

class 13

Mudit

```
library(BiocManager)  
library(DESeq2)
```

Loading required package: S4Vectors

Loading required package: stats4

Loading required package: BiocGenerics

Attaching package: 'BiocGenerics'

The following objects are masked from 'package:stats':

IQR, mad, sd, var, xtabs

The following objects are masked from 'package:base':

anyDuplicated, aperm, append, as.data.frame, basename, cbind,
colnames, dirname, do.call, duplicated, eval, evalq, Filter, Find,
get, grep, grepl, intersect, is.unsorted, lapply, Map, mapply,
match, mget, order, paste, pmax, pmax.int, pmin, pmin.int,
Position, rank, rbind, Reduce, rownames, sapply, saveRDS, setdiff,
table, tapply, union, unique, unsplit, which.max, which.min

Attaching package: 'S4Vectors'

The following object is masked from 'package:utils':

findMatches

The following objects are masked from 'package:base':

expand.grid, I, unname

Loading required package: IRanges

Loading required package: GenomicRanges

Loading required package: GenomeInfoDb

Loading required package: SummarizedExperiment

Loading required package: MatrixGenerics

Loading required package: matrixStats

Attaching package: 'MatrixGenerics'

The following objects are masked from 'package:matrixStats':

colAlls, colAnyNAs, colAnys, colAvgsPerRowSet, colCollapse,
colCounts, colCummaxs, colCummins, colCumprods, colCumsums,
colDiffs, colIQRDiffs, colIQRs, colLogSumExps, colMadDiffs,
colMads, colMaxs, colMeans2, colMedians, colMins, colOrderStats,
colProds, colQuantiles, colRanges, colRanks, colSdDiffs, colSds,
colSums2, colTabulates, colVarDiffs, colVars, colWeightedMads,
colWeightedMeans, colWeightedMedians, colWeightedSds,
colWeightedVars, rowAlls, rowAnyNAs, rowAnys, rowAvgsPerColSet,
rowCollapse, rowCounts, rowCummaxs, rowCummins, rowCumprods,
rowCumsums, rowDiffs, rowIQRDiffs, rowIQRs, rowLogSumExps,
rowMadDiffs, rowMads, rowMaxs, rowMeans2, rowMedians, rowMins,
rowOrderStats, rowProds, rowQuantiles, rowRanges, rowRanks,
rowSdDiffs, rowSds, rowSums2, rowTabulates, rowVarDiffs, rowVars,
rowWeightedMads, rowWeightedMeans, rowWeightedMedians,
rowWeightedSds, rowWeightedVars

Loading required package: Biobase

Welcome to Bioconductor

Vignettes contain introductory material; view with
'browseVignettes()'. To cite Bioconductor, see
'citation("Biobase")', and for packages 'citation("pkgname")'.

Attaching package: 'Biobase'

The following object is masked from 'package:MatrixGenerics':

rowMedians

The following objects are masked from 'package:matrixStats':

anyMissing, rowMedians

Today we are going to work with some bulk RNAseq data from Himes et al., where airway smooth muscle cells were treated with dexamethasone, a synthetic glucocorticoid steroid with anti-inflammatory effects

3. Import countData and colData

```
# Complete the missing code
counts <- read.csv("airway_scaledcounts.csv", row.names=1)
metadata <- read.csv("airway_metadata.csv")
```

Now, take a look at the head of each.

```
head(counts)
```

	SRR1039508	SRR1039509	SRR1039512	SRR1039513	SRR1039516
ENSG000000000003	723	486	904	445	1170
ENSG000000000005	0	0	0	0	0
ENSG000000000419	467	523	616	371	582
ENSG000000000457	347	258	364	237	318

ENSG00000000460	96	81	73	66	118
ENSG00000000938	0	0	1	0	2
	SRR1039517	SRR1039520	SRR1039521		
ENSG00000000003	1097	806	604		
ENSG00000000005	0	0	0		
ENSG00000000419	781	417	509		
ENSG00000000457	447	330	324		
ENSG00000000460	94	102	74		
ENSG00000000938	0	0	0		

```
head(metadata)
```

	id	dex	celltype	geo_id
1	SRR1039508	control	N61311	GSM1275862
2	SRR1039509	treated	N61311	GSM1275863
3	SRR1039512	control	N052611	GSM1275866
4	SRR1039513	treated	N052611	GSM1275867
5	SRR1039516	control	N080611	GSM1275870
6	SRR1039517	treated	N080611	GSM1275871

Q1. How many genes are in this dataset?

38694 genes there are `nrow(counts)` genes in this dataset

```
nrow(counts)
```

```
[1] 38694
```

Q2. How many 'control' cell lines do we have?

4 controls

```
sum(metadata$dex == "control")
```

```
[1] 4
```

```
a <- metadata[metadata[, "dex"] == "control",]
nrow(a)
```

```
[1] 4
```

I want to compare “control” vs “treated”

1. Lets split the “counts” into `control.counts` and `treated.counts`

```
head(counts)
```

	SRR1039508	SRR1039509	SRR1039512	SRR1039513	SRR1039516
ENSG00000000003	723	486	904	445	1170
ENSG00000000005	0	0	0	0	0
ENSG000000000419	467	523	616	371	582
ENSG000000000457	347	258	364	237	318
ENSG000000000460	96	81	73	66	118
ENSG000000000938	0	0	1	0	2

	SRR1039517	SRR1039520	SRR1039521
ENSG00000000003	1097	806	604
ENSG00000000005	0	0	0
ENSG000000000419	781	417	509
ENSG000000000457	447	330	324
ENSG000000000460	94	102	74
ENSG000000000938	0	0	0

```
treat_ind <- metadata$dex == "treated"
contr_ind <- metadata$dex == "control"

metadata[treat_ind,]
```

	id	dex	celltype	geo_id
2	SRR1039509	treated	N61311	GSM1275863
4	SRR1039513	treated	N052611	GSM1275867
6	SRR1039517	treated	N080611	GSM1275871
8	SRR1039521	treated	N061011	GSM1275875

```
metadata[contr_ind,]
```

	id	dex	celltype	geo_id
1	SRR1039508	control	N61311	GSM1275862
3	SRR1039512	control	N052611	GSM1275866
5	SRR1039516	control	N080611	GSM1275870
7	SRR1039520	control	N061011	GSM1275874

```
control.counts <- counts[,contr_ind]
```

```
treat.counts <- counts[,treat_ind]
```

Let's calculate the mean counts per gene for "control" and "treated" - then we can compare.

