Analysis summary of project New Folder

Report generated by: robot, Thu Nov 15 15:48:16 CST 2018
using RobiNA version 1.2.4_build656

The analysis was run using count data imported from a precomputedcounts table file: APEX5_Hiseq_Hisat_Counts_V7.txt
Between the samples defined in the imported counts table, the following contrastshave been computed:

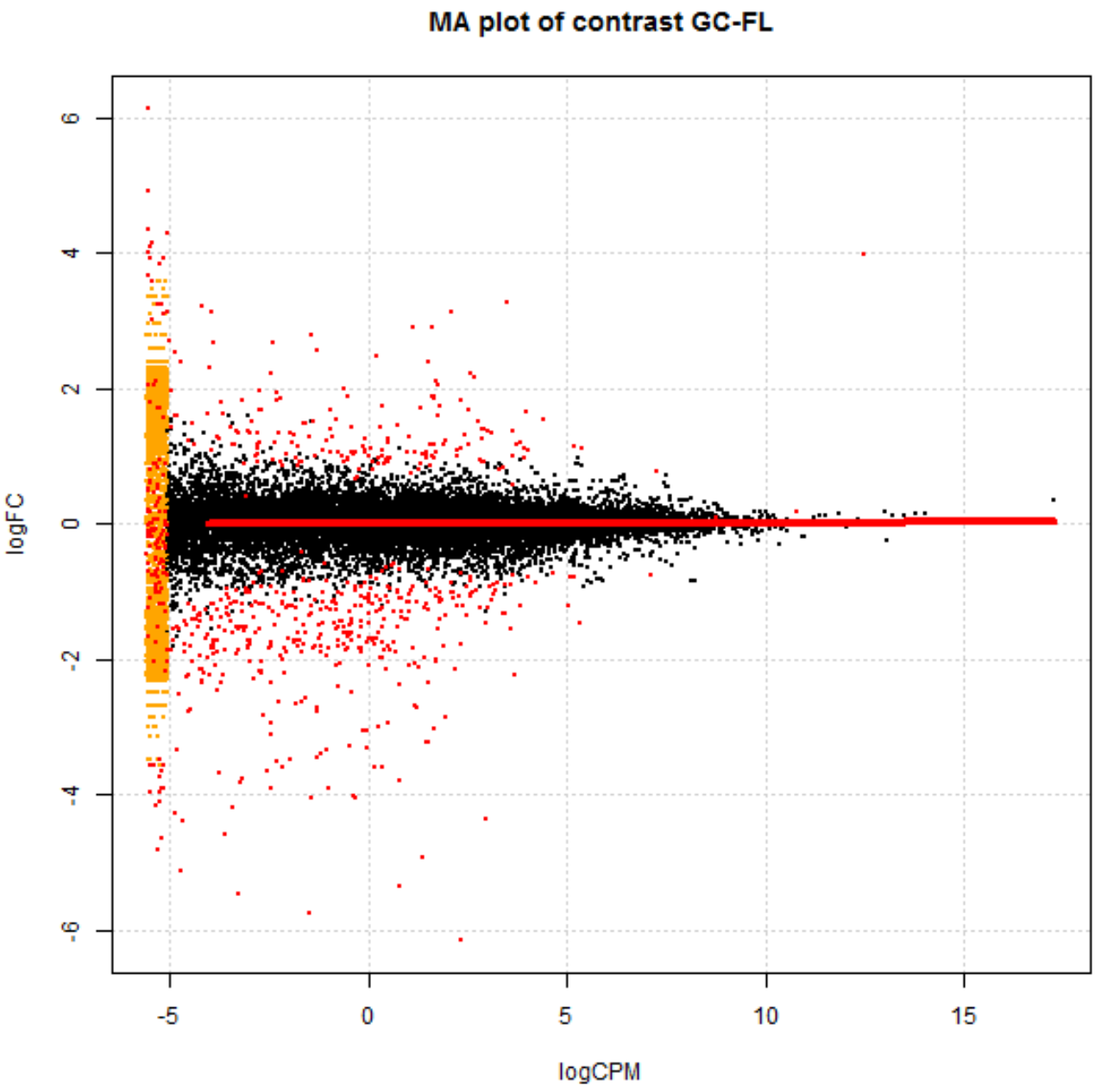
GC-FL

The normalization and statistical evaluation of differential gene expression has been performed using edgeR (Robinson et al., 2010) with a p-value cut-off of 0.05 and using the Benjamini-Hochberg (1995) method for multiple testing correction.
The raw data was normalized according to the default procedure of the differential expression analysis package used.The dispersion was estimated using the auto setting

MA plots of each comparison

