

# Long gene analysis in gene expression datasets

*Ayush T Raman, Liu Lab, Baylor College of Medicine*

*2018-04-21*

## Contents

Section 1: Introduction . . . . .	1
Section 2: Libraries and all functions . . . . .	1
Section 3: Figure 1 and Supplementary Figure 2-4 . . . . .	1
Section 4: Figure 2 . . . . .	32
Section 5: Figure 3 and Supplementary Figure 5 . . . . .	41
Section 6: Figure 4 and Supplementary Fig. 6-9 . . . . .	51
Section 7: Figure 5 and Supplementary Fig. 10 . . . . .	79
Section 8: Session info . . . . .	89

## Section 1: Introduction

The code and the graphs are attached below.

## Section 2: Libraries and all functions

```
rm(list = ls())

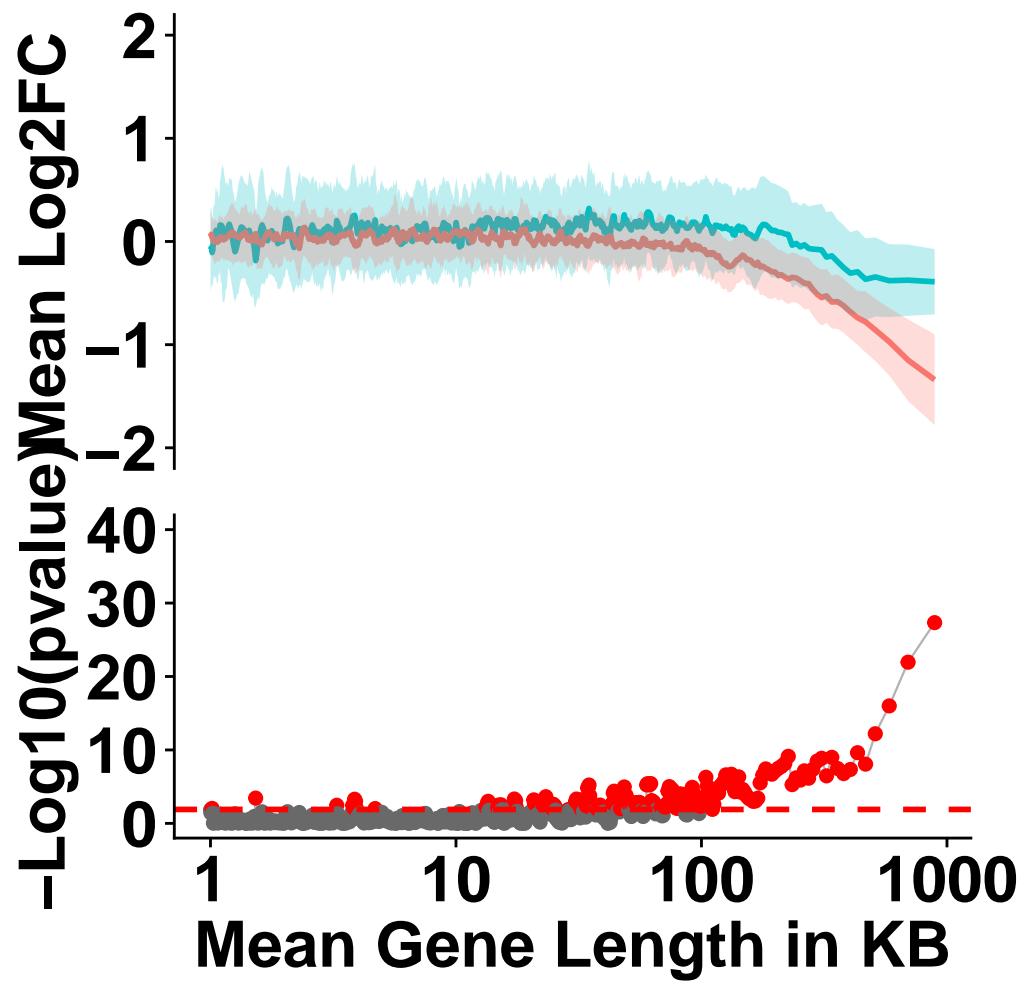
## functions and set the working directory
source("libraries.R")
options(warn=-1)
```

## Section 3: Figure 1 and Supplementary Figure 2-4

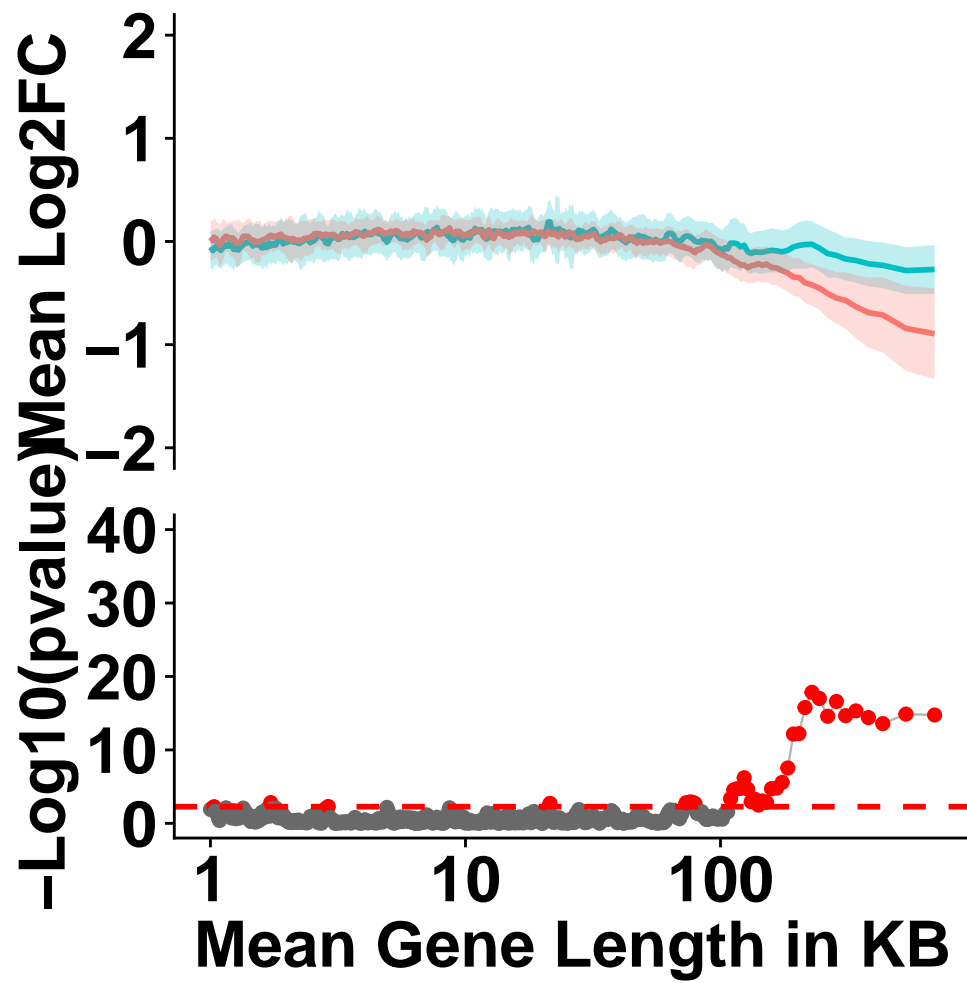
```
cat("\n\n Printing Figure 1 \n\n")

##
##
## Printing Figure 1
figure1()

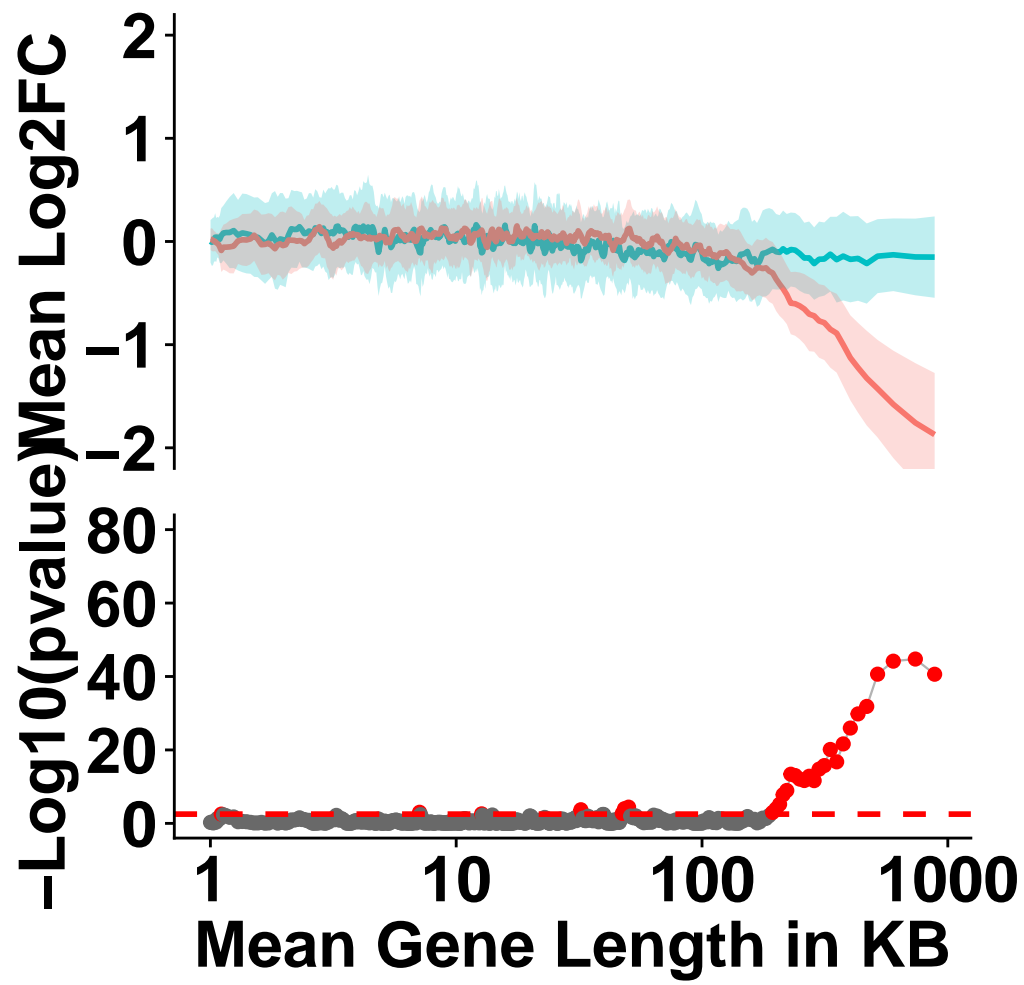
##
##
## Fig. 1(A)--Cultured Cortical Neurons; (RNA-Seq; King et al.)
```



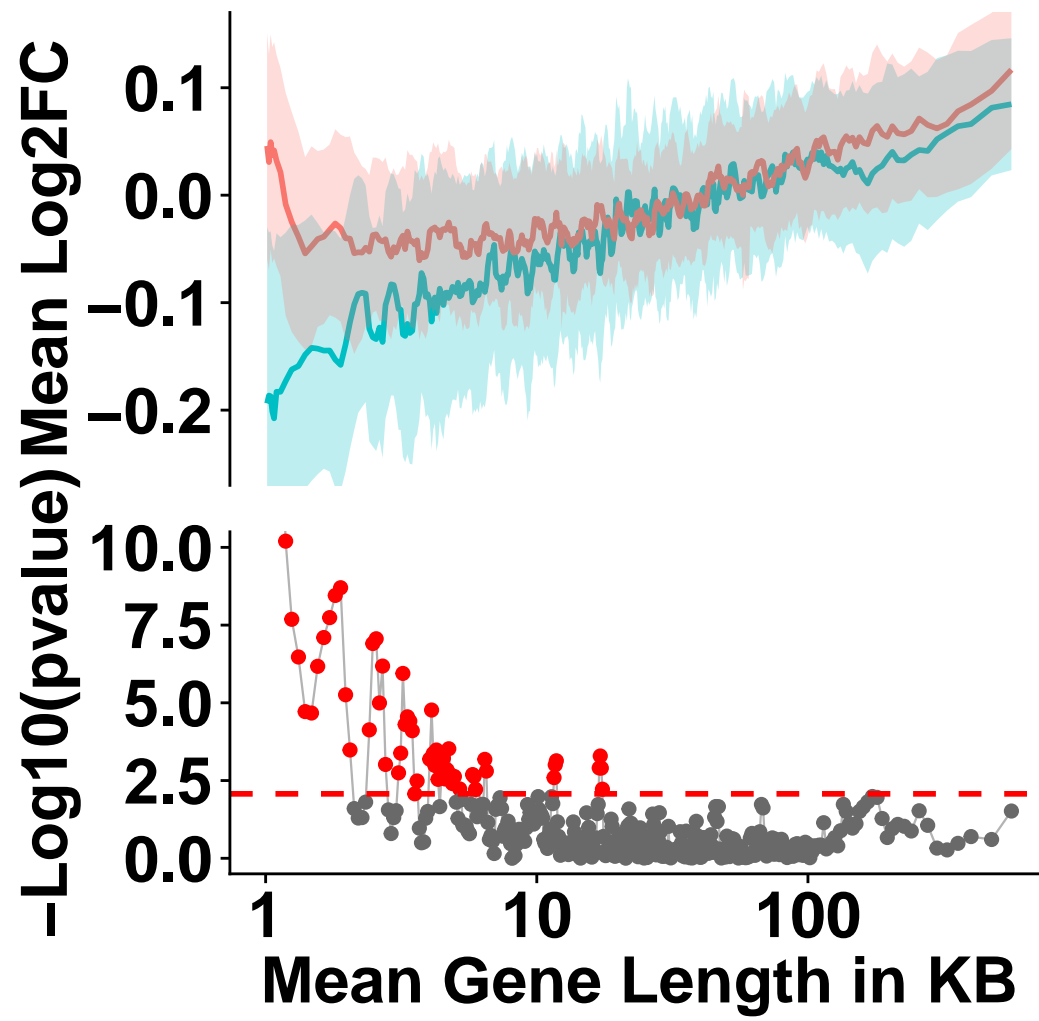
##  
 ##  
 ## Fig. 1(A) -- Cultured Cortical Neurons; (Array; King et al.)



##  
 ##  
 ## Fig. 1(A)--Cultured Cortical Neurons; (RNA-seq; Mabb et al.)



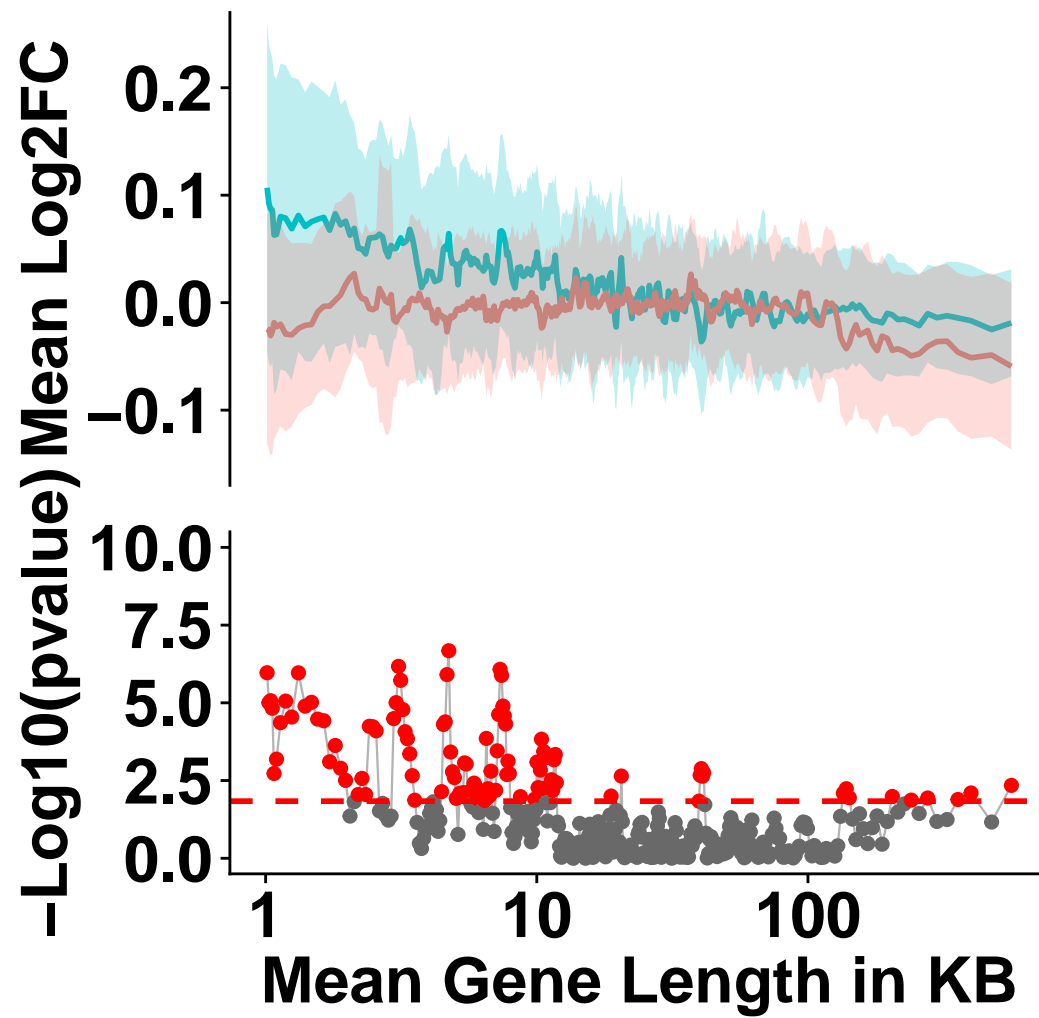
##  
 ##  
 ## Fig. 1(B) -- Amygdala (KO/WT)



