

HDAC inhibitor, ITF2357, reveals reduction of inflammation, invasive and mesenchymal pathways as diastolic function is restored.

By Harrison Smith

# outline

Post Filtered GSEA

MDS-plot

EdgeR significant genes

NES vs FDR plots with relevant pathways highlighted

Heatmap

Jaccard similarity scatter plot

# GSEA Filtration

This step is done to remove reads from noncoding RNA

This allows analysis to consist solely of gene expression.

41 noncoding RNA reads removed and output was used for further analysis

snoRNA, miRNA, siRNA, etc.

# GSEA post filtration (ITF vs DOCA\_8)

## Enrichment in phenotype: DOCA\_8 (1 samples)

- 2248 / 3302 gene sets are upregulated in phenotype **DOCA\_8**
- 1020 gene sets are significant at FDR < 25%
- 429 gene sets are significantly enriched at nominal pvalue < 1%
- 769 gene sets are significantly enriched at nominal pvalue < 5%
- [Snapshot](#) of enrichment results
- Detailed [enrichment results in html](#) format
- Detailed [enrichment results in excel](#) format (tab delimited text)
- [Guide to](#) interpret results

## Enrichment in phenotype: ITF\_8 (1 samples)

- 1054 / 3302 gene sets are upregulated in phenotype **ITF\_8**
- 146 gene sets are significant at FDR < 25%
- 94 gene sets are significantly enriched at nominal pvalue < 1%
- 188 gene sets are significantly enriched at nominal pvalue < 5%
- [Snapshot](#) of enrichment results
- Detailed [enrichment results in html](#) format
- Detailed [enrichment results in excel](#) format (tab delimited text)
- [Guide to](#) interpret results

Up in DOCA while down in ITF

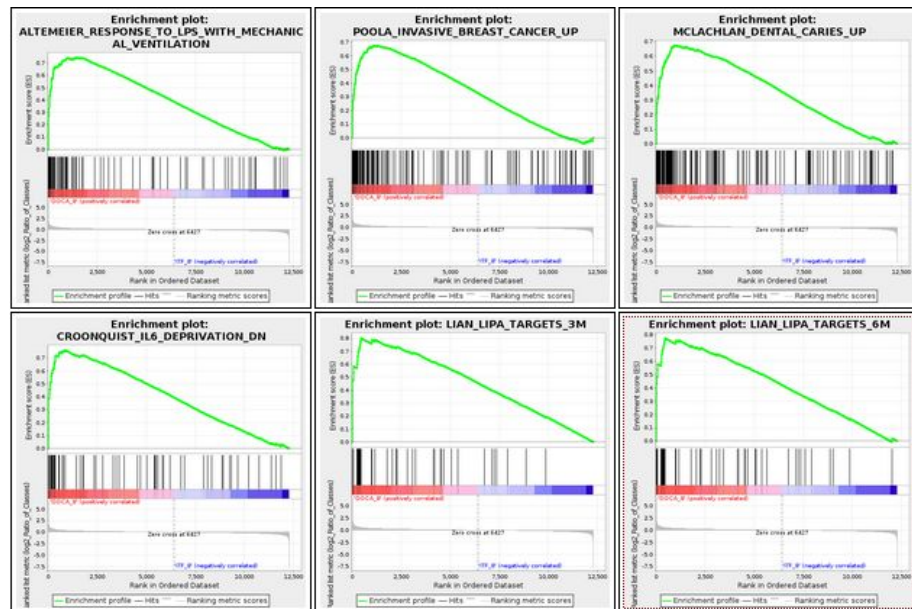
1. **Altemeier\_response\_LPS\_with\_mechanical\_ventilation**
2. **Poola\_invasive\_breast\_cancer\_up**
3. **Mchlachlan\_dental\_carries\_up**
4. **Croonquist\_il6\_deprivation**
5. **Lian\_lipa\_targets\_3m**

Up in ITF while down in DOCA

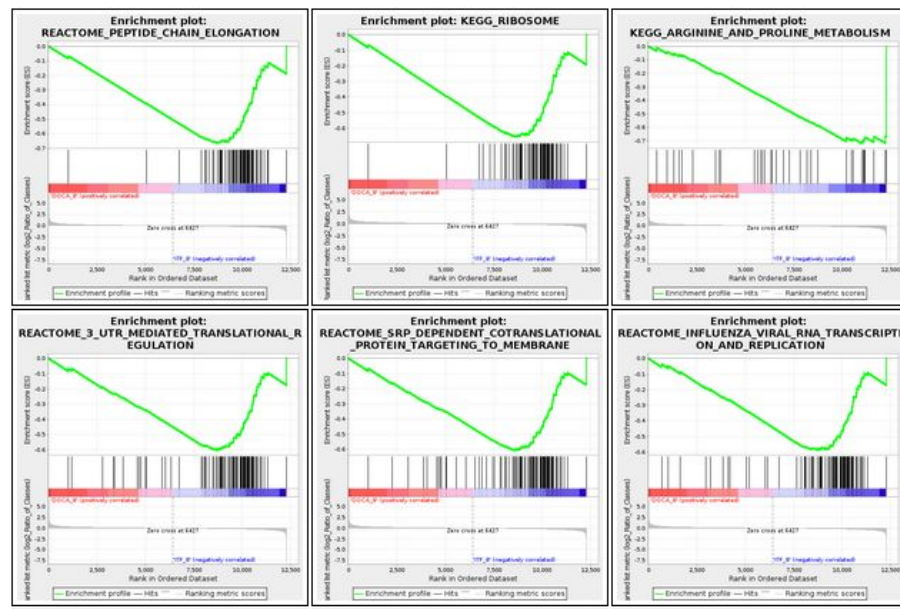
1. **Reactome\_peptide\_elongation**
2. **Kegg\_ribosome**
3. **Kegg\_arginine\_proline\_metabolism**
4. **Reactome\_3\_UTR\_mediated\_translation\_regulation**
5. **Reactome\_srp\_dependent\_cotranslational\_protein\_targeting\_to\_membrane**

# DOCA 8 vs ITF\_8

Up in DOCA while down in ITF 8



Up in ITF while down in DOCA\_8



# GSEA post filtration (DOCA\_8 vs Sham\_8)

## Enrichment in phenotype: DOCA\_8 (1 samples)

- 2384 / 3302 gene sets are upregulated in phenotype **DOCA\_8**
- 593 gene sets are significant at FDR < 25%
- 277 gene sets are significantly enriched at nominal pvalue < 1%
- 571 gene sets are significantly enriched at nominal pvalue < 5%
- [Snapshot](#) of enrichment results
- Detailed [enrichment results in html](#) format
- Detailed [enrichment results in excel](#) format (tab delimited text)
- [Guide to](#) interpret results