Let's call it control.mean and treat.mean

I can use the apply function to apply mean over the rows or columns of any data.frame

```
control.mean <- apply(control.counts, 1, mean)
treated.mean <- apply(treat.counts, 1, mean)
#control.mean
#mean(counts[,treat_ind])
```

Put these together for ease of book-keeping

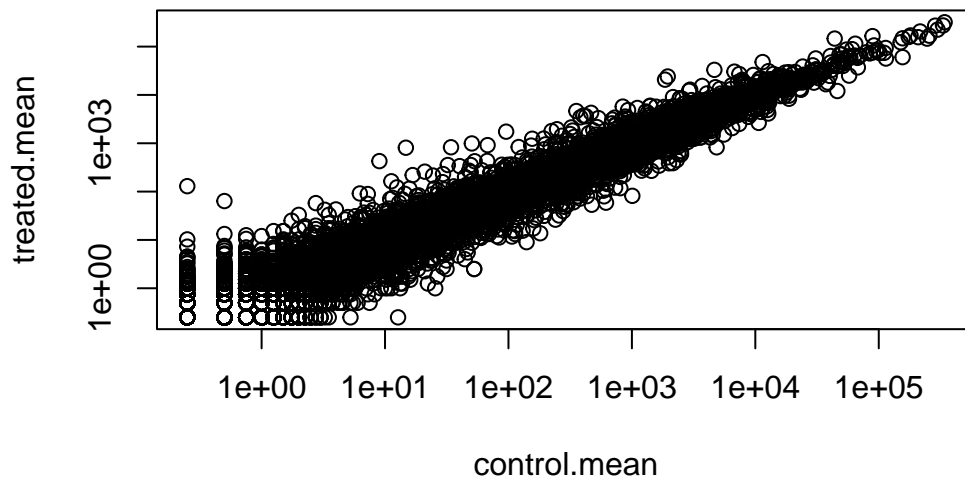
```
meancounts <- data.frame(control.mean, treated.mean)
head(meancounts)
```

	control.mean	treated.mean
ENSG000000000003	900.75	658.00
ENSG000000000005	0.00	0.00
ENSG000000000419	520.50	546.00
ENSG000000000457	339.75	316.50
ENSG000000000460	97.25	78.75
ENSG000000000938	0.75	0.00

```
plot(meancounts, log = "xy")
```

Warning in xy.coords(x, y, xlabel, ylabel, log): 15032 x values <= 0 omitted from logarithmic plot

Warning in xy.coords(x, y, xlabel, ylabel, log): 15281 y values <= 0 omitted from logarithmic plot



we most often use log2 transforms here because it makes math easier

Let's calculate the log2 fold change and add it to our wee table `meancounts`

```
meancounts$log2fc <- log2(meancounts$treated.mean/meancounts$control.mean)
head(meancounts)
```

	control.mean	treated.mean	log2fc
ENSG000000000003	900.75	658.00	-0.45303916
ENSG000000000005	0.00	0.00	NaN
ENSG000000000419	520.50	546.00	0.06900279
ENSG000000000457	339.75	316.50	-0.10226805
ENSG000000000460	97.25	78.75	-0.30441833
ENSG000000000938	0.75	0.00	-Inf

filter out all genes with zero counts in either control or treated:

```
to.rm <- rowSums(meancounts[, 1:2] == 0) > 0
mycounts <- meancounts[!to.rm, ]
```

```
nrow(mycounts)
```

```
[1] 21817
```

How many “down” regulated genes do we have at the common log2 fold change value of -2

```
to.rm_1 <- mycounts$log2fc < -2  
mycounts[to.rm_1, ]
```

	control.mean	treated.mean	log2fc
ENSG00000015520	32.00	6.00	-2.415037
ENSG00000019186	26.50	1.75	-3.920566
ENSG00000025423	295.00	54.25	-2.443020
ENSG00000028277	88.25	22.00	-2.004093
ENSG00000029559	1.25	0.25	-2.321928
ENSG00000049246	405.00	93.00	-2.122619
ENSG00000049247	1.25	0.25	-2.321928
ENSG00000052344	2.25	0.25	-3.169925
ENSG00000054179	3.00	0.25	-3.584963
ENSG00000064201	30.00	6.50	-2.206451
ENSG00000066468	52.50	11.50	-2.190684
ENSG00000072182	2.00	0.25	-3.000000
ENSG00000072694	2.25	0.50	-2.169925
ENSG00000075290	2.00	0.25	-3.000000
ENSG00000075340	8.00	1.25	-2.678072
ENSG00000077498	1.25	0.25	-2.321928
ENSG00000078549	2.25	0.50	-2.169925
ENSG00000083782	1.50	0.25	-2.584963
ENSG00000084710	43.50	9.00	-2.273018
ENSG00000085552	5.50	0.50	-3.459432
ENSG00000086967	4.25	0.50	-3.087463
ENSG00000092969	372.25	92.50	-2.008747
ENSG00000095110	1.50	0.25	-2.584963
ENSG00000095585	45.75	6.50	-2.815260
ENSG00000099194	2503.50	604.75	-2.049536
ENSG00000099282	1.50	0.25	-2.584963
ENSG00000100302	60.75	14.00	-2.117458
ENSG00000100784	300.75	75.00	-2.003602
ENSG00000101278	3.25	0.50	-2.700440
ENSG00000101440	1.75	0.25	-2.807355

ENSG00000103742	629.25	143.75	-2.130071
ENSG00000105989	617.00	59.00	-3.386484
ENSG00000106003	30.25	5.00	-2.596935
ENSG00000106484	1657.75	343.50	-2.270844
ENSG00000108379	39.00	4.75	-3.037475
ENSG00000108602	18.25	3.25	-2.489385
ENSG00000108688	24.25	6.00	-2.014950
ENSG00000108700	31.50	2.50	-3.655352
ENSG00000108830	366.00	80.00	-2.193772
ENSG00000109511	2.00	0.25	-3.000000
ENSG00000112559	5.00	0.75	-2.736966
ENSG00000113525	2.75	0.25	-3.459432
ENSG00000114547	1.25	0.25	-2.321928
ENSG00000115353	1.25	0.25	-2.321928
ENSG00000116703	1.50	0.25	-2.584963
ENSG00000116711	194.50	31.50	-2.626346
ENSG00000116991	240.25	26.25	-3.194147
ENSG00000117148	11.50	2.00	-2.523562
ENSG00000118729	25.50	1.00	-4.672425
ENSG00000119714	113.25	13.25	-3.095447
ENSG00000122679	16.00	1.50	-3.415037
ENSG00000122877	31.50	3.75	-3.070389
ENSG00000123405	23.25	3.75	-2.632268
ENSG00000123610	585.25	104.00	-2.492470
ENSG00000123689	80.25	12.00	-2.741467
ENSG00000124134	50.50	8.75	-2.528928
ENSG00000124194	2.50	0.50	-2.321928
ENSG00000124766	2307.25	447.50	-2.366215
ENSG00000125848	337.50	60.25	-2.485854
ENSG00000126016	731.50	120.75	-2.598835
ENSG00000126878	161.25	25.75	-2.646655
ENSG00000128165	113.00	15.25	-2.889442
ENSG00000128285	12.75	0.25	-5.672425
ENSG00000128342	273.25	55.00	-2.312718
ENSG00000128510	300.50	53.00	-2.503301
ENSG00000128594	36.50	8.50	-2.102362
ENSG00000128606	2151.50	387.75	-2.472144
ENSG00000129159	4.25	0.50	-3.087463
ENSG00000129993	5.25	0.75	-2.807355
ENSG00000130032	93.50	20.75	-2.171855
ENSG00000130287	3.25	0.75	-2.115477
ENSG00000130427	1.50	0.25	-2.584963
ENSG00000130487	13.75	1.75	-2.974005

