

Brundle Development

Andrew Holding

11/7/2017

Comparison of simple normalisation methods

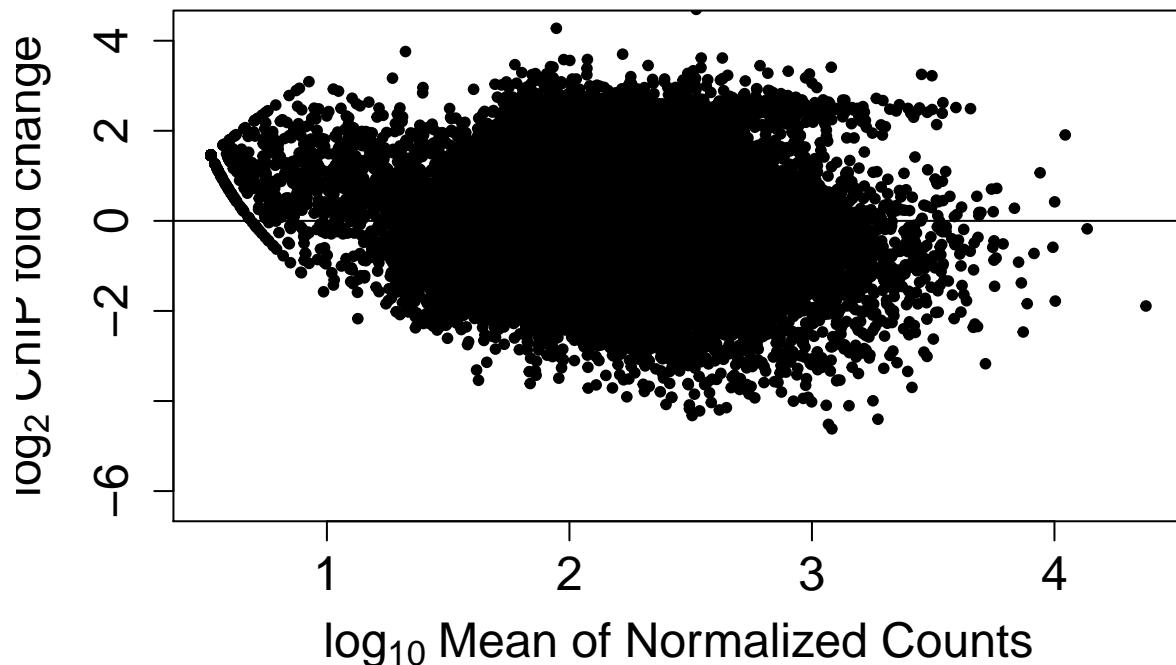
```
## Loading required package: GenomicRanges
## Loading required package: stats4
## Loading required package: BiocGenerics
## Loading required package: parallel
##
## Attaching package: 'BiocGenerics'
## The following objects are masked from 'package:parallel':
## 
##     clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##     clusterExport, clusterMap, parApply, parCapply, parLapply,
##     parLapplyLB, parRapply, parSapply, parSapplyLB
## The following objects are masked from 'package:stats':
## 
##     IQR, mad, sd, var, xtabs
## The following objects are masked from 'package:base':
## 
##     anyDuplicated, append, as.data.frame, cbind, colMeans,
##     colnames, colSums, do.call, duplicated, eval, evalq, Filter,
##     Find, get, grep, grepl, intersect, is.unsorted, lapply,
##     lengths, Map, mapply, match, mget, order, paste, pmax,
##     pmax.int, pmin, pmin.int, Position, rank, rbind, Reduce,
##     rowMeans, rownames, rowSums, sapply, setdiff, sort, table,
##     tapply, union, unique, unsplit, which, which.max, which.min
## Loading required package: S4Vectors
##
## Attaching package: 'S4Vectors'
## The following object is masked from 'package:base':
## 
##     expand.grid
## Loading required package: IRanges
## Loading required package: GenomeInfoDb
## Loading required package: SummarizedExperiment
## Loading required package: Biobase
## Welcome to Bioconductor
##
##     Vignettes contain introductory material; view with
##     'browseVignettes()'. To cite Bioconductor, see
```

```

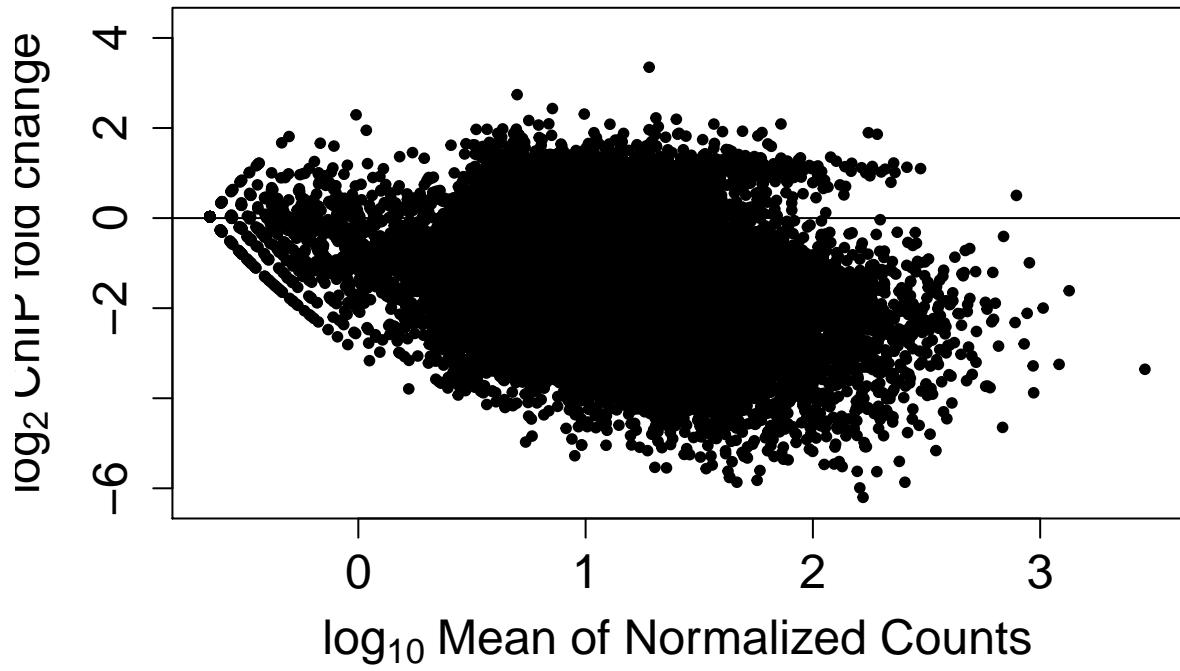
##      'citation("Biobase")', and for packages 'citation("pkgname")'.
## Loading required package: DelayedArray
## Loading required package: matrixStats
##
## Attaching package: 'matrixStats'
## The following objects are masked from 'package:Biobase':
## 
##     anyMissing, rowMedians
##
## Attaching package: 'DelayedArray'
## The following objects are masked from 'package:matrixStats':
## 
##     colMaxs, colMins, colRanges, rowMaxs, rowMins, rowRanges
## The following object is masked from 'package:base':
## 
##     apply
##