## Enrichment in phenotype: Sham\_8 (1 samples)

- 918 / 3302 gene sets are upregulated in phenotype **Sham\_8**
- 58 gene sets are significantly enriched at FDR < 25%
- 46 gene sets are significantly enriched at nominal pvalue < 1%
- 131 gene sets are significantly enriched at nominal pvalue < 5%
- [Snapshot](#) of enrichment results
- Detailed [enrichment results in html](#) format
- Detailed [enrichment results in excel](#) format (tab delimited text)
- [Guide to](#) interpret results

Up in DOCA while down in Sham

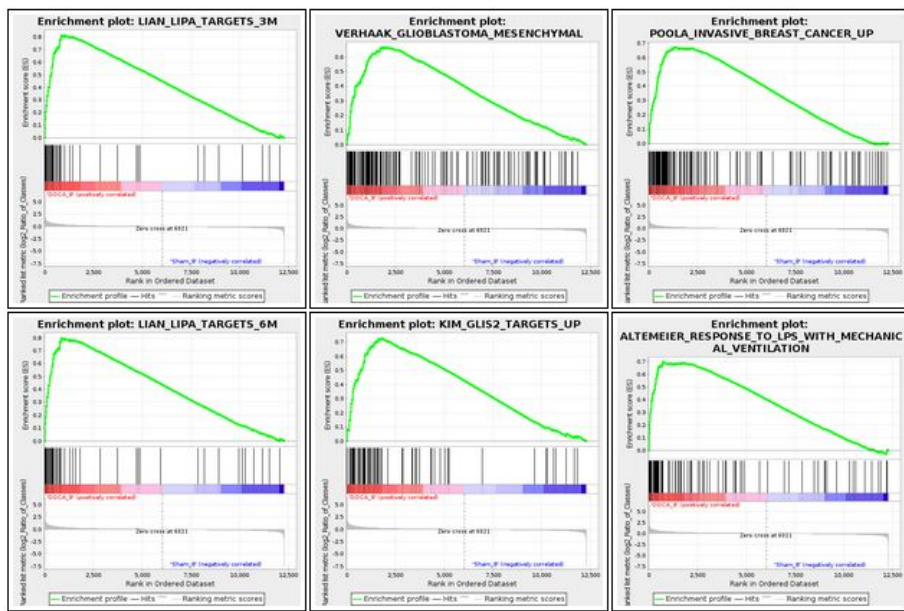
1. **Lian\_Lipa\_targets\_3m**
2. **Verhaak\_glioblastoma\_mesenchymal**
3. **Poola\_invasive\_brea\_\_cancer\_up**
4. **Lian\_Lipa\_targets\_6m**
5. **Kim\_GLI52\_targets up**
6. **Altemeier\_response\_to\_LP\_S\_with\_mechanical\_ventilation**

Up in Sham while down in DOCA

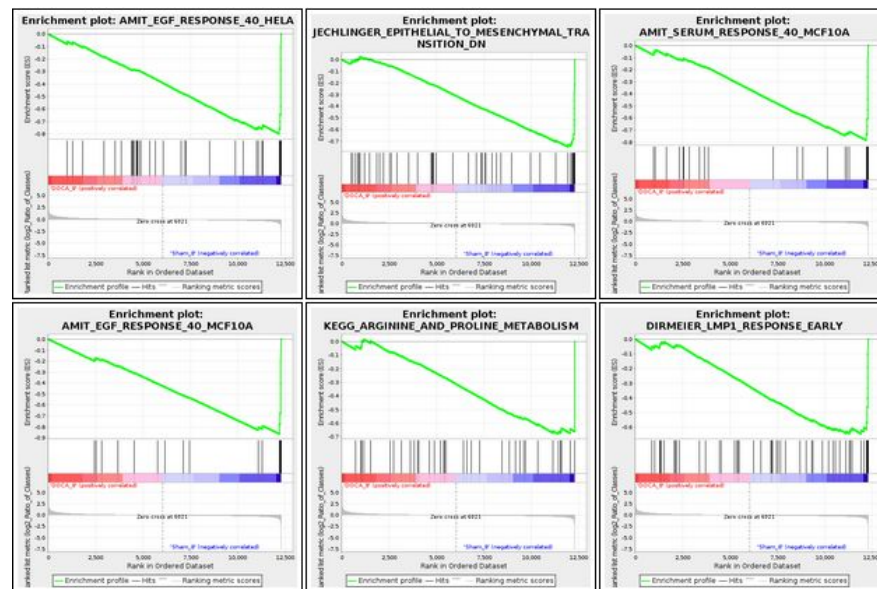
1. **AMIT\_EGF\_response\_40\_HELA**
2. **Jechlinger\_epithelial\_to\_mesenchymal\_transition\_dn**
3. **amit\_serum\_response\_40\_MCF10A**
4. **AMIT\_EGF\_response\_40\_MCF10A**
5. **KEGG\_arginine\_proline\_metabolism**

# DOCA\_8 vs Sham\_8

Up in DOCA while down in Sham



Up in Sham while down in Doca





# GSEA post filtration (ITF vs Sham\_8)

## Enrichment in phenotype: ITF\_8 (1 samples)

- 1653 / 3302 gene sets are upregulated in phenotype **ITF\_8**
- 167 gene sets are significant at FDR < 25%
- 121 gene sets are significantly enriched at nominal pvalue < 1%
- 257 gene sets are significantly enriched at nominal pvalue < 5%
- [Snapshot](#) of enrichment results
- Detailed [enrichment results in html](#) format
- Detailed [enrichment results in excel](#) format (tab delimited text)
- [Guide to](#) interpret results

## Enrichment in phenotype: Sham\_8 (1 samples)

- 1649 / 3302 gene sets are upregulated in phenotype **Sham\_8**
- 473 gene sets are significant at FDR < 25%
- 212 gene sets are significantly enriched at nominal pvalue < 1%
- 421 gene sets are significantly enriched at nominal pvalue < 5%
- [Snapshot](#) of enrichment results
- Detailed [enrichment results in html](#) format
- Detailed [enrichment results in excel](#) format (tab delimited text)
- [Guide to](#) interpret results

Up in ITF while down in Sham

1. **Reactome\_Peptide\_chain\_elongation**
2. **reactome\_3\_UTR\_mediated\_translation\_regulation**
3. **Kegg\_ribosome**
4. **ABBUD\_LIF\_signaling\_1\_up**
5. **Reactome\_influenza\_viral\_replication\_and\_transcription**