ENSG00000130612	3.25	0.75	-2.115477
ENSG00000130762	2.25	0.25	-3.169925
ENSG00000131126	2.25	0.50	-2.169925
ENSG00000131242	92.25	15.75	-2.550197
ENSG00000131389	1267.50	303.00	-2.064596
ENSG00000131771	159.00	17.25	-3.204358
ENSG00000132622	116.75	20.50	-2.509727
ENSG00000132854	10.25	1.50	-2.772590
ENSG00000134253	245.00	57.00	-2.103748
ENSG00000134321	16.50	2.25	-2.874469
ENSG00000134757	3.25	0.50	-2.700440
ENSG00000135517	1.75	0.25	-2.807355
ENSG00000135697	4.50	0.75	-2.584963
ENSG00000136541	3.25	0.50	-2.700440
ENSG00000136999	806.25	117.75	-2.775500
ENSG00000138316	1439.00	269.00	-2.419389
ENSG00000139269	31.25	6.50	-2.265345
ENSG00000139610	2.50	0.50	-2.321928
ENSG00000140600	19.75	3.75	-2.396890
ENSG00000141161	1.25	0.25	-2.321928
ENSG00000141469	130.25	15.50	-3.070943
ENSG00000142185	1.25	0.25	-2.321928
ENSG00000143061	13.50	3.25	-2.054448
ENSG00000143171	1.25	0.25	-2.321928
ENSG00000144550	2.50	0.25	-3.321928
ENSG00000144596	6.00	1.25	-2.263034
ENSG00000145103	2.25	0.50	-2.169925
ENSG00000145721	2.75	0.50	-2.459432
ENSG00000145777	277.75	39.50	-2.813862
ENSG00000146006	140.25	9.00	-3.961932
ENSG00000146038	12.50	3.00	-2.058894
ENSG00000146250	595.25	79.75	-2.899939
ENSG00000146592	1197.50	271.00	-2.143661
ENSG00000147481	1.25	0.25	-2.321928
ENSG00000147488	2.50	0.25	-3.321928
ENSG00000147676	2.50	0.50	-2.321928
ENSG00000149403	82.00	16.50	-2.313158
ENSG00000150471	10.00	1.75	-2.514573
ENSG00000150594	69.50	13.25	-2.391021
ENSG00000151948	2.25	0.25	-3.169925
ENSG00000152128	2.25	0.50	-2.169925
ENSG00000153012	13.50	1.75	-2.947533
ENSG00000154162	2.25	0.50	-2.169925

ENSG00000155897	44.25	4.00	-3.467606
ENSG00000157303	8.00	1.50	-2.415037
ENSG00000157368	95.00	20.75	-2.194816
ENSG00000158055	3.00	0.25	-3.584963
ENSG00000159712	1.25	0.25	-2.321928
ENSG00000160791	3.00	0.50	-2.584963
ENSG00000160963	1.25	0.25	-2.321928
ENSG00000162552	4.50	1.00	-2.169925
ENSG00000162692	1007.75	81.75	-3.623775
ENSG00000162728	1.75	0.25	-2.807355
ENSG00000162738	15.50	3.00	-2.369234
ENSG00000163347	5.25	0.50	-3.392317
ENSG00000163394	1999.25	385.00	-2.376529
ENSG00000163491	265.00	55.25	-2.261946
ENSG00000163554	29.50	2.75	-3.423211
ENSG00000164007	1.25	0.25	-2.321928
ENSG00000164266	1.50	0.25	-2.584963
ENSG00000164691	4.00	0.75	-2.415037
ENSG00000164742	618.00	111.75	-2.467332
ENSG00000165272	257.25	32.00	-3.007027
ENSG00000165309	5.75	1.00	-2.523562
ENSG00000165495	32.75	5.50	-2.573991
ENSG00000165694	1.50	0.25	-2.584963
ENSG00000165794	1.25	0.25	-2.321928
ENSG00000166670	15.50	1.75	-3.146841
ENSG00000168298	1.75	0.25	-2.807355
ENSG00000168334	2.50	0.50	-2.321928
ENSG00000168398	10650.50	2638.00	-2.013405
ENSG00000168634	1.75	0.25	-2.807355
ENSG00000168811	70.00	17.25	-2.020759
ENSG00000168843	141.75	31.75	-2.158520
ENSG00000169245	2.75	0.25	-3.459432
ENSG00000169297	15.75	2.50	-2.655352
ENSG00000169752	17.50	3.25	-2.428843
ENSG00000169783	1.50	0.25	-2.584963
ENSG00000169884	3.00	0.25	-3.584963
ENSG00000170965	10.50	1.00	-3.392317
ENSG00000171346	5.25	1.25	-2.070389
ENSG00000171695	1.25	0.25	-2.321928
ENSG00000172061	381.50	71.75	-2.410632
ENSG00000173110	24.25	3.75	-2.693022
ENSG00000174473	3.25	0.50	-2.700440
ENSG00000174837	5.25	1.00	-2.392317

ENSG00000175489	19.50	4.00	-2.285402
ENSG00000175928	24.00	4.25	-2.497500
ENSG00000176020	7.25	0.75	-3.273018
ENSG00000176402	2.00	0.25	-3.000000
ENSG00000176654	1.25	0.25	-2.321928
ENSG00000176771	197.25	38.25	-2.366494
ENSG00000177108	1.75	0.25	-2.807355
ENSG00000177301	8.25	1.75	-2.237039
ENSG00000177508	8.25	0.75	-3.459432
ENSG00000177570	165.25	40.75	-2.019778
ENSG00000177614	286.25	48.25	-2.568675
ENSG00000177994	2.25	0.50	-2.169925
ENSG00000178695	4860.25	824.00	-2.560314
ENSG00000179270	3.25	0.25	-3.700440
ENSG00000179813	7.50	0.50	-3.906891
ENSG00000180245	1.75	0.25	-2.807355
ENSG00000181634	46.75	11.00	-2.087463
ENSG00000181656	1.75	0.25	-2.807355
ENSG00000181790	2.50	0.50	-2.321928
ENSG00000181984	2.75	0.25	-3.459432
ENSG00000182574	1.25	0.25	-2.321928
ENSG00000182674	11.75	2.00	-2.554589
ENSG00000182732	10.50	2.25	-2.222392
ENSG00000183090	4.50	0.50	-3.169925
ENSG00000183092	12.50	1.75	-2.836501
ENSG00000183311	2489.50	581.75	-2.097385
ENSG00000183395	5.25	0.25	-4.392317
ENSG00000183454	26.50	4.75	-2.479993
ENSG00000183508	167.00	28.75	-2.538214
ENSG00000184261	3.00	0.50	-2.584963
ENSG00000184564	84.00	11.25	-2.900464
ENSG00000184599	2.50	0.25	-3.321928
ENSG00000185306	12.00	2.00	-2.584963
ENSG00000186265	2.50	0.50	-2.321928
ENSG00000186472	2.25	0.50	-2.169925
ENSG00000187753	2.25	0.50	-2.169925
ENSG00000187848	3.75	0.75	-2.321928
ENSG00000187944	3.75	0.50	-2.906891
ENSG00000188176	106.75	11.75	-3.183503
ENSG00000188501	46.75	5.50	-3.087463
ENSG00000188582	3.00	0.50	-2.584963
ENSG00000188730	2.50	0.25	-3.321928
ENSG00000188833	2.00	0.25	-3.000000