```

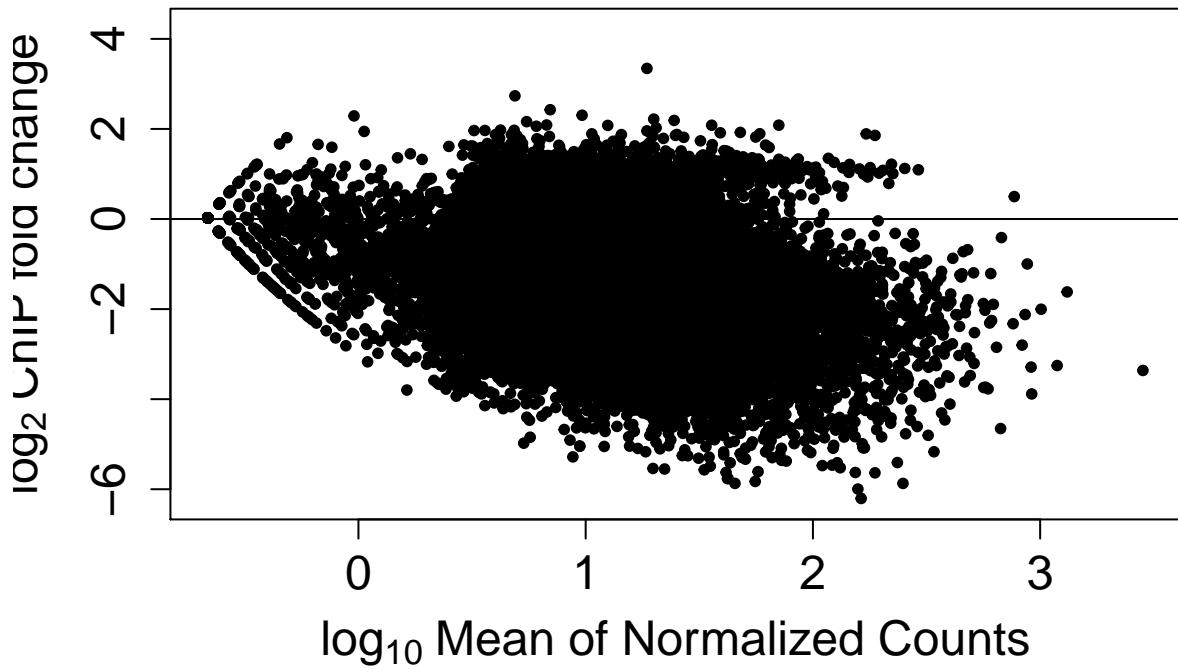
RPM reads in peaks



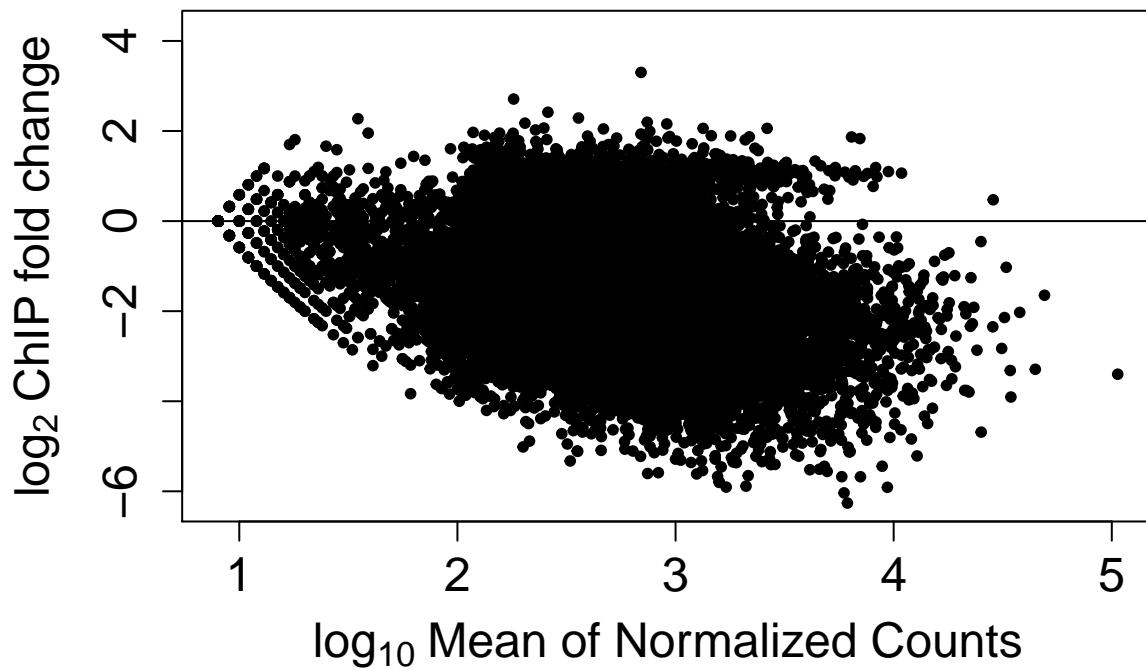
RPM aligned reads



RPM total reads

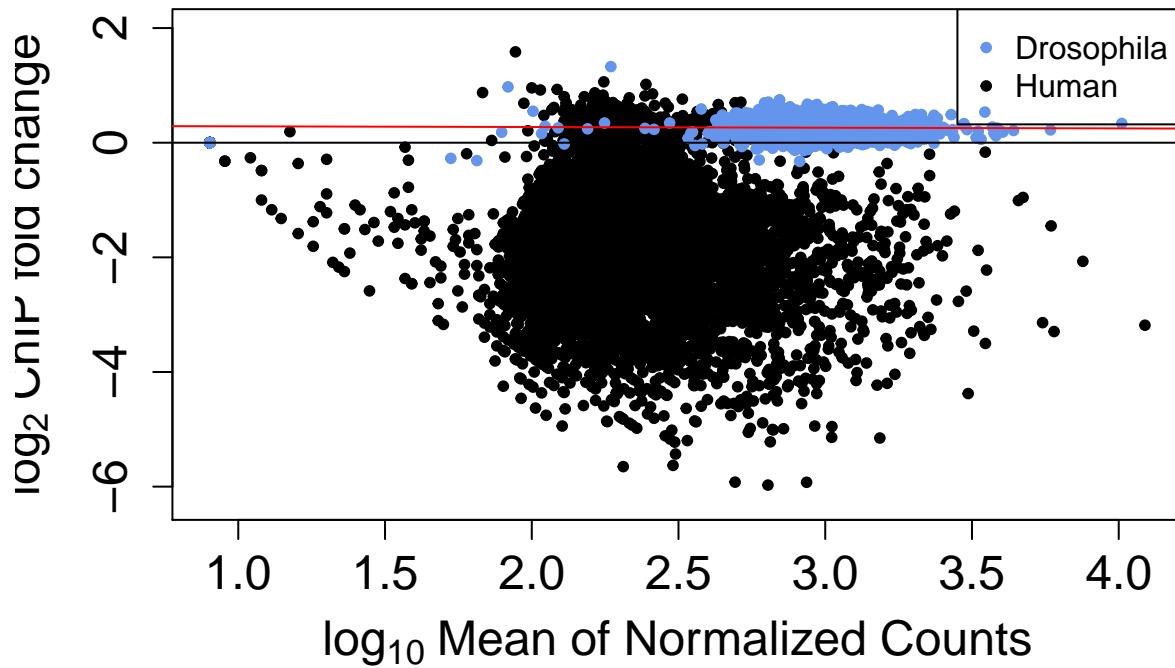


Raw counts in peaks

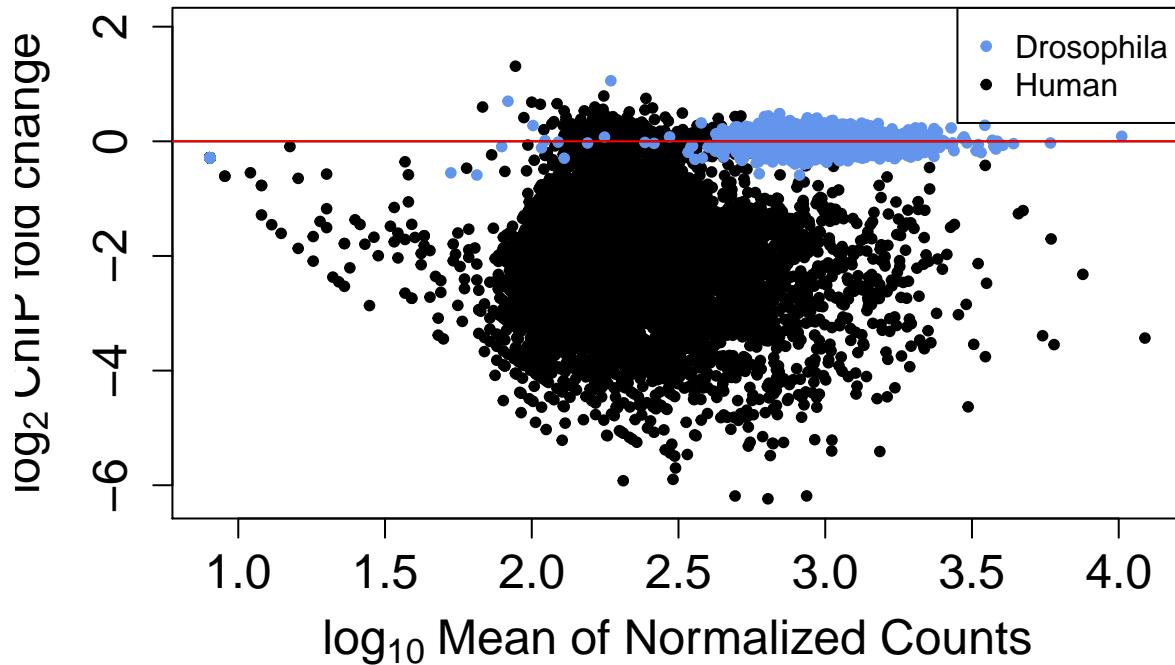


MA Plot of H2AV normalization

RPM aligned reads



Counts normalized by H2av Distribution

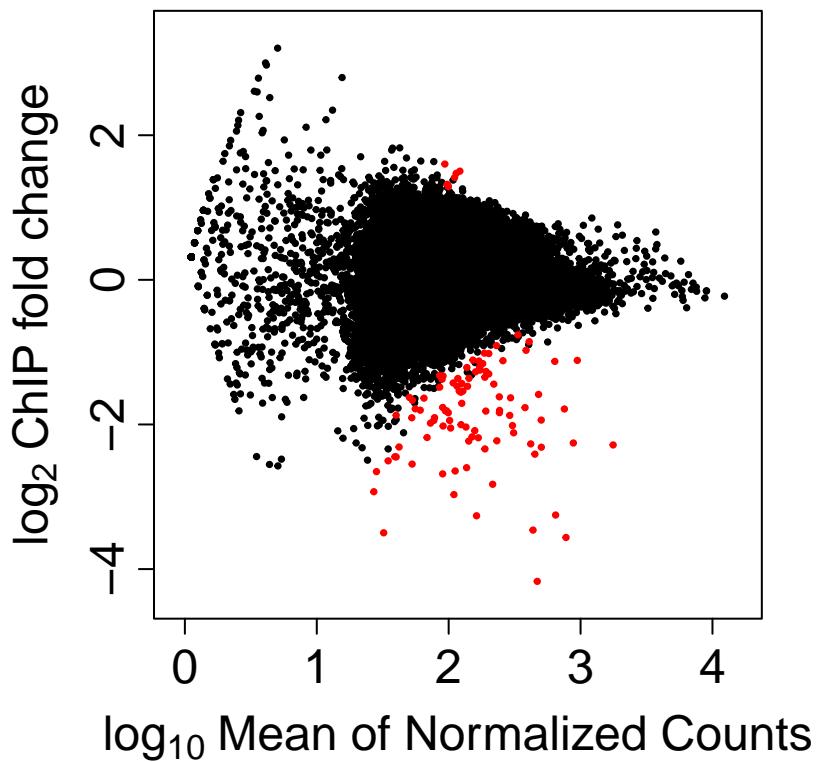


DeSeq Results

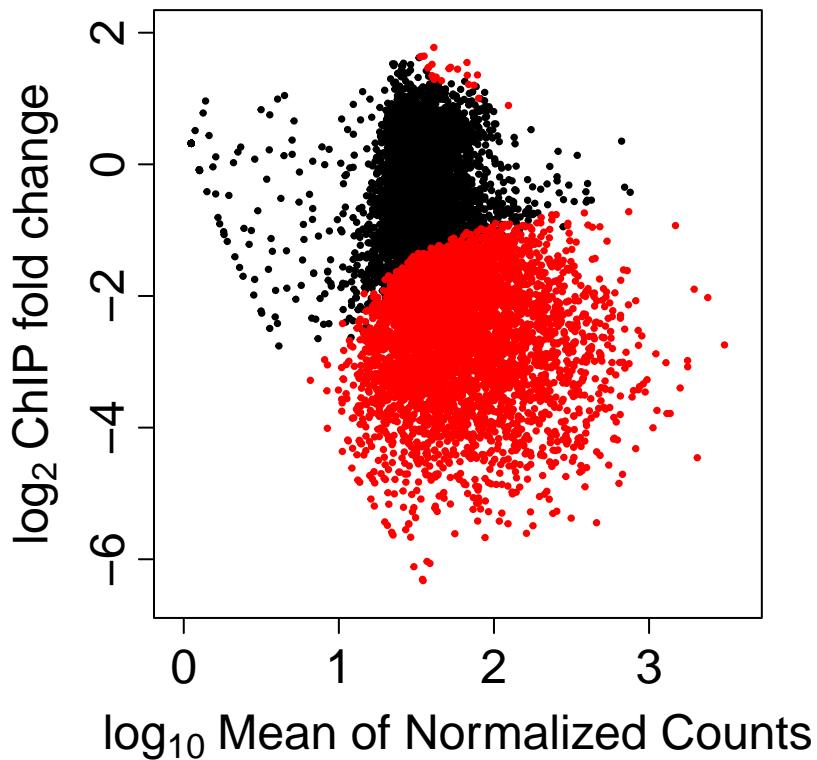
```
## Loading required package: Biostrings
## Loading required package: XVector
##
## Attaching package: 'Biostrings'
## The following object is masked from 'package:DelayedArray':
## 
##     type
## The following object is masked from 'package:base':
## 
##     strsplit
## <anonymous>: local variable 'treatment_fit' assigned but may not be used
## <anonymous>: no visible binding for global variable 'jg.controlSampleSheet'
## <anonymous>: no visible binding for global variable 'jg.controlSampleSheet'
## <anonymous>: no visible binding for global variable 'jg.controlSampleSheet'
## <anonymous>: no visible binding for global variable 'jg.experimentSampleSheet'
## <anonymous>: local variable 'jg.sampleIds' assigned but may not be used
## converting counts to integer mode
## gene-wise dispersion estimates
## mean-dispersion relationship
## final dispersion estimates
## converting counts to integer mode
```

```
## gene-wise dispersion estimates  
## mean-dispersion relationship  
## final dispersion estimates
```