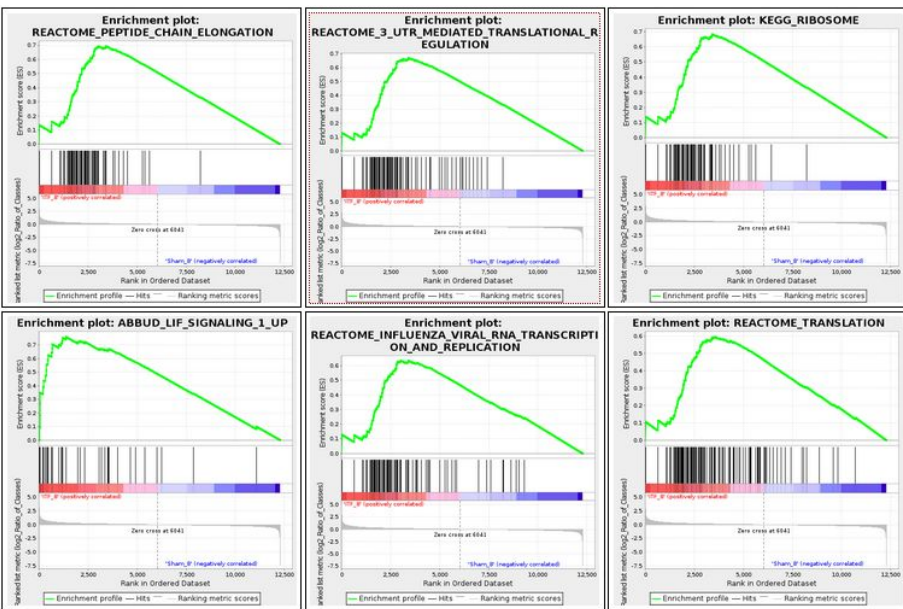
Up in Sham while down in ITF

1. **Sotiriou\_breast\_cancer\_grade\_1\_vs\_3\_UP**
2. **lee\_early\_T\_lymphocyte\_up**
3. **Burton\_adipogenesis\_3**
4. **Whiteford\_pediatric\_cancer\_markers**
5. **rosty\_cervical\_cancer\_proliferation\_cluster**

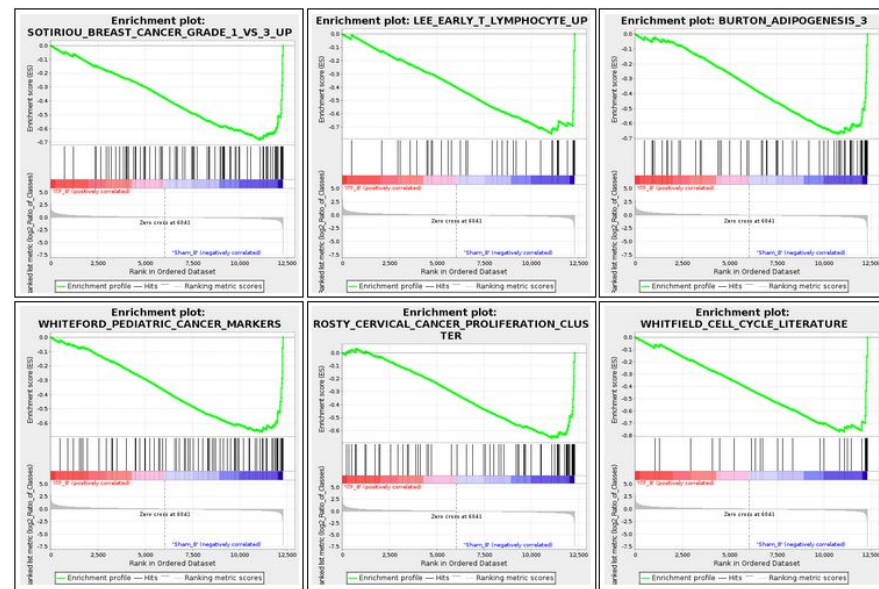


# ITF\_8 vs Sham\_8

Up in ITF while down in Sham



Up in Sham while down in ITF



# MDS-plot

Multidimensional scaling plot of the RNA samples in which distances correspond to leading log-fold-changes between each pair of RNA samples.

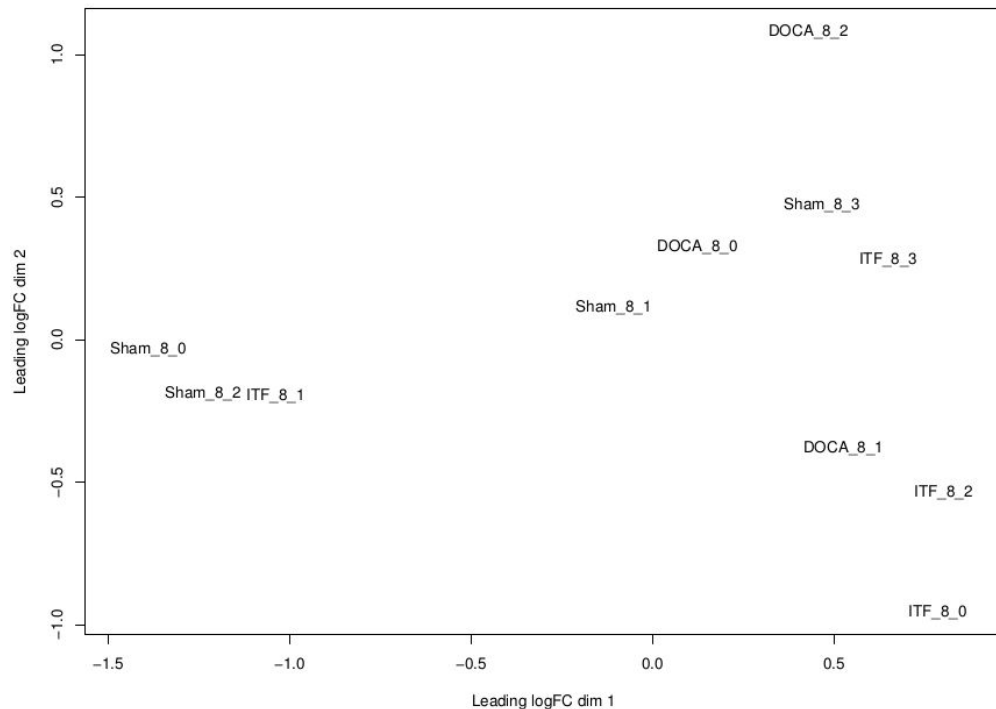
Examine for outliers

---

Noncoding RNA was removed

Normalization was made with `calcNormFactors` function

Low expressed genes were filtered out if there were less than 1 count in at least 3 replicates