ENSG00000196218	16.00	2.00 -3.000000
ENSG00000196517	432.50	80.50 -2.425639
ENSG00000196660	2.50	0.50 -2.321928
ENSG00000196932	130.75	30.25 -2.111804
ENSG00000197744	2.25	0.50 -2.169925
ENSG00000198734	6.00	0.50 -3.584963
ENSG00000203618	11.25	1.00 -3.491853
ENSG00000203697	3.75	0.50 -2.906891
ENSG00000204060	9.25	1.00 -3.209453
ENSG00000204335	13.25	1.25 -3.405992
ENSG00000204390	7.25	1.75 -2.050626
ENSG00000204439	42.00	10.00 -2.070389
ENSG00000204882	3.50	0.50 -2.807355
ENSG00000206172	18.50	2.50 -2.887525
ENSG00000206287	52.75	2.50 -4.399171
ENSG00000206463	1.25	0.25 -2.321928
ENSG00000206549	1.50	0.25 -2.584963
ENSG00000212643	10.00	2.25 -2.152003
ENSG00000212664	4.25	0.50 -3.087463
ENSG00000213174	1.25	0.25 -2.321928
ENSG00000213830	1.25	0.25 -2.321928
ENSG00000214210	1.50	0.25 -2.584963
ENSG00000214814	472.75	53.50 -3.143467
ENSG00000215529	2.25	0.50 -2.169925
ENSG00000215912	8.50	2.00 -2.087463
ENSG00000216754	1.25	0.25 -2.321928
ENSG00000218175	3.00	0.50 -2.584963
ENSG00000219133	1.25	0.25 -2.321928
ENSG00000219881	1.25	0.25 -2.321928
ENSG00000220685	3.00	0.50 -2.584963
ENSG00000223668	1.50	0.25 -2.584963
ENSG00000223687	22.25	1.25 -4.153805
ENSG00000224864	3.50	0.75 -2.222392
ENSG00000225217	2.75	0.50 -2.459432
ENSG00000226498	3.00	0.50 -2.584963
ENSG00000226624	1.75	0.25 -2.807355
ENSG00000227176	1.25	0.25 -2.321928
ENSG00000227315	21.00	2.25 -3.222392
ENSG00000227582	3.25	0.50 -2.700440
ENSG00000227868	1.25	0.25 -2.321928
ENSG00000228057	1.25	0.25 -2.321928
ENSG00000228289	1.75	0.25 -2.807355
ENSG00000228665	1.25	0.25 -2.321928

ENSG00000228921	1.25	0.25	-2.321928
ENSG00000228998	4.50	0.75	-2.584963
ENSG00000229207	1.50	0.25	-2.584963
ENSG00000229808	2.50	0.50	-2.321928
ENSG00000230295	4.50	0.75	-2.584963
ENSG00000230305	1.25	0.25	-2.321928
ENSG00000230447	1.25	0.25	-2.321928
ENSG00000231561	2.25	0.50	-2.169925
ENSG00000232075	1.50	0.25	-2.584963
ENSG00000232111	3.75	0.75	-2.321928
ENSG00000232414	2.50	0.50	-2.321928
ENSG00000232953	1.50	0.25	-2.584963
ENSG00000233132	3.25	0.50	-2.700440
ENSG00000233239	1.25	0.25	-2.321928
ENSG00000233430	1.50	0.25	-2.584963
ENSG00000233673	1.50	0.25	-2.584963
ENSG00000234258	7.25	1.75	-2.050626
ENSG00000235107	52.75	2.50	-4.399171
ENSG00000235109	139.00	32.00	-2.118941
ENSG00000235205	2.00	0.25	-3.000000
ENSG00000235297	1.75	0.25	-2.807355
ENSG00000235708	1.75	0.25	-2.807355
ENSG00000235974	1.25	0.25	-2.321928
ENSG00000236047	2.75	0.50	-2.459432
ENSG00000237039	1.50	0.25	-2.584963
ENSG00000237161	2.00	0.25	-3.000000
ENSG00000237977	2.25	0.50	-2.169925
ENSG00000240821	2.00	0.25	-3.000000
ENSG00000240914	1.25	0.25	-2.321928
ENSG00000241146	2.25	0.50	-2.169925
ENSG00000241166	3.25	0.75	-2.115477
ENSG00000241431	1.75	0.25	-2.807355
ENSG00000241735	2.25	0.25	-3.169925
ENSG00000241782	1.50	0.25	-2.584963
ENSG00000242156	1.50	0.25	-2.584963
ENSG00000242457	2.50	0.50	-2.321928
ENSG00000243627	1.50	0.25	-2.584963
ENSG00000243742	59.75	11.00	-2.441435
ENSG00000244018	2.00	0.25	-3.000000
ENSG00000244255	108.25	15.75	-2.780943
ENSG00000244313	4.25	0.50	-3.087463
ENSG00000248290	5.50	1.00	-2.459432
ENSG00000248909	1.25	0.25	-2.321928

ENSG00000250090	3.00	0.25 -3.584963
ENSG00000250424	5.75	1.25 -2.201634
ENSG00000250645	3.75	0.75 -2.321928
ENSG00000251495	1.75	0.25 -2.807355
ENSG00000253333	1.50	0.25 -2.584963
ENSG00000253543	1.25	0.25 -2.321928
ENSG00000253620	1.25	0.25 -2.321928
ENSG00000253934	1.75	0.25 -2.807355
ENSG00000254415	1.50	0.25 -2.584963
ENSG00000254483	1.50	0.25 -2.584963
ENSG00000254701	11.75	1.50 -2.969626
ENSG00000254778	10.50	2.50 -2.070389
ENSG00000254875	4.50	0.50 -3.169925
ENSG00000255583	1.25	0.25 -2.321928
ENSG00000256407	1.25	0.25 -2.321928
ENSG00000257198	3.00	0.50 -2.584963
ENSG00000257243	1.50	0.25 -2.584963
ENSG00000258142	5.75	1.25 -2.201634
ENSG00000258186	2.00	0.25 -3.000000
ENSG00000258628	2.25	0.50 -2.169925
ENSG00000258788	5.25	1.25 -2.070389
ENSG00000260139	2.50	0.25 -3.321928
ENSG00000260326	2.00	0.25 -3.000000
ENSG00000260342	66.75	13.25 -2.332775
ENSG00000260571	1.25	0.25 -2.321928
ENSG00000261581	1.75	0.25 -2.807355
ENSG00000262633	2.75	0.50 -2.459432
ENSG00000262874	2.25	0.25 -3.169925
ENSG00000263606	1.50	0.25 -2.584963
ENSG00000265763	22.00	3.25 -2.758992
ENSG00000265818	12.00	1.75 -2.777608
ENSG00000266265	5.00	1.00 -2.321928
ENSG00000266853	1.50	0.25 -2.584963
ENSG00000266995	1.50	0.25 -2.584963
ENSG00000267091	1.75	0.25 -2.807355
ENSG00000268083	2.50	0.50 -2.321928
ENSG00000268223	1.50	0.25 -2.584963
ENSG00000268529	1.25	0.25 -2.321928
ENSG00000268643	11.25	1.75 -2.684498
ENSG00000268861	7.00	1.00 -2.807355
ENSG00000268870	1.50	0.25 -2.584963
ENSG00000268951	1.25	0.25 -2.321928
ENSG00000270775	2.25	0.50 -2.169925