Fold-change in CTCF binding



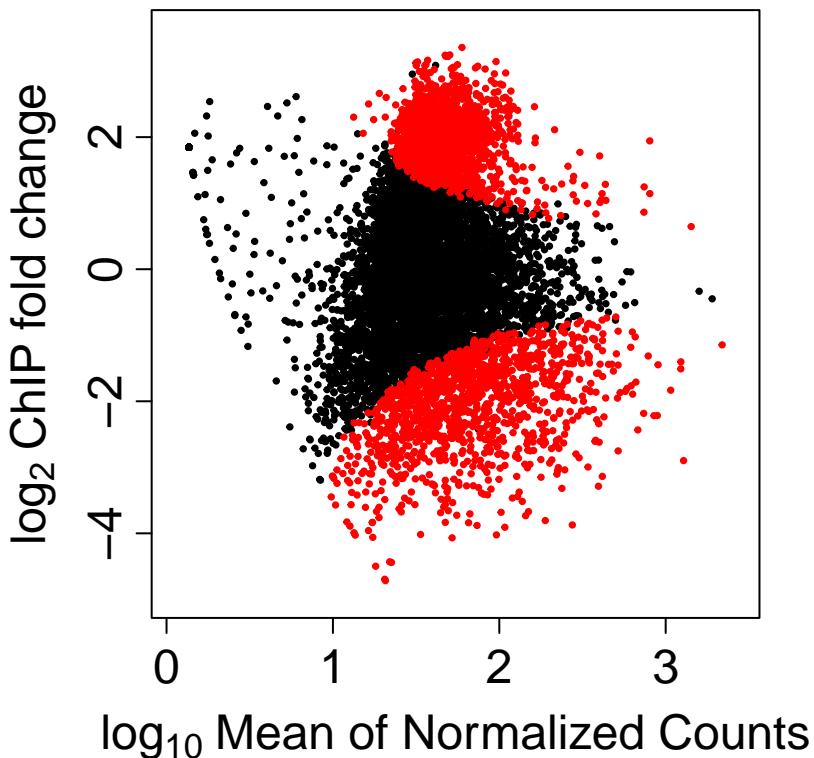
Fold-change in ER binding



CTCF Binding

```
## converting counts to integer mode  
## gene-wise dispersion estimates  
## mean-dispersion relationship  
## final dispersion estimates
```

Fold-change in ER binding (no correction)

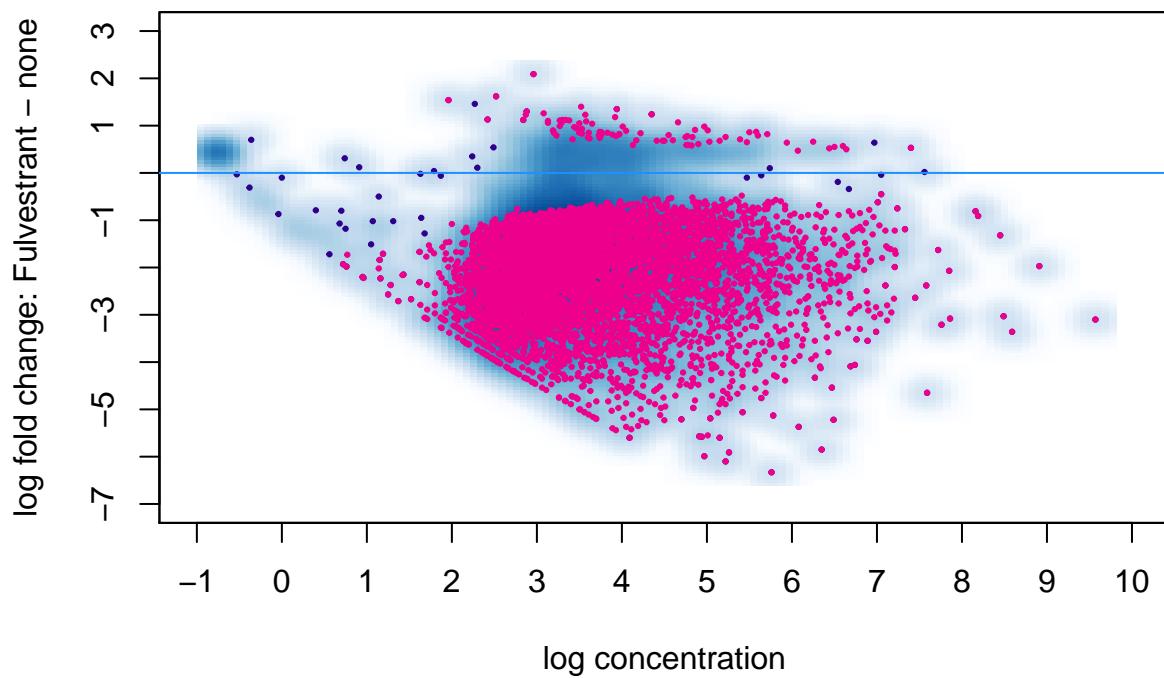


H2av With Diff Bind

```
## converting counts to integer mode
## gene-wise dispersion estimates
## mean-dispersion relationship
## final dispersion estimates
## converting counts to integer mode
## gene-wise dispersion estimates
## mean-dispersion relationship
## -- note: fitType='parametric', but the dispersion trend was not well captured by the
##   function: y = a/x + b, and a local regression fit was automatically substituted.
##   specify fitType='local' or 'mean' to avoid this message next time.
## final dispersion estimates
## converting counts to integer mode
## gene-wise dispersion estimates
## mean-dispersion relationship
## final dispersion estimates
## converting counts to integer mode
```

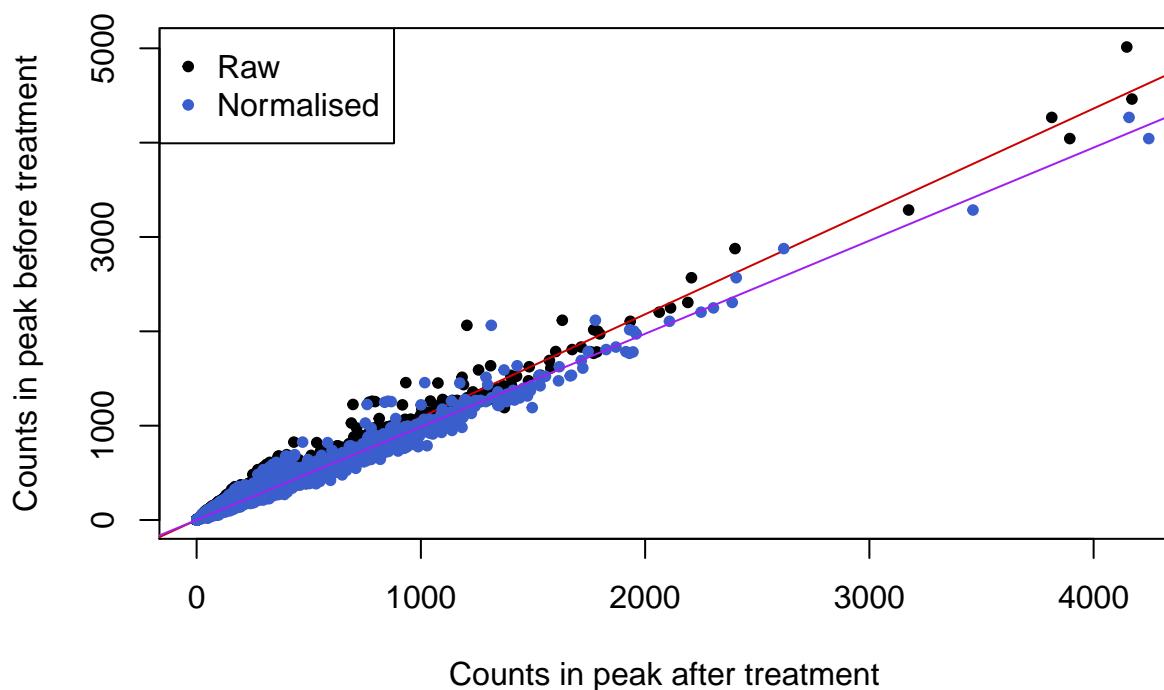
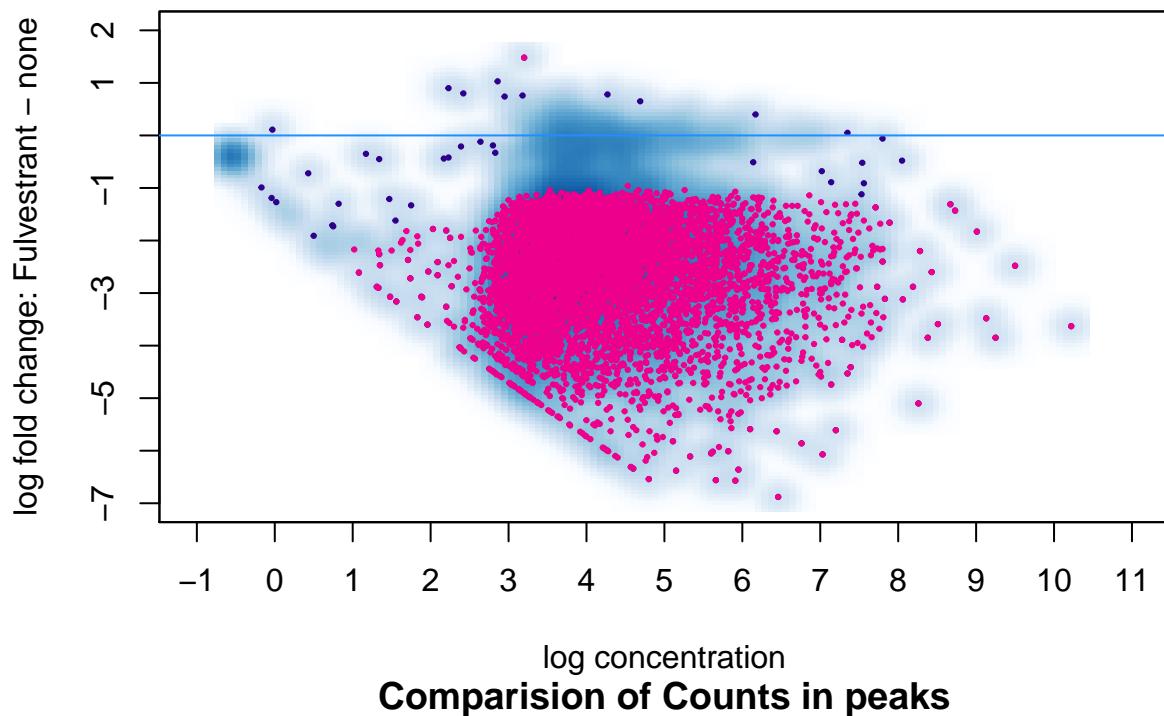
```
## gene-wise dispersion estimates  
## mean-dispersion relationship  
## final dispersion estimates
```

Binding Affinity: Fulvestrant vs. none (5521 FDR < 0.050)



```
## converting counts to integer mode  
## gene-wise dispersion estimates  
## mean-dispersion relationship  
## final dispersion estimates
```

Binding Affinity: Fulvestrant vs. none (5576 FDR < 0.050)