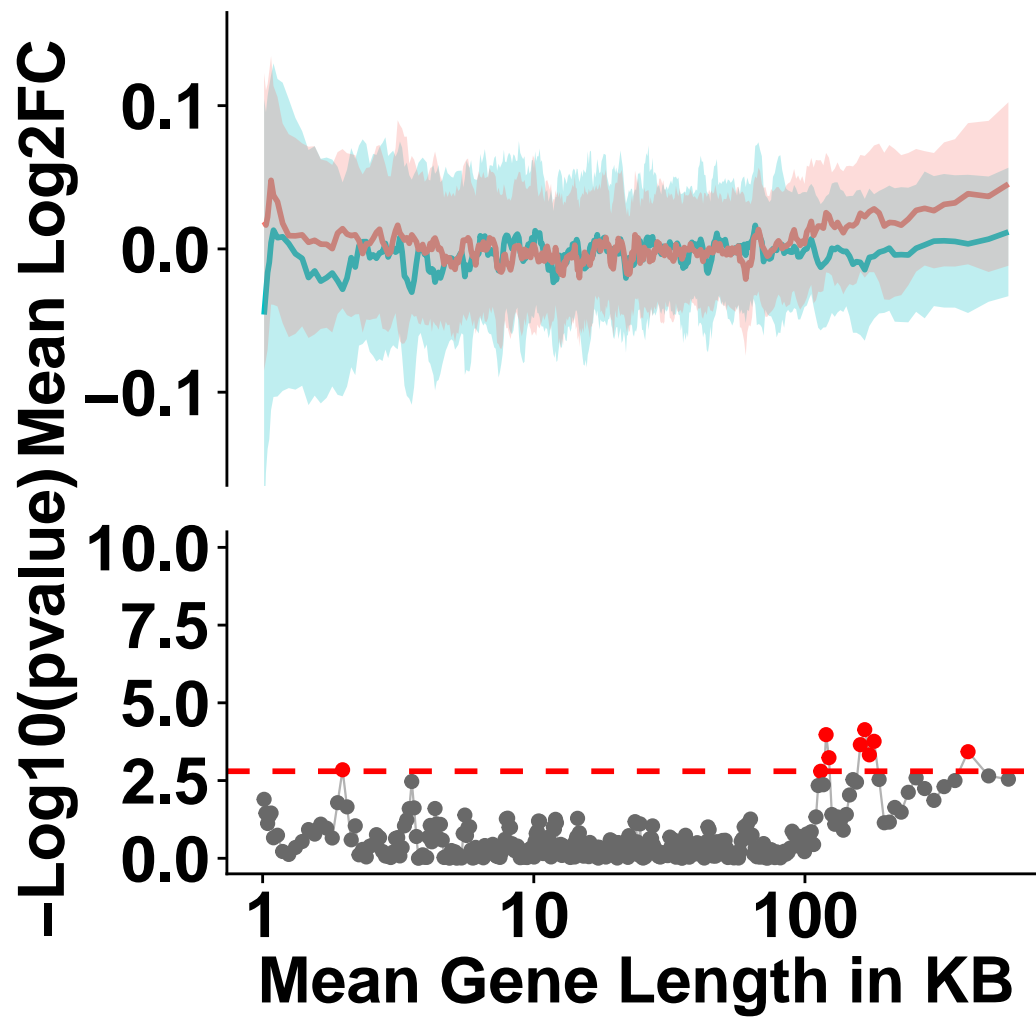
##

##

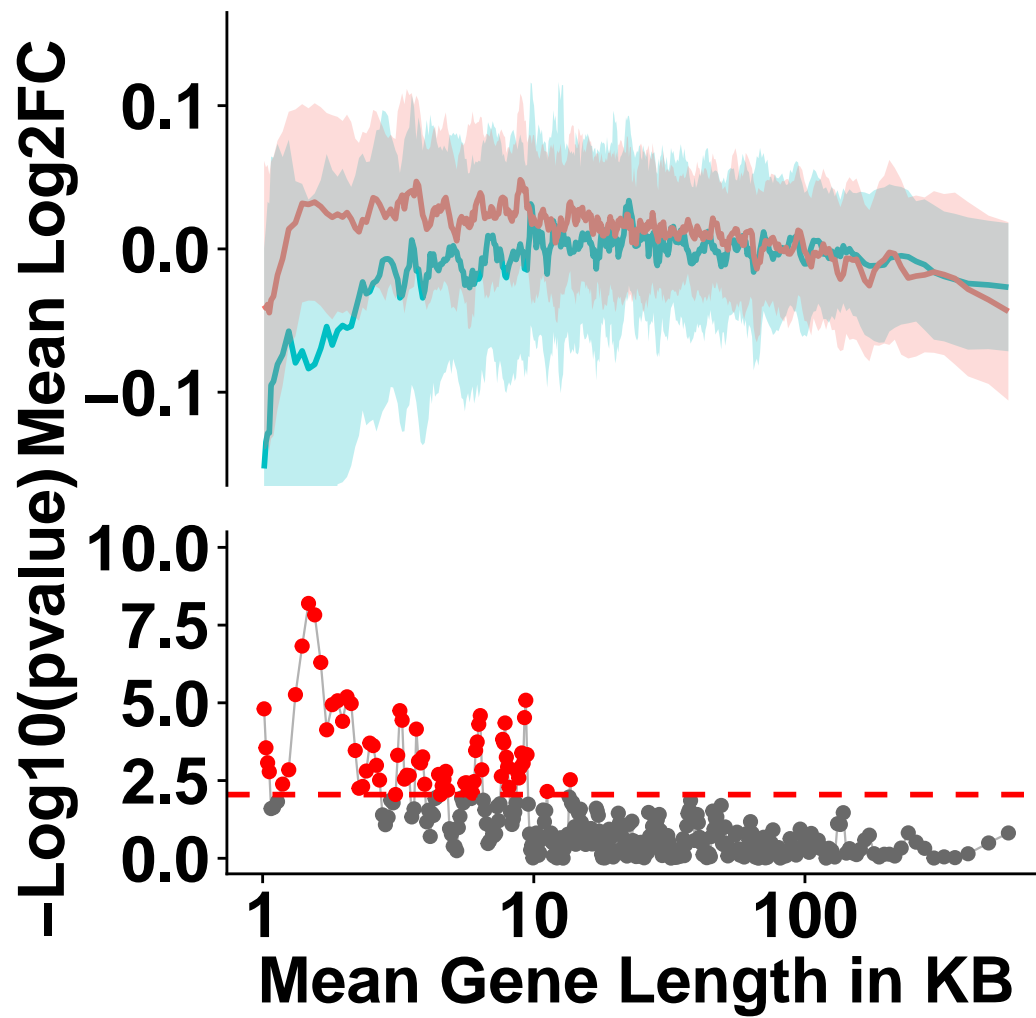
## Fig. 1(C) -- Amygdala (Tg/WT)