ENSG00000270906	2.00	0.25	-3.000000
ENSG00000270986	2.75	0.50	-2.459432
ENSG00000271153	4.50	0.50	-3.169925
ENSG00000271779	1.75	0.25	-2.807355
ENSG00000271890	1.50	0.25	-2.584963
ENSG00000273079	5.50	1.25	-2.137504
ENSG00000273336	1.25	0.25	-2.321928
ENSG00000273398	10.00	2.25	-2.152003
ENSG00000273715	13.75	1.25	-3.459432
ENSG00000273851	2.00	0.25	-3.000000
ENSG00000273933	3.50	0.25	-3.807355
ENSG00000274378	2.25	0.50	-2.169925
ENSG00000274443	6.00	1.25	-2.263034
ENSG00000275153	58.00	11.50	-2.334419
ENSG00000275181	3.50	0.50	-2.807355
ENSG00000275282	1.25	0.25	-2.321928
ENSG00000275833	1.75	0.25	-2.807355
ENSG00000276375	2.50	0.50	-2.321928
ENSG00000276644	4.75	1.00	-2.247928
ENSG00000277125	5.50	1.25	-2.137504
ENSG00000277397	10.25	1.75	-2.550197
ENSG00000277614	75.75	17.75	-2.093427
ENSG00000277669	13.25	1.00	-3.727920
ENSG00000277977	4.50	1.00	-2.169925
ENSG00000278373	2.25	0.50	-2.169925
ENSG00000278599	1.50	0.25	-2.584963
ENSG00000279062	1.25	0.25	-2.321928
ENSG00000279483	180.75	13.75	-3.716492
ENSG00000280221	2.50	0.25	-3.321928
ENSG00000282090	1.75	0.25	-2.807355
ENSG00000282144	10.00	1.00	-3.321928
ENSG00000282432	2.25	0.25	-3.169925
ENSG00000282501	2.25	0.50	-2.169925
ENSG00000282832	2.50	0.50	-2.321928
ENSG00000282913	1.50	0.25	-2.584963
ENSG00000282970	1.25	0.25	-2.321928

```
#mycounts <- meancounts[to.rm_1, ]
```

Q. How many “up” at $\log_2\text{fc} > +2$


```
to.rm_2 <- mycounts$log2fc > 2
mycounts[to.rm_2, ]
```

	control.mean	treated.mean	log2fc
ENSG00000004799	270.50	1429.25	2.401558
ENSG00000006788	2.75	19.75	2.844349
ENSG00000008438	0.50	2.75	2.459432
ENSG00000011677	0.50	2.25	2.169925
ENSG00000015413	0.50	3.00	2.584963
ENSG00000015592	0.50	2.25	2.169925
ENSG00000046653	323.00	2126.50	2.718875
ENSG00000070190	0.50	3.00	2.584963
ENSG00000070388	3.50	17.50	2.321928
ENSG00000074317	0.25	1.75	2.807355
ENSG00000086717	1.00	4.50	2.169925
ENSG00000088320	1.00	4.75	2.247928
ENSG00000090382	0.25	1.75	2.807355
ENSG00000096060	354.75	4646.00	3.711115
ENSG00000099337	315.50	1556.25	2.302362
ENSG00000099998	199.50	948.25	2.248879
ENSG00000100033	7.25	89.00	3.617752
ENSG00000100276	0.25	1.25	2.321928
ENSG00000101342	6.25	91.50	3.871844
ENSG00000101347	1951.25	24068.25	3.624661
ENSG00000101977	2.75	11.50	2.064130
ENSG00000102554	48.00	271.25	2.498517
ENSG00000102760	247.25	1951.25	2.980356
ENSG00000103196	930.50	4864.75	2.386288
ENSG00000104059	0.50	2.25	2.169925
ENSG00000107485	0.50	2.25	2.169925
ENSG00000108242	0.25	1.50	2.584963
ENSG00000109906	14.75	808.75	5.776907
ENSG00000110786	0.25	1.25	2.321928
ENSG00000112499	0.25	1.25	2.321928
ENSG00000112936	404.50	3740.25	3.208923
ENSG00000115009	0.50	2.25	2.169925
ENSG00000115468	3.75	16.75	2.159199
ENSG00000115474	0.50	3.75	2.906891
ENSG00000116285	1107.50	5372.25	2.278220
ENSG00000118194	2.50	12.00	2.263034
ENSG00000119508	88.25	359.50	2.026324
ENSG00000120129	843.50	5737.25	2.765899

ENSG00000122592	0.75	3.25	2.115477
ENSG00000124557	0.25	1.25	2.321928
ENSG00000125144	6.25	28.75	2.201634
ENSG00000125148	1405.25	5970.25	2.086965
ENSG00000125872	0.75	4.75	2.662965
ENSG00000126803	152.75	895.50	2.551521
ENSG00000127324	3.25	42.00	3.691878
ENSG00000127954	34.25	826.75	4.593275
ENSG00000128045	16.75	219.50	3.711988
ENSG00000128052	1.25	6.00	2.263034
ENSG00000128917	11.25	163.50	3.861294
ENSG00000130294	0.25	2.00	3.000000
ENSG00000131864	0.25	1.25	2.321928
ENSG00000132518	1.50	8.00	2.415037
ENSG00000133169	2.00	8.25	2.044394
ENSG00000133980	0.50	2.50	2.321928
ENSG00000134160	0.50	2.75	2.459432
ENSG00000134243	2153.50	9028.50	2.067803
ENSG00000134709	0.75	5.75	2.938599
ENSG00000135312	1.50	10.25	2.772590
ENSG00000135604	7.00	33.00	2.237039
ENSG00000135821	4656.00	33031.50	2.826680
ENSG00000136237	24.50	214.75	3.131804
ENSG00000137673	0.25	10.25	5.357552
ENSG00000138075	0.25	2.50	3.321928
ENSG00000138483	0.75	10.50	3.807355
ENSG00000138823	1.50	8.25	2.459432
ENSG00000138829	6770.25	30340.00	2.163940
ENSG00000139055	6.50	29.00	2.157541
ENSG00000139540	0.50	3.00	2.584963
ENSG00000140465	0.50	5.00	3.321928
ENSG00000140807	61.00	276.75	2.181702
ENSG00000140955	0.25	1.25	2.321928
ENSG00000141401	239.00	987.25	2.046405
ENSG00000143127	53.25	424.75	2.995761
ENSG00000144010	0.25	2.25	3.169925
ENSG00000144852	0.25	1.25	2.321928
ENSG00000145390	2964.75	12845.25	2.115253
ENSG00000145569	8.00	41.75	2.383704
ENSG00000148677	0.75	12.50	4.058894
ENSG00000149043	0.25	1.50	2.584963
ENSG00000149294	0.50	5.25	3.392317
ENSG00000150907	244.75	1062.50	2.118082

ENSG00000152463	2.00	16.75	3.066089
ENSG00000152503	1.50	7.50	2.321928
ENSG00000152583	95.75	1737.50	4.181597
ENSG00000152669	0.25	2.50	3.321928
ENSG00000152779	14.50	89.75	2.629859
ENSG00000154734	11456.75	47855.25	2.062479
ENSG00000157150	49.50	355.00	2.842319
ENSG00000157152	11.00	70.00	2.669851
ENSG00000157502	1.25	8.25	2.722466
ENSG00000157514	424.00	3615.75	3.092159
ENSG00000158022	0.75	5.50	2.874469
ENSG00000158246	25.25	116.75	2.209067
ENSG00000159167	330.50	1830.00	2.469121
ENSG00000160838	0.50	2.25	2.169925
ENSG00000162267	2.75	16.75	2.606658
ENSG00000162630	11.50	74.25	2.690757
ENSG00000163017	520.75	2163.25	2.054538
ENSG00000163083	67.75	915.00	3.755479
ENSG00000163251	18.75	83.50	2.154886
ENSG00000163283	0.25	1.25	2.321928
ENSG00000163661	3710.00	15357.25	2.049429
ENSG00000163884	50.50	991.00	4.294530
ENSG00000163995	1.00	12.00	3.584963
ENSG00000164122	7.50	36.25	2.273018
ENSG00000164142	2.25	17.50	2.959358
ENSG00000165443	0.25	4.00	4.000000
ENSG00000165730	0.75	5.25	2.807355
ENSG00000165995	121.00	835.75	2.788064
ENSG00000166741	2907.75	12172.75	2.065680
ENSG00000167549	695.00	3086.25	2.150770
ENSG00000167641	67.25	431.25	2.680918
ENSG00000167769	0.25	1.25	2.321928
ENSG00000168309	21.00	258.50	3.621703
ENSG00000168481	2.00	33.00	4.044394
ENSG00000168702	1.25	6.50	2.378512
ENSG00000169271	29.25	156.75	2.421957
ENSG00000170214	32.25	354.75	3.459432
ENSG00000170323	2.25	19.25	3.096862
ENSG00000171051	0.75	6.00	3.000000
ENSG00000171819	9.00	427.25	5.569012
ENSG00000172362	0.25	1.25	2.321928
ENSG00000173432	0.25	4.25	4.087463
ENSG00000173838	12.50	122.25	3.289834