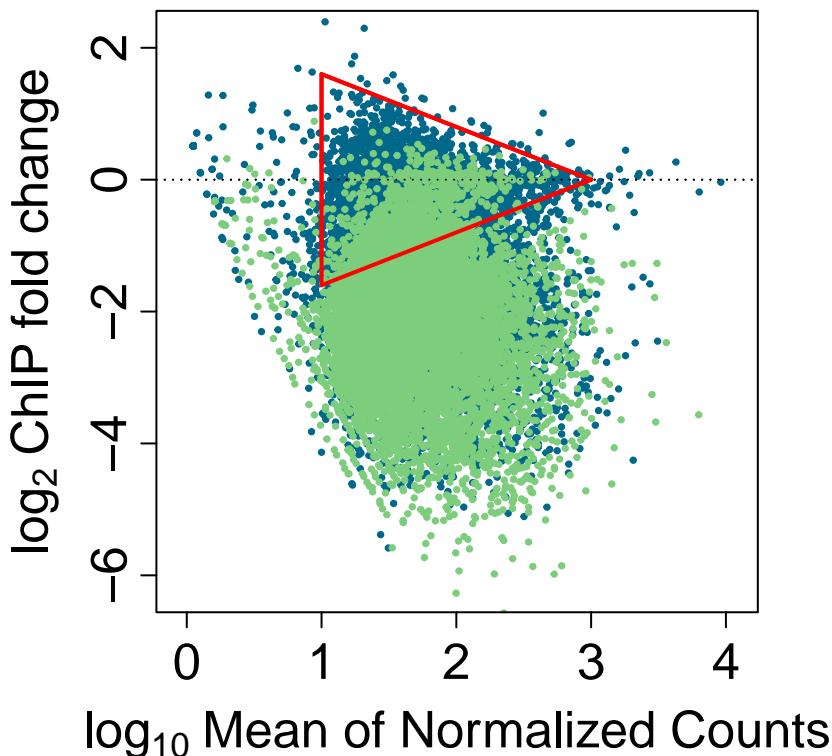
```
## rowMeans(jg.controlCountsTreated)
##                               1.090441
## converting counts to integer mode
## gene-wise dispersion estimates
```

```

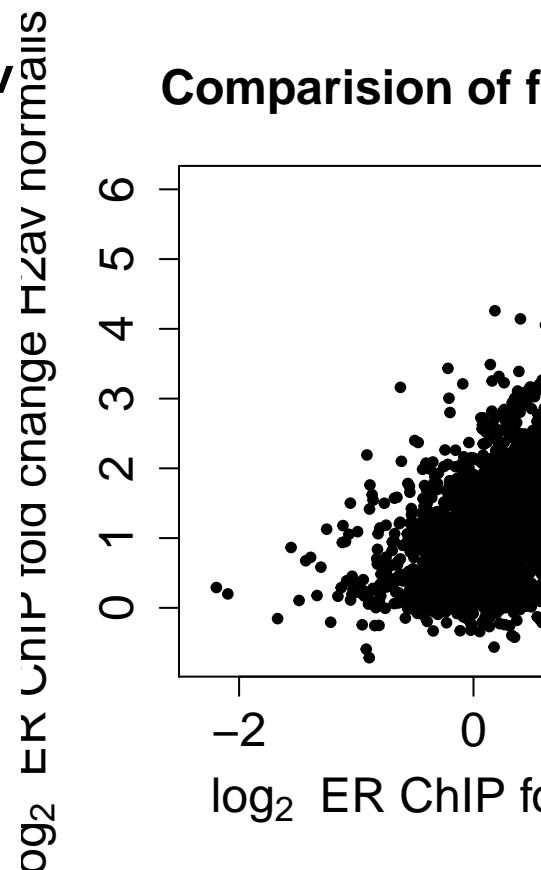
## mean-dispersion relationship
## final dispersion estimates
## converting counts to integer mode
## gene-wise dispersion estimates
## mean-dispersion relationship
## -- note: fitType='parametric', but the dispersion trend was not well captured by the
## function: y = a/x + b, and a local regression fit was automatically substituted.
## specify fitType='local' or 'mean' to avoid this message next time.
## final dispersion estimates
## converting counts to integer mode
## gene-wise dispersion estimates
## mean-dispersion relationship
## final dispersion estimates
## converting counts to integer mode
## gene-wise dispersion estimates
## mean-dispersion relationship
## final dispersion estimates
## Loading required package: RColorBrewer

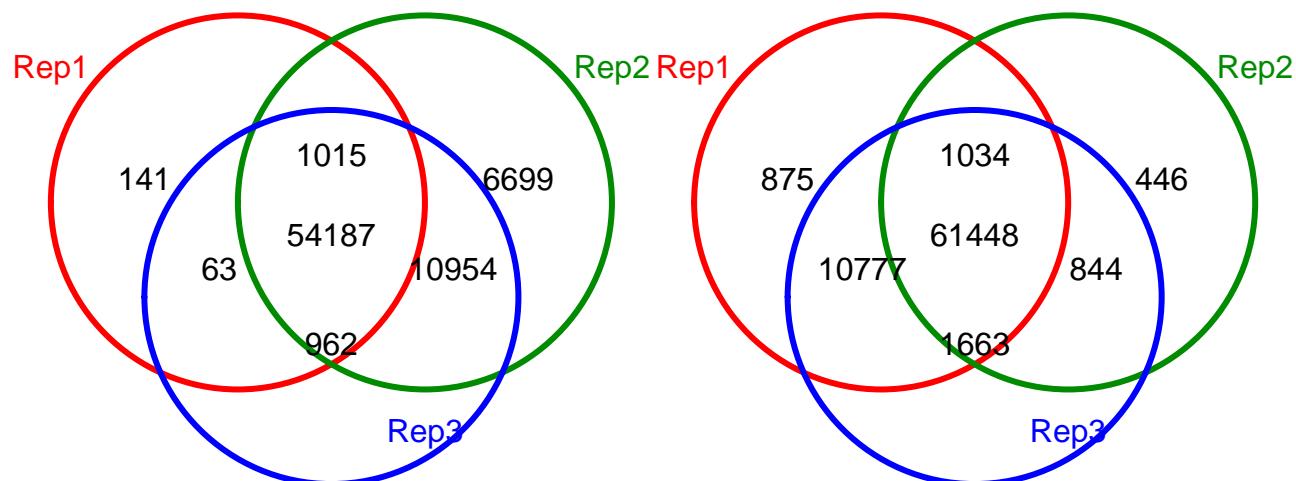
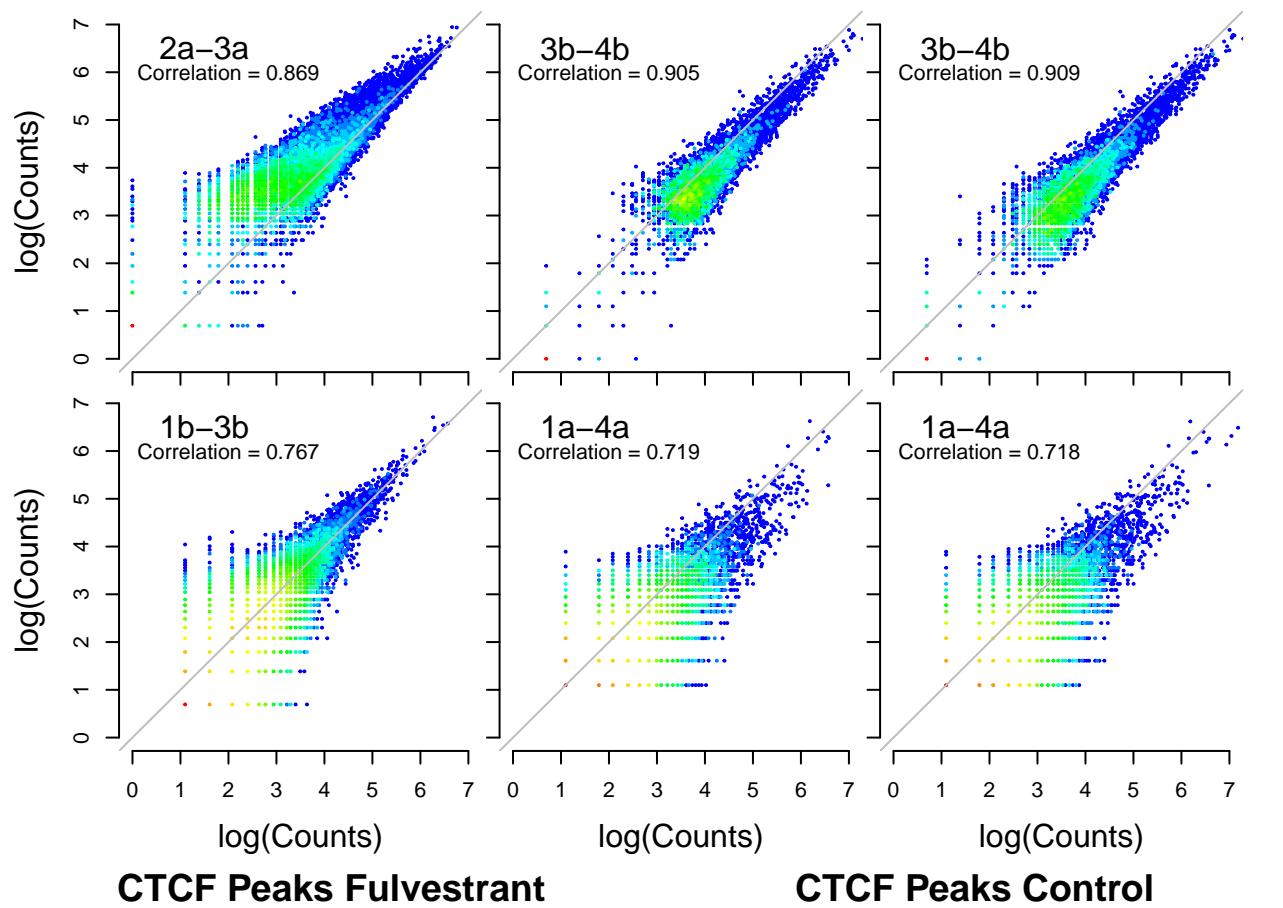
```