##  
 ##  
 ## Fig. 1(B) -- Cerebellum (KO/WT)

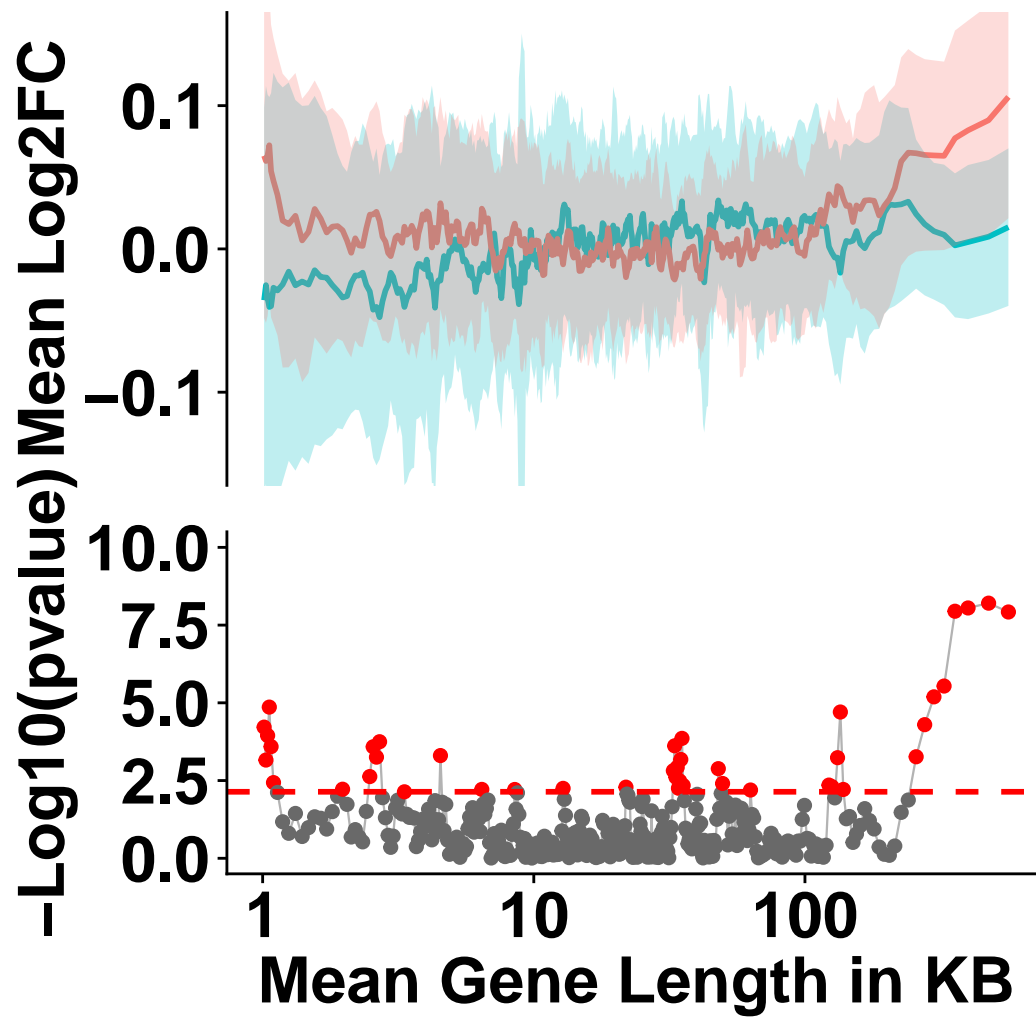


##  
 ##  
 ## Fig. 1(C) -- Cerebellum (Tg/WT)

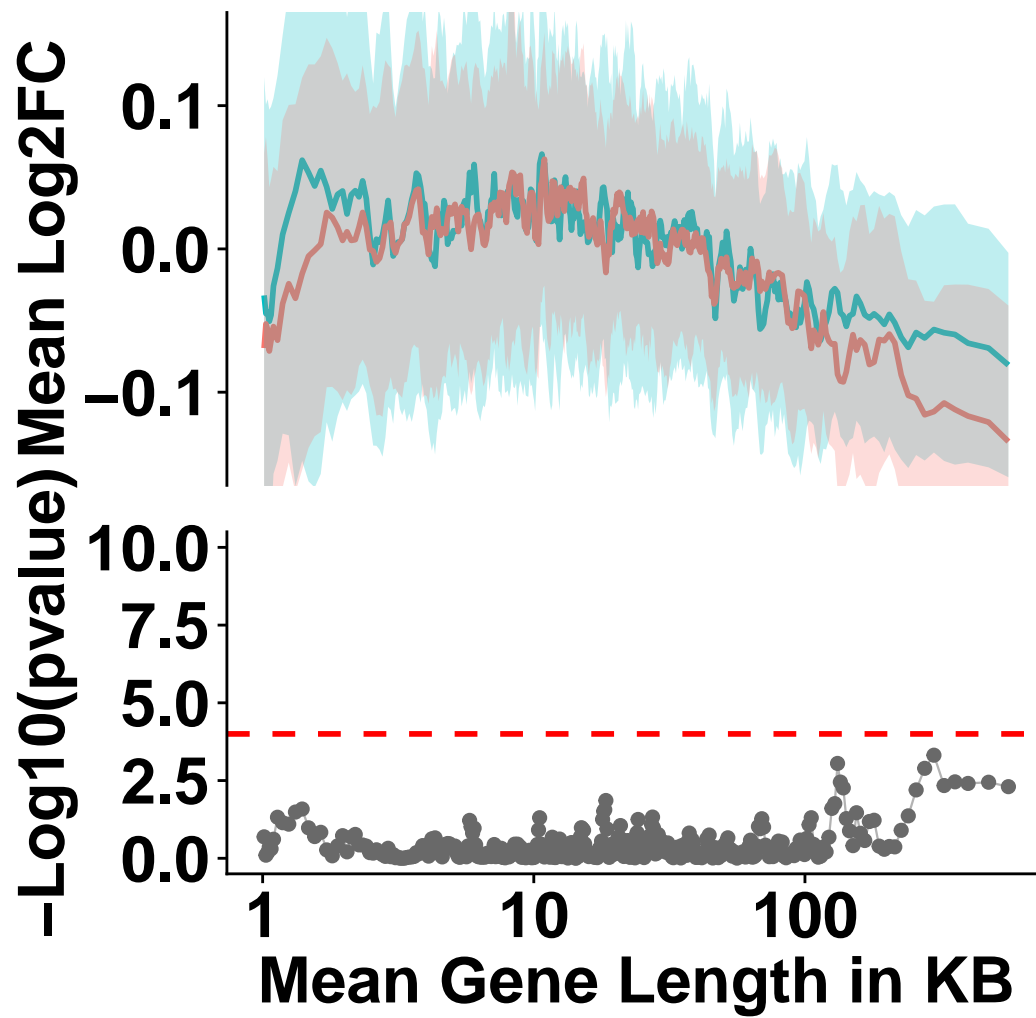


##  
 ##  
 ## Fig. 1(B) -- Hypothalamus (KO/WT)

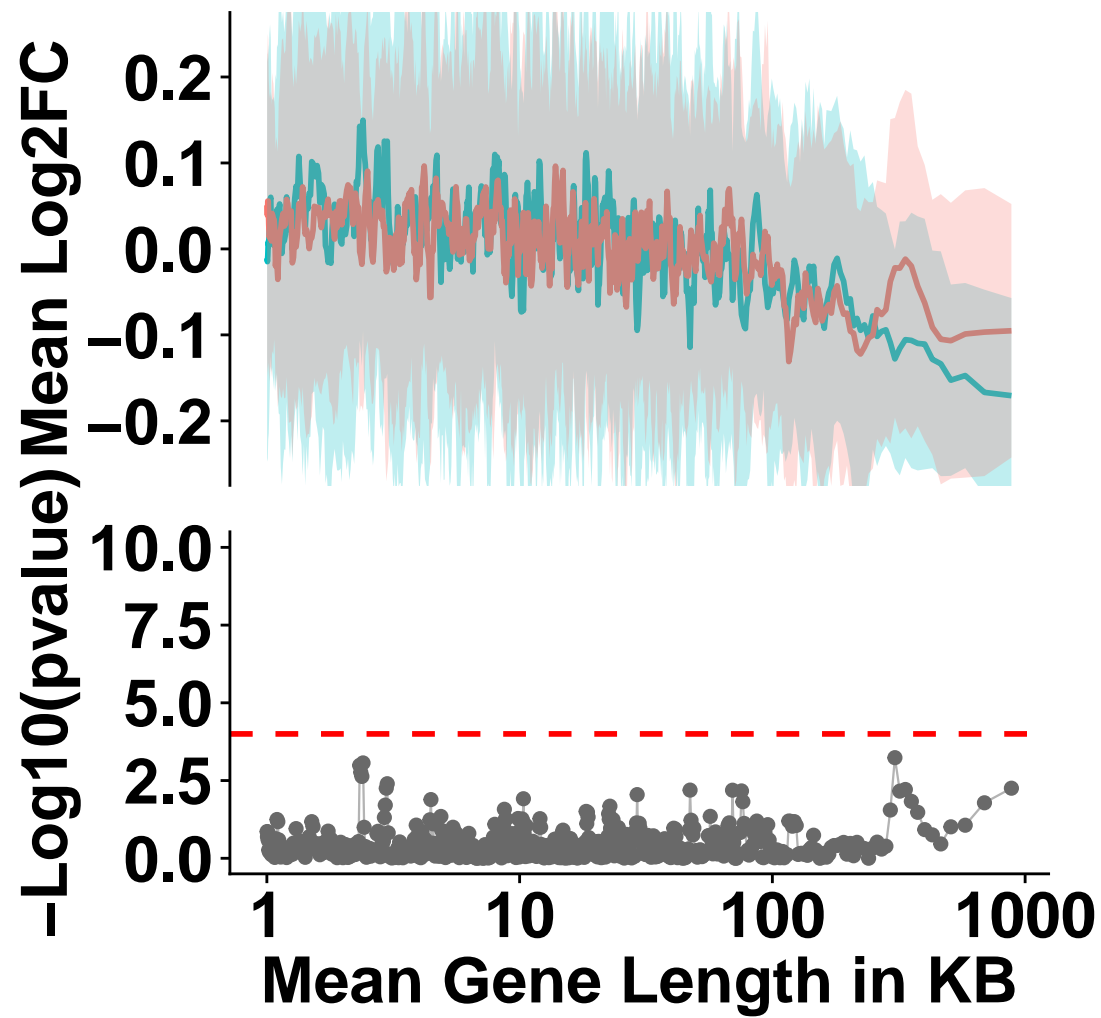




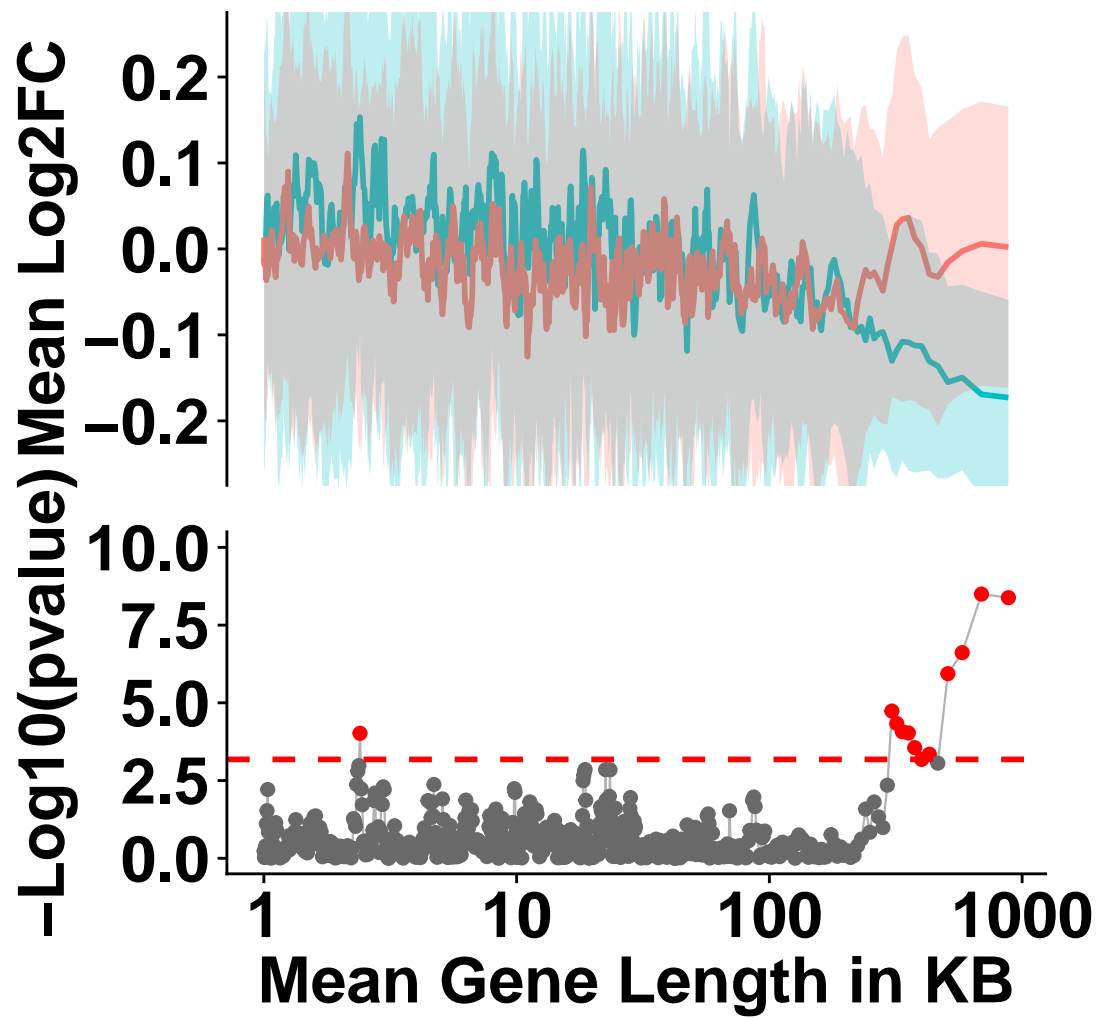
##  
 ##  
 ## Fig. 1(C) -- Hypothalamus (Tg/WT)



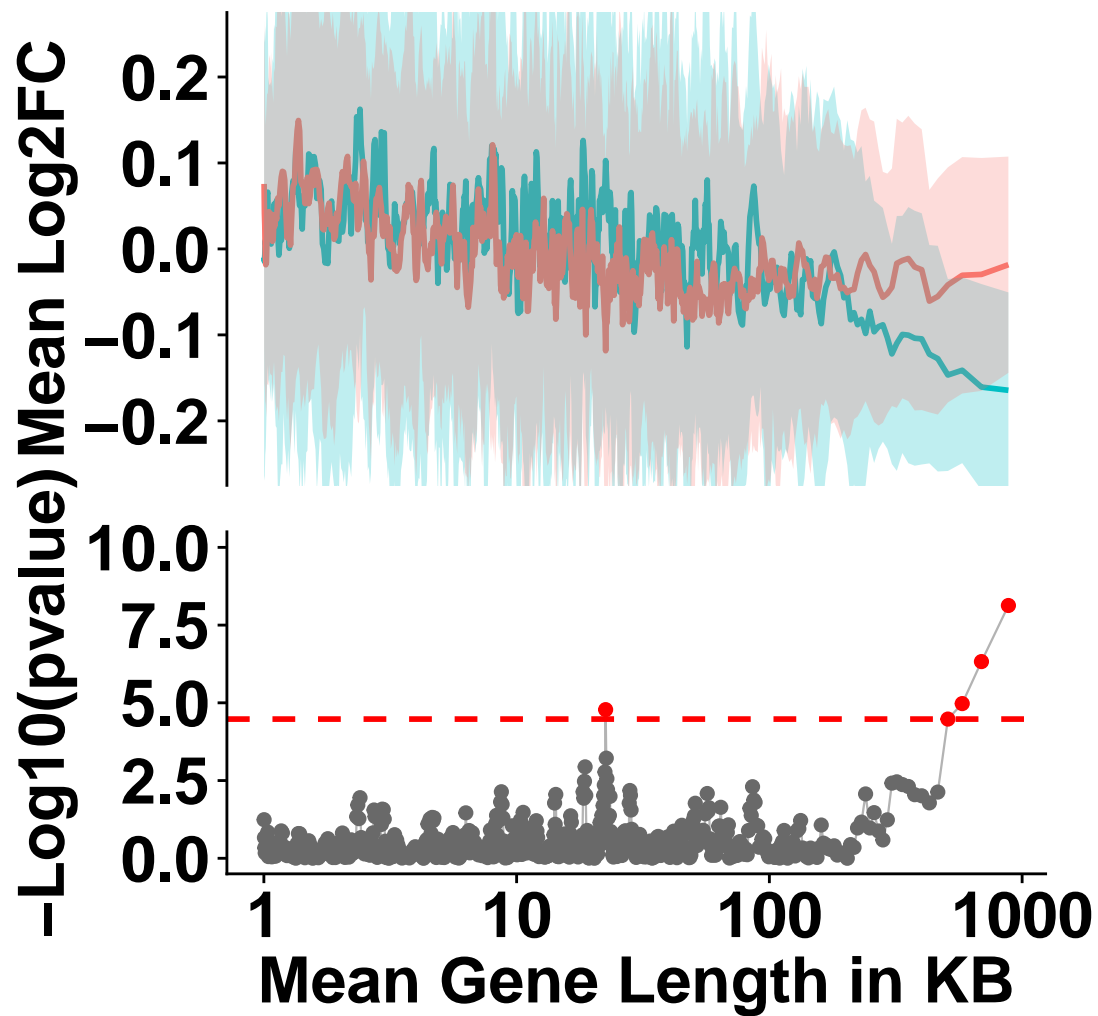
##  
 ##  
 ## Fig. 1(D) -- R106W Excitatory Neurons Female



##  
 ##  
 ## Fig. 1(D) -- R106W MUT Excitatory Neurons Female



##  
 ##  
 ## Fig. 1(D) -- T158M MUT Excitatory Neurons Female



```
cat("\n\n Printing Supplementary Fig. 2 \n\n")
```

```
##
```

```
##
```

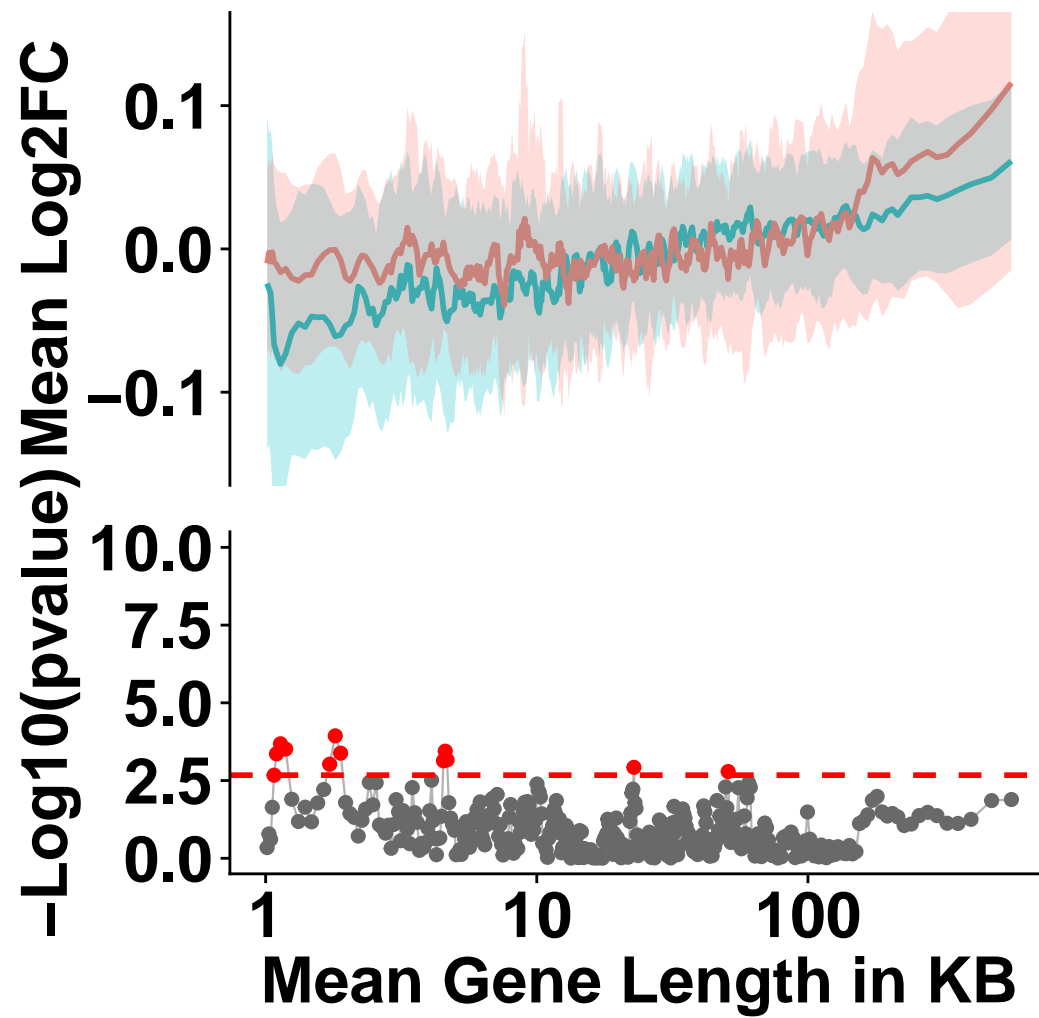
```
## Printing Supplementary Fig. 2
```

```
figureS2()
```

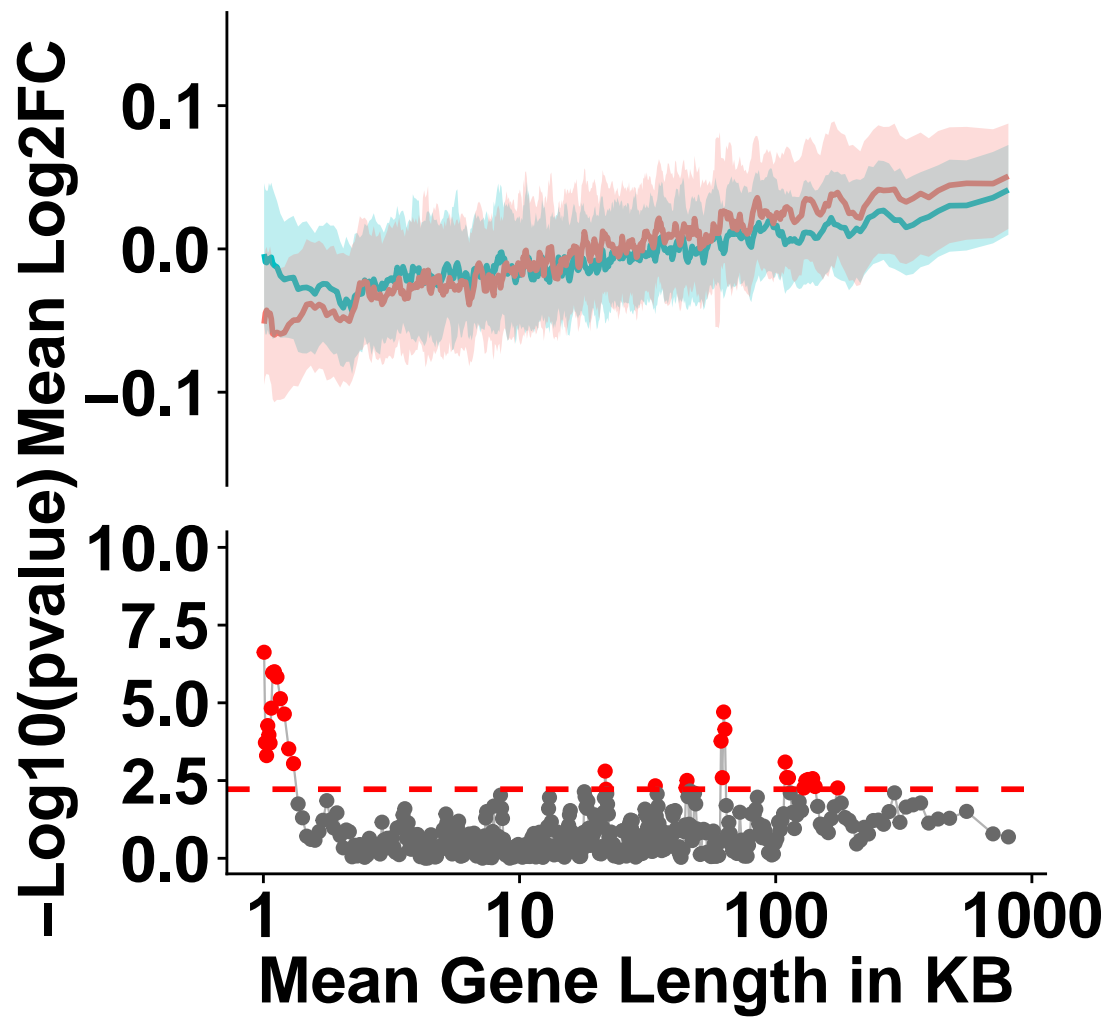
```
##
```

```
##
```

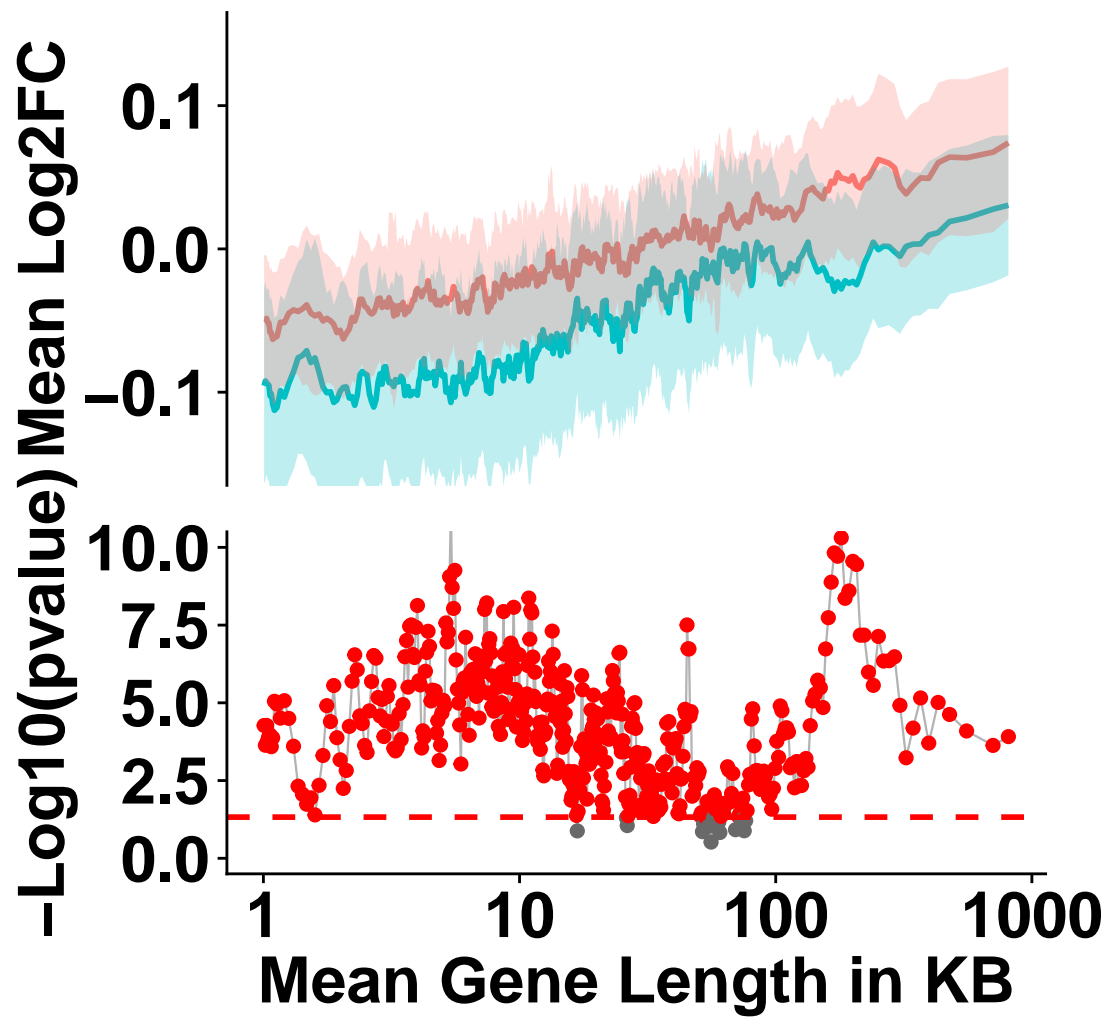
```
## Supplemnetary Fig. 2(A)-- Striatum
```



