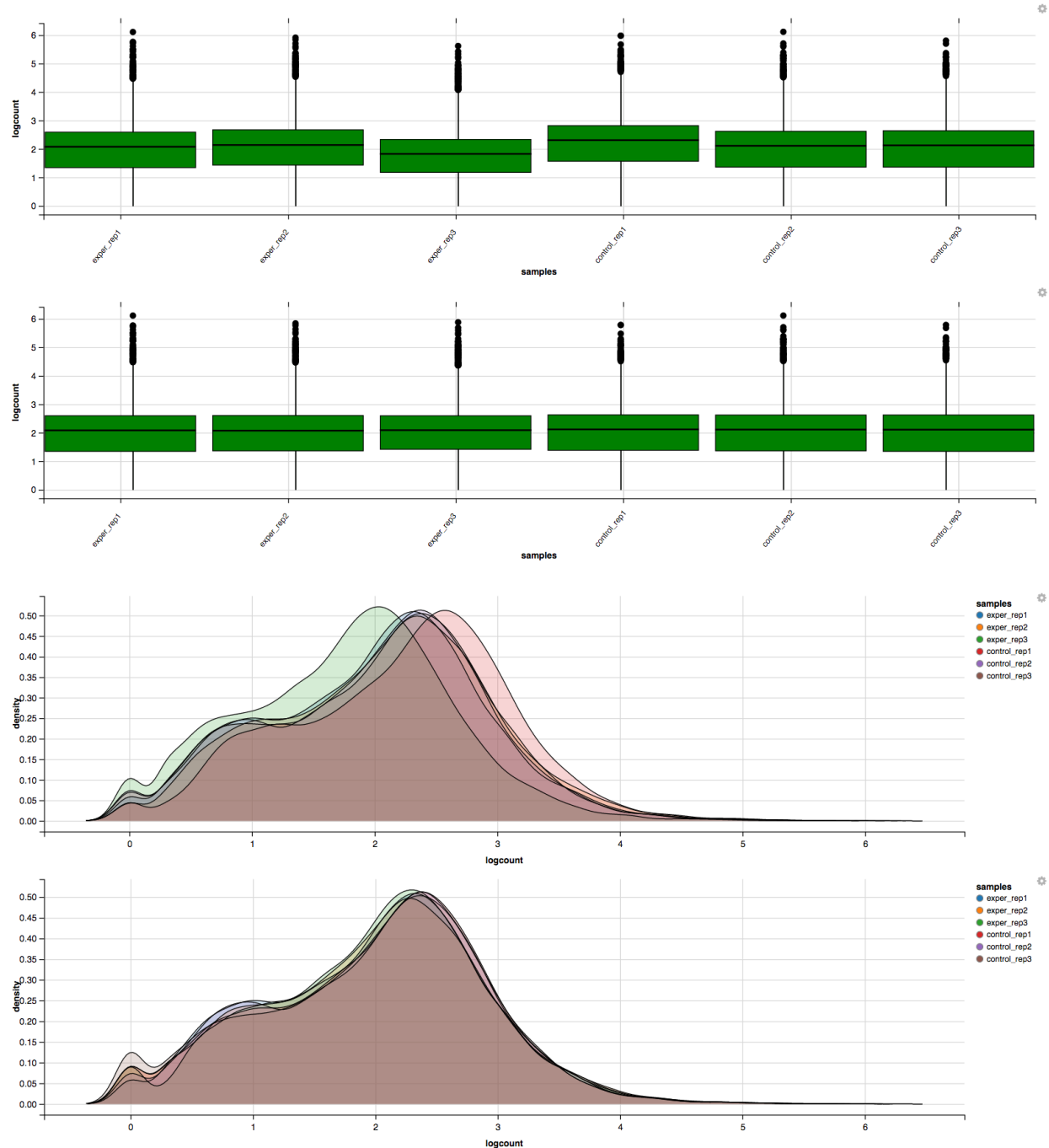
most-varied

 Download Data









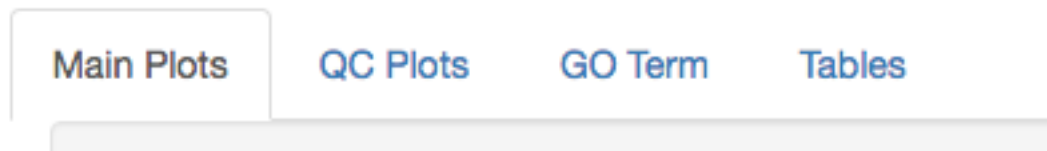
You can also view specific tables of your input data for each type of dataset available and search for a specific geneset by inputting a comma-separated list of genes or regex terms to search for in the search box within the left panel. To view these tables, you must select the tab labeled ‘Tables’ as well as the dataset from the dropdown menu on the left panel.

If you ever want to change your file/condition parameters, or even add a new set of comparisons, you can always return to the ‘Data Prep’ to change and resubmit your data. To completely start over, you can also hit the ‘Reset’ button on the ‘Data Prep’ page.

Once you are happy with your dataset and you have selected your conditions within the ‘DE Analysis’ section, you can then hit ‘Submit!’ to begin.

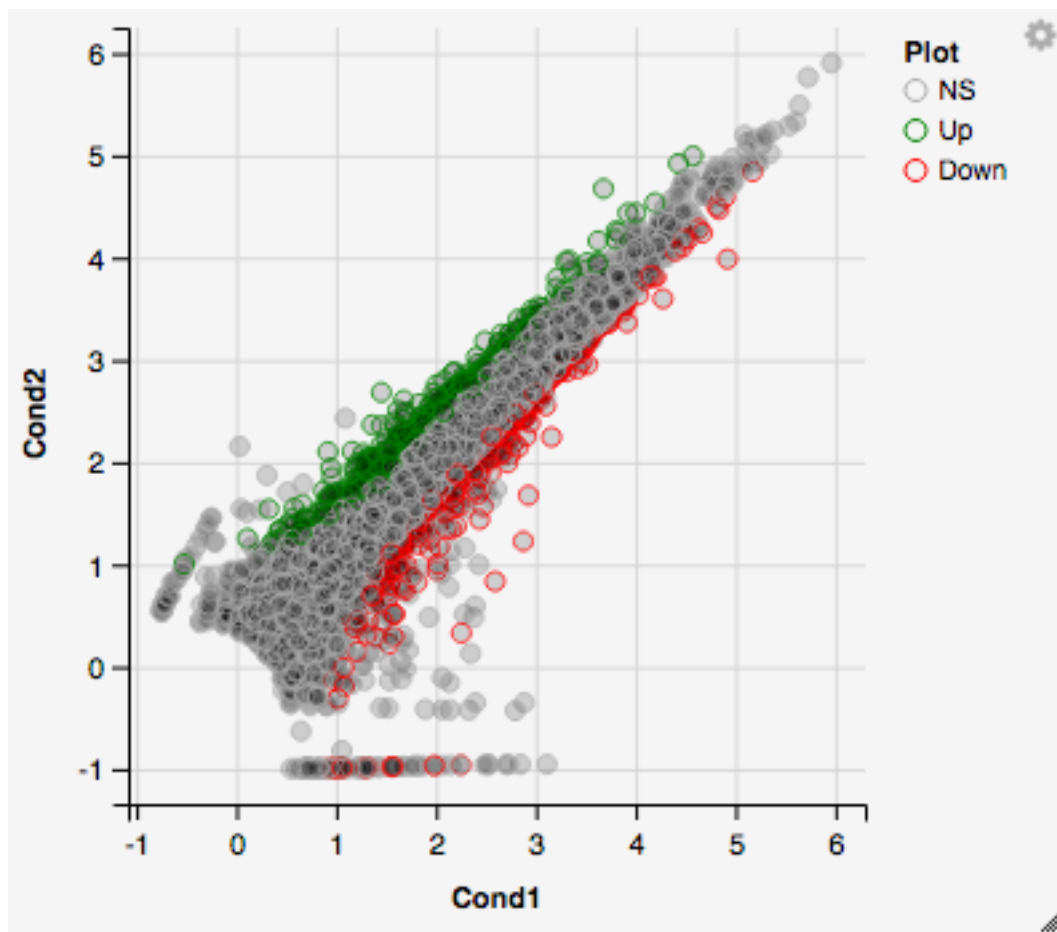
## 1.2 The Main Plots

After clicking on the ‘Submit!’ button, DESeq2 will analyze your comparisons and store the results into separate data tables. Shiny will then allow you to access this data, with multiple interactive features, at the click of a button. It is important to note that the resulting data produced from DESeq2 is normalized. Upon finishing the DESeq2 analysis, a tab based menu will appear with multiple options.