Comparision between CTCF and H2av

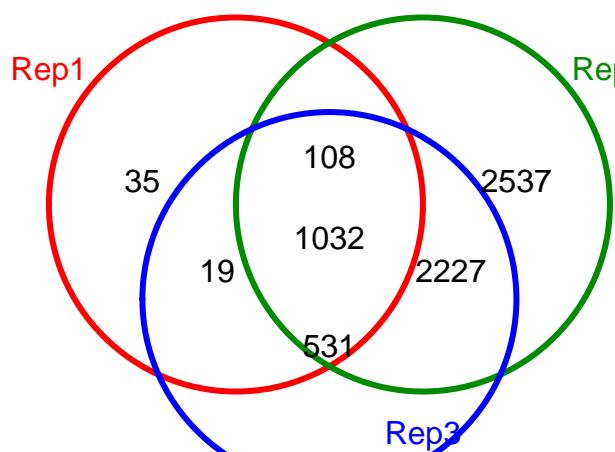


Comparision of f

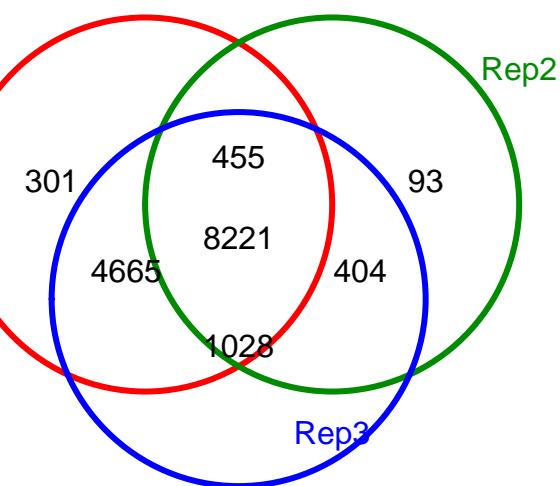




ER Peaks Fulvestrant

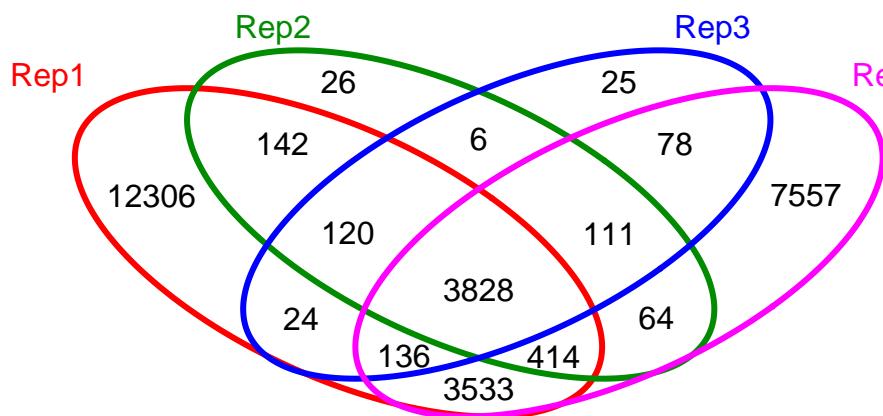


ER Peaks Control

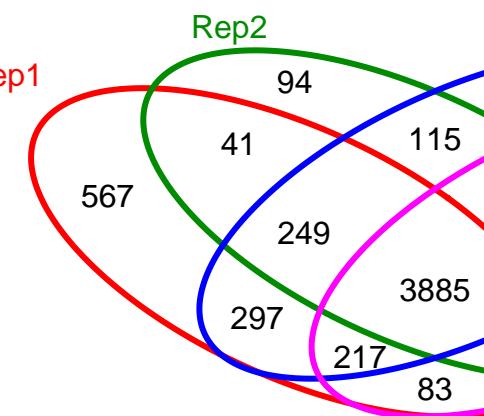


H2av Peaks Fulvestrant

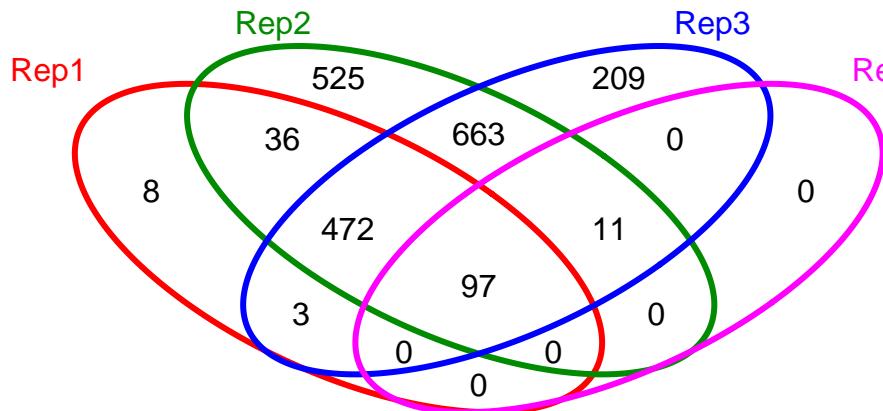
H2av Peaks Control



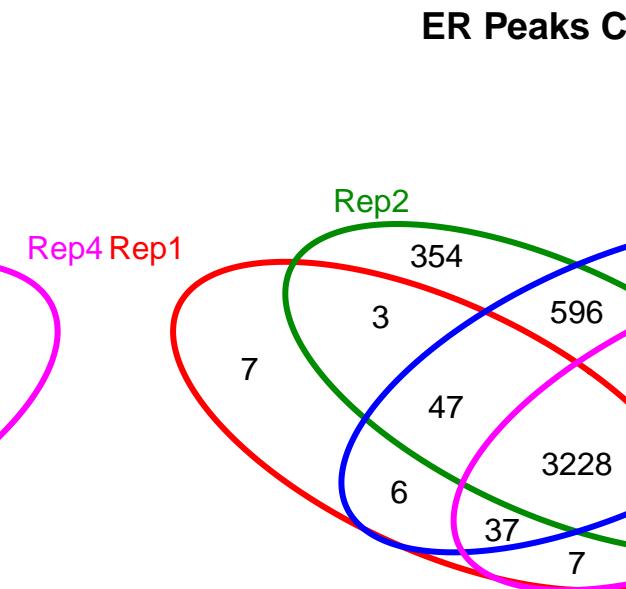
Rep4 Rep1



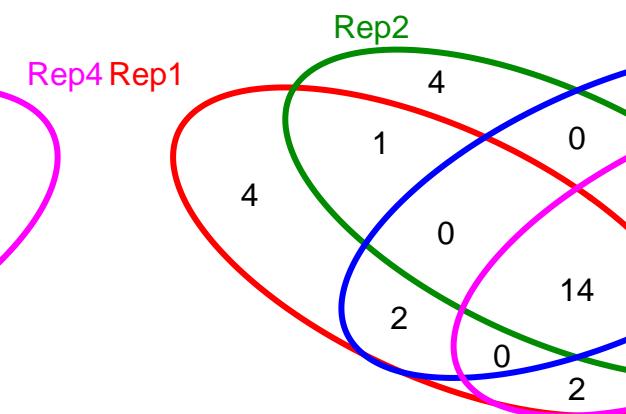
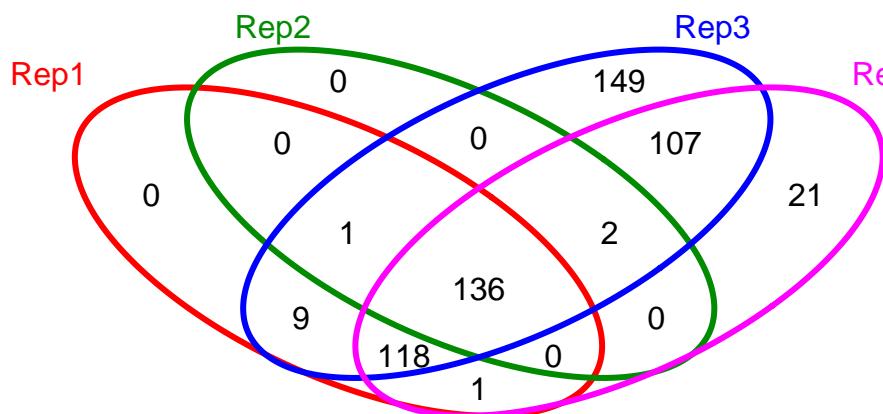
ER Peaks Fulvestrant