##  
 ##  
 ## Supplementary Fig. 2(B) -- Hippocampus (4 weeks)

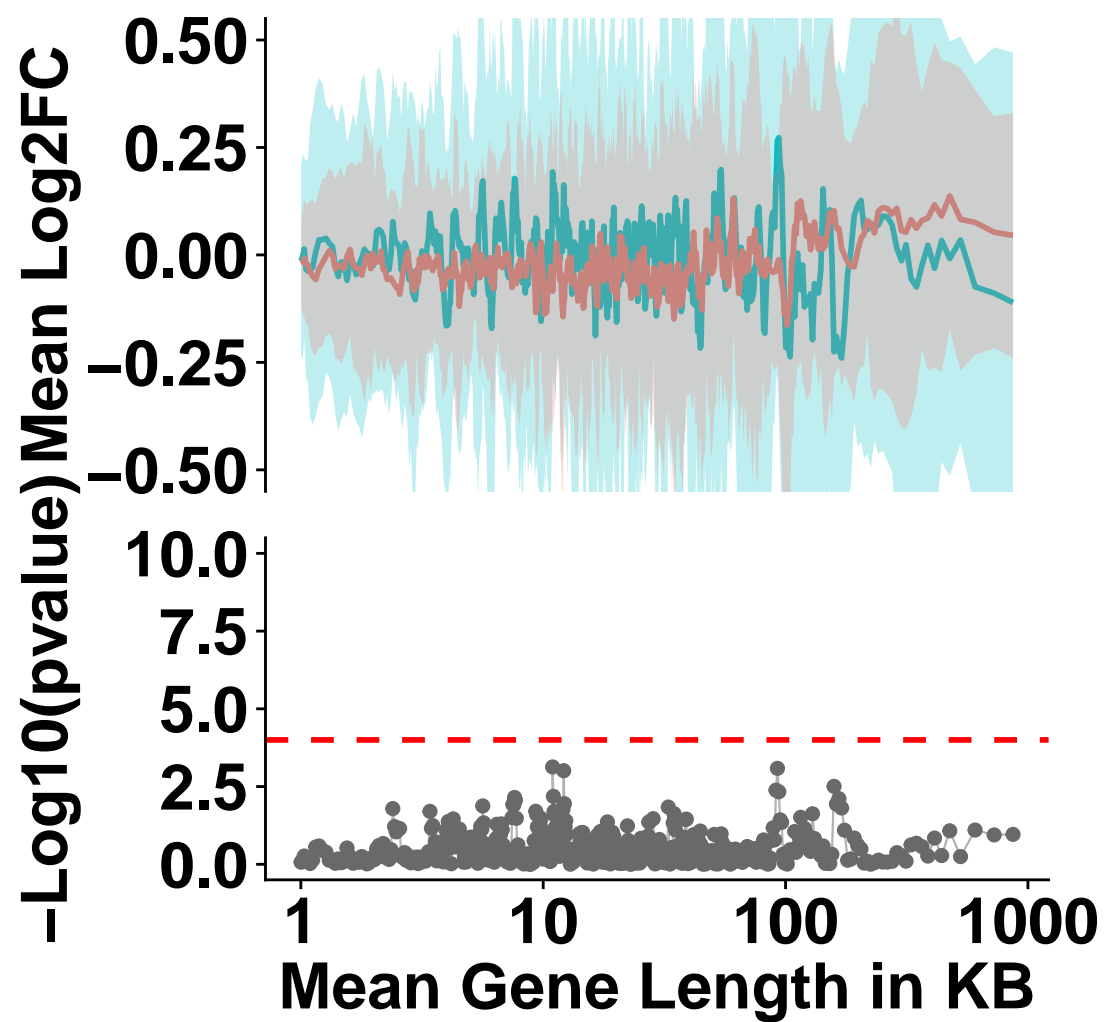


##  
 ##  
 ## Supplementary Fig. 2(C) -- Hippocampus (9 weeks)

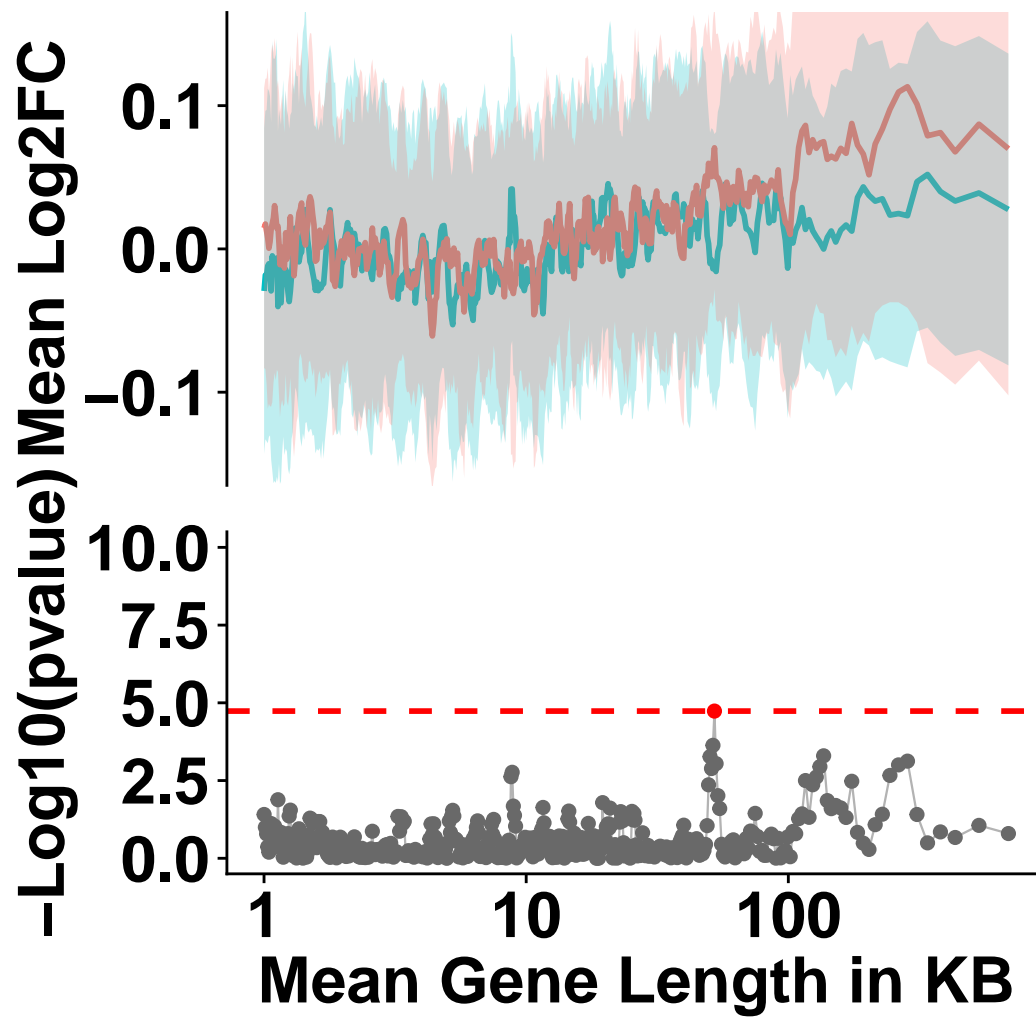


##  
 ##  
 ## Supplementary Fig. 2(D) -- Visual Cortex (RNA-seq)

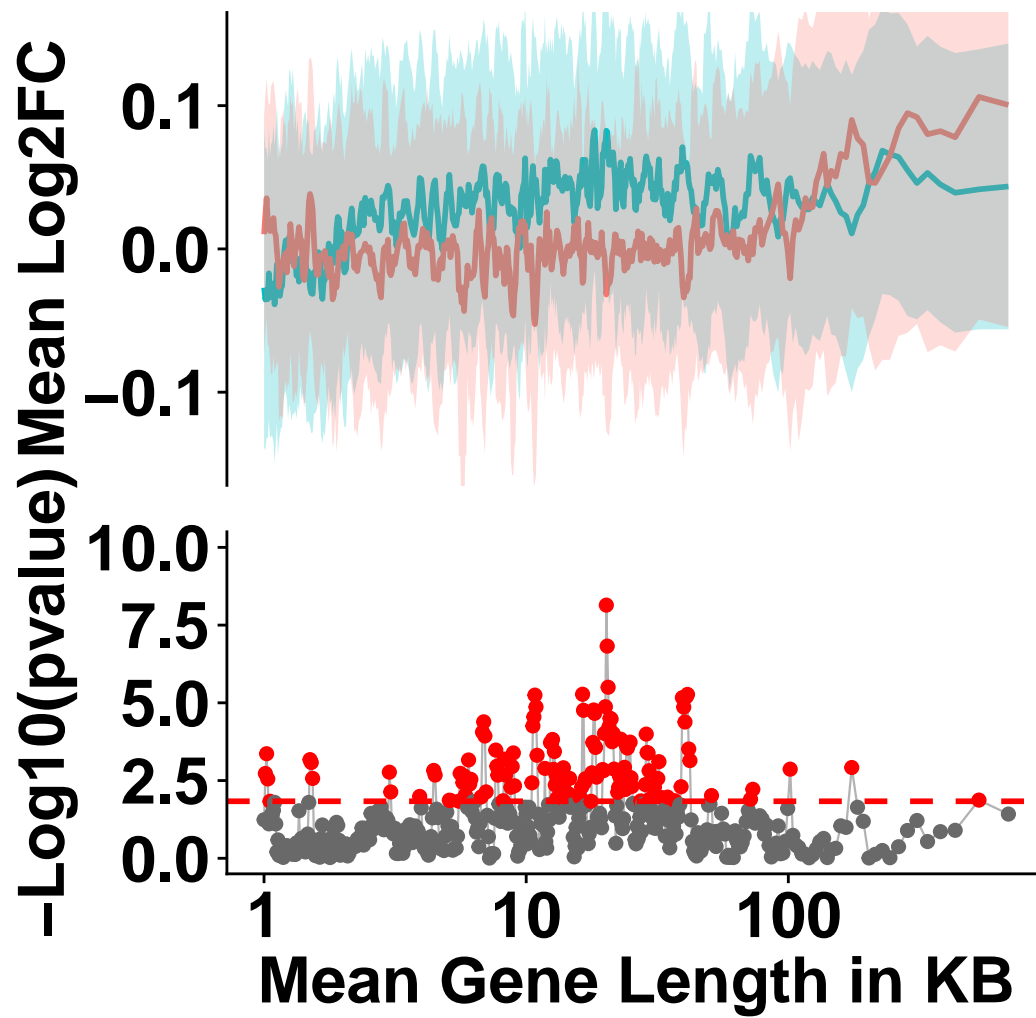




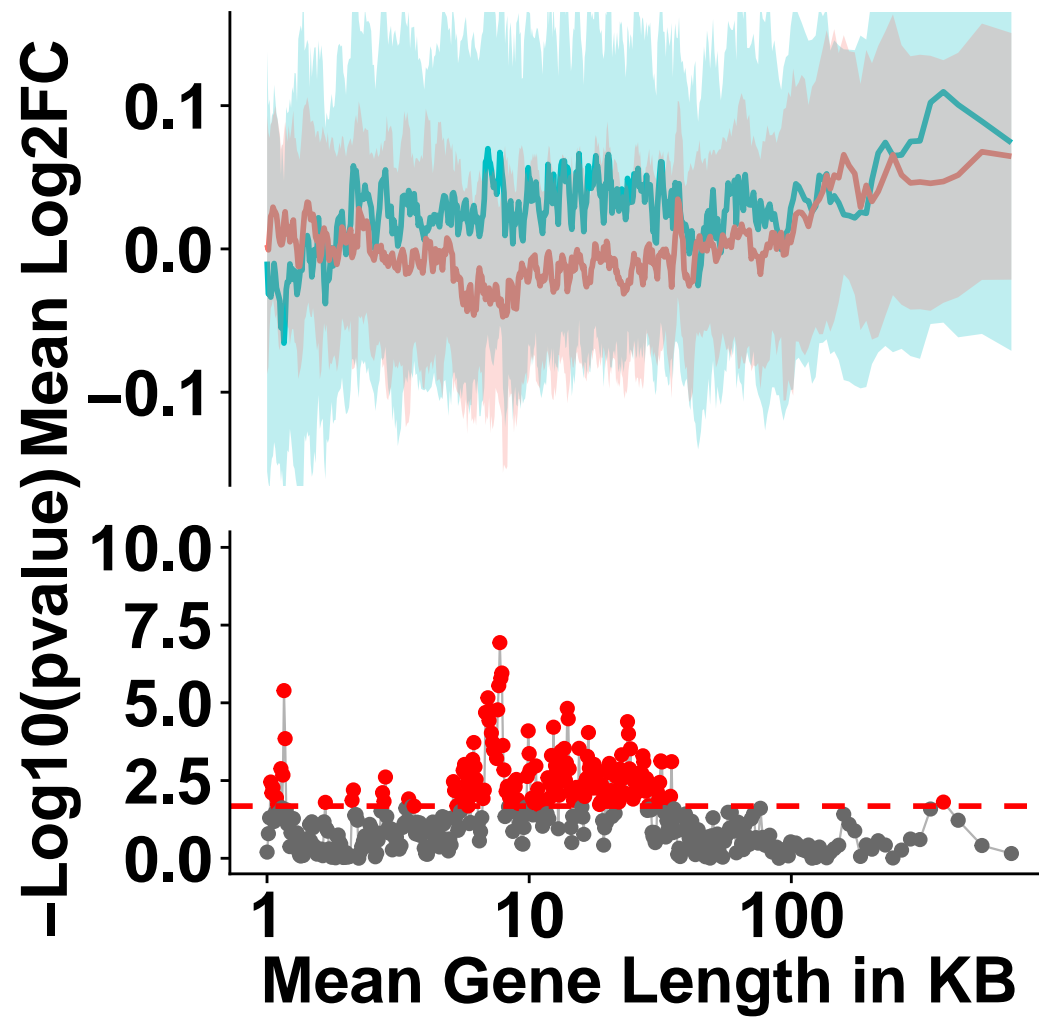
##  
 ##  
 ## Supplementary Fig. 2(E) -- Locus Coeruleus (P22)



##  
 ##  
 ## Supplementary Fig. 2(F) -- Locus Coeruleus



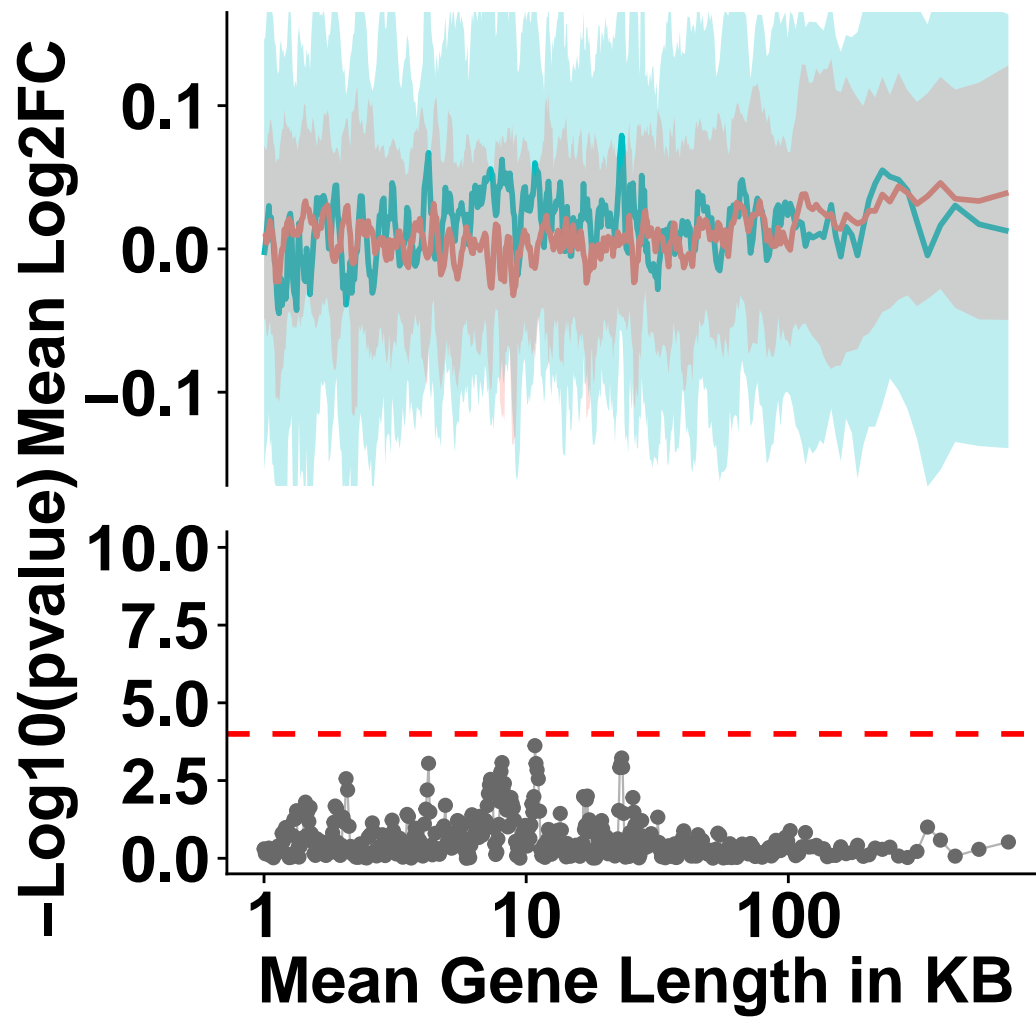
##  
 ##  
 ## Supplementary Fig. 2(G) -- FS, Motor Cortex