The MA plots show the log2 fold change (M; logFC) plotted versus the average expression strength (A; LogConc) for each of the comparisons that was computed. Usually, these scatter plots show a trumpet-like shape which is attributed to the fact that genes with a lower expression signal strength are more strongly affected by noise than strongly expressed genes.

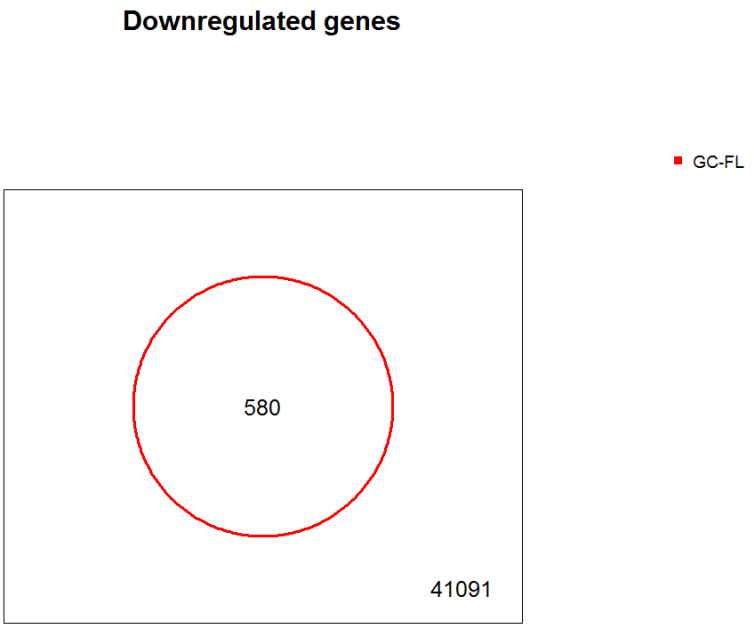
According to the assumption that under most experimental conditions the bulk of the genes of an organism are not responding differentially, the cloud of points should be centered around a log fold change of 0.
Genes that were called significantly differentially expressed are shown in red.



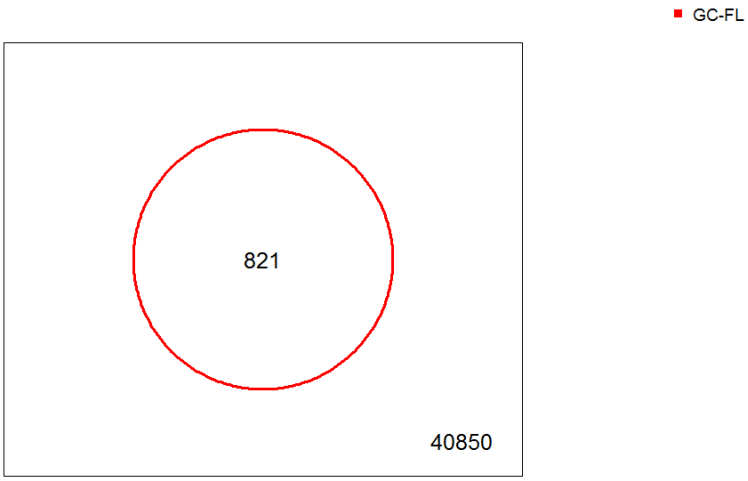
Venn Diagrams

Venn diagrams visualize the amount of genes that were called significantly differentially expressed in each comparison. The conditions are represented by circles. Genes that show a significant reponse to more than one condition are plotted in the overlapping areas while the amount of not significantly changed genes is given in the lower right corner of the plots.