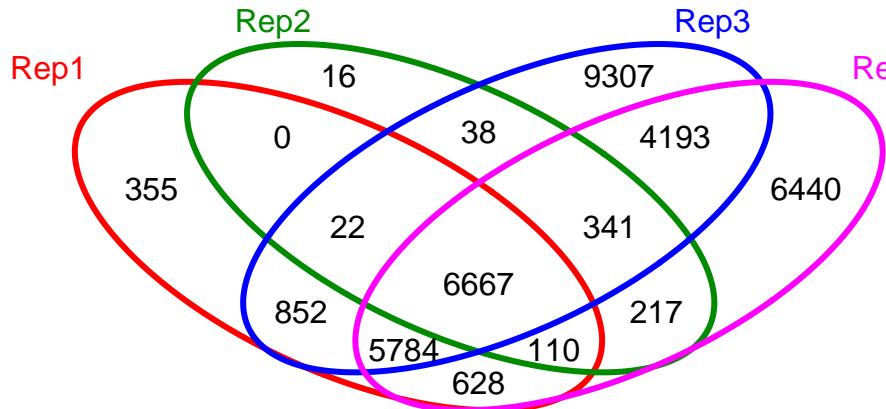
Mouse ER Peaks Fulvestrant



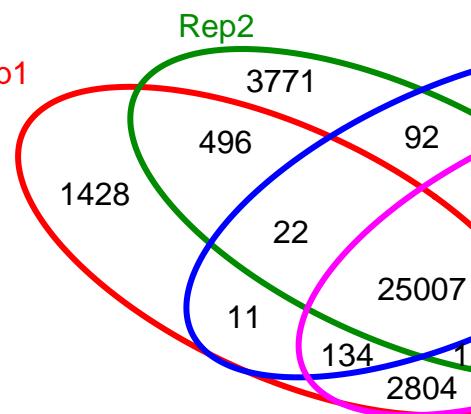
Mouse ER Peaks C



ER Peaks Fulvestrant



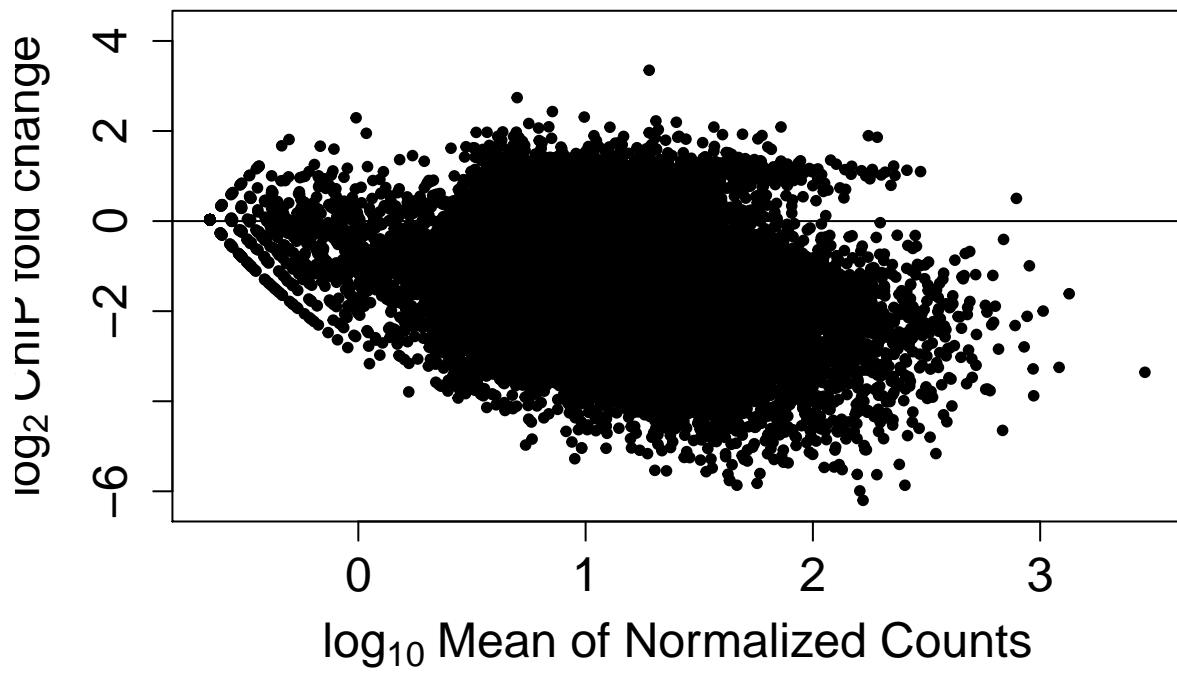
ER Peaks C



```
##   ICI_4 ICI_1 Control_4 ICI_3 ICI_2 Control_1 Control_3 Control_2
## 1    1    1        1    1    1        1        1        1
## 2    1    1        1    1    1        1        1        1
## 3    1    1        1    1    1        1        1        1
## 4    1    1        3    1    1        6        1        1
## 5    1    2        1    3    1        3        1        1
## 6    1    1        1    1    1        1        1        1

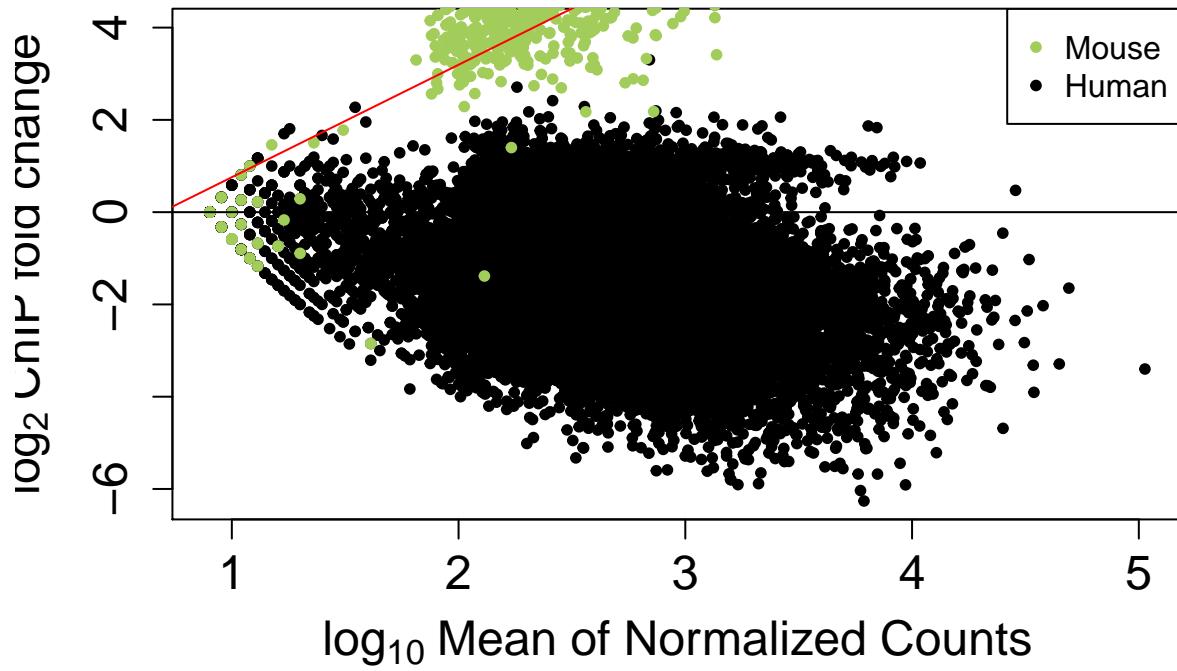
##   ICI_4 ICI_1 Control_4 ICI_3 ICI_2 Control_1 Control_3 Control_2
## 1   36     8        2   37   39        2        1        1
## 2   41    23        6   70   53       12        5        5
## 3   15     7        3   36   27        1        3        2
## 4   20    12        1   55   37        1        7        3
## 5   35    19        3   76   38        2        1        3
## 6   24    13        2   53   34        2        3        1
```