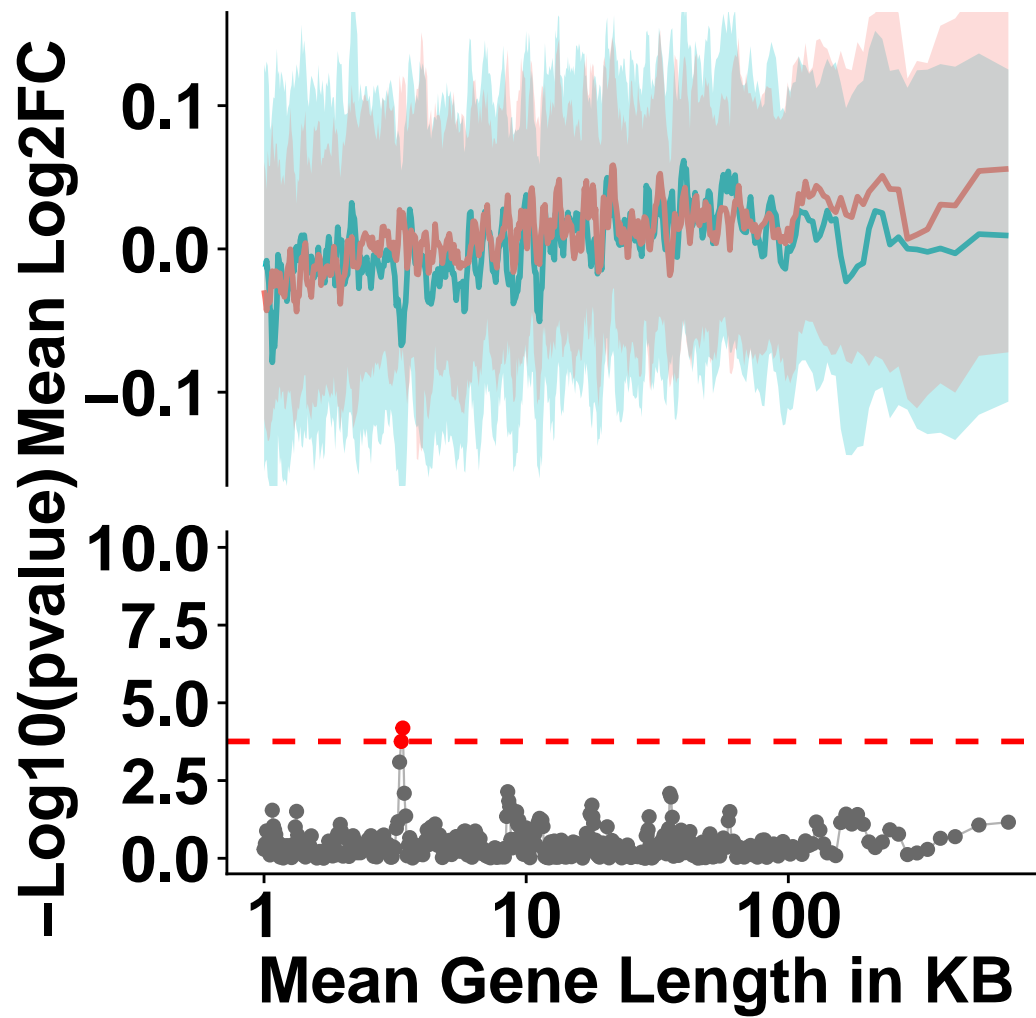
##

##

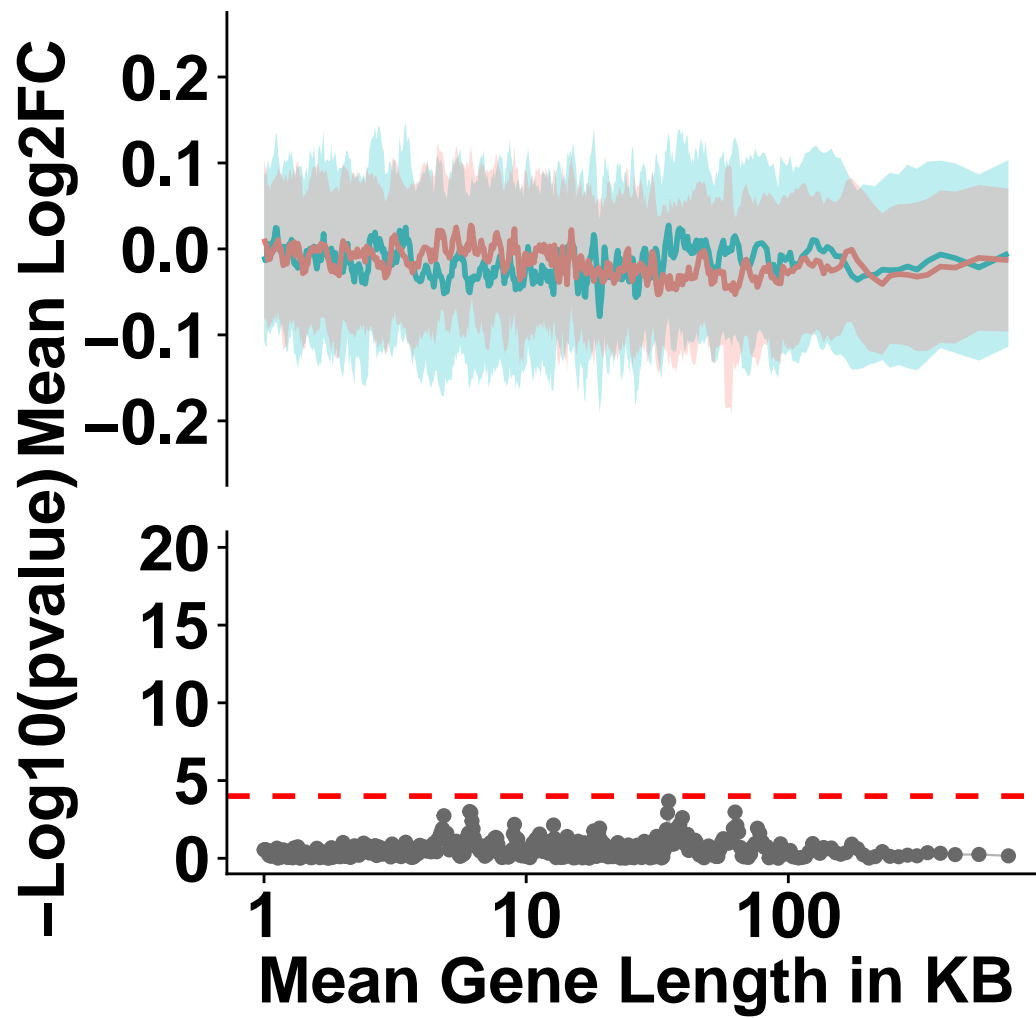
## Supplementary Fig. 2(H) -- Purkinje Cells Cerebellum



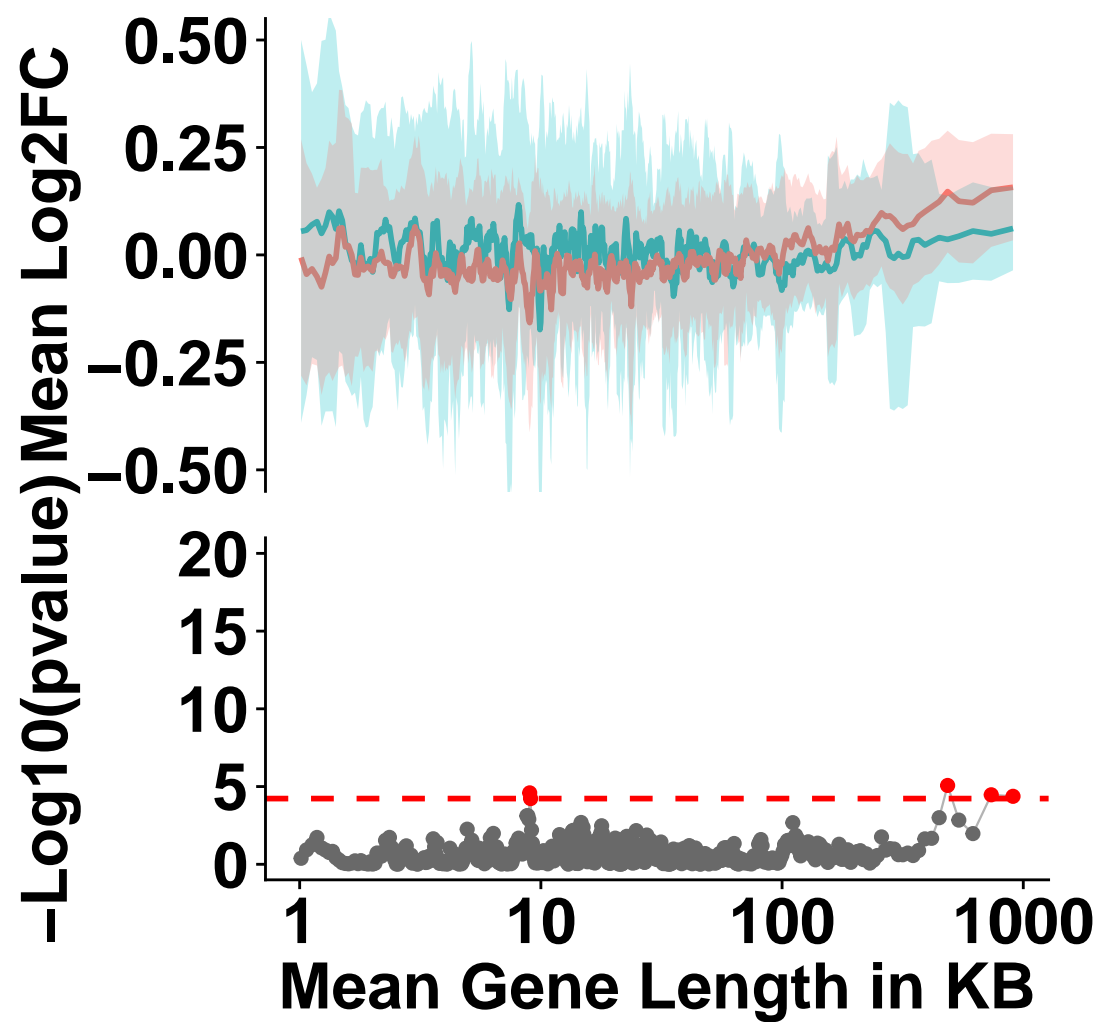
##  
 ##  
 ## Supplementary Fig. 2(I) -- PN, Motor Cortex



##  
 ##  
 ## Supplementary Fig. 2(J) -- Callosal Projection Neurons

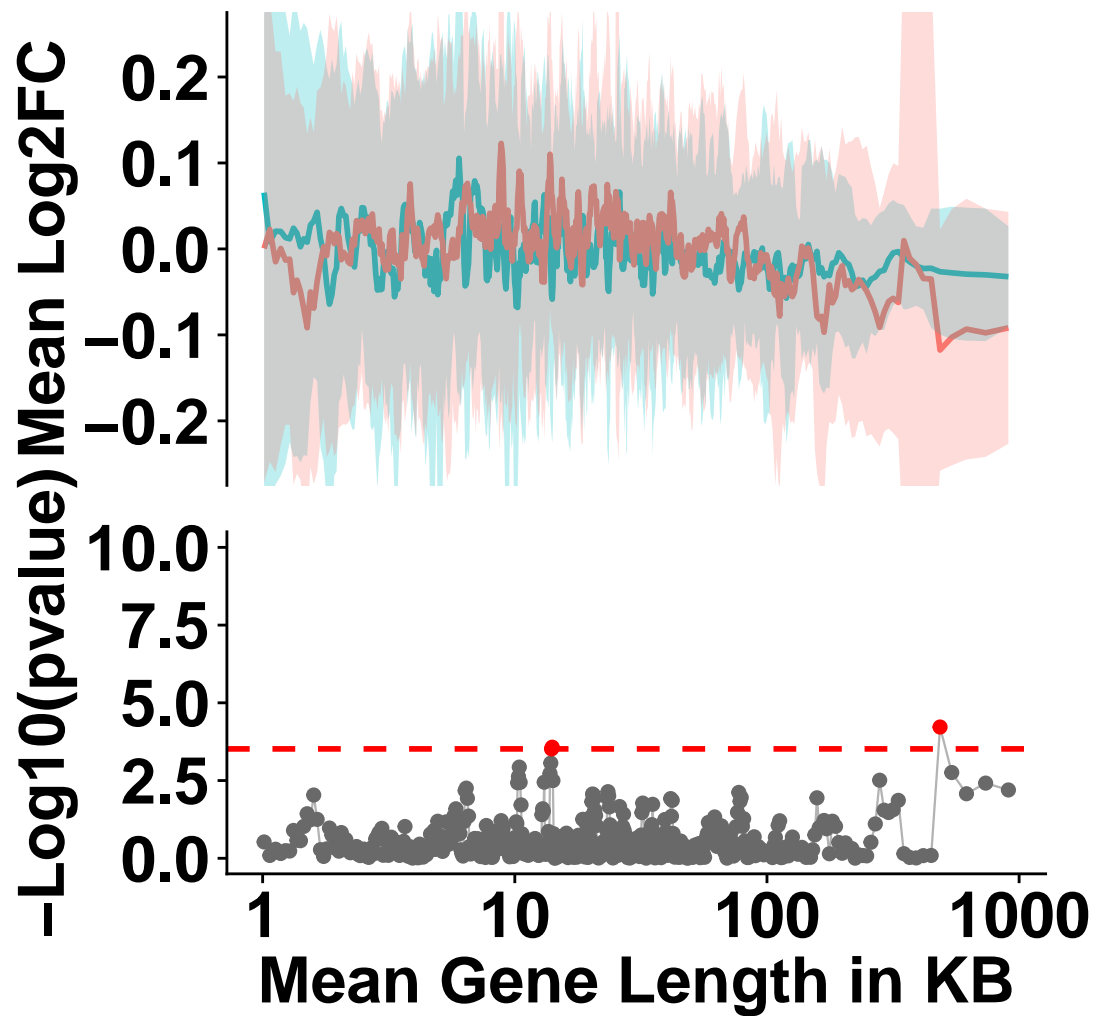


##  
 ##  
 ## Printing Supp. Figure 2(K) -- Hypothalamus (KO -- RNA-Seq)



```
##
##
## Printing Supp. Figure 2(L) -- Hypothalamus (Tg -- RNA-Seq)
```





```
cat("\n\n Printing Supplementary Fig. 3 \n\n")
```

```
##
```

```
##
```