# Significant Genes ITF vs DOCA\_8

Quasi-likelihood

FDR < 5%

F-tests

FDR < 5%

DE : 16

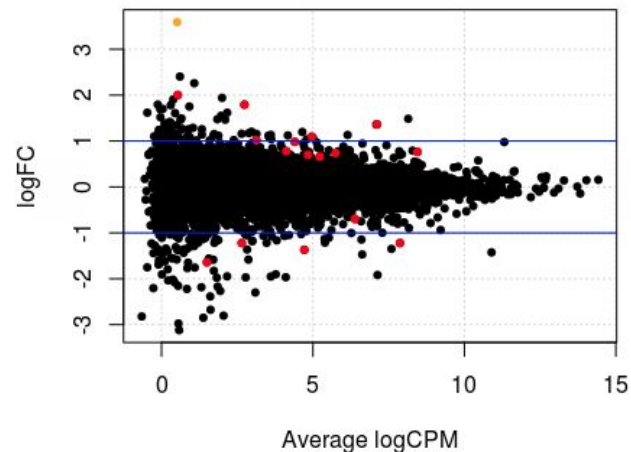
Down : 5

Not sig : 12865

Up : 11

		logFC	logCPM	F	PValue	FDR
	Atp2a3	0.9809536	4.3999189	96.03623	6.188393e-07	0.004081963
	Aldh1b1	1.0941972	4.9616910	95.59539	6.337960e-07	0.004081963
	Tmem35	1.7915514	2.7386168	85.06549	1.157170e-06	0.004968502
	Ndr4	1.3615932	7.1143424	77.92189	1.810203e-06	0.005829305
	Zcchc5	-1.6412121	1.5007543	54.06038	1.107981e-05	0.027589322
	Itga5	-0.6967933	6.3867524	51.86239	1.352714e-05	0.027589322
	Abca6	1.0162224	3.1336378	58.12253	1.499303e-05	0.027589322
	Unc13b	0.6979311	4.8375191	68.61441	1.740095e-05	0.028017703
	Casq1	-1.3690827	4.7141439	45.99234	2.391166e-05	0.033827303
	Gnai1	0.7775335	4.1235101	45.79804	2.956036e-05	0.033827303
	Rcan1	-1.2229359	7.8728259	43.92973	2.962875e-05	0.033827303
	Srgap1	-1.2222318	2.6485730	43.35045	3.151367e-05	0.033827303
	Hopx	0.7620590	8.4516245	41.67801	3.779946e-05	0.037453449
	Syt7	0.7376020	5.7454806	40.64439	4.242127e-05	0.038869018
	Fbxo44	2.0027960	0.5379083	40.07241	4.526320e-05	0.038869018
	Ephx1	0.6594507	5.2244194	37.75689	5.930888e-05	0.047747359

Plot smear highlighting FDR < 5%



LogCPM = log of counts per million (relative abundance)

# DE genes Sham\_8 vs DOCA\_8

Quasi-likelihood

F-tests

14% > FDR > 9% (Lowest 10 FDR)

Plot highlighting 14% > FDR > 9%

FDR < 5%

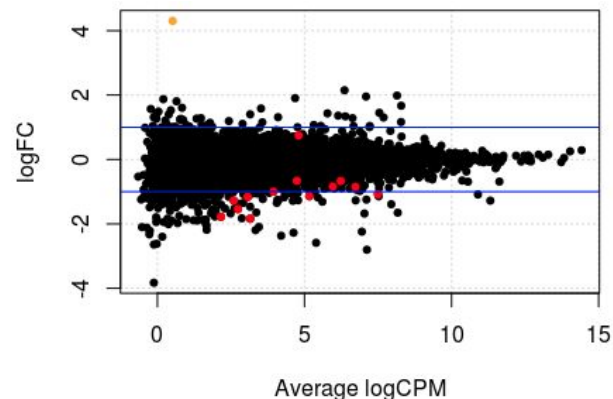
DE : 0

Down : 0

Not sig : 12881

Up : 0

		logFC	logCPM	F	PValue	FDR
	Aqp8	-1.8379283	3.155949	52.43605	1.283196e-05	0.09725679
	Ppap2a	-0.6689293	6.232766	44.67169	2.740444e-05	0.09725679
	Cysltr1	-1.5406994	2.738515	43.95435	2.955164e-05	0.09725679
	Nabp1	-0.8516406	6.731138	41.39075	3.902186e-05	0.09725679
	S100a6	-0.6605108	4.741672	40.81285	4.162457e-05	0.09725679
	Scx	-1.1624964	3.068195	40.06481	4.530244e-05	0.09725679
	AI593442	-1.7786363	2.160467	38.52486	5.414709e-05	0.09963839
	Bex1	-1.2742797	2.593221	34.50306	8.871637e-05	0.13885189
	Capg	-0.9998858	3.946951	32.56539	1.143675e-04	0.13885189
	Uck2	-1.1391830	5.167354	32.19783	1.201679e-04	0.13885189