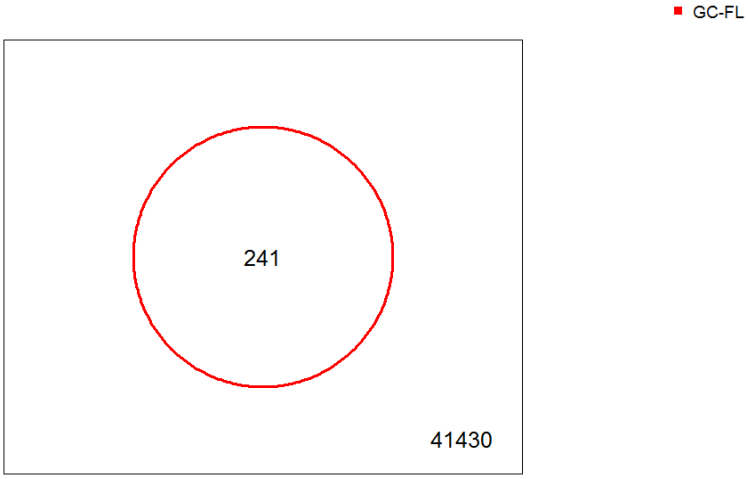
Venn diagrams allow a simple and quick overview of the impact of the treatments on the gene expression profile and also the specificity of the responses