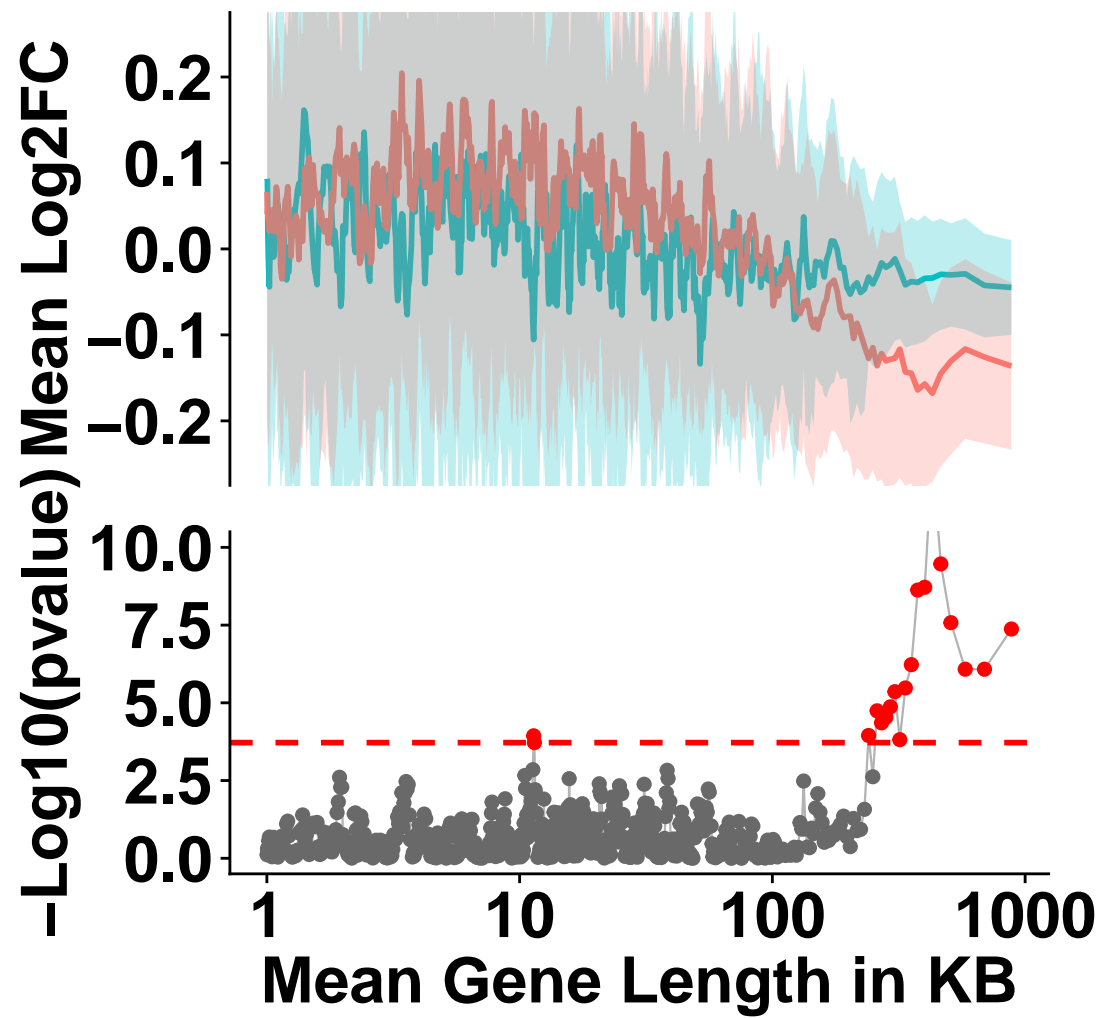
```
## Printing Supplementary Fig. 3
```

```
figureS3()
```

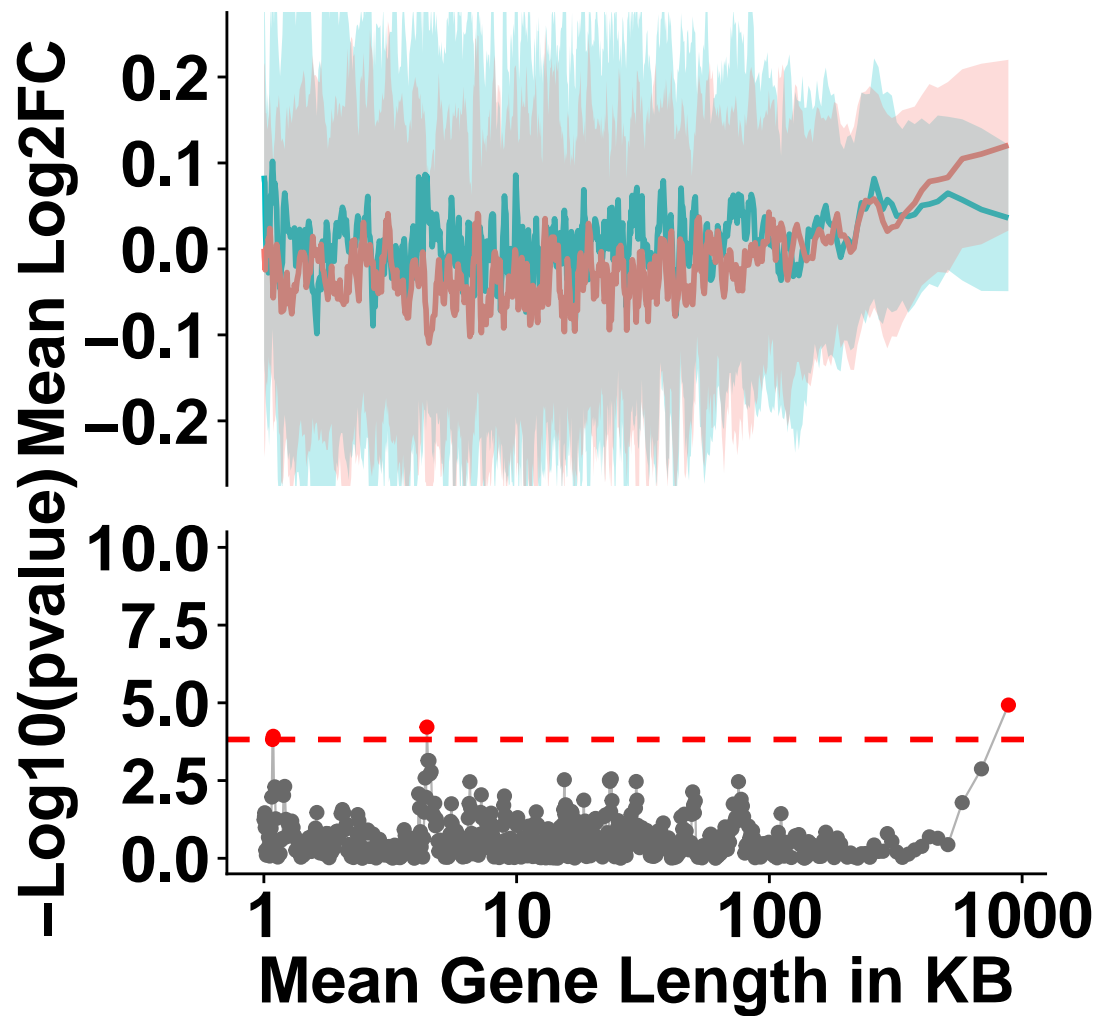
```
##
```

```
##
```

```
## Supplemnetary Fig. 3 -- GRO-seq
```



```
##
##
##  Supplementary Fig. 3 -- Whole Cell-seq
```



```
cat("\n\n Printing Supplementary Fig. 4 \n\n")
```

```
##
```

```
##
```

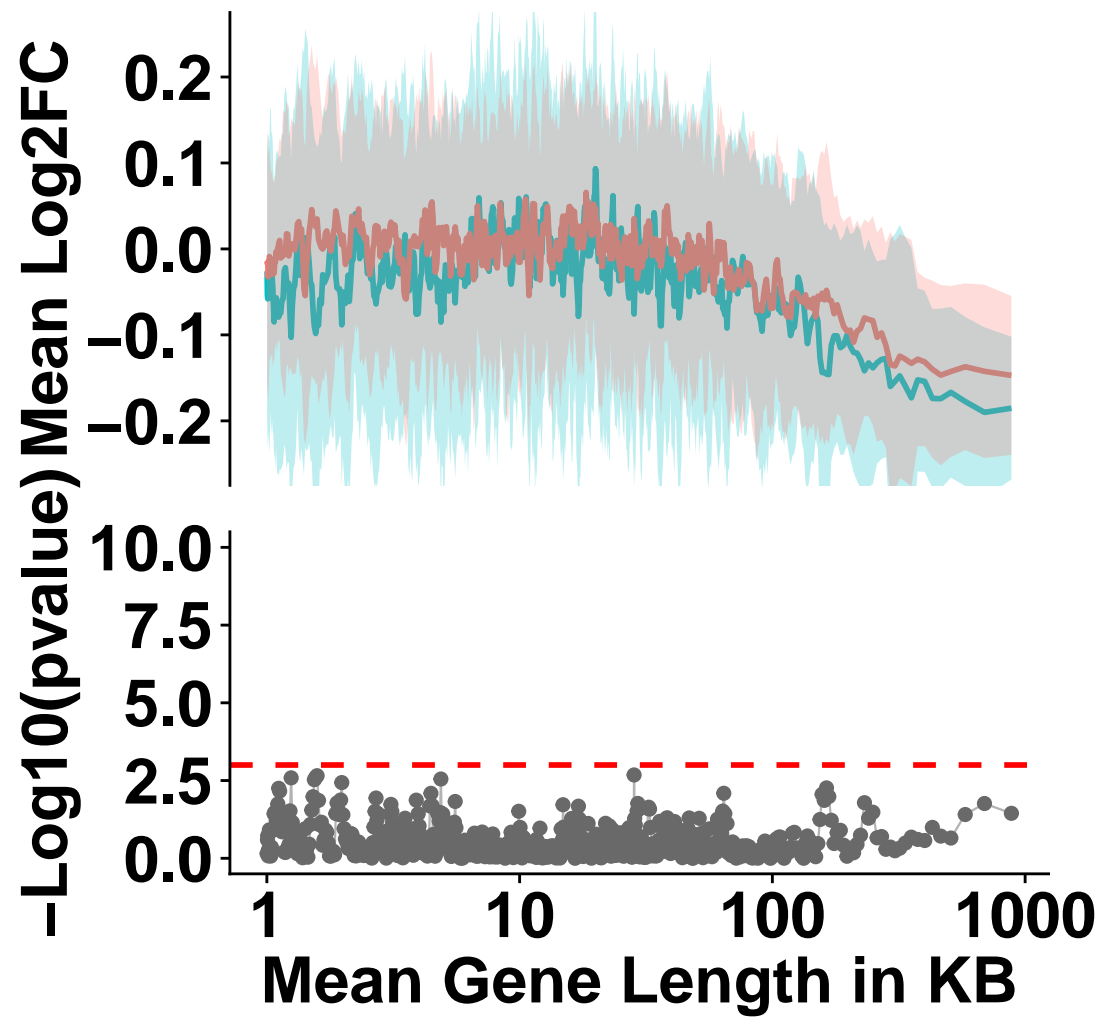
```
## Printing Supplementary Fig. 4
```

```
figureS4()
```

```
##
```

```
##
```

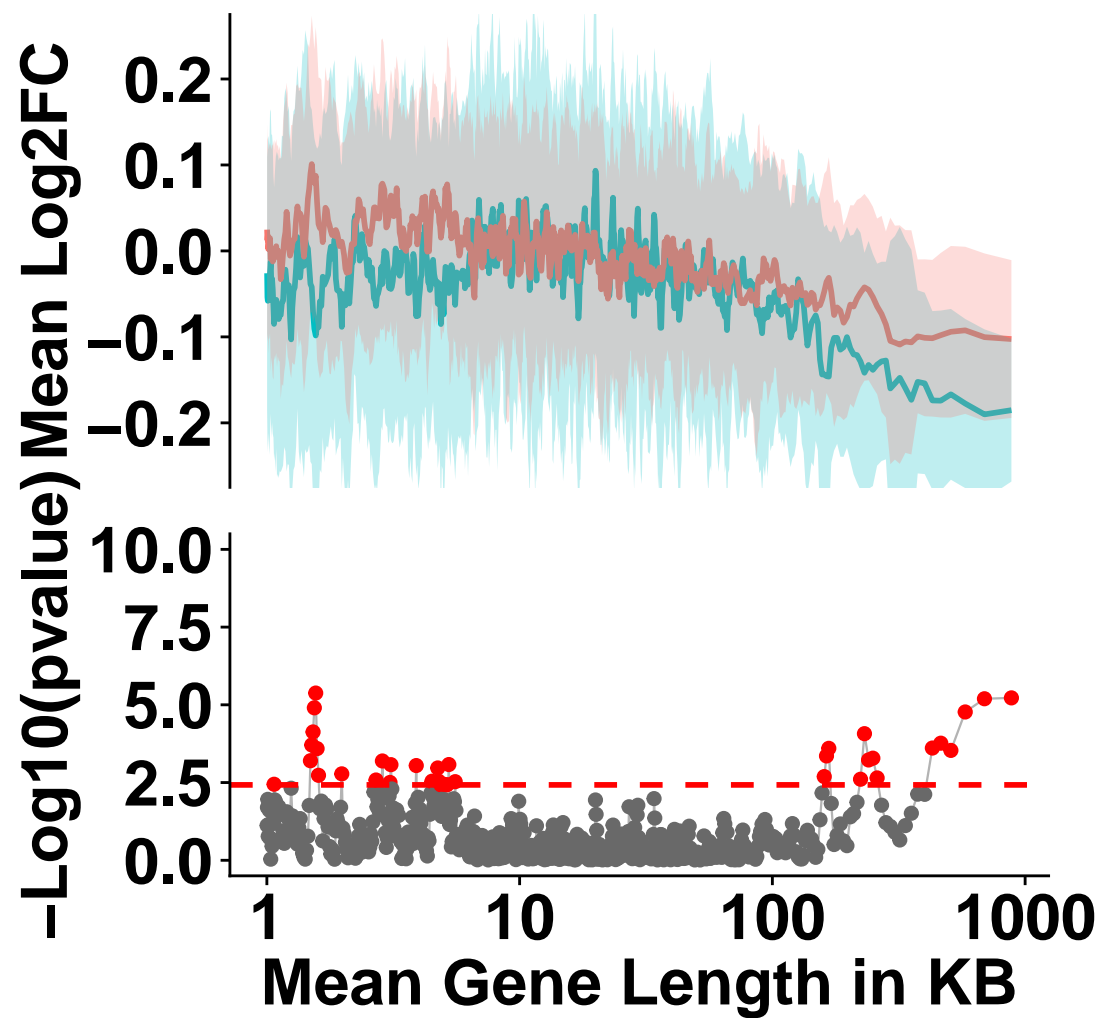
```
## Supplemnetary Fig. 4(A) -- R106W Excitatory Nuclear Male
```



##

##

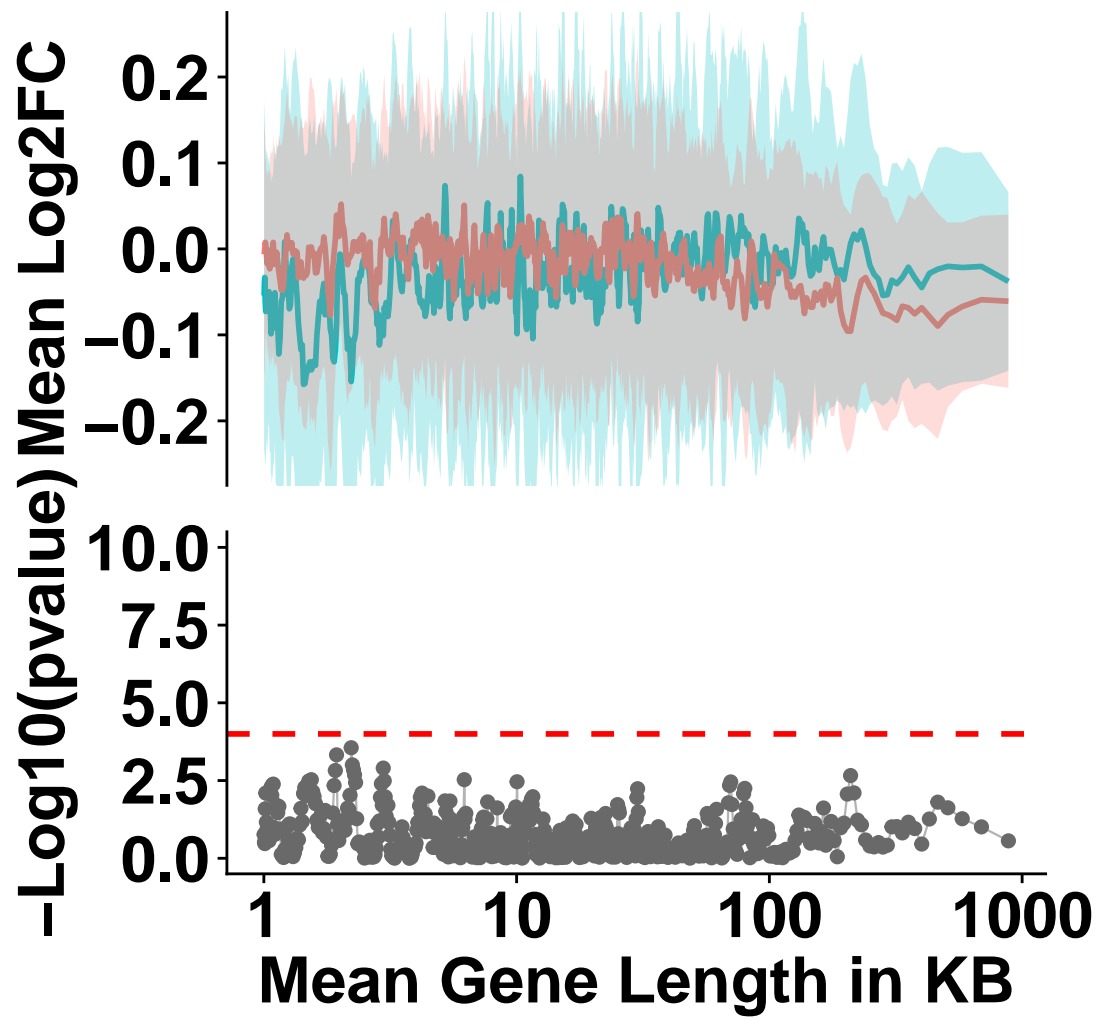
## Supplementary Fig. 4(B) -- T158M Excitatory Nuclear Male