# DE genes Sham\_8 vs ITF (page 1)

Quasi-likelihood

F-tests

FDR < 5%

DE : 57

Down : 34

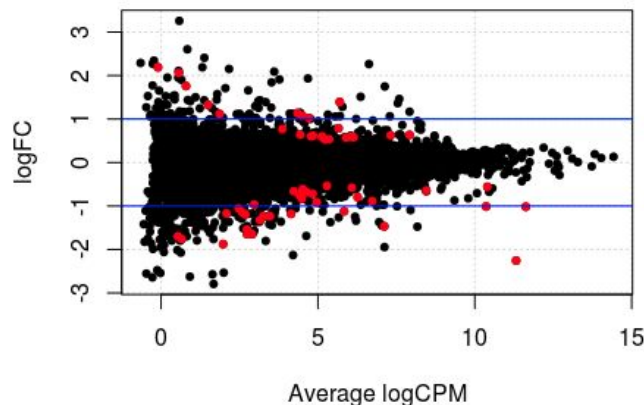
Not sig : 12824

Up : 23

FDR < 5%

	logFC	logCPM	F	PValue	FDR
Ndr4	-1.4661999	7.1143424	109.19512	3.164266e-07	0.003473385
Zdhc2	-1.2328919	3.2587733	98.61158	5.393036e-07	0.003473385
Tmem35	-1.6449003	2.7386168	89.40006	8.962835e-07	0.003848342
Aldh1b1	-0.9145876	4.9616910	81.52380	1.438322e-06	0.004631758
Atp2a3	-0.7401273	4.3999189	66.80575	3.924577e-06	0.008518885
Golm1	-0.7332571	4.4353202	75.04297	4.228754e-06	0.008518885
Adap1	-1.1961580	2.6859960	64.62029	4.629469e-06	0.008518885
Sema3g	-0.6887445	4.6541405	59.24200	7.101931e-06	0.011434997
Neat1	-1.0119301	10.3610662	52.00481	1.335050e-05	0.018818043
Nabp1	-0.8864366	6.7311381	50.84988	1.486509e-05	0.018818043
Ccl24	1.7594875	0.7946260	49.42683	1.701760e-05	0.018818043
Cysl1r1	-1.5416296	2.7385149	49.11853	1.753098e-05	0.018818043
Zfp763	-1.0557977	2.4729515	80.81475	2.194503e-05	0.021744153
Thrsp	-1.1212199	5.8339193	45.37229	2.548205e-05	0.022023853
Micall1	0.6256516	7.3160110	91.27033	2.564691e-05	0.022023853
Dusp8	1.3921124	5.6952983	44.22088	2.873153e-05	0.022062743
Dkk3	-1.6473840	2.8988437	43.07527	3.245821e-05	0.022062743
Abat	-0.8310251	4.4856429	42.98297	3.278243e-05	0.022062743

Plot smear highlighting FDR < 5%





# DE genes Sham\_8 vs ITF (page 2)

FDR < 5%

	logFC	logCPM	F	PValue	FDR		logFC	logCPM	F	PValue	FDR
Tspan12	-0.5777827	6.0798361	57.12162	3.400198e-05	0.022062743	Adcy1	-0.9655515	2.9702321	32.95036	1.086415e-04	0.034763071
Cmss1	0.5973145	5.1337336	42.57651	3.425626e-05	0.022062743	Nppa	-2.2554166	11.3172222	32.90644	1.092775e-04	0.034763071
Spata13	1.0753158	4.5020469	41.28713	3.947416e-05	0.024212701	Speg	0.6356154	7.9211481	32.74762	1.116140e-04	0.034763071
Dcakd	0.5893901	5.1286894	79.42182	4.295289e-05	0.025148915	Asun	0.6408131	4.4333095	32.64265	1.131904e-04	0.034763071
Slc25a23	0.5707594	5.8977054	37.45547	6.149132e-05	0.031100814	Ddah1	-0.7186920	4.8302823	32.63218	1.133490e-04	0.034763071
Pde4b	0.7854414	5.6560214	37.28108	6.279703e-05	0.031100814	Myadml2	0.5834192	6.1393745	31.99393	1.235345e-04	0.037005770
Scn4a	1.1349532	4.3454582	37.03753	6.467543e-05	0.031100814	Ikbip	-0.6086967	4.5157597	31.48286	1.324714e-04	0.038781003
Fbxo44	-1.6961795	0.5379083	36.98808	6.506486e-05	0.031100814	Bex1	-1.1452365	2.5932208	31.23237	1.371280e-04	0.039037731
Zcchc5	1.3238057	1.5007543	38.78017	6.747075e-05	0.031100814	Sec31b	0.6164885	4.8437377	31.11326	1.394096e-04	0.039037731
Ctsh	-0.6602754	4.2320840	54.15655	6.760522e-05	0.031100814	Ctsz	-0.5361596	5.2863643	51.31404	1.523544e-04	0.041754830
Krt80	-1.1711937	2.0851766	35.90618	7.431542e-05	0.033008859	Aqp8	-1.3236798	3.1559491	29.97765	1.635596e-04	0.043059603
Hopx	-0.6427315	8.4516245	35.56719	7.752599e-05	0.033166739	Reep1	0.5335126	5.3557862	29.96731	1.638010e-04	0.043059603
4930402H24Rik	0.5956070	6.0447004	35.33486	7.982058e-05	0.033166739	Ddn	-1.7587647	0.6583224	29.68149	1.706450e-04	0.043961556
Cpxm2	-0.7961910	6.2654744	34.40525	8.983716e-05	0.034763071	Gm7367	-1.1804340	4.1428243	29.16125	1.839844e-04	0.045758514
Magix	0.6003633	4.7819081	34.14150	9.294314e-05	0.034763071	Btbd11	1.1148179	1.8658037	29.02987	1.875444e-04	0.045758514
Samd1	0.7702225	3.8648113	33.74777	9.781779e-05	0.034763071	Eef1a1	-0.5577974	10.4001841	34.17368	1.925377e-04	0.045758514
Prnd	2.0611714	0.5658076	33.64399	9.915231e-05	0.034763071	Akap8l	0.5290462	5.2244186	49.95858	1.944277e-04	0.045758514
Ankrd1	-1.0173598	11.6258574	33.26269	1.042423e-04	0.034763071	Casq1	1.0223382	4.7141439	28.75068	1.953822e-04	0.045758514
Cd79b	2.1899621	-0.1001089	32.95971	1.085067e-04	0.034763071	Lrrc52	-1.8757211	1.9769874	28.16730	2.130368e-04	0.049002257
						Fstl4	-1.2330561	3.4734355	27.92172	2.210252e-04	0.049947814

# Pathways highlighted

NES = normalized enrichment score

Here we show a plot comparing NES vs the significant (FDR)

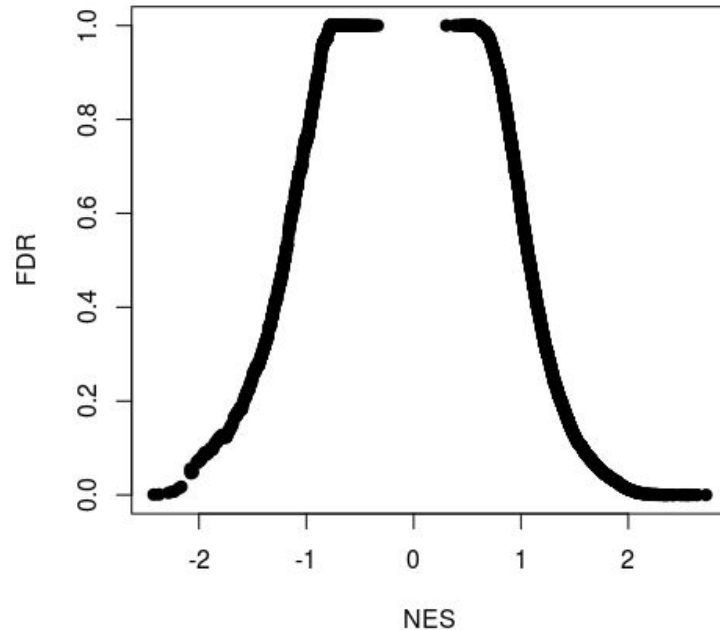
- + NES = up regulation (DOCA)
- NES = up regulation (ITF)

Important ones are located towards the bottom right and left.