RPM aligned reads



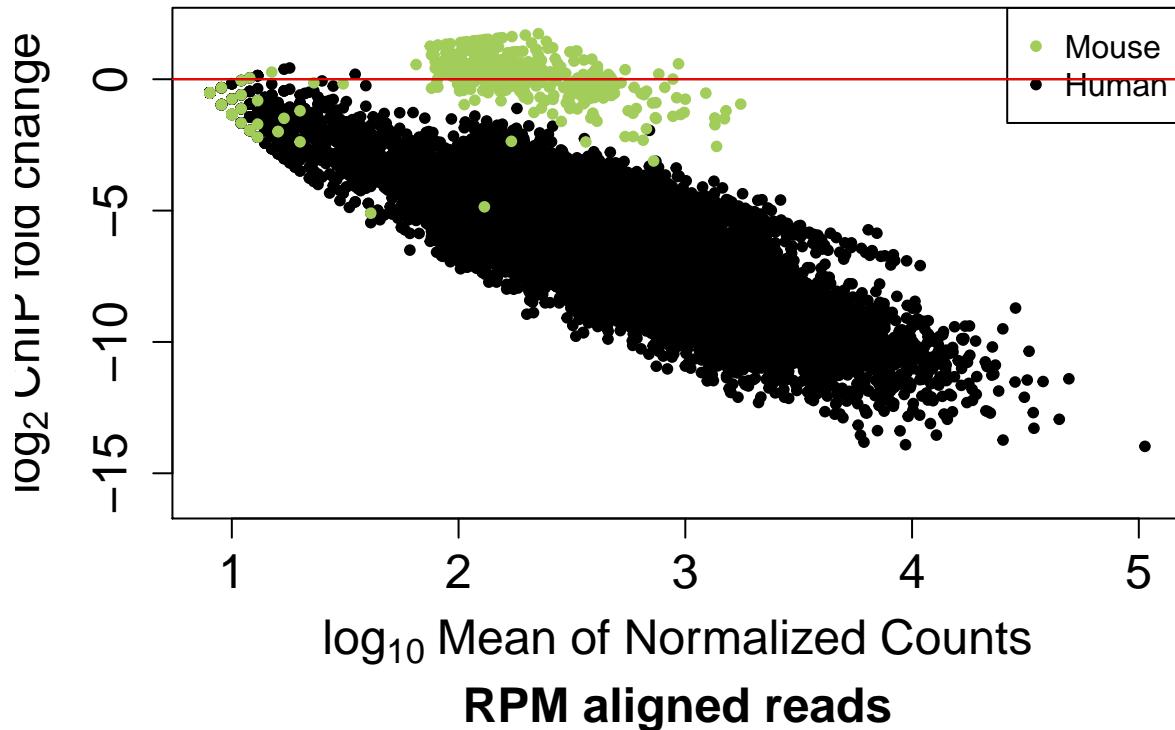
\log_{10} Mean of Normalized Counts

RPM aligned reads

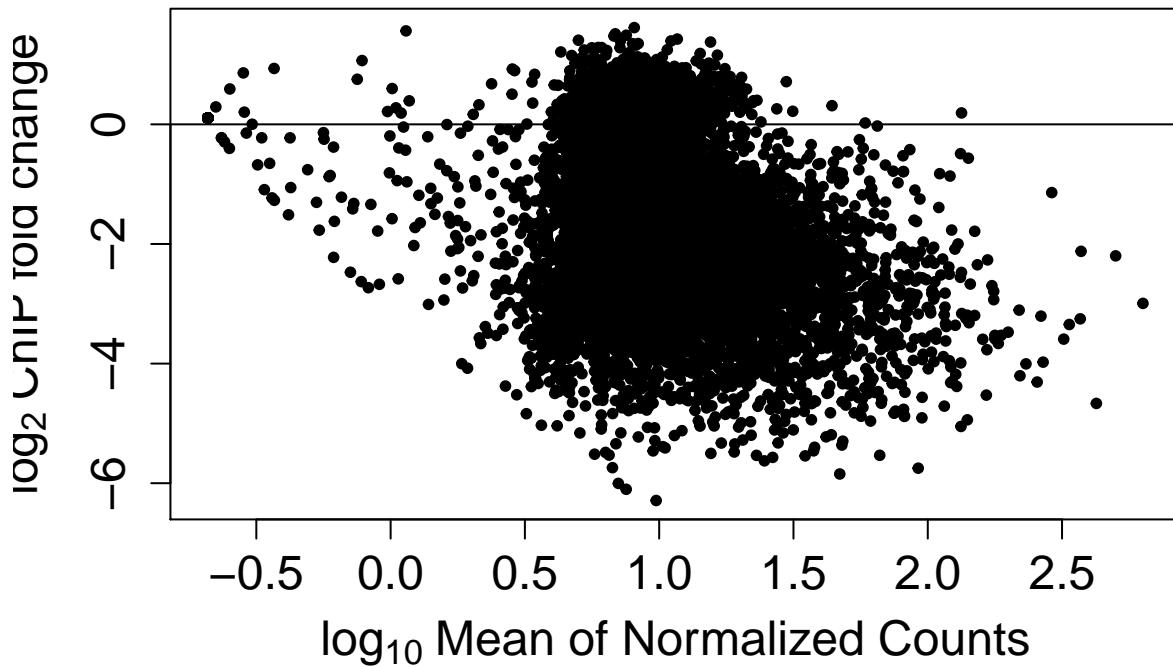


\log_{10} Mean of Normalized Counts

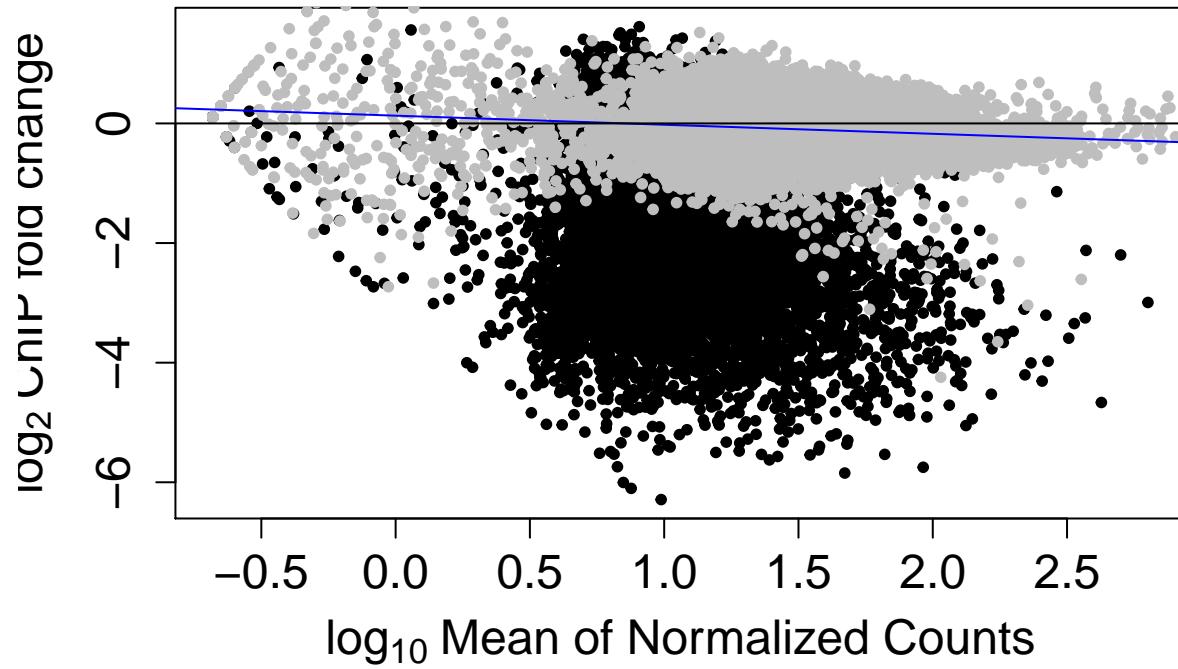
Counts normalized by Mouse Distribution



RPM aligned reads

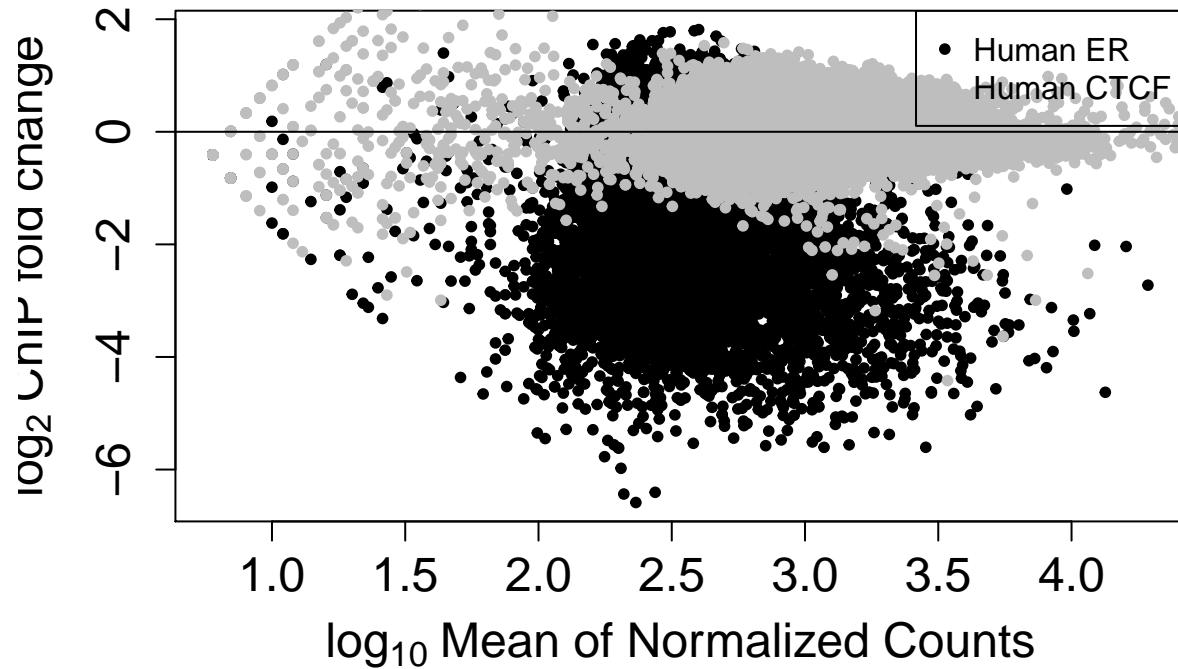


RPM aligned reads



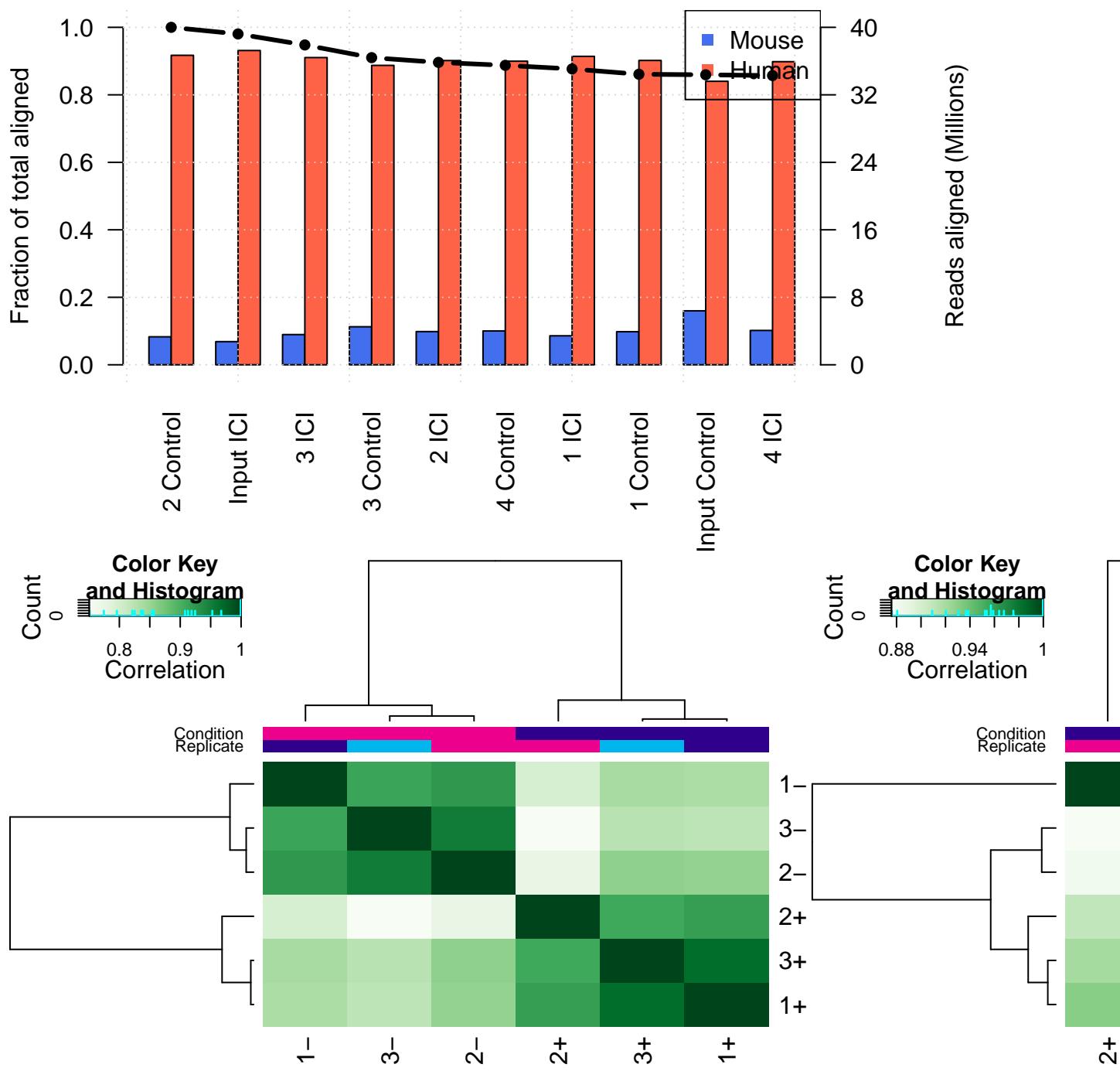
$\log_{10} \text{Mean of Normalized Counts}$