##

##

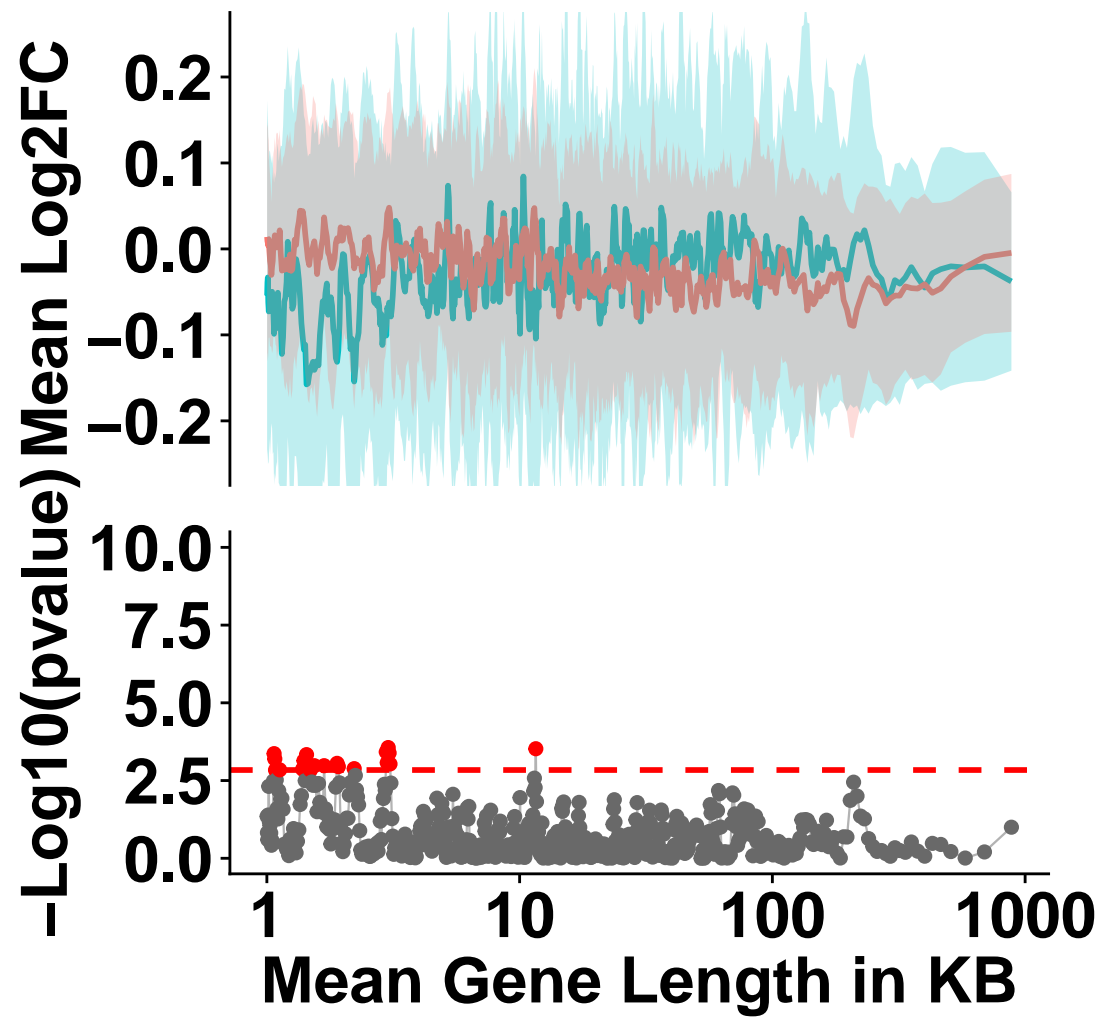
## Supplementary Fig. 4(C) -- R106W Inhibitory Nuclear Male



##

##

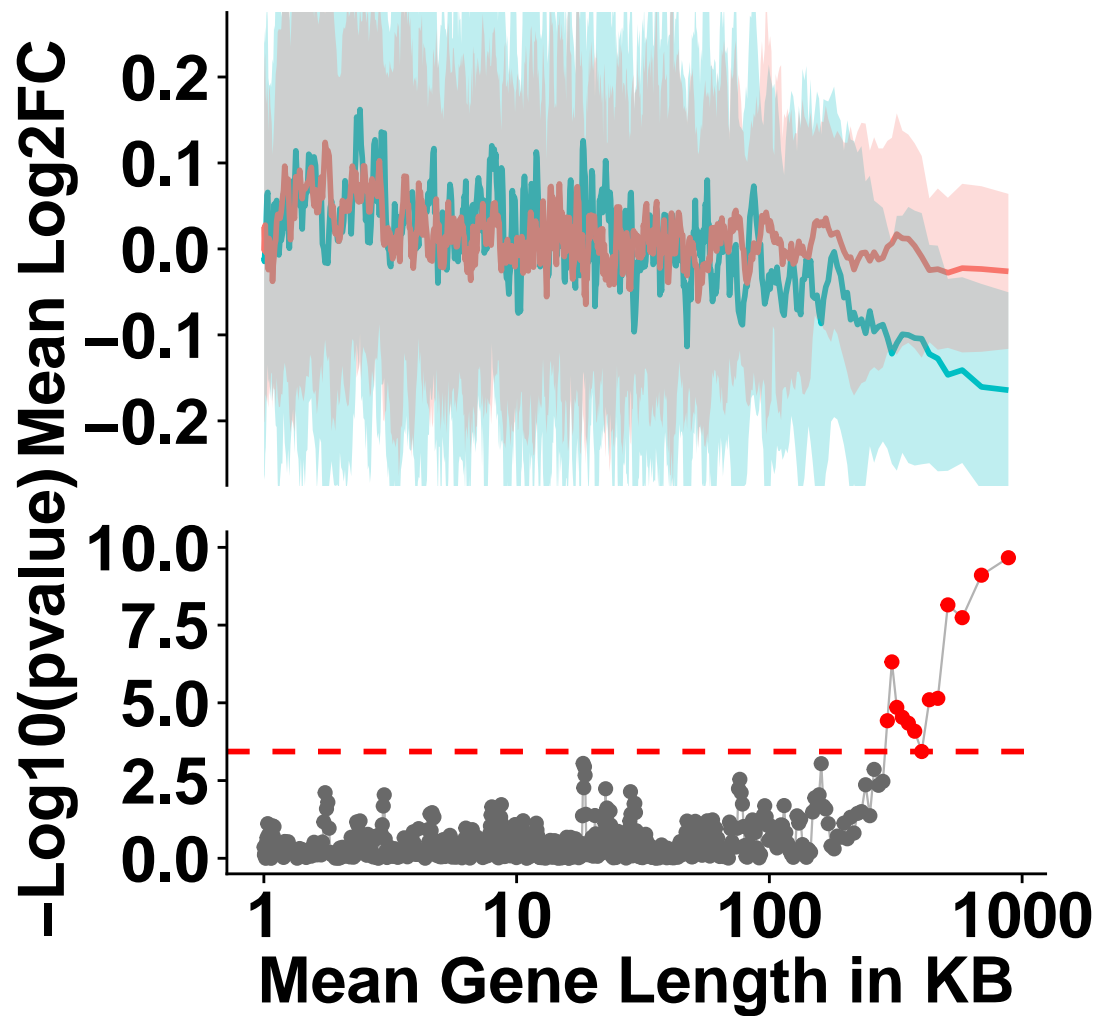
## Supplementary Fig. 4(D) -- T158M Inhibitory Nuclear Male



##

##

## Supplementary Fig. 4(E)--T158M-WT Excitatory Nuclear Female



Section 4: Figure 2

```
cat("\n\n Printing Figure 2 \n\n")
```

```
##
```

```
##
```

```
## Printing Figure 2
```

```
figure2()
```

```
##
```

```
##
```

```
## Fig. 2(A) -- Lowry's Human RTT in vitro Dataset
```

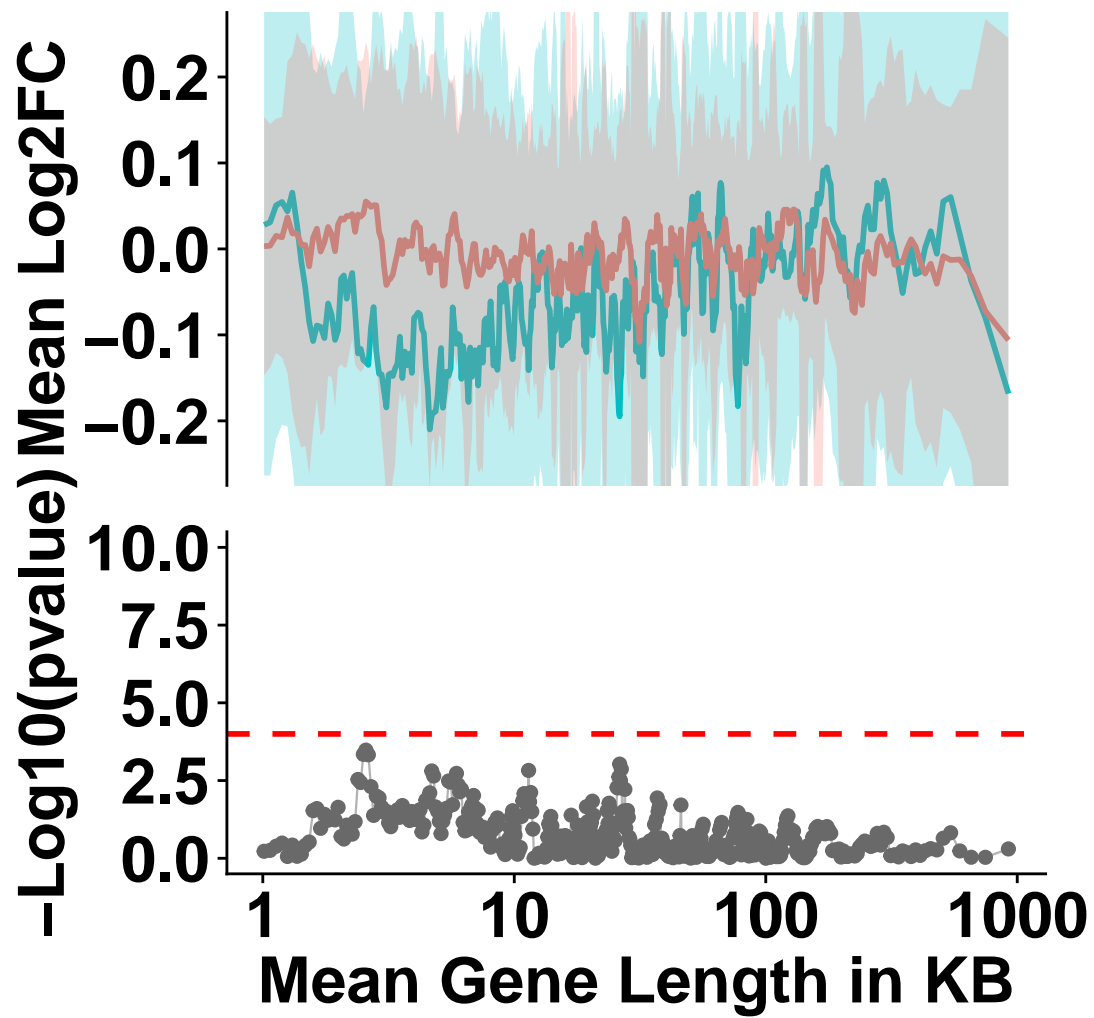
```
##
```

```
##
```

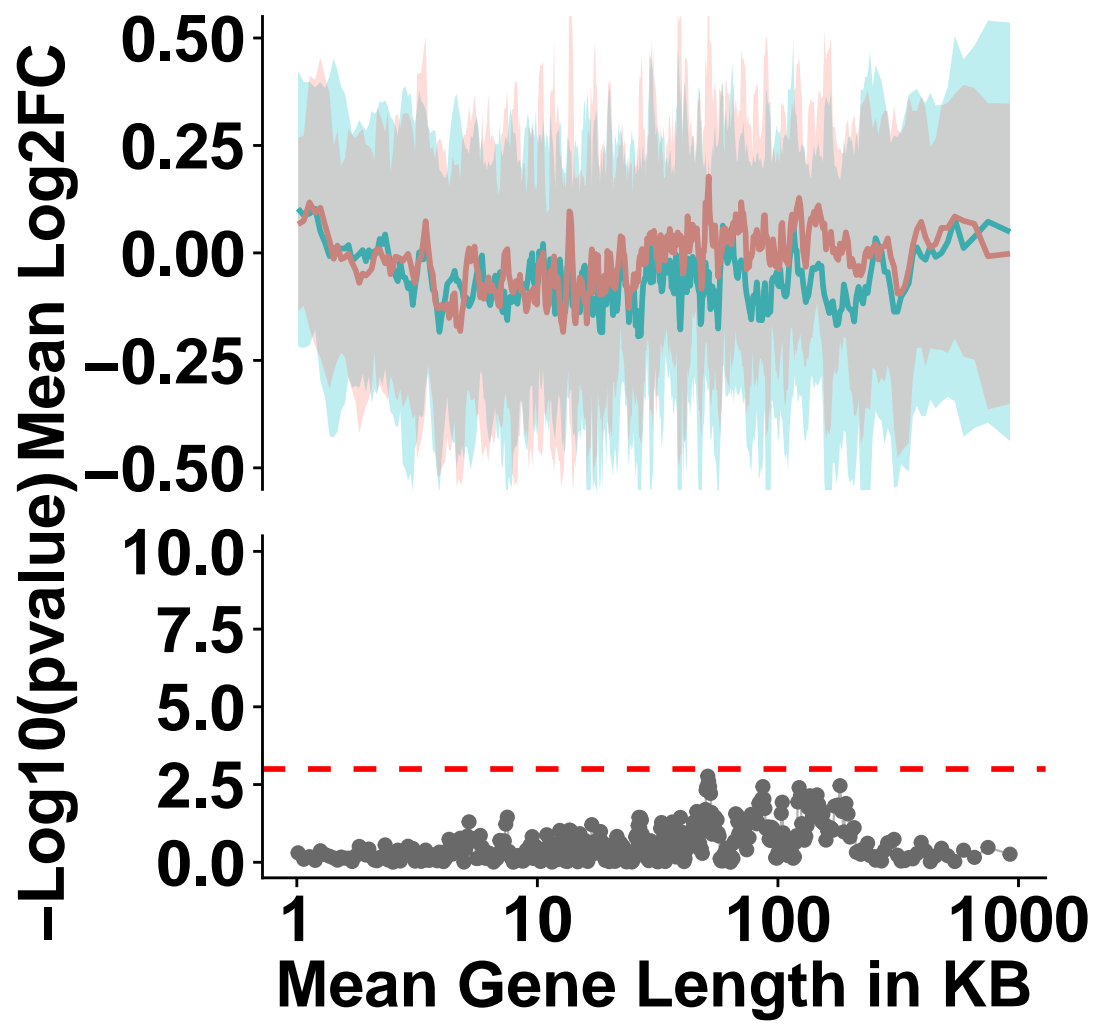
```
##
```

```
## iPSC Dataset (RTT/WT)
```