I wanted to include highlighted paths in one plot but I am a newbie at R

## All Biological pathways

DOCA\_8 and ITF\_8 NES vs. FDR





# Pathways highlighted 2

The following slides will indicate association with relevant pathways.

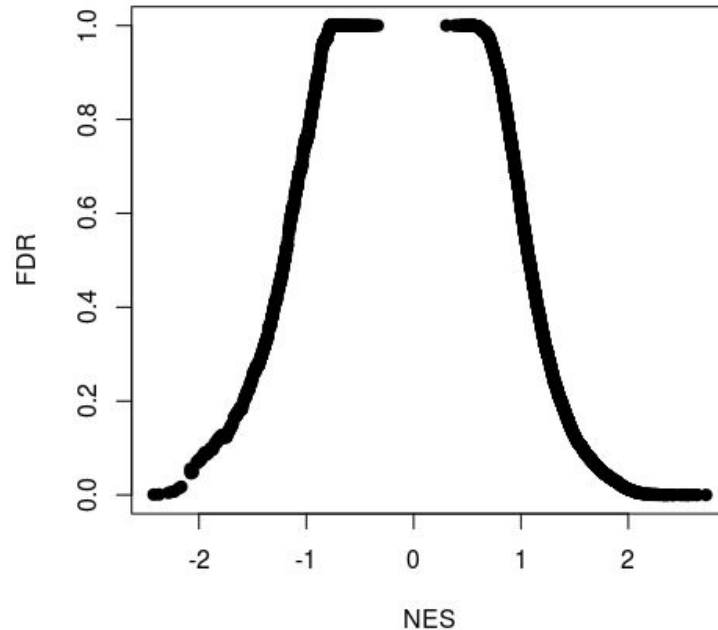
I selected all pathways with the words, “Mesenchymal, inflammatory, and invasive” in the title of each gene set.

I also googled the top 100 NES values in DOCA and top 100 NES values in ITF to see if any were relevant and included them as well.

Some related gene sets just did not have those keywords in the title but I included relevant ones

## All Biological pathways

DOCA\_8 and ITF\_8 NES vs. FDR



# Pathways highlighted

From this plot we can see that desired pathways are much more abundant in the DOCA control group with  $FDR < 5\%$

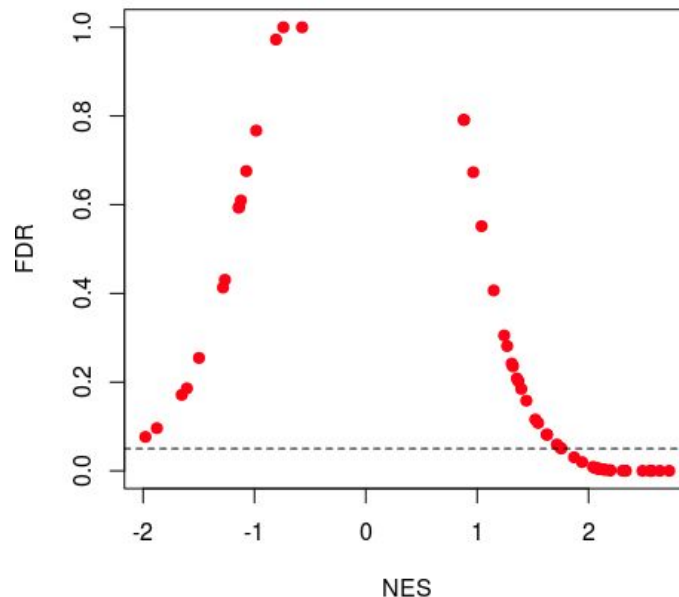
ITF enriched paths are shown on the left and DOCA paths are shown on the right.

The dashed lines indicates an ideal significance.

$FDR < .05$  means less than 5% that the result was due to random chance.

**All desired pathways  
(mesenchymal, inflammatory,  
invasive)**

DOCA\_8 and ITF\_8 NES vs. FDR



# Pathways highlighted

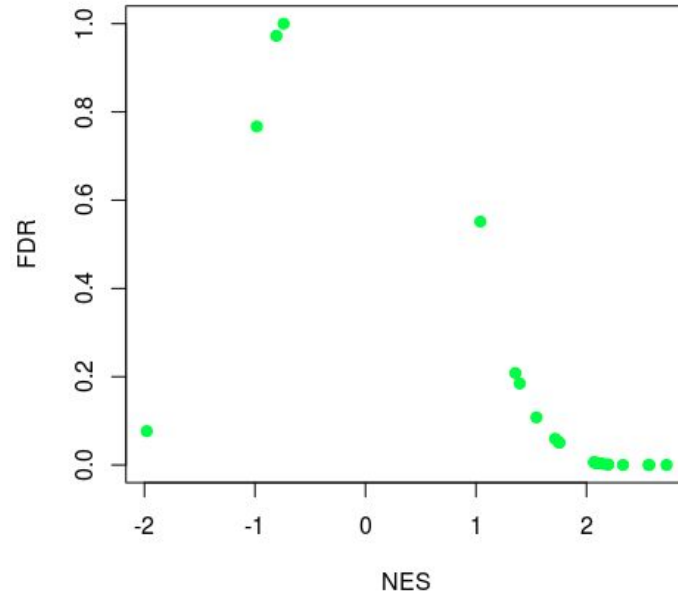
From this plot we can see that desired pathways are much more abundant in the DOCA control group with  $FDR < 5\%$

ITF enriched paths are shown on the left and DOCA paths are shown on the right.

$FDR < .05$  means less than 5% that the result was due to random chance.