Significantly regulated genes



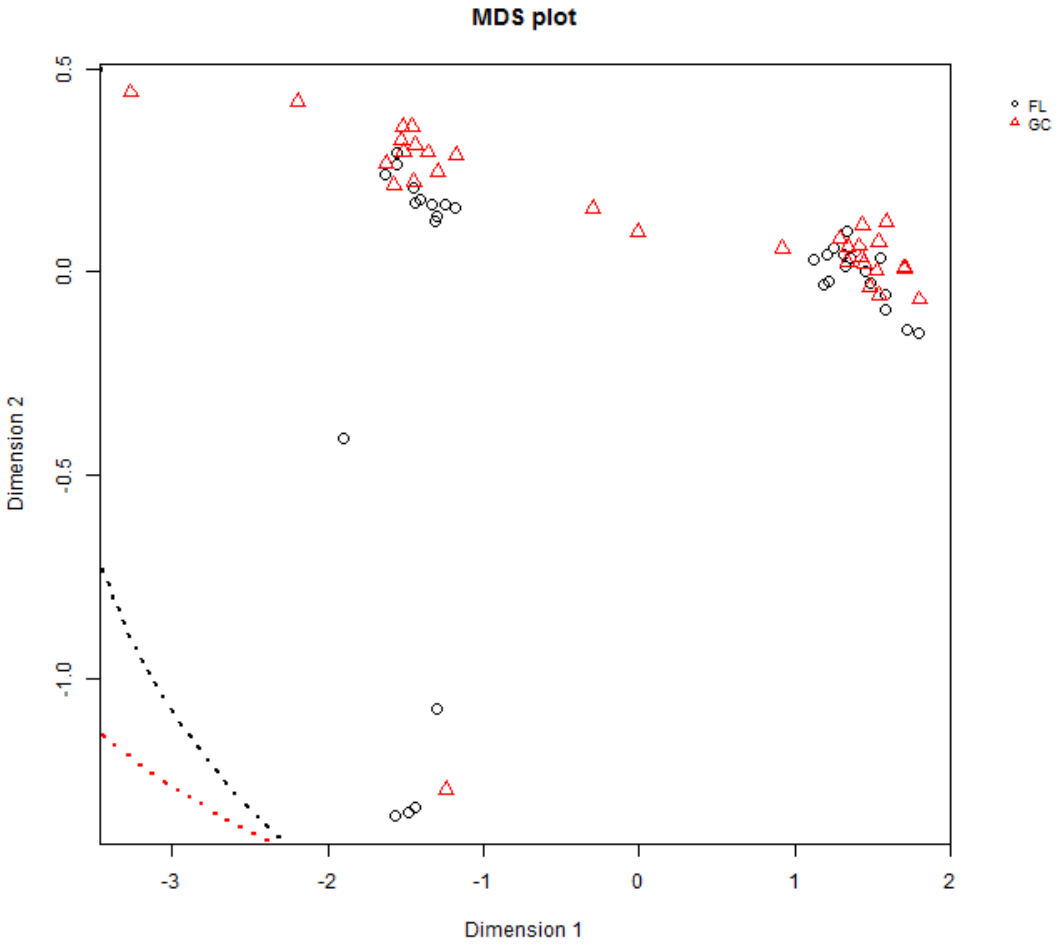
Upregulated genes



Multi-dimensional scaling (MDS) plot

The MDS or principal coordinate plot visualizes the distances between the RNA-Seq libraries in the experiment. To compute the points, a set of 500 tags(genes) that have the largest variation bewteen the libraries (i.e. the largest tagwise dispersion when treating all libraries as one experimental group) is selected. The distance between each pair of libraries is equivalent to the square root of the common disperion between these two libraries (using the top 500 genes.

Hence, the MDS plot gives an insight into the structure of the experiment - libraries that were generated on biological replicates of the same treatment should cluster together



Top 10 differentially expressed genes tables for each contrast

Top differentially expressed genes: full_table_GC-FL.txt

Identifier	logFC	logCPM	PValue	FDR
AT3G41768.1	3.97148938212694	12.4578905460649	3.12008285063524e-106	1.30016972468821e-101
AT3G41768.1	3.97148938212694	12.4578905460649	3.12008285063524e-106	1.30016972468821e-101
AT3G41768.1	3.97148938212694	12.4578905460649	3.12008285063524e-106	1.30016972468821e-101
AT3G41768.1	3.97148938212694	12.4578905460649	3.12008285063524e-106	1.30016972468821e-101
AT3G41768.1	3.97148938212694	12.4578905460649	3.12008285063524e-106	1.30016972468821e-101
AT3G41979.1	3.27970507797367	3.47117824427969	1.08513706435839e-96	2.26093733044392e-92
AT3G41979.1	3.27970507797367	3.47117824427969	1.08513706435839e-96	2.26093733044392e-92
AT3G41979.1	3.27970507797367	3.47117824427969	1.08513706435839e-96	2.26093733044392e-92
AT3G41979.1	3.27970507797367	3.47117824427969	1.08513706435839e-96	2.26093733044392e-92
AT3G41979.1	3.27970507797367	3.47117824427969	1.08513706435839e-96	2.26093733044392e-92
AT2G01010.1	3.12726493485294	2.10773086933339	1.02495940436629e-84	1.42370277797826e-80
AT2G01010.1	3.12726493485294	2.10773086933339	1.02495940436629e-84	1.42370277797826e-80
AT2G01010.1	3.12726493485294	2.10773086933339	1.02495940436629e-84	1.42370277797826e-80
AT2G01010.1	3.12726493485294	2.10773086933339	1.02495940436629e-84	1.42370277797826e-80
AT2G01010.1	3.12726493485294	2.10773086933339	1.02495940436629e-84	1.42370277797826e-80
AT1G26390.1	-6.133134882633487	2.32049443099168	2.17713432609492e-79	2.26808411256754e-75
AT1G26390.1	-6.133134882633487	2.32049443099168	2.17713432609492e-79	2.26808411256754e-75
AT1G26390.1	-6.133134882633487	2.32049443099168	2.17713432609492e-79	2.26808411256754e-75
AT1G26390.1	-6.133134882633487	2.32049443099168	2.17713432609492e-79	2.26808411256754e-75
AT1G26390.1	-6.133134882633487	2.32049443099168	2.17713432609492e-79	2.26808411256754e-75
AT2G01020.1	2.167692305955116	2.644132608667217	8.14163887316008e-71	6.78540466966908e-67
AT2G01020.1	2.167692305955116	2.644132608667217	8.14163887316008e-71	6.78540466966908e-67