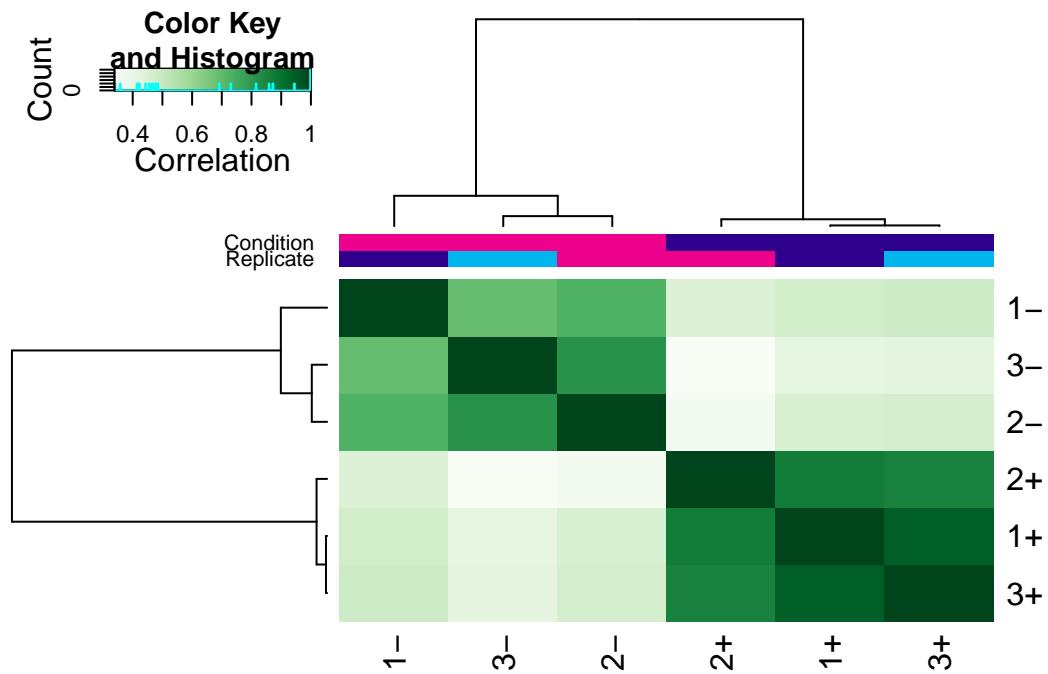
Normalised to Human CTCF



$\log_{10} \text{Mean of Normalized Counts}$

Relative Read Alignment in samples

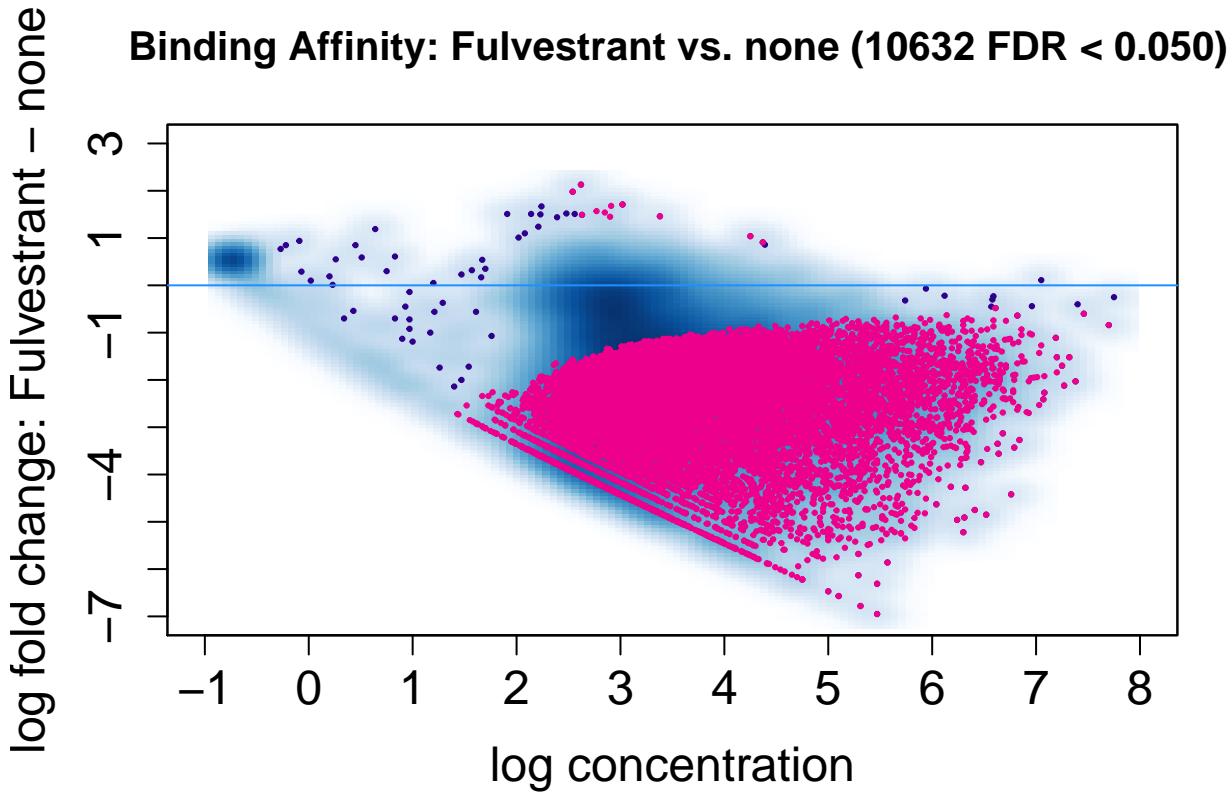




```

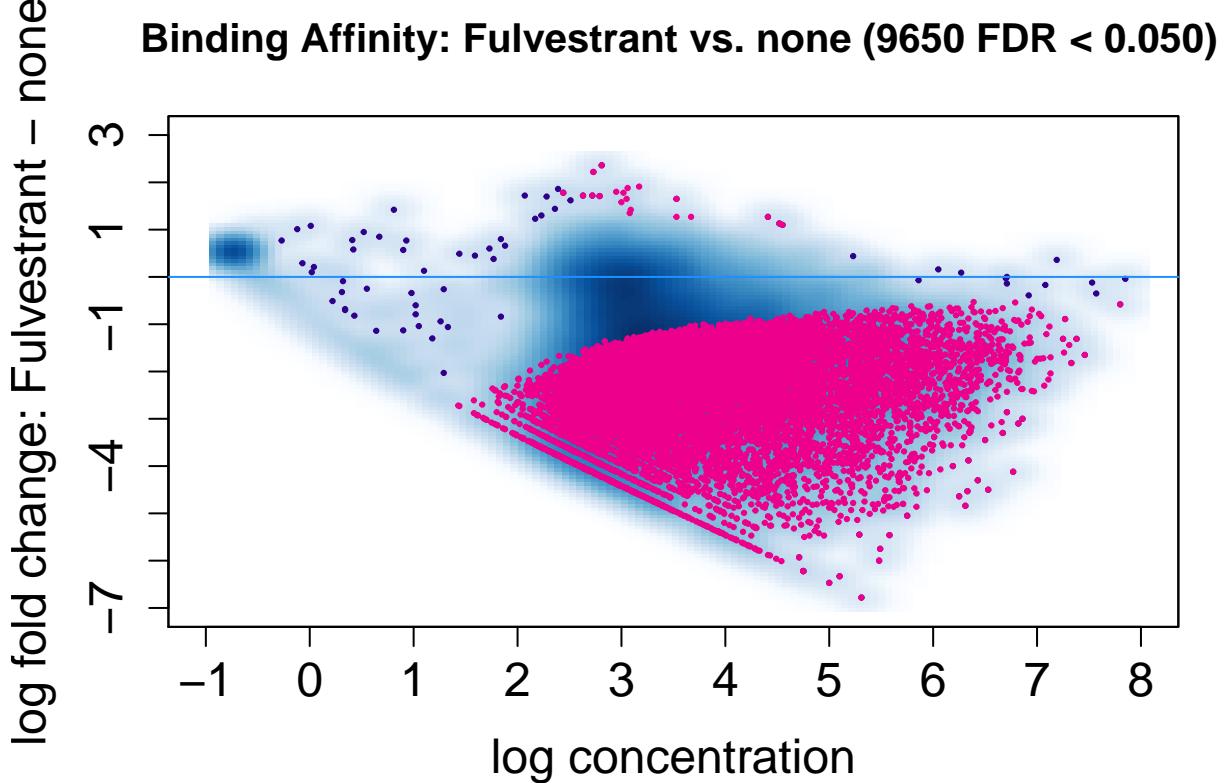
## converting counts to integer mode
## gene-wise dispersion estimates
## mean-dispersion relationship
## final dispersion estimates

```

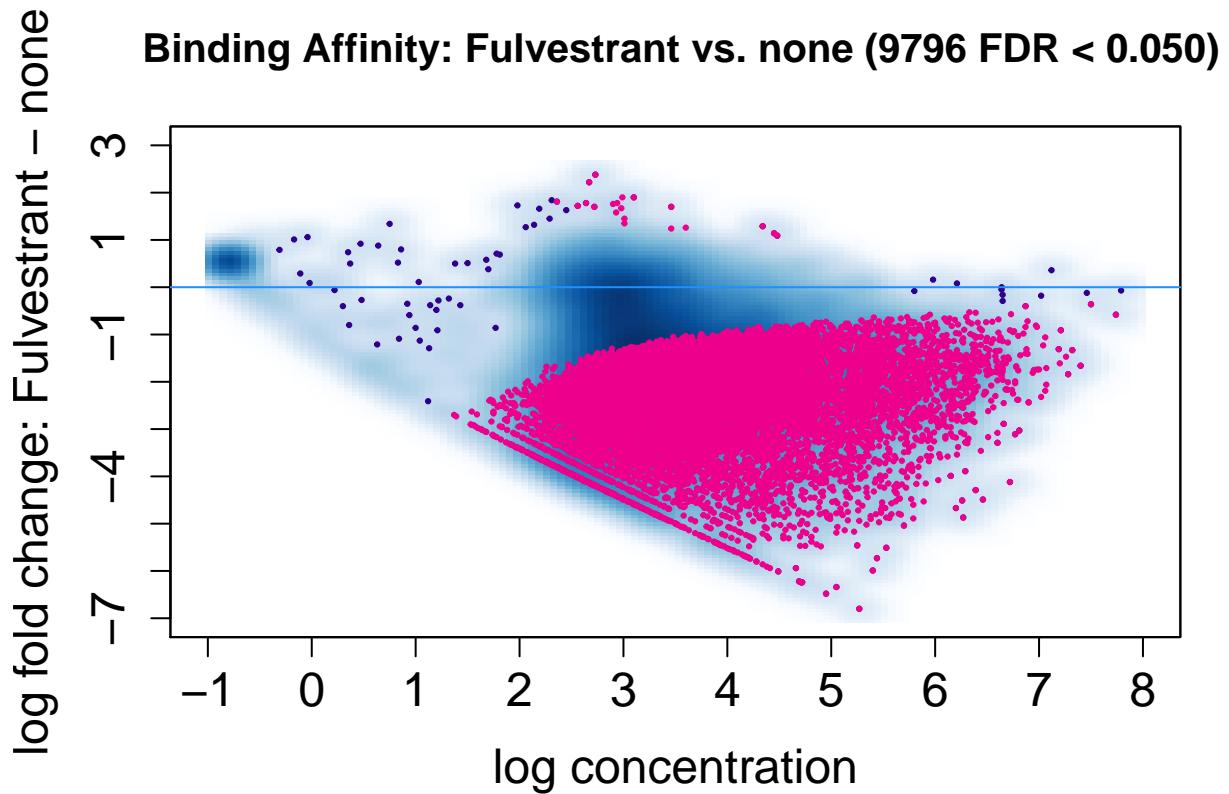


```
## converting counts to integer mode
```

```
## gene-wise dispersion estimates  
## mean-dispersion relationship  
## final dispersion estimates
```



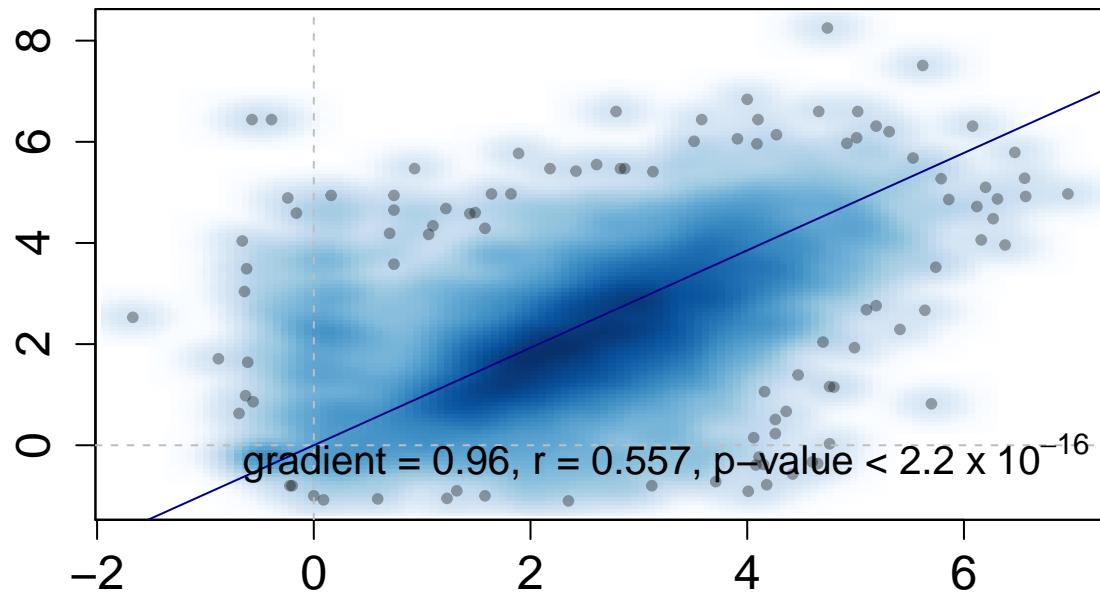
```
## converting counts to integer mode  
## gene-wise dispersion estimates  
## mean-dispersion relationship  
## final dispersion estimates  
## converting counts to integer mode  
## gene-wise dispersion estimates  
## mean-dispersion relationship  
## final dispersion estimates  
## converting counts to integer mode  
## gene-wise dispersion estimates  
## mean-dispersion relationship  
## final dispersion estimates
```



```
## converting counts to integer mode  
## gene-wise dispersion estimates  
## mean-dispersion relationship  
## final dispersion estimates  
## converting counts to integer mode  
## gene-wise dispersion estimates  
## mean-dispersion relationship  
## final dispersion estimates
```

change normalized using CTCF control

Comparision of normalisation methods



Fold change normalized using Drosophila/H2av spike-in

```
## converting counts to integer mode
## gene-wise dispersion estimates
## mean-dispersion relationship
## final dispersion estimates
## converting counts to integer mode
## gene-wise dispersion estimates
## mean-dispersion relationship
## final dispersion estimates
## converting counts to integer mode
## gene-wise dispersion estimates
## mean-dispersion relationship
## final dispersion estimates
```

Log(FoldChange) in ER Binding Cross–Normalised to CTCI

Comparision of Fold–Change between Cross–Normalised to Xenogenic Spike–in