## Inflammatory Pathways

DOCA\_8 and ITF\_8 NES vs. FDR



# Pathways highlighted

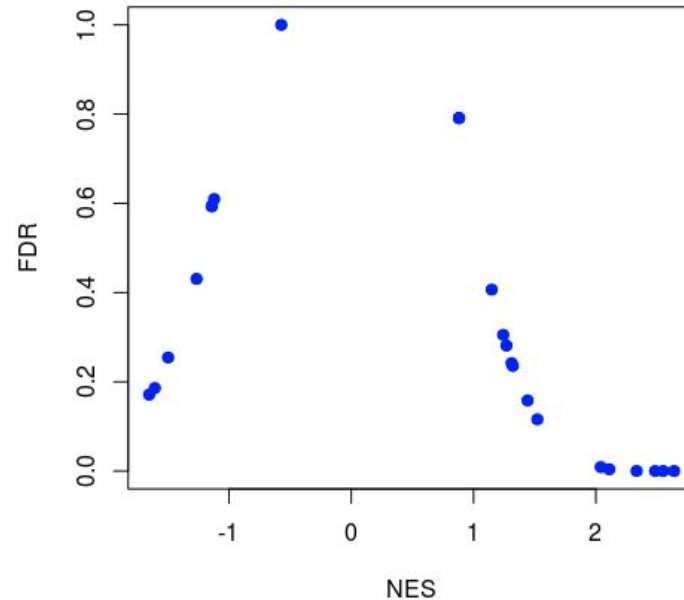
From this plot we can see that desired pathways are much more abundant in the DOCA control group with  $FDR < 5\%$

ITF enriched paths are shown on the left and DOCA paths are shown on the right.

$FDR < .05$  means less than 5% that the result was due to random chance.

## Invasive Pathways

DOCA\_8 and ITF\_8 NES vs. FDR



# Pathways highlighted

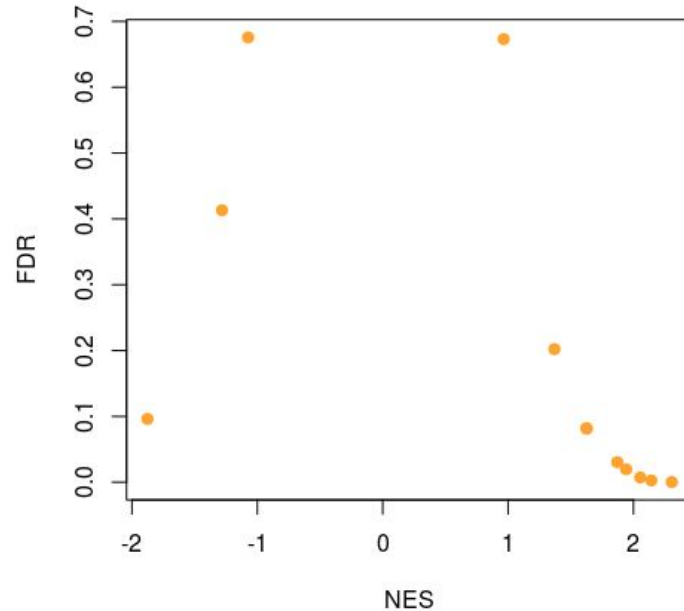
From this plot we can see that desired pathways are much more abundant in the DOCA control group with  $FDR < 5\%$

ITF enriched paths are shown on the left and DOCA paths are shown on the right.

$FDR < .05$  means less than 5% that the result was due to random chance.

## Mesenchymal Pathways

DOCA\_8 and ITF\_8 NES vs. FDR



# HeatMap

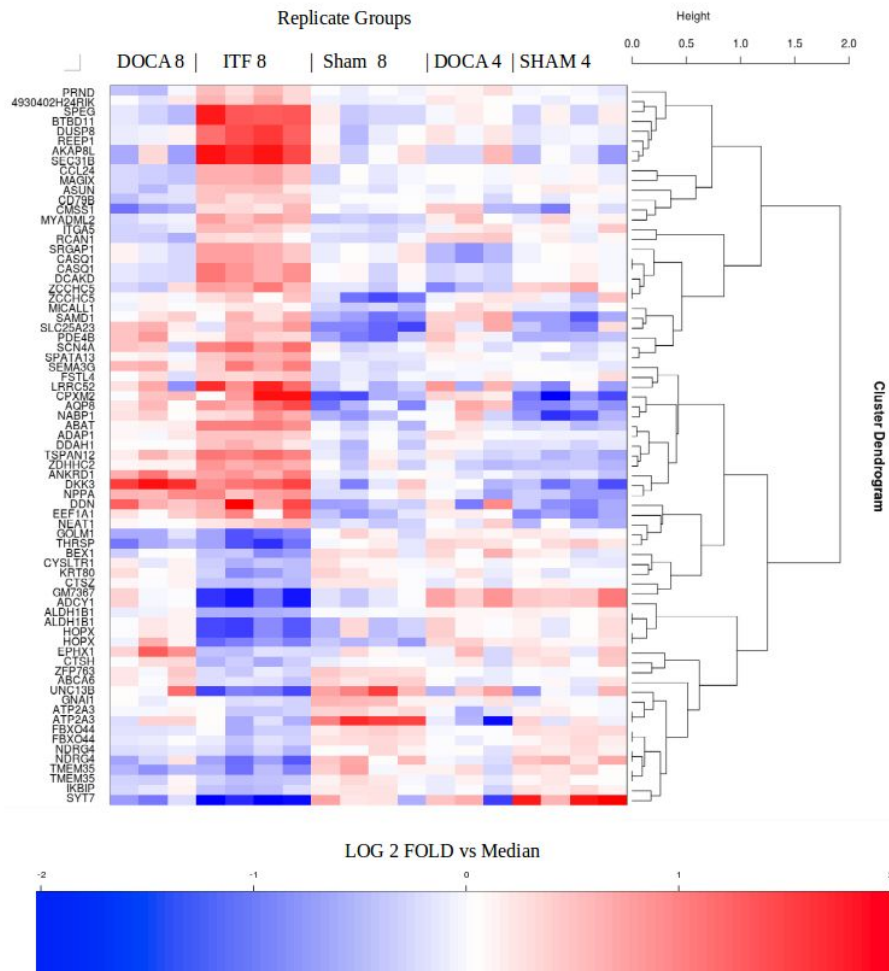
Here's a visual representation of the selected DEG from edgeRs  
Calculated FDR, and LogFC

Using the fpkm raw expression  
And the gene list from above the  
heatmap was generated.

Light or blank colors indicated no  
significant change of expression.

The logCPM might be a better  
values to use as expression  
counts which I will do soon.

Here we see a significant change  
based on drug treatment.