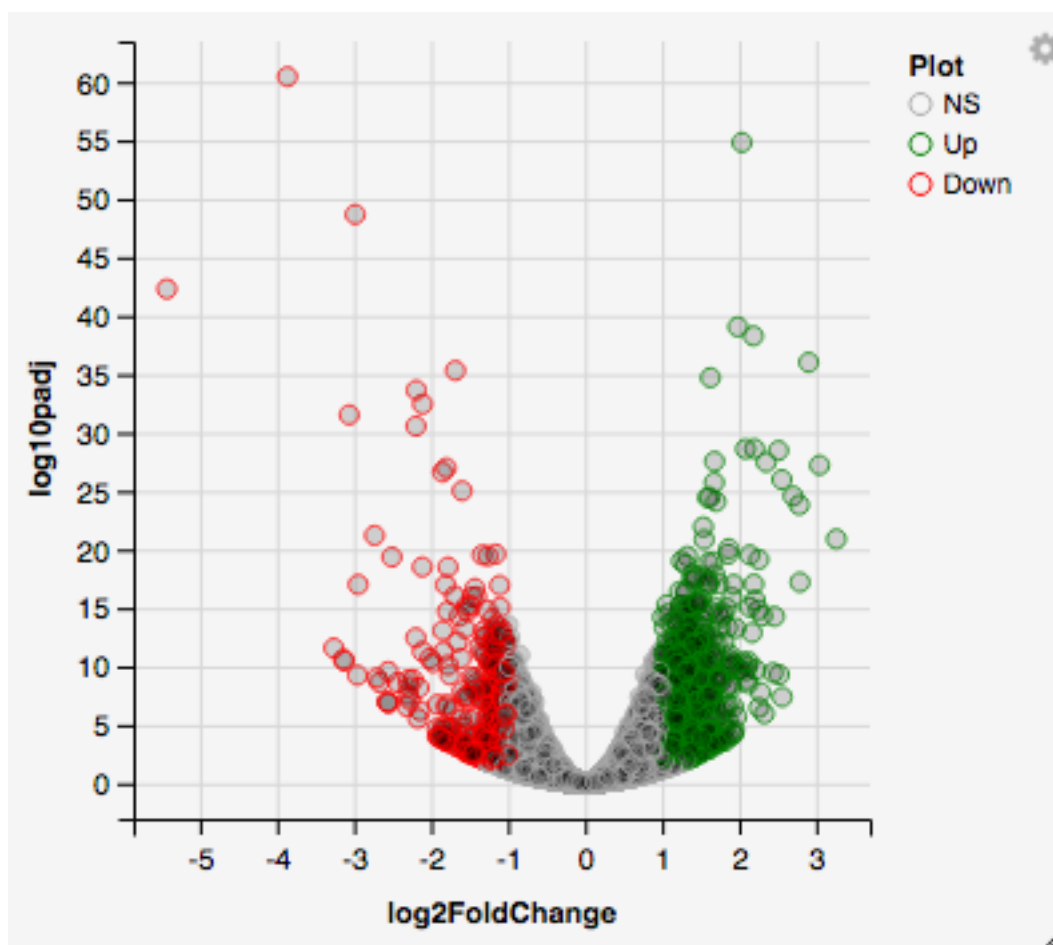


The first tab, the ‘Main Plots’ section, is where you will be able to view the interactive results plots. Plot choices include:

Scatter plot

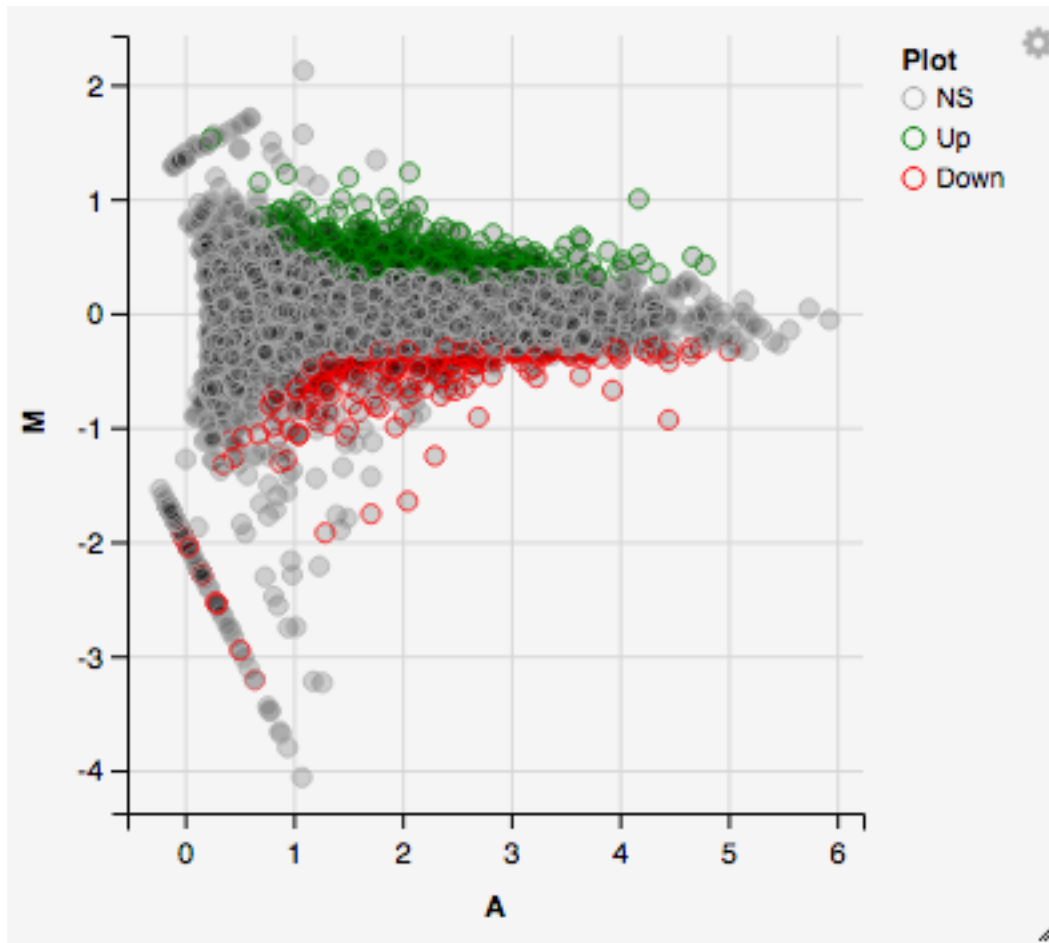


Volcano plot

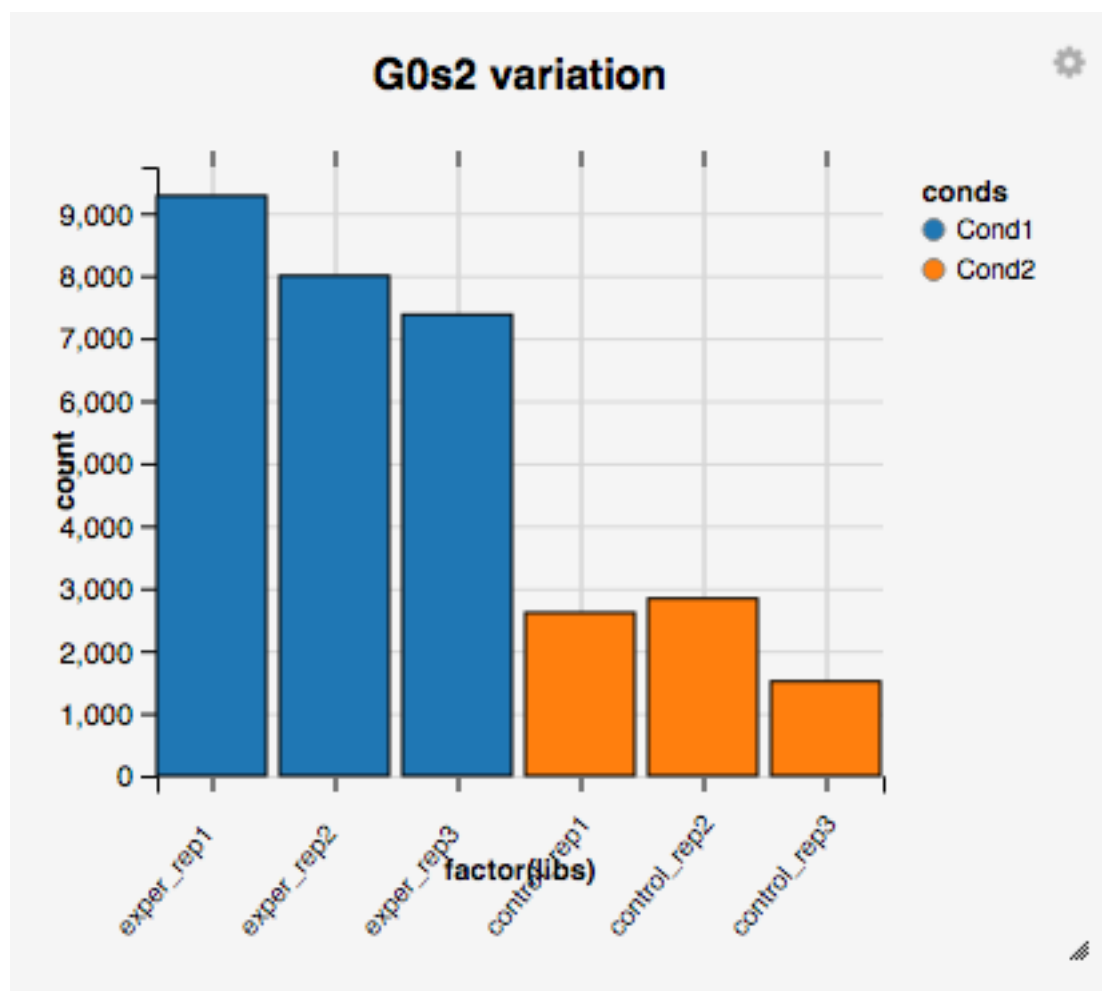


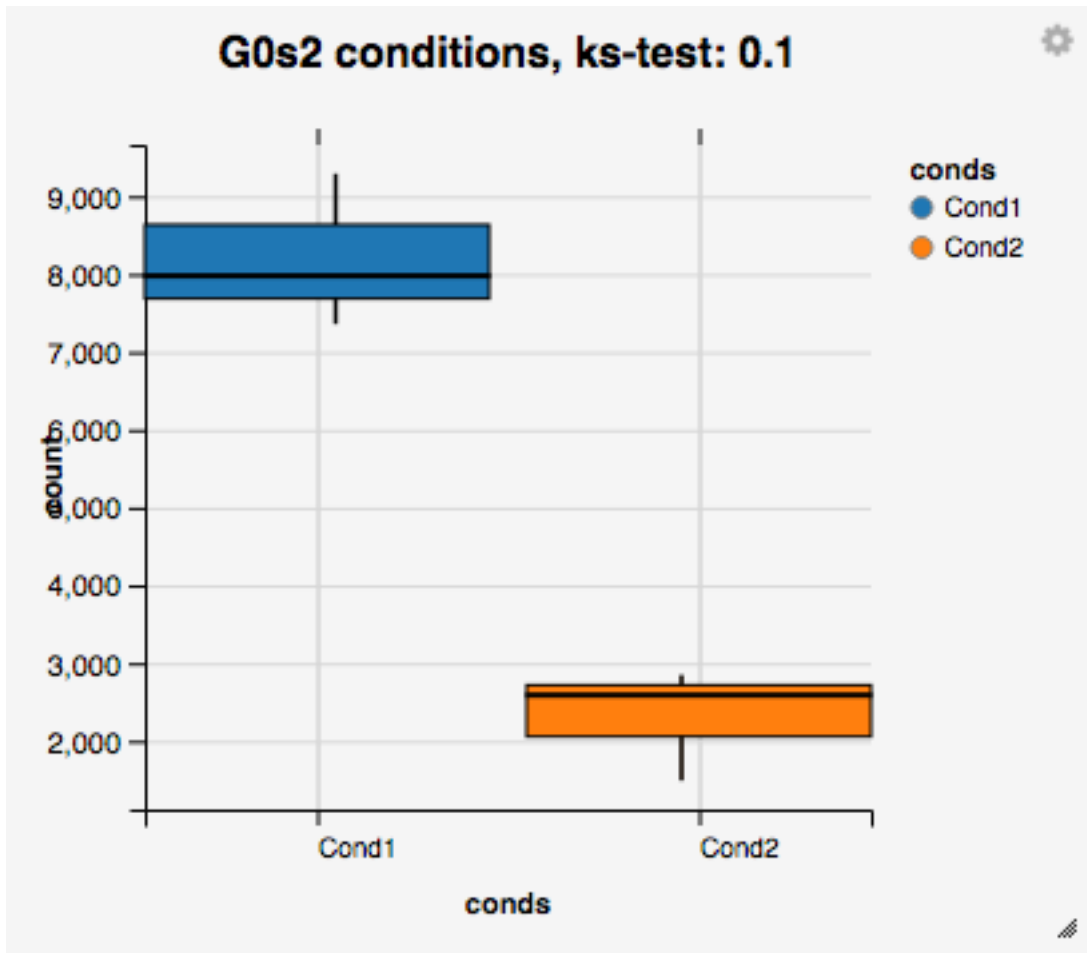
MA plot



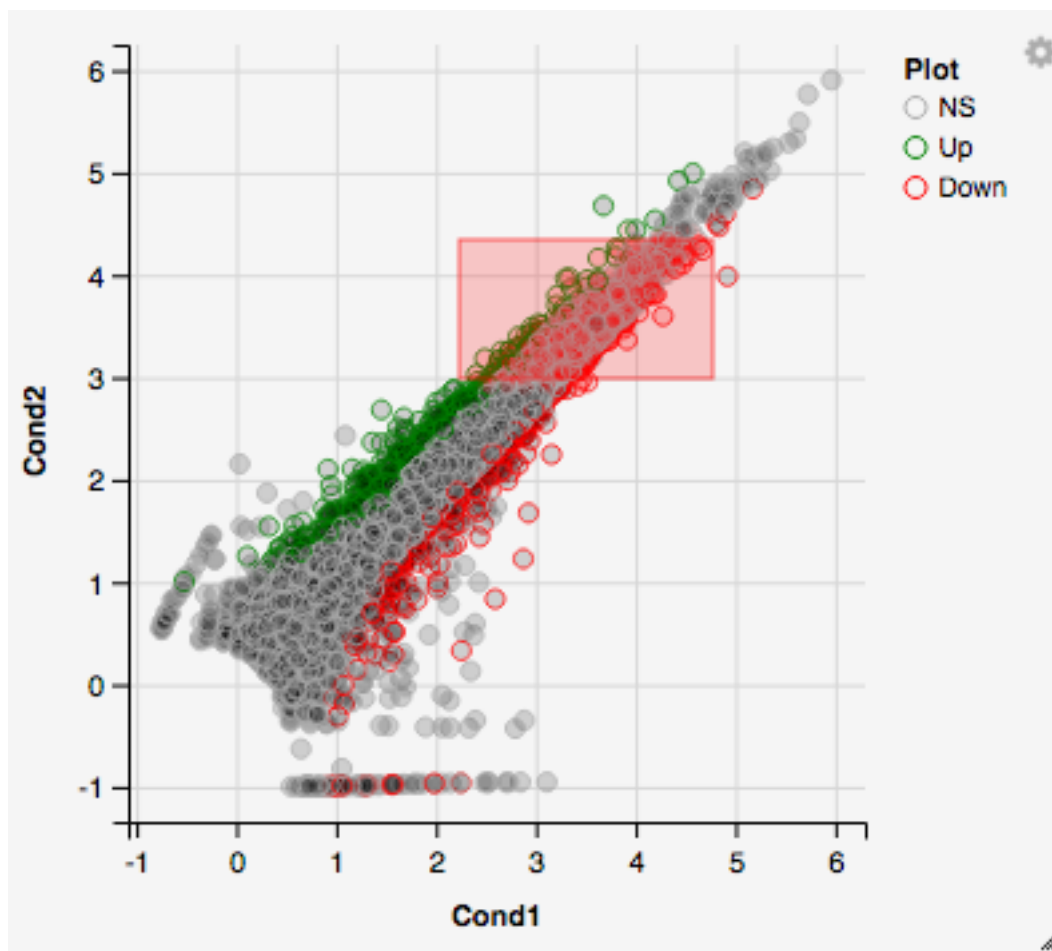


You can hover over the scatterplot points to display more information about the point selected. A few bargraphs will be generated for the user to view as soon as a scatterplot point is hovered over.

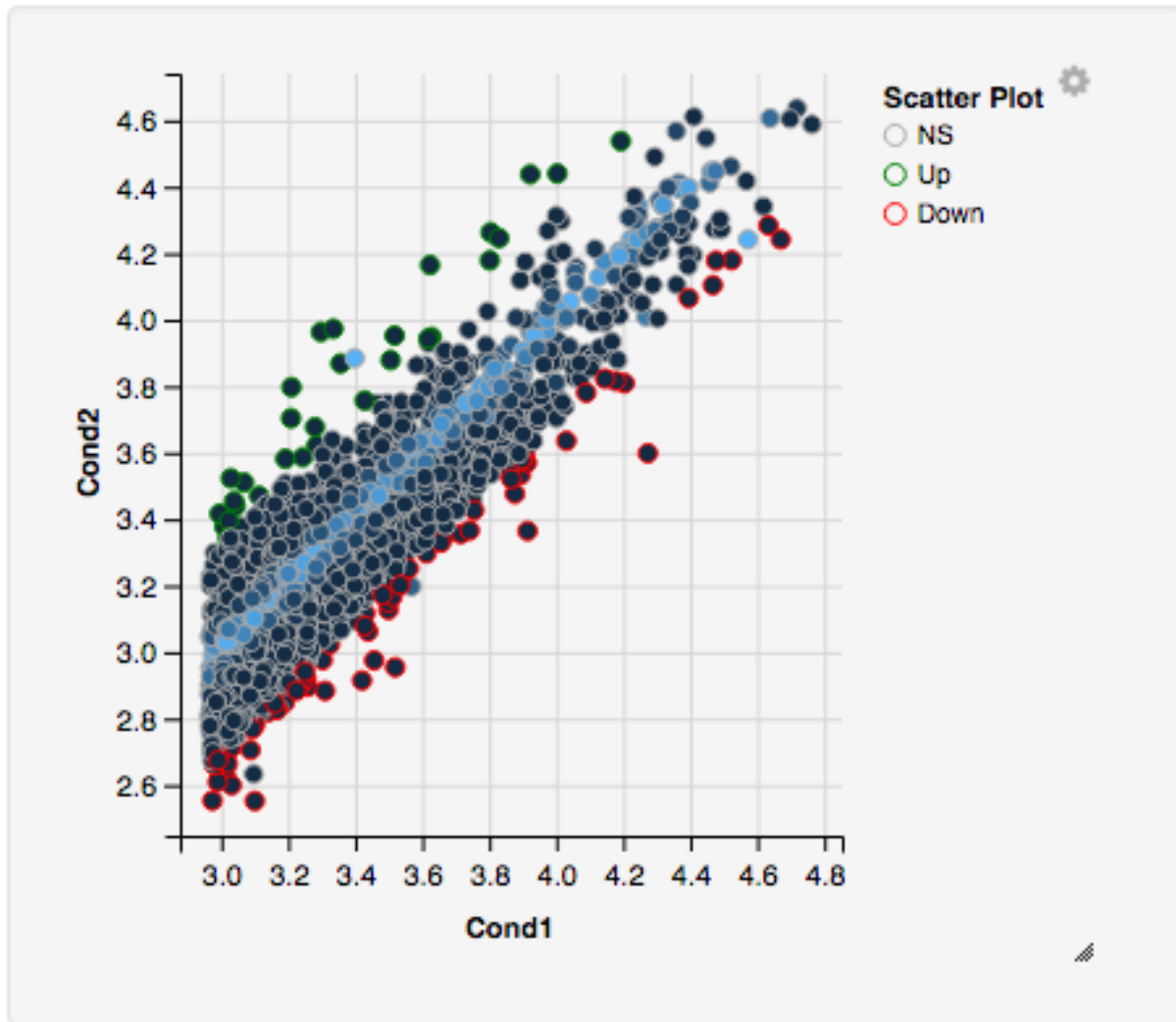




You can also select a specific region within the scatter plot and zoom in on the selected window.



Once you've selected a specific region, a new scatterplot of the selected area will appear on the right