ENSG00000174453	0.50	2.50	2.321928
ENSG00000174576	1.00	5.00	2.321928
ENSG00000174697	29.50	230.75	2.967544
ENSG00000175294	0.25	2.00	3.000000
ENSG00000177283	49.50	252.00	2.347923
ENSG00000178015	6.75	36.75	2.444785
ENSG00000178715	0.50	2.25	2.169925
ENSG00000179094	193.00	1249.50	2.694678
ENSG00000179300	15.75	115.00	2.868210
ENSG00000179593	0.25	129.50	9.016808
ENSG00000180613	0.25	1.75	2.807355
ENSG00000182836	7.25	55.75	2.942919
ENSG00000183169	1.50	9.50	2.662965
ENSG00000184156	5.75	31.00	2.430634
ENSG00000185313	0.25	1.25	2.321928
ENSG00000185352	0.50	6.75	3.754888
ENSG00000185432	1404.50	5875.25	2.064594
ENSG00000185436	0.25	2.00	3.000000
ENSG00000185522	1.75	8.25	2.237039
ENSG00000185950	881.00	3570.50	2.018912
ENSG00000186919	1.75	8.00	2.192645
ENSG00000187193	40.25	336.75	3.064617
ENSG00000187288	3.50	33.00	3.237039
ENSG00000188916	18.00	80.75	2.165465
ENSG00000189221	467.25	4267.00	3.190956
ENSG00000189366	0.50	3.25	2.700440
ENSG00000196136	155.75	1240.75	2.993908
ENSG00000196826	5.50	22.25	2.016302
ENSG00000197381	725.00	3388.25	2.224487
ENSG00000197769	44.50	200.75	2.173523
ENSG00000198624	536.25	3314.75	2.627922
ENSG00000204965	0.75	3.25	2.115477
ENSG00000205089	2.75	14.00	2.347923
ENSG00000205358	0.25	1.50	2.584963
ENSG00000205364	13.75	69.25	2.332382
ENSG00000206285	10.75	44.50	2.049469
ENSG00000211445	1856.00	20692.25	3.478822
ENSG00000213018	0.25	1.50	2.584963
ENSG00000217783	0.25	1.25	2.321928
ENSG00000219088	0.25	3.25	3.700440
ENSG00000219565	1.50	8.75	2.544321
ENSG00000219642	0.50	3.25	2.700440
ENSG00000221852	0.75	4.00	2.415037

ENSG00000221866	223.75	1791.25	3.001007
ENSG00000221887	1.50	8.50	2.502500
ENSG00000223638	0.50	7.00	3.807355
ENSG00000224650	0.25	1.50	2.584963
ENSG00000224668	0.50	2.25	2.169925
ENSG00000224712	1.50	7.50	2.321928
ENSG00000225419	0.25	1.25	2.321928
ENSG00000227141	0.50	5.00	3.321928
ENSG00000227694	0.50	2.75	2.459432
ENSG00000228612	0.25	1.50	2.584963
ENSG00000228725	0.25	1.25	2.321928
ENSG00000229292	0.25	2.25	3.169925
ENSG00000229332	0.25	2.50	3.321928
ENSG00000230018	13.00	58.75	2.176077
ENSG00000230245	0.25	1.50	2.584963
ENSG00000230341	0.25	4.50	4.169925
ENSG00000230395	0.50	2.50	2.321928
ENSG00000230524	0.25	2.25	3.169925
ENSG00000230712	2.75	11.75	2.095157
ENSG00000231550	0.25	1.25	2.321928
ENSG00000231991	1.00	5.00	2.321928
ENSG00000232133	1.75	8.25	2.237039
ENSG00000234000	0.25	4.50	4.169925
ENSG00000235022	1.00	6.25	2.643856
ENSG00000235579	0.25	1.25	2.321928
ENSG00000236060	0.25	1.25	2.321928
ENSG00000236512	0.25	1.25	2.321928
ENSG00000236801	1.50	6.25	2.058894
ENSG00000236967	1.00	6.25	2.643856
ENSG00000237506	0.25	2.75	3.459432
ENSG00000240520	0.25	1.50	2.584963
ENSG00000240935	0.25	1.50	2.584963
ENSG00000241713	0.25	7.25	4.857981
ENSG00000242441	0.50	6.00	3.584963
ENSG00000243444	5.00	29.75	2.572890
ENSG00000244301	3.25	14.00	2.106915
ENSG00000248180	0.25	1.50	2.584963
ENSG00000248769	0.25	1.75	2.807355
ENSG00000249936	1.25	7.00	2.485427
ENSG00000249967	2.50	13.25	2.405992
ENSG00000250197	0.25	1.25	2.321928
ENSG00000250803	0.25	1.75	2.807355
ENSG00000250942	0.25	1.25	2.321928

ENSG00000251056	0.50	2.75	2.459432
ENSG00000254294	0.25	1.25	2.321928
ENSG00000254332	0.50	3.25	2.700440
ENSG00000255967	0.50	3.00	2.584963
ENSG00000256861	4.00	42.75	3.417853
ENSG00000257150	2.00	10.25	2.357552
ENSG00000258741	0.25	3.50	3.807355
ENSG00000260170	1.75	10.00	2.514573
ENSG00000261862	0.75	4.00	2.415037
ENSG00000265713	0.25	1.25	2.321928
ENSG00000267472	0.50	2.50	2.321928
ENSG00000267733	0.50	2.25	2.169925
ENSG00000269823	1.50	15.25	3.345775
ENSG00000270578	0.50	2.75	2.459432
ENSG00000270689	0.50	7.50	3.906891
ENSG00000271134	0.25	1.25	2.321928
ENSG00000271321	0.25	1.50	2.584963
ENSG00000271581	0.25	1.25	2.321928
ENSG00000273259	382.75	3441.25	3.168458
ENSG00000274183	0.25	1.50	2.584963
ENSG00000274892	0.25	1.25	2.321928
ENSG00000274933	0.25	1.75	2.807355
ENSG00000274944	2.75	57.75	4.392317
ENSG00000276469	0.25	1.75	2.807355
ENSG00000277196	0.50	63.75	6.994353
ENSG00000277354	3.50	22.00	2.652077
ENSG00000277399	0.50	13.25	4.727920
ENSG00000277983	6.25	28.75	2.201634
ENSG00000277984	0.25	1.50	2.584963
ENSG00000278186	0.75	5.00	2.736966
ENSG00000278516	3.50	16.00	2.192645
ENSG00000278608	4.75	26.50	2.479993
ENSG00000278817	1.75	25.00	3.836501
ENSG00000278843	0.75	7.00	3.222392
ENSG00000279767	1.50	15.25	3.345775
ENSG00000279986	0.75	3.75	2.321928
ENSG00000280148	0.25	1.50	2.584963
ENSG00000280641	1.25	15.25	3.608809
ENSG00000281527	0.50	2.50	2.321928
ENSG00000282003	9.75	39.50	2.018379

Q. Do we trust these results? Is there anything missing?

We are missing the stats!!!