Identifier	logFC	logCPM	PValue	FDR
AT2G01020.1	2.16769230595116	2.64413260867217	8.14163887316008e-71	6.78540466966908e-67
AT2G01020.1	2.16769230595116	2.64413260867217	8.14163887316008e-71	6.78540466966908e-67
AT2G01020.1	2.16769230595116	2.64413260867217	8.14163887316008e-71	6.78540466966908e-67
AT4G31970.1	-4.925496190700799	1.351639484144376	7.76811311578834e-65	5.39508402746693e-61
AT4G31970.1	-4.925496190700799	1.351639484144376	7.76811311578834e-65	5.39508402746693e-61
AT4G31970.1	-4.925496190700799	1.351639484144376	7.76811311578834e-65	5.39508402746693e-61
AT4G31970.1	-4.925496190700799	1.351639484144376	7.76811311578834e-65	5.39508402746693e-61
AT4G31970.1	-4.925496190700799	1.351639484144376	7.76811311578834e-65	5.39508402746693e-61
AT2G41810.1	-5.353702672333343	0.761093367043858	1.84346138294832e-55	1.09741256126914e-51
AT2G41810.1	-5.353702672333343	0.761093367043858	1.84346138294832e-55	1.09741256126914e-51
AT2G41810.1	-5.353702672333343	0.761093367043858	1.84346138294832e-55	1.09741256126914e-51
AT2G41810.1	-5.353702672333343	0.761093367043858	1.84346138294832e-55	1.09741256126914e-51
AT2G41810.1	-5.353702672333343	0.761093367043858	1.84346138294832e-55	1.09741256126914e-51
AT5G38910.1	-5.741969236663189	-1.487285181815977	8.63375146377749e-50	4.49721321558884e-46
AT5G38910.1	-5.741969236663189	-1.487285181815977	8.63375146377749e-50	4.49721321558884e-46
AT5G38910.1	-5.741969236663189	-1.487285181815977	8.63375146377749e-50	4.49721321558884e-46
AT5G38910.1	-5.741969236663189	-1.487285181815977	8.63375146377749e-50	4.49721321558884e-46
AT5G38910.1	-5.741969236663189	-1.487285181815977	8.63375146377749e-50	4.49721321558884e-46
AT2G30750.1	-4.343349494622301	2.955726664779595	9.82500040566085e-50	4.54908435449215e-46

Identifier	logFC	logCPM	PValue	FDR
AT2G30750.1	-4.343349494622301	2.955726664795595	9.825000405660085e-50	4.549084354492215e-46
AT2G30750.1	-4.343349494622301	2.955726664795595	9.825000405660085e-50	4.549084354492215e-46
AT2G30750.1	-4.343349494622301	2.955726664795595	9.825000405660085e-50	4.549084354492215e-46
AT2G30750.1	-4.343349494622301	2.955726664795595	9.825000405660085e-50	4.549084354492215e-46

Literature

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139-140

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B, 57, 289–300.

R session information

```
R version 2.15.0 (2012-03-30), i386-pc-mingw32|
Locale: LC_COLLATE=English_United States.1252|, LC_CTYPE=English_United
States.1252|, LC_MONETARY=English_United States.1252|, LC_NUMERIC=C|,
LC_TIME=English_United States.1252|
Base packages: base, datasets, graphics, grDevices, methods,
stats, utils
Other packages: edgeR 2.6.12, limma 3.12.3
```