You also have a wide array of options when it comes to fold change cut-off levels, padj cut-off values, which comparison set to use, and dataset of genes to analyze.

**Main Plots:**

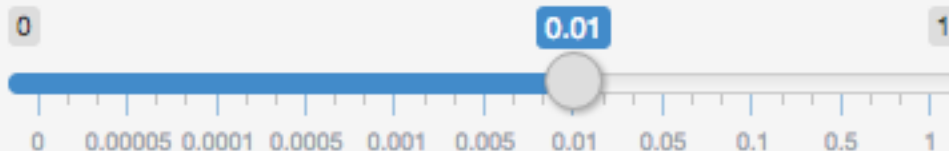
- ☒ Scatter  
☐ VolcanoPlot  
☐ MAPlot

**Submit!****Choose a dataset:**

up+down ▼

 Download Data**Search**

Regular expressions can be used Ex: ^Al => Al..., Al\$ => ...al

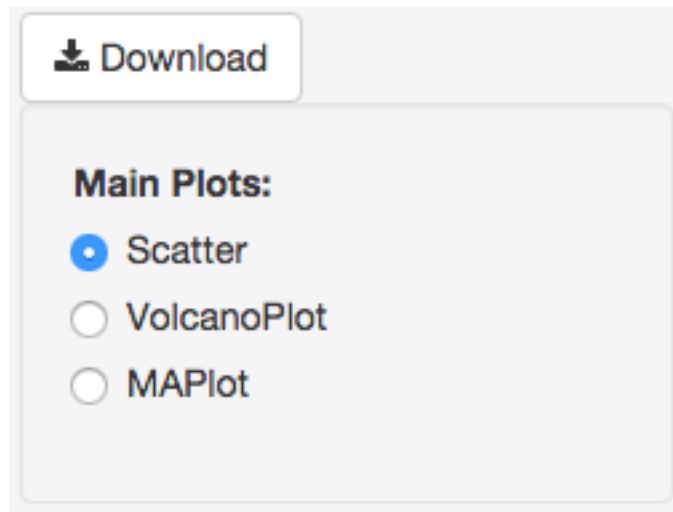
**Filter****padj value cut off****or padj**

0.01

It is important to note that when conducting multiple comparisons, the comparisons are labeled based on the order that they are input. If you don't remember which samples are in your current comparison you can always view the samples in each condition at the top of the main plots.