Genes from selected pathways (not sure if this is necessary since GSEA exists)

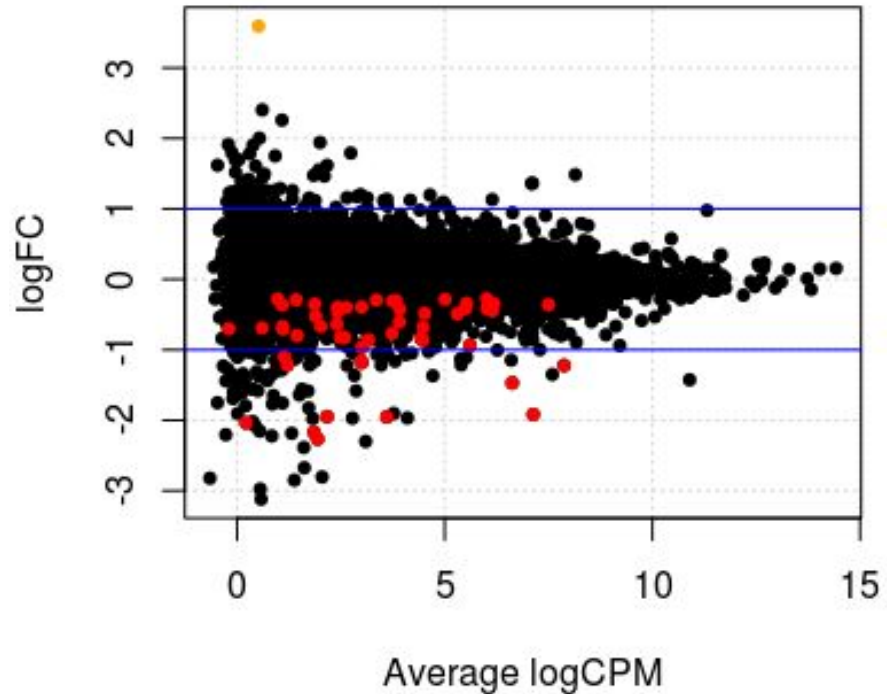
MA plot of I8vsD8

These are the leading edge subset

(genes enriched in this pathway)

Indicates the downregulation in the ITF group

Alteimer\_response\_to\_mechanical\_ventilation





DE genes with highlighted pathways

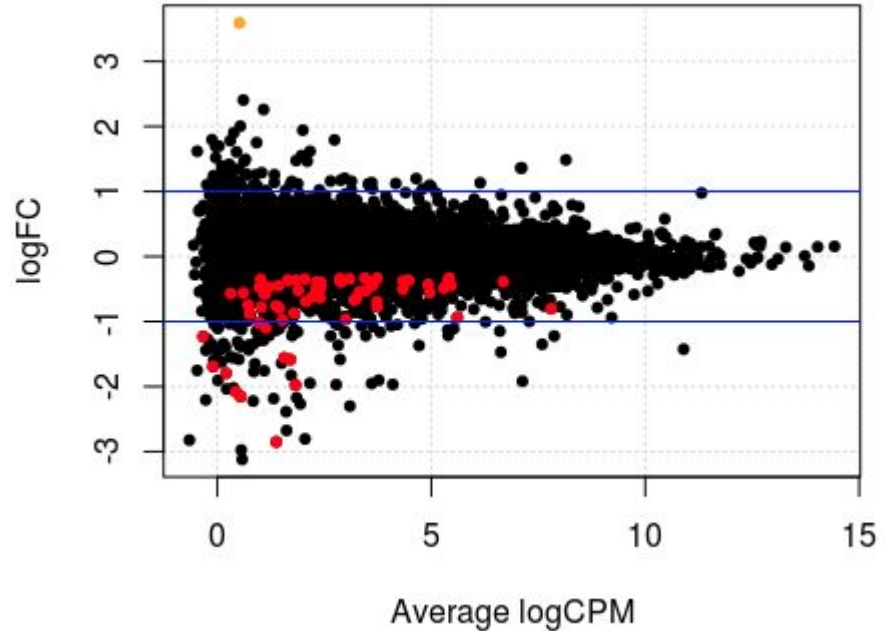
MA plot of I8vsD8

These are the leading edge subset

(genes enriched in this pathway)

Indicates the downregulation in the ITF group

Poola\_Invasive\_Breast\_Cancer



DE genes with highlighted pathways

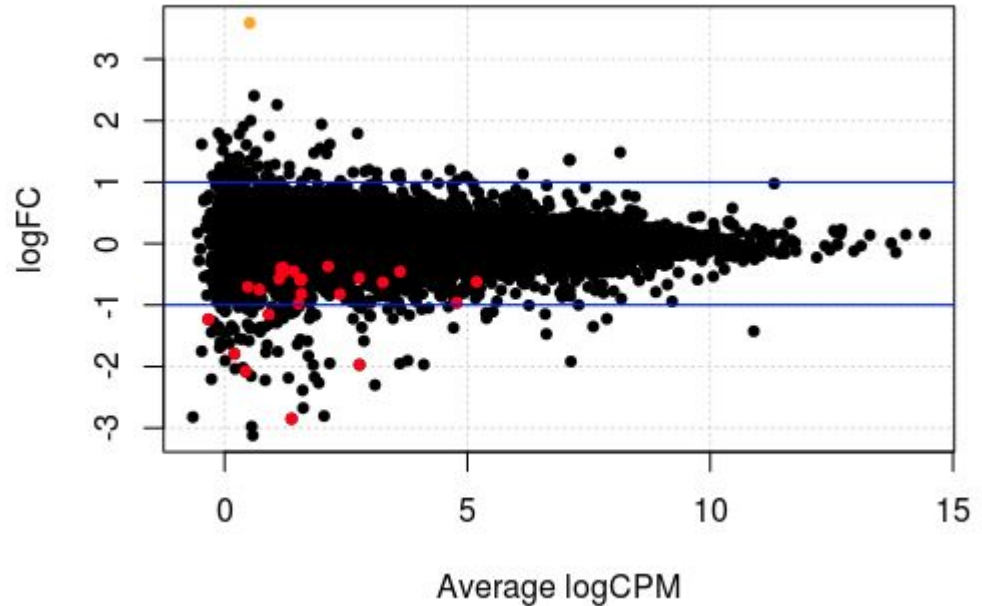
MA plot of I8vsD8

These are the leading edge subset

(genes enriched in this pathway)

Indicates the downregulation in the ITF group

Croonquist\_IL6\_deprivation



Genes with highlighted pathways

MA plot of I8vsD8

These are the leading edge subset

(genes enriched in this pathway)

Indicates the downregulation in the ITF group

Lian\_Lipa\_Targets