DESeq analysis

```
#!/ message: false  
library(DESeq2)
```

DESeq, like many BioConductor packages, wants our input data in a very specific format.

```
dds <- DESeqDataSetFromMatrix(countData = counts,  
                              colData = metadata,  
                              design = ~dex)
```

converting counts to integer mode

Warning in DESeqDataSet(se, design = design, ignoreRank): some variables in design formula are characters, converting to factors

The main function in DESeq2 is called DESeq().

```
dds <- DESeq(dds)
```

estimating size factors

estimating dispersions

gene-wise dispersion estimates

mean-dispersion relationship

final dispersion estimates

fitting model and testing

```
res <- results(dds)
```

```
head(res)
```

log2 fold change (MLE): dex treated vs control

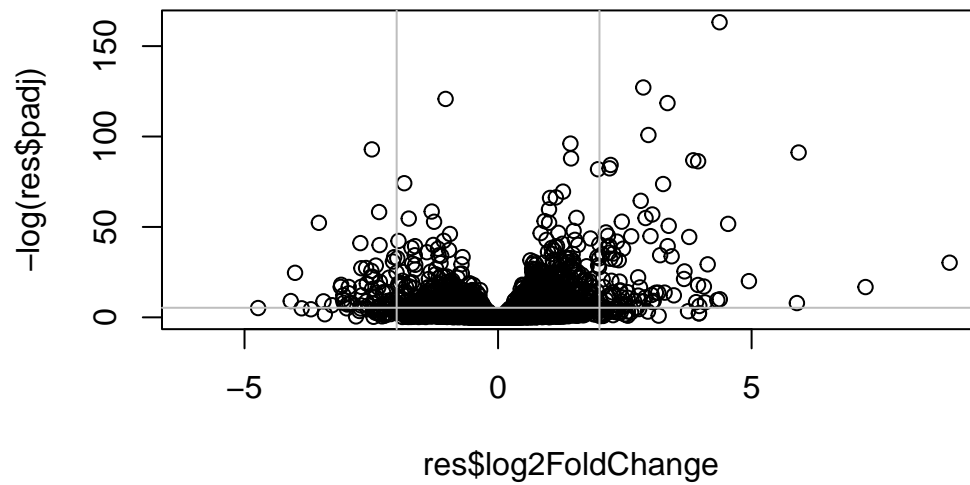
Wald test p-value: dex treated vs control

DataFrame with 6 rows and 6 columns

	baseMean	log2FoldChange	lfcSE	stat	pvalue
	<numeric>	<numeric>	<numeric>	<numeric>	<numeric>
ENSG000000000003	747.194195	-0.3507030	0.168246	-2.084470	0.0371175
ENSG000000000005	0.000000	NA	NA	NA	NA
ENSG000000000419	520.134160	0.2061078	0.101059	2.039475	0.0414026
ENSG000000000457	322.664844	0.0245269	0.145145	0.168982	0.8658106
ENSG000000000460	87.682625	-0.1471420	0.257007	-0.572521	0.5669691
ENSG000000000938	0.319167	-1.7322890	3.493601	-0.495846	0.6200029
	padj				
	<numeric>				
ENSG000000000003	0.163035				
ENSG000000000005	NA				
ENSG000000000419	0.176032				
ENSG000000000457	0.961694				
ENSG000000000460	0.815849				
ENSG000000000938	NA				

A common overview figure plots the logFC vs P-value

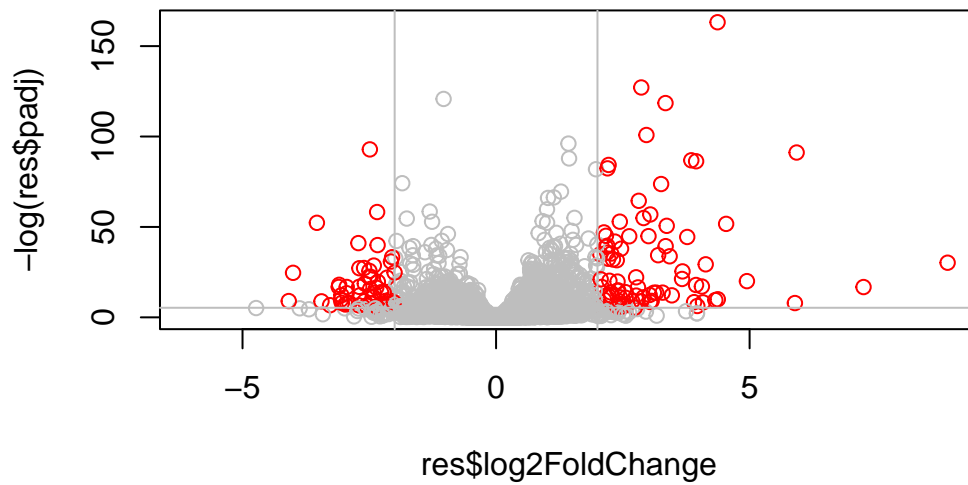
```
plot(res$log2FoldChange, -log(res$padj))
abline(v=c(-2, 2), col="gray")
abline(h = -log(0.005), col="gray")
```

```
mycols <- rep("gray", nrow(res))

mycols[res$log2FoldChange > 2] <- "red"
mycols[res$log2FoldChange < -2] <- "red"
mycols[res$padj > 0.005] <- "gray"
#mycols[res$log2FoldChange > -2 & res$log2FoldChange < 2] <- "green"

plot(res$log2FoldChange, -log(res$padj), col = mycols)
abline(v=c(-2, 2), col="gray")
abline(h = -log(0.005), col="gray")
```



```
write.csv(res, file = "myresults.csv")
```

Gene annotation

```
library("AnnotationDbi")
library("org.Hs.eg.db")
```

```
res$symbol <- mapIds(org.Hs.eg.db,
                     keys=row.names(res), # Our genenames
                     keytype="ENSEMBL",   # The format of our genenames
                     column="SYMBOL",     # The new format we want to add
                     multiVals="first")
```

'select()' returned 1:many mapping between keys and columns

```
head(res)
```

log2 fold change (MLE): dex treated vs control

Wald test p-value: dex treated vs control

DataFrame with 6 rows and 7 columns

	baseMean	log2FoldChange	lfcSE	stat	pvalue
	<numeric>	<numeric>	<numeric>	<numeric>	<numeric>
ENSG000000000003	747.194195	-0.3507030	0.168246	-2.084470	0.0371175
ENSG000000000005	0.000000	NA	NA	NA	NA
ENSG000000000419	520.134160	0.2061078	0.101059	2.039475	0.0414026
ENSG000000000457	322.664844	0.0245269	0.145145	0.168982	0.8658106
ENSG000000000460	87.682625	-0.1471420	0.257007	-0.572521	0.5669691
ENSG000000000938	0.319167	-1.7322890	3.493601	-0.495846	0.6200029
	padj	symbol			
	<numeric>	<character>			
ENSG000000000003	0.163035	TSPAN6			
ENSG000000000005	NA	TNMD			
ENSG000000000419	0.176032	DPM1			
ENSG000000000457	0.961694	SCYL3			
ENSG000000000460	0.815849	FIRRM			
ENSG000000000938	NA	FGR			