**Cond1:** exper\_rep1,exper\_rep2,exper\_rep3 vs. **Cond2:** control\_rep1,control\_rep2,control\_rep3

If you can select the type of plot at the bottom of the filter tab.



You can download the results in CSV or TSV format by selecting your 'File type' and clicking the 'download' button once you've ran DESeq. You can also download the plot or graphs themselves by clicking on the gear in the upper-left corner of each plot or graph.

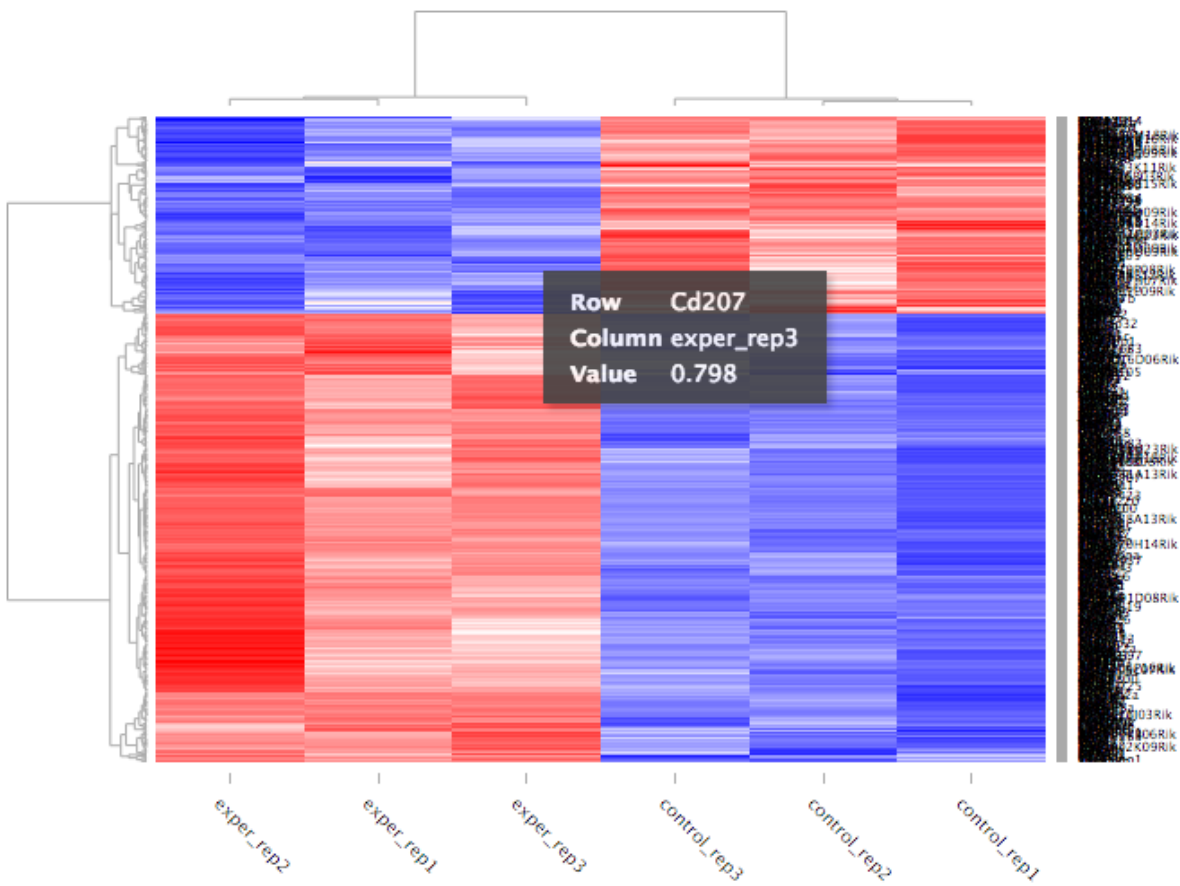
## 1.3 Quality Control Plots

Selecting the 'QC Plots' tab will take you to the quality control plots section. These QC plots are very similar to the QC plots shown before running DESeq and the dataset being used here depends on the one you select in the left panel. In addition to the all-to-all plot shown within the previous QC analysis, users can also view a heatmap, PCA, IQR, and density plots of their analyzed data by selecting the proper plot on the left menu. You can also choose the appropriate clustering and distance method you would like to use for this heatmap just about the plot just like in the previous QC section.

For additional information about the clustering methods used, you can consult [this website](#).

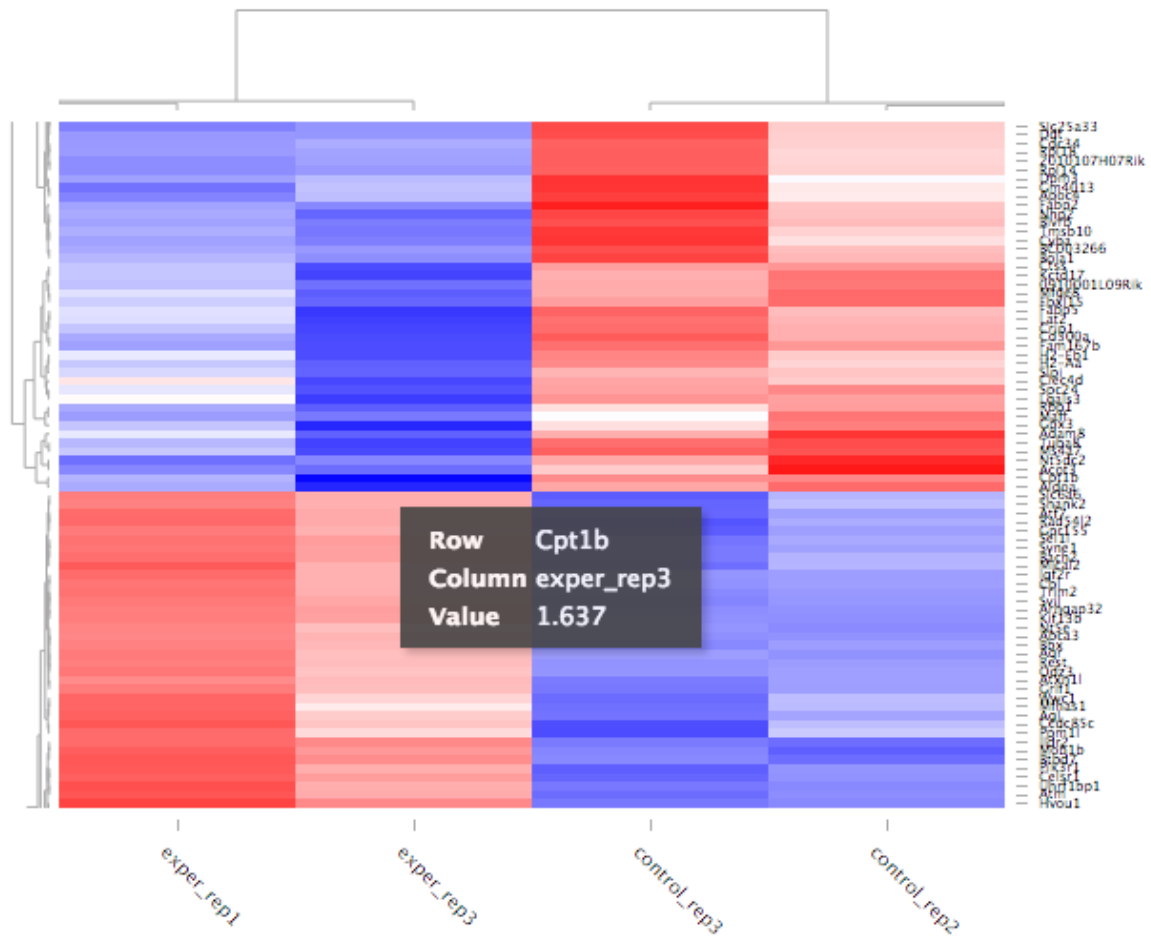
For additional information about the distance methods used, you can consult [here](#).

For distances other than 'cor', the distance function defined will be  $(1 - (\text{the correlation between samples}))$ . Each qc plot also has options to adjust the plot height and width, as well as a download button for a pdf output located above each plot. For the Heatmap, you can also view an interactive session of the heatmap by selecting the 'Interactive' checkbox before submitting your heatmap request. Make sure that before selecting the interactive heatmap option that your dataset being used is 'Up+down'. Just like in the Main Plots, you can click and drag to create a selection. To select a specific portion of the heatmap, make sure to highlight the middle of the heatmap gene box in order to fully select a specific gene. This selection can be used later within the GO Term plots for specific queries on your selection.



Your selection will also zoom in for better viewing resolution.





## 1.4 Heat Maps

The heatmap is a great way to analyze replicate results of genes all in one simple plot (Figure 17). Users have the option to change the clustering method used as well as the distance method used to display their heatmap. In addition, you can also change the size of the heatmap produced and adjust the p-adjust and fold change cut off for this plot as well. Once all of the parameters have been set, click the ‘Submit!’ button at the bottom of the left menu to generate your heatmap.

## Used clustering and linkage methods in heatmap

- **complete:** Complete-linkage clustering is one of the linkage method used in hierarchical clustering. In each step of clustering, closest cluster pairs are always merged up to a specified distance threshold. Distance between clusters for complete link clustering is the maximum of the distances between the members of the clusters.
- **ward D2:** Ward method aims to find compact and spherical clusters. The distance between two clusters is calculated by the sum of squared deviations from points to centroids. “ward.D2” method uses criterion (Murtagh and Legendre 2014) to minimize ward clustering method. The only difference ward.D2 and ward is the dissimilarities from ward method squared before cluster updating. This method tends to be sensitive to the outliers.
- **single:** Distance between clusters for single linkage is the minimum of the distances between the members of

the clusters.

- **average:** Distance between clusters for average linkage is the average of the distances between the members of the clusters.
- **mcquitty:** mcquitty linkage is when two clusters are joined, the distance of the new cluster to any other cluster is calculated by the average of the distances of the soon to be joined clusters to that other cluster.
- **median:** This is a different averaging method that uses the median instead of the mean. It is used to reduce the effect of outliers.
- **centroid:** The distance between cluster pairs is defined as the Euclidean distance between their centroids or means.


## Used distance methods in heatmap

- **cor:**  $1 - \text{cor}(x)$  are used to define the dissimilarity between samples. It is less sensitive to the outliers and scaling.
- **euclidean:** It is the most common use of distance. It is sensitive to the outliers and scaling. It is defined as the square root of the sum of the square differences between gene counts.
- **maximum:** The maximum distance between two samples is the sum of the maximum expression value of the corresponding genes.
- **manhattan:** The Manhattan distance between two samples is the sum of the differences of their corresponding genes.
- **canberra:** Canberra distance is similar to the Manhattan distance and it is a special form of the Minkowski distance. The difference is that the absolute difference between the gene counts of the two genes is divided by the sum of the absolute counts prior to summing.
- **minkowsky:** It is generalized form of euclidean distance.

You can also select to view an interactive version of the heatmap by clicking on the 'Interactive' checkbox on the left panel under the height and width options. Selecting this feature changes the heatmap into an interactive version with two colors, allowing you to select specific genes to be compared within the GO term plots. In order to use the interactive heatmap selection within your GO term query, you must use either the up+down dataset or the most varied dataset for the heatmap display.

## 1.5 GO Term Plots

The next tab, 'GO Term', takes you to the ontology comparison portion of DEBrowser. From here you can select the standard dataset options such as p-adjust value, fold change cut off value, which comparison set to use, and which dataset to use on the left menu. In addition to these parameters, you also can choose from the 4 different ontology plot options: 'enrichGO', 'enrichKEGG', 'Disease', and 'compareCluster'. Selecting one of these plot options queries their specific databases with your current DESeq results.

 Download

**Go Plots:**

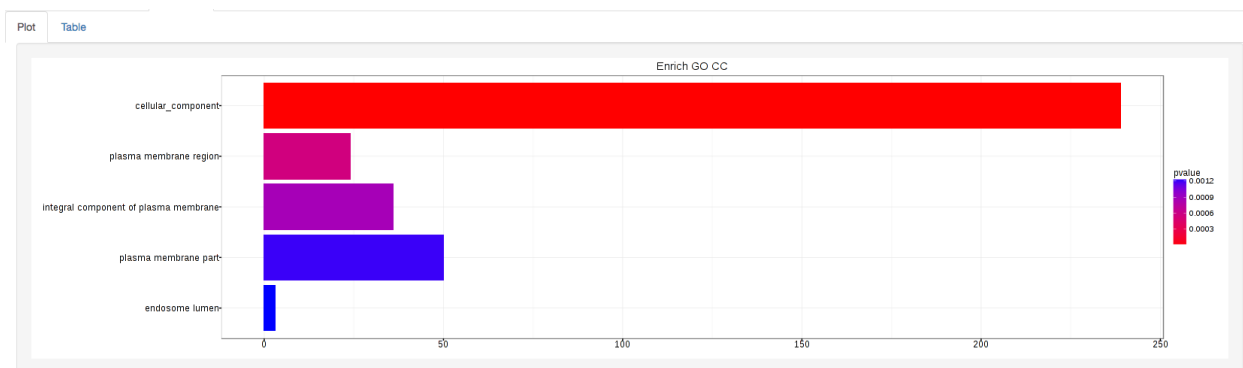
- ☒ enrichGO
- ☐ enrichKEGG
- ☐ Disease
- ☐ compareClusters

Your GO plots include:

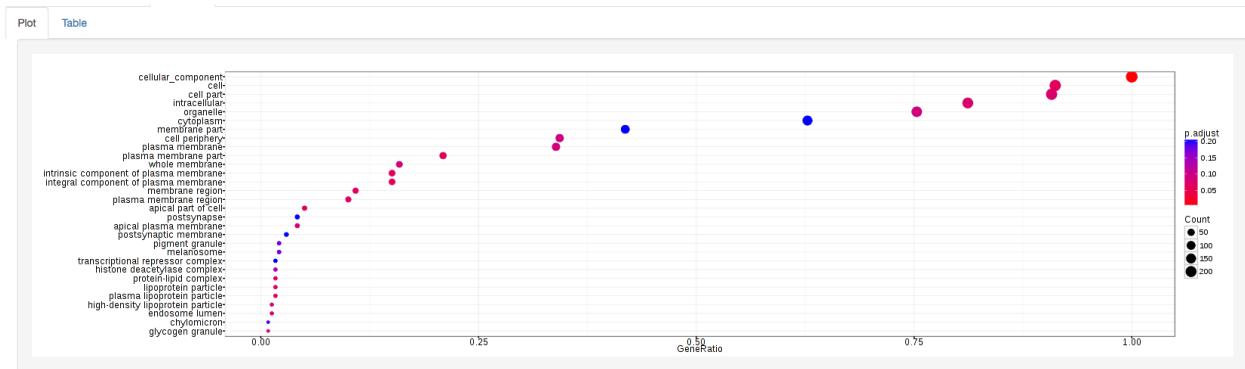
- enrichGO - use enriched GO terms
- enrichKEGG - use enriched KEGG terms
- Disease - enriched for diseases
- compareClusters - comparison of your clustered data

The types of plots you will be able to generate include:

Summary plot:



GOdotplot:



Changing the type of ontology to use will also produce custom parameters for that specific ontology at the bottom of the left option panel.