Pathway analysis

A quick KEGG pathway analysis with the **gage** function.

```
library(pathview)
```

```
#####
Pathview is an open source software package distributed under GNU General
Public License version 3 (GPLv3). Details of GPLv3 is available at
http://www.gnu.org/licenses/gpl-3.0.html. Particullary, users are required to
formally cite the original Pathview paper (not just mention it) in publications
or products. For details, do citation("pathview") within R.
```

The pathview downloads and uses KEGG data. Non-academic uses may require a KEGG license agreement (details at <http://www.kegg.jp/kegg/legal.html>).

```
#####
```

```
library(gage)
```

```
library(gageData)

data(kegg.sets.hs)

# Examine the first 2 pathways in this kegg set for humans
head(kegg.sets.hs, 2)
```

```
$`hsa00232 Caffeine metabolism`
[1] "10"    "1544" "1548" "1549" "1553" "7498" "9"

$`hsa00983 Drug metabolism - other enzymes`
[1] "10"      "1066"   "10720"  "10941"  "151531" "1548"   "1549"   "1551"
[9] "1553"    "1576"   "1577"   "1806"   "1807"   "1890"   "221223" "2990"
[17] "3251"    "3614"   "3615"   "3704"   "51733"  "54490"  "54575"  "54576"
[25] "54577"   "54578"  "54579"  "54600"  "54657"  "54658"  "54659"  "54963"
[33] "574537"  "64816"  "7083"   "7084"   "7172"   "7363"   "7364"   "7365"
[41] "7366"    "7367"   "7371"   "7372"   "7378"   "7498"   "79799"  "83549"
[49] "8824"    "8833"   "9"      "978"
```

I need to speak ENTREZID so I can check KEGG pathway overlap as KEGG uses ENTREZ format IDs

```
res$entrez <- mapIds(org.Hs.eg.db,
                     keys=row.names(res), # Our genenames
                     keytype="ENSEMBL",   # The format of our genenames
                     column="ENTREZID",   # The new format we want to add
                     multiVals="first")
```

'select()' returned 1:many mapping between keys and columns

I can now use the **gage** function to check for overlap with known KEGG pathways

```
foldchanges <- res$log2FoldChange

names(foldchanges) <- res$entrez

head(foldchanges)
```

```
      7105      64102      8813      57147      55732      2268
-0.35070302      NA  0.20610777  0.02452695 -0.14714205 -1.73228897
```

```
# Get the results
keggres = gage(foldchanges, gsets=kegg.sets.hs)
```

```
attributes(keggres)
```

```
$names
[1] "greater" "less"    "stats"
```

```
head(keggres$less, 3)
```

		p.geomean	stat.mean	p.val
hsa05332	Graft-versus-host disease	0.0004250461	-3.473346	0.0004250461
hsa04940	Type I diabetes mellitus	0.0017820293	-3.002352	0.0017820293
hsa05310	Asthma	0.0020045888	-3.009050	0.0020045888

		q.val	set.size	exp1
hsa05332	Graft-versus-host disease	0.09053483	40	0.0004250461
hsa04940	Type I diabetes mellitus	0.14232581	42	0.0017820293
hsa05310	Asthma	0.14232581	29	0.0020045888

```
pathview(gene.data=foldchanges, pathway.id="hsa05310")
```

'select()' returned 1:1 mapping between keys and columns

Info: Working in directory /Users/muditg19/Downloads/BGGN213/class 13

Info: Writing image file hsa05310.pathview.png

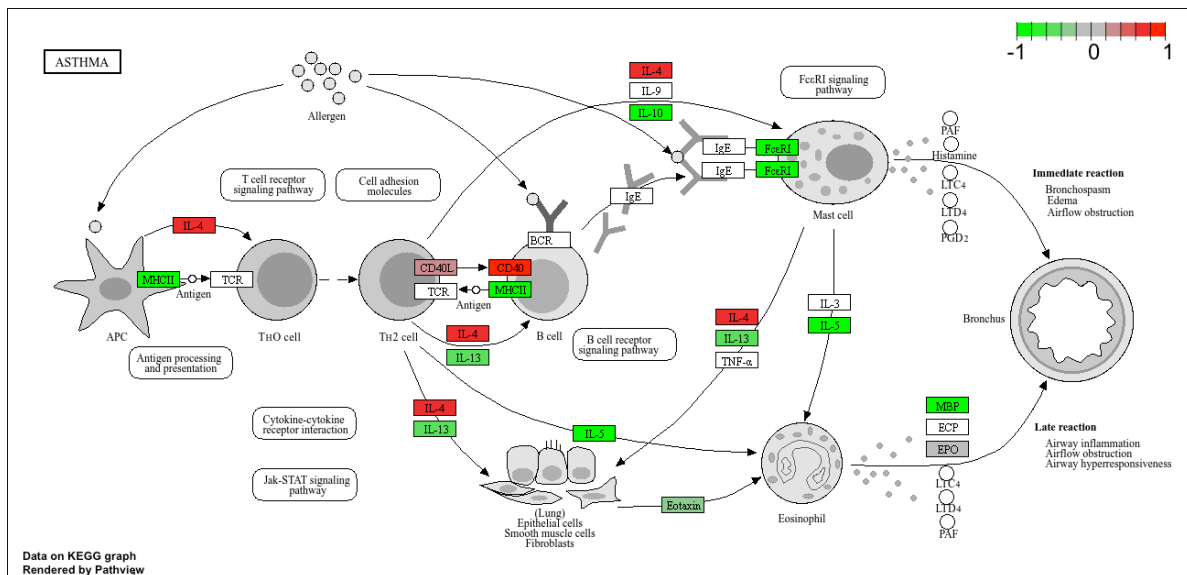


Figure 1: A pathway figure

4. Toy differential gene expression

Note that the control samples are SRR1039508, SRR1039512, SRR1039516, and SRR1039520. This bit of code will first find the sample id for those labeled control. Then calculate the mean counts per gene across these samples:

```
control <- metadata[metadata[, "dex"] == "control", ]
control.counts <- counts[ , control$id]
control.mean <- rowSums( control.counts ) / 4
head(control.mean)
```

```
ENSG000000000003  ENSG000000000005  ENSG000000000419  ENSG000000000457  ENSG000000000460
                900.75                0.00                520.50                339.75                97.25
ENSG000000000938
                0.75
```

An alternative way to do this same thing using the dplyr package from the tidyverse is shown below. Which do you prefer and why?

```
library(dplyr)
```

```
Attaching package: 'dplyr'
```

The following object is masked from 'package:AnnotationDbi':

select

The following object is masked from 'package:Biobase':

combine

The following object is masked from 'package:matrixStats':

count

The following objects are masked from 'package:GenomicRanges':

intersect, setdiff, union

The following object is masked from 'package:GenomeInfoDb':

intersect

The following objects are masked from 'package:IRanges':

collapse, desc, intersect, setdiff, slice, union

The following objects are masked from 'package:S4Vectors':

first, intersect, rename, setdiff, setequal, union

The following objects are masked from 'package:BiocGenerics':

combine, intersect, setdiff, union

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

```
control <- metadata %>% filter(dex=="control")
control.counts <- counts %>% select(control$id)
control.mean <- rowSums(control.counts)/4
head(control.mean)
```

```
ENSG000000000003 ENSG000000000005 ENSG000000000419 ENSG000000000457 ENSG000000000460
          900.75           0.00           520.50           339.75           97.25
ENSG000000000938
          0.75
```