Once you have adjusted all of your parameters, you may hit the submit button in the top right and then wait for the results to show on screen!

## 1.6 Data Tables

The lasttab at the top of the screen displays various different data tables. These datatables include:

- All Detected
- Up Regulated
- Down Regulated
- Up+down Regulated
- Selected scatterplot points
- Most varied genes
- Comparison differences

[Main Plots](#)
[QC Plots](#)
[GO Term](#)
[Tables](#)

Show **10** entries

Search:

	exper_rep1	exper_rep2	exper_rep3	control_rep1	control_rep2	control_rep3	padj	log2FoldChange	pvalue	foldChange	log10padj
Syng1	7.135273	25.37579	38.917939	7.871420	2.031122e+00	2.798633	2.670881e-03	-1.479199	5.713910e-04	0.3586880	2.573346
Pkp1	23.310344	25.34808	21.862658	5.015264	6.055972e+00	2.796967	6.137991e-04	-1.577093	1.019709e-04	0.3351565	3.211974
AK206687	4.065426	15.31221	9.379557	0.000000	6.271097e-03	0.000000	7.288816e-04	-1.654479	1.248060e-04	0.3176525	3.137343
Rps6ka3	15.280315	12.32422	18.450445	47.740842	7.445891e+01	33.702584	3.250601e-04	1.421371	4.866031e-05	2.6783986	3.488036
D330041H03Rik	80.087331	86.16724	28.416393	6.998361	1.009576e+01	2.745741	2.391427e-09	-2.429347	7.245195e-11	0.1856495	8.621343
Nr2c2	86.363440	24.88276	56.424369	229.305017	1.912994e+02	201.460333	3.523672e-08	1.669908	1.513202e-09	3.1819421	7.453005
Zfhx2	83.320057	54.94153	47.574396	219.240720	2.003584e+02	189.868073	6.859476e-10	1.547058	1.754693e-11	2.9222062	9.163709
Fcgbp	13.176789	13.96867	18.962477	3.278481	1.019879e+00	3.796185	4.634870e-03	-1.412922	1.087597e-03	0.3755504	2.333962
Ly6d	339.117011	334.96468	408.621931	113.570277	1.349614e+02	88.423299	4.973157e-15	-1.628726	3.766736e-17	0.3233737	14.303368
C78339	101.615743	51.38191	130.392025	327.724952	3.262240e+02	294.855319	2.535484e-10	1.612138	5.688760e-12	3.0570451	9.595939

Showing 1 to 10 of 355 entries

Previous
1
2
3
4
5
...
36
Next

All of the tables tables, except the Comparisons table, contain the following information:

- ID - The specific gene ID
- Sample Names - The names of the samples given and they're corresponding tmm normalized counts
- Conditions - The log averaged values
- padj - adjusted value
- log2FoldChange - The Log2 fold change
- foldChange - The fold change
- log10padj - The log 10 padjusted value

The Comparisons table generates values based on the number of comparisons you have conducted. For each pairwise comparison, these values will be generated:

- Values for each sample used
- foldChange of comparison A vs B
- pvalue of comparison A vs B
- padj value of comparison A vs B

Main Plots
QC Plots
GO Term
Tables

Show 10 entries
Search:

	exper_rep1	exper_rep2	exper_rep3	control_rep1	control_rep2	control_rep3	foldChange.C1.vs.C2	padjC1.vs.C2	foldChange.C3.vs.C4	padjC3.vs.C4
1110007C09Rik	145.00	211.01	77.00	75.00	52.00	67.00	0.3693497	1.057546e-08	0.3416683	2.589037e-06
1810044D09Rik	73.00	90.00	36.00	32.00	27.00	22.00	0.3562541	3.755707e-06	0.3801446	9.042321e-04
2010003K11Rik	885.00	826.00	403.00	375.00	271.39	334.00	0.3617310	1.140383e-16	0.3338943	1.195500e-11
2410076I21Rik	34.97	64.00	21.97	2.94	13.85	16.00	0.3390859	5.020396e-04	0.3149882	1.752537e-03
2610005L07Rik	248.96	154.06	127.61	1061.71	632.64	468.95	2.6469646	2.756012e-09	3.0490664	2.185518e-08
3930402G23Rik	69.00	49.00	20.00	12.00	8.00	2.00	0.2028349	2.516955e-08	0.2433717	1.730217e-05
4930470H14Rik	36.74	42.01	20.92	242.11	139.65	100.34	3.1135333	1.197587e-08	3.3840596	3.607676e-07
4930565N06Rik	6.01	1.00	0.00	30.01	19.00	12.00	2.6527298	5.414730e-03	2.2782137	4.006262e-02
4932438A13Rik	246.00	234.00	130.00	1434.43	754.31	678.00	3.1780469	1.199287e-19	3.4064309	2.772699e-14
4933411K16Rik	4.00	5.00	5.00	56.96	41.00	15.00	3.1104165	4.043031e-04	3.9322051	1.179784e-04

Showing 1 to 10 of 440 entries
Previous
1
2
3
4
5
...
44
Next

You can further customize and filter each specific table a multitude of ways. For unique table or dataset options, select the type of table dataset you would like to customize on the left panel under ‘Choose a dataset’ to view it’s additional options. All of the tables have a built in search function at the top right of the table and you can further sort the table by column by clicking on the column header you wish to sort by. The ‘Search’ box on the left panel allows for multiple searches via a comma-seperated list. You can additionally use regex terms such as “^al” or “\*lm” for even more advanced searching. This search will be applied to wherever you are within DEBrowser, including both the plots and the tables.

---

## Local Install Guide

---

### 2.1 Quick Local Install

Running these simple command will launch the DEBrowser within your local machine.

Before you start; First, you will have to install R and/or RStudio. (On Fedora/Red Hat/CentOS, these packages have to be installed; openssl-devel, libxml2-devel, libcurl-devel, libpng-devel)

You can download the source code or the tar file for DEBrowser [here](#).

**Installation instructions from source:**

1. Install the required dependencies by running the following commands in R or RStudio.

```
source("http://www.bioconductor.org/biocLite.R")
biocLite("debrowser")
```

2. Start R and load the library

```
library(debrowser)
```

3. Start DE browser

```
startDEBrowser()
```

Once you run 'startDEBrowser()' shiny will launch a web browser with your local version of DEBrowser running and ready to use!

For more information about DEBrowser, please visit our Quick-start Guide or our DESeq/DEBrowser section within [readthedocs](#).





---

## DESeq/DEBrowser

---

This guide contains a brief description of DESeq2 used within the DEBrowser

### 3.1 Introduction

Differential gene expression analysis has become an increasingly popular tool in determining and viewing up and/or down expressed genes between two sets of samples. The goal of Differential gene expression analysis is to find genes or transcripts whose difference in expression, when accounting for the variance within condition, is higher than expected by chance. [DESeq2](#) is an R package available via Bioconductor and is designed to normalize count data from high-throughput sequencing assays such as RNA-Seq and test for differential expression (Love et al. 2014). For more information on the DESeq2 algorithm, you can visit [this website](#). With multiple parameters such as padjust values, log fold changes, and plot styles, altering plots created with your DE data can be a hassle as well as time consuming. The Differential Expression Browser uses DESeq2, EdgeR, and Limma coupled with shiny to produce real-time changes within your plot queries and allows for interactive browsing of your DESeq results. In addition to DESeq analysis, DEBrowser also offers a variety of other plots and analysis tools to help visualize your data even further.

### 3.2 DESeq2

For the details please check the user guide. [DESeq2 userguide](#)

DESeq2 performs multiple steps in order to analyze the data you've provided for it. The first step is to indicate the condition that each column (experiment) in the table represent. You can group multiple samples into one condition column. DESeq2 will compute the probability that a gene is differentially expressed (DE) for ALL genes in the table. It outputs both a nominal and a multiple hypothesis corrected p-value (p<sub>adj</sub>) using a negative binomial distribution.

### 3.3 Un-normalized counts

DESeq2 requires count data as input obtained from RNA-Seq or another high-throughput sequencing experiment in the form of matrix values. Here we convert un-integer values to integer to be able to run DESeq2. The matrix values should be un-normalized, since DESeq2 model internally corrects for library size. So, transformed or normalized values such as counts scaled by library size should not be used as input. Please use edgeR or limma for normalized counts.

## 3.4 Used parameters for DESeq2

- **fitType:** either “parametric”, “local”, or “mean” for the type of fitting of dispersions to the mean intensity. See `estimateDispersions` for description.
- **betaPrior:** whether or not to put a zero-mean normal prior on the non-intercept coefficients. See `nbinomWaldTest` for description of the calculation of the beta prior. By default, the beta prior is used only for the Wald test, but can also be specified for the likelihood ratio test.
- **testType:** either “Wald” or “LRT”, which will then use either Wald significance tests (defined by `nbinomWaldTest`), or the likelihood ratio test on the difference in deviance between a full and reduced model formula (defined by `nbinomLRT`)
- **rowsum.filter:** regions/genes/isoforms with total count (across all samples) below this value will be filtered out

## 3.5 EdgeR

For the details please check the user guide. [EdgeR userguide](#).

## 3.6 Used parameters for EdgeR

- **Normalization:** Calculate normalization factors to scale the raw library sizes. Values can be “TMM”, “RLE”, “upperquartile”, “none”.
- **Dispersion:** either a numeric vector of dispersions or a character string indicating that dispersions should be taken from the data object.
- **testType:** `exactTest` or `glmLRT`. `exactTest`: Computes p-values for differential abundance for each gene between two samples, conditioning on the total count for each gene. The counts in each group are assumed to follow a binomial distribution. `glmLRT`: Fits a negative binomial generalized log-linear model to the read counts for each gene and conducts gene-wise statistical tests.
- **rowsum.filter:** regions/genes/isoforms with total count (across all samples) below this value will be filtered out

## 3.7 Limma

For the details please check the user guide. [Limma userguide](#).

Limma is a package to analyse of microarray or RNA-Seq data. If data is normalized with spike-in or any other scaling, transformation or normalization method, Limma can be ideal. In that case, prefer limma rather than DESeq2 or EdgeR.

## 3.8 Used parameters for Limma

- **Normalization:** Calculate normalization factors to scale the raw library sizes. Values can be “TMM”, “RLE”, “upperquartile”, “none”.
- **Fit Type:** fitting method; “ls” for least squares or “robust” for robust regression
- **Norm. Bet. Arrays:** Normalization Between Arrays; Normalizes expression intensities so that the intensities or log-ratios have similar distributions across a set of arrays.

- **rowsum.filter:** regions/genes/isoforms with total count (across all samples) below this value will be filtered out

## 3.9 DEBrowser

DEBrowser utilizes [Shiny](#), a R based application development tool that creates a wonderful interactive user interface (UI) combined with all of the computing prowess of R. After the user has selected the data to analyze and has used the shiny UI to run DESeq2, the results are then input to DEBrowser. DEBrowser manipulates your results in a way that allows for interactive plotting by which changing padj or fold change limits also changes the displayed graph(s). For more details about these plots and tables, please visit our quickstart guide for some helpful tutorials.

For comparisons against other popular data visualization tools, see the table below.

	DEBrowser	MeV	Chipster	Galaxy	CummeRBund
Easy Install	✓	✓	✗	✓	✓
Full R Support	✓	✗	✗	✗	✓
DESeq2	✓	✓	✓	✓	✗
EDGeR	✓	✓	✓	✓	✗
Limma	✓	✓	✗	✓	✗
CuffDiff	✗	✗	✓	✓	✓
Web Interface	✓	✓	✓	✓	✗
Local Install	✓	✓	✓	✗	✓
Gene Searching with Regular Expression	✓	✓	✗	✗	✗
Interactive Heat Map	✓	✗	✓	✗	✗
Interactive Expression Scatter Plots	✓	✓	✓	✗	✗
JSON Object Reading Capabilities	✓	✗	✗	✗	✗
Interactive All-to-All Scatter Plots	✓	✗	✗	✗	✗
Multiple DE Tool Comparison Tables	✓	✗	✗	✗	✗
Built-in GO Term Analysis	✓	✓	✓	✓	✗
Cluster Selection on Heatmap for GO Term Analysis	✓	✗	✗	✗	✗
Interactive Visualization of Gene Highlighting	✓	✓	✓	✗	✗
Interactive Visualization of Biological Variation/Condition Comparison	✓	✗	✗	✗	✗

For more information on the programs compared against DEBrowser, please visit these pages:

- [MeV](#)
- [Chipster](#)
- [Galaxy](#)
- [CummeRbund](#)

## 3.10 References

1. Anders,S. et al. (2014) HTSeq - A Python framework to work with high-throughput sequencing data.
2. Chang,W. et al. (2016) shiny: Web Application Framework for R.
3. Chang,W. and Wickham,H. (2015) ggvis: Interactive Grammar of Graphics.
4. Giardine,B. et al. (2005) Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.*, 15, 1451–1455.
5. Howe,E.A. et al. (2011) RNA-Seq analysis in MeV. *Bioinformatics*, 27, 3209–3210.
6. Kallio,M.A. et al. (2011) Chipster: user-friendly analysis software for microarray and other high-throughput data. *BMC Genomics*, 12, 507.
7. Li,B. and Dewey,C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, 12, 323.
8. Love,M.I. et al. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, 15, 550.
9. Reese,S.E. et al. (2013) A new statistic for identifying batch effects in high-throughput genomic data that uses guided principal component analysis. *Bioinformatics*, 29, 2877–2883.
10. Reich,M. et al. (2006) GenePattern 2.0. *Nat. Genet.*, 38, 500–501.
11. Risso,D. et al. (2014) Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat. Biotechnol.*, 32, 896–902.
12. Ritchie,M.E. et al. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.*, 43, e47–e47.
13. Trapnell,C. et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.*, 7, 562–578.
14. Vernia,S. et al. (2014) The PPAR $\alpha$ -FGF21 hormone axis contributes to metabolic regulation by the hepatic JNK signaling pathway. *Cell Metab.*, 20, 512–525.
15. Murtagh, Fionn and Legendre, Pierre (2014). Ward’s hierarchical agglomerative clustering method: which algorithms implement Ward’s criterion? *Journal of Classification* 31 (forthcoming).

---

## DEBrowser Additions

---

In the future, we are going to add:

- Venn Diagrams to compare overlapping differentially expressed genes in different condition comparison results.
- Increase in the number of used clustering methods.
- Batch effect correction method(s).
- GO term analysis gene lists will be added for found GO categories.

This page will be updated as new capabilities are added to DEBrowser and/or if we begin new advancements.



---

## Indices and tables

---

- `genindex`
- `modindex`
- `search`