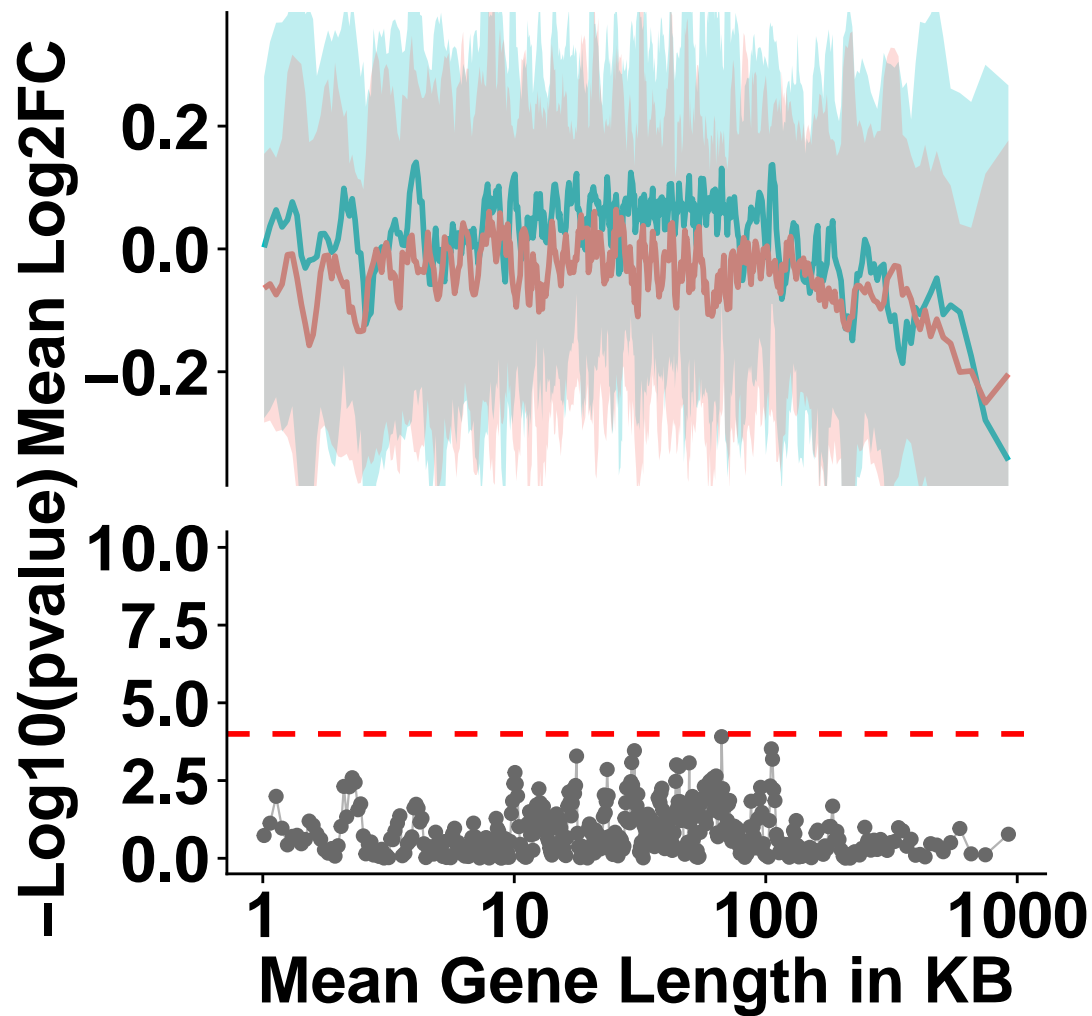




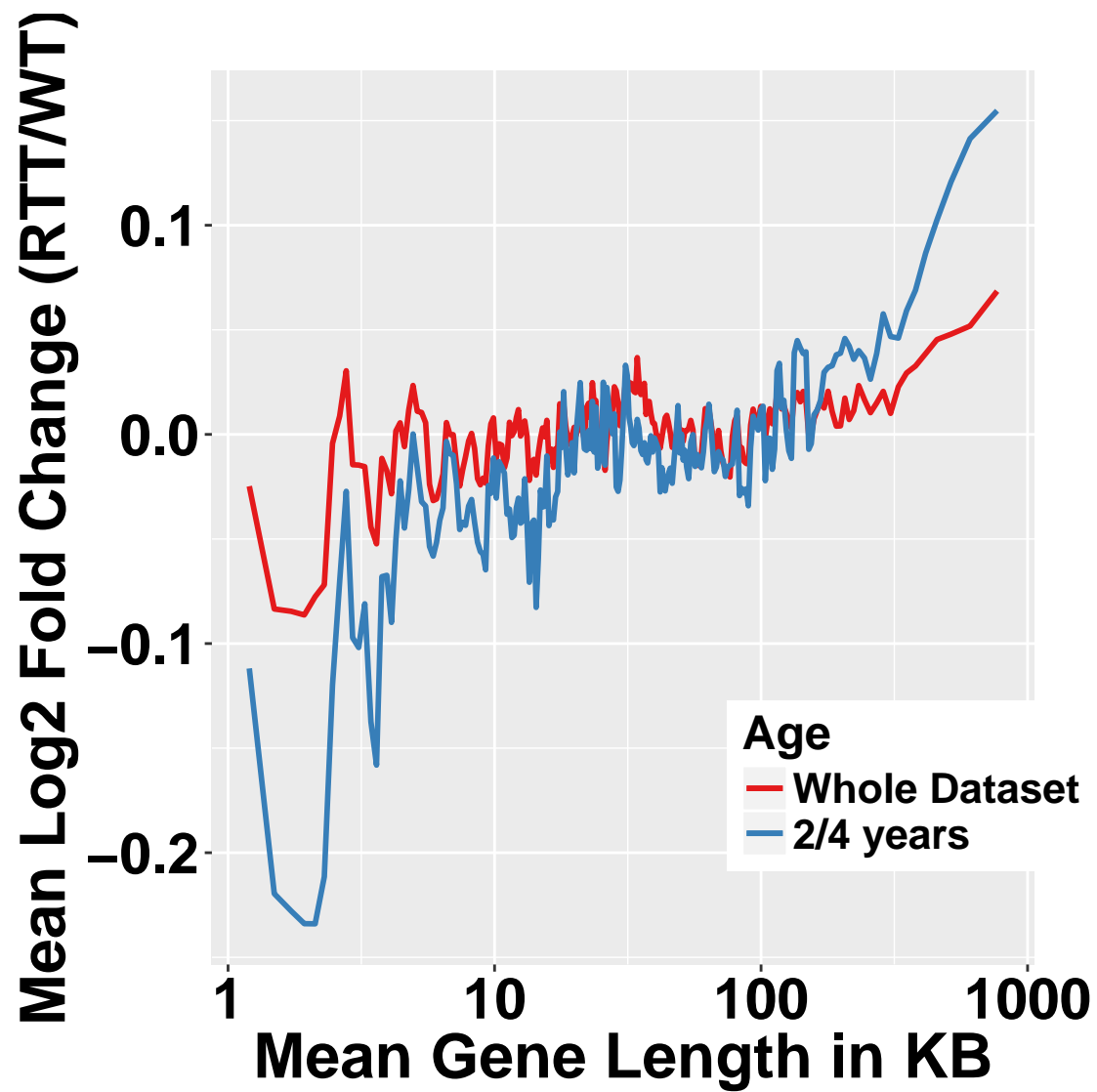
##  
 ##  
 ## NPC Dataset (RTT/WT)



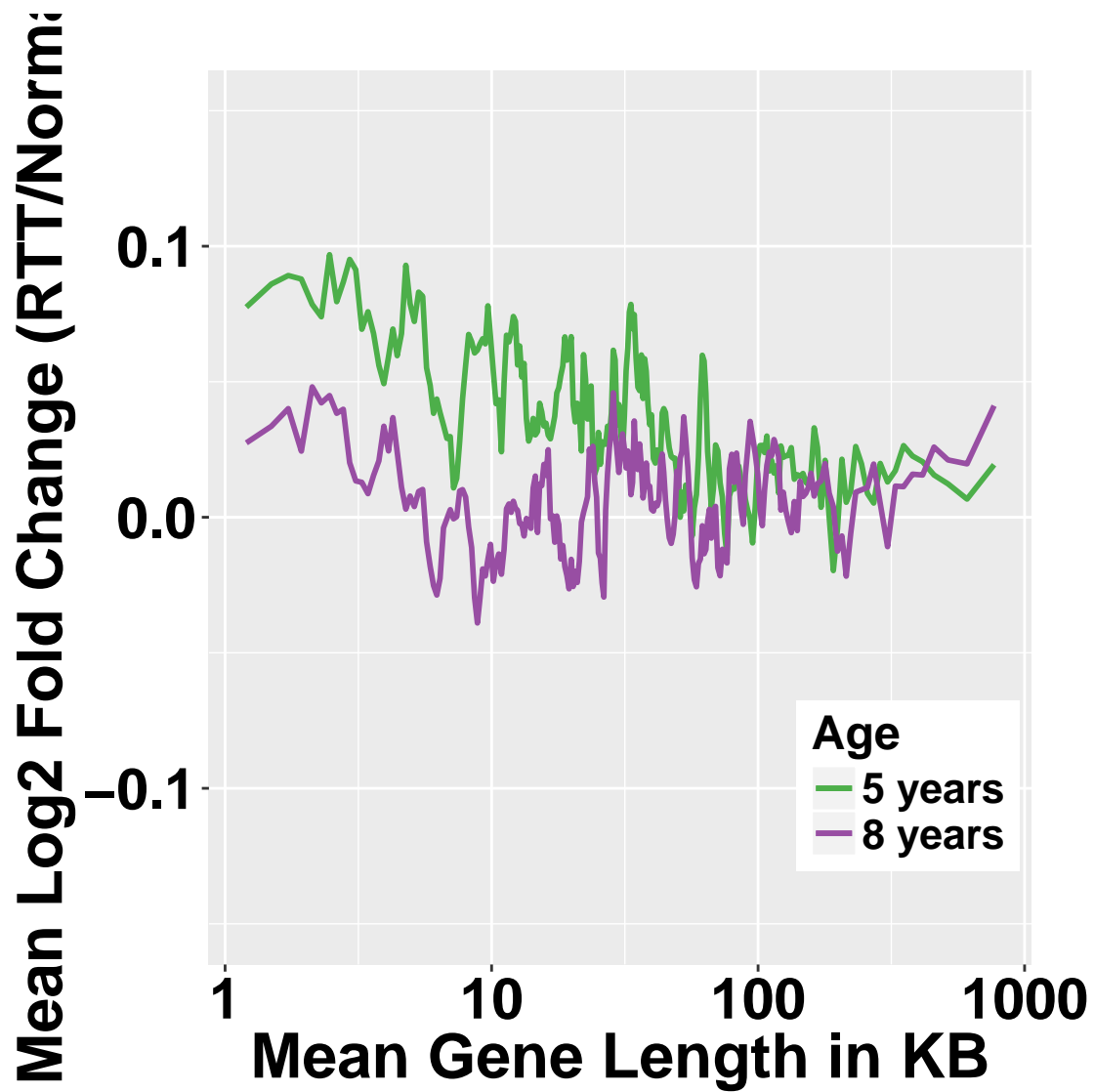
##  
##  
## Neuron Dataset (RTT/WT)



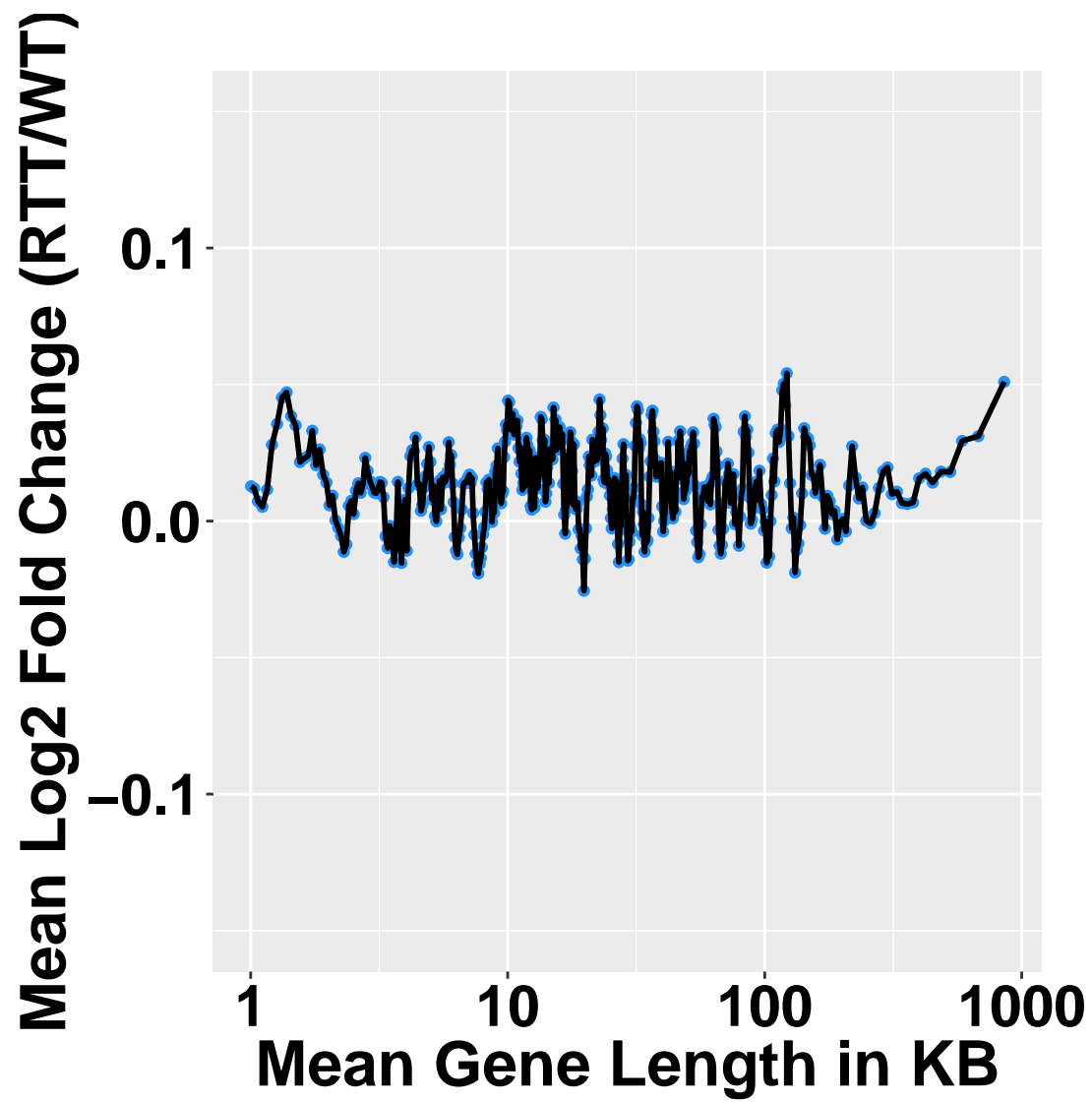
```
##
##
## Fig. 2(B) -- Deng's Dataset (RTT/WT)
##
## Running Deng's Mecp2 Dataset -- Human Dataset
## Reading in : ../dat/GEO/GSE6955_RAW/GSM160306.CEL.gz
## Reading in : ../dat/GEO/GSE6955_RAW/GSM160307.CEL.gz
## Reading in : ../dat/GEO/GSE6955_RAW/GSM160308.CEL.gz
## Reading in : ../dat/GEO/GSE6955_RAW/GSM160309.CEL.gz
## Reading in : ../dat/GEO/GSE6955_RAW/GSM160310.CEL.gz
## Reading in : ../dat/GEO/GSE6955_RAW/GSM160311.CEL.gz
## Background correcting
## Normalizing
## Calculating Expression
##
##
## Printing Figure 4(A)
```



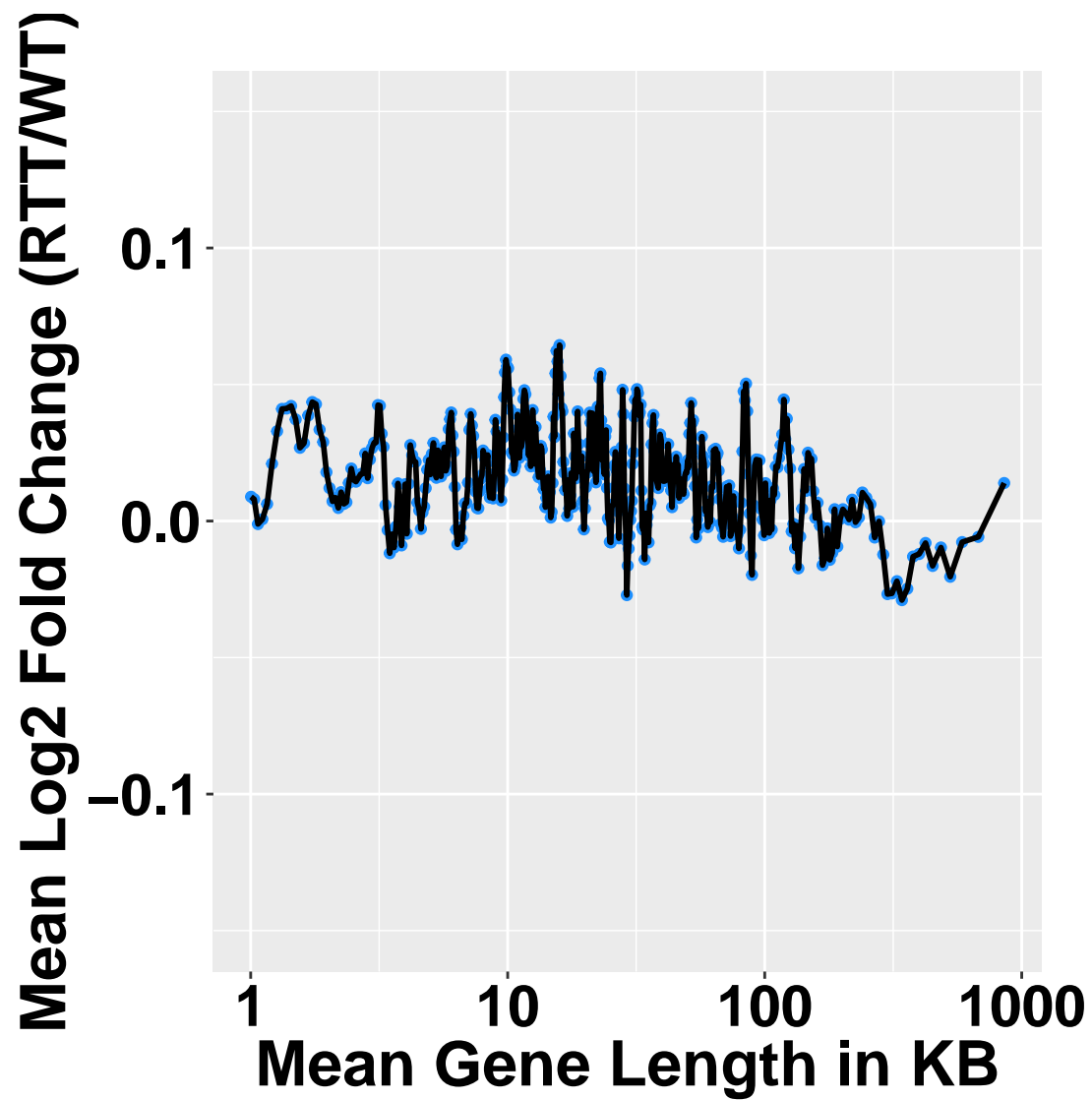
```
##  
##  
## Printing Figure 4(B)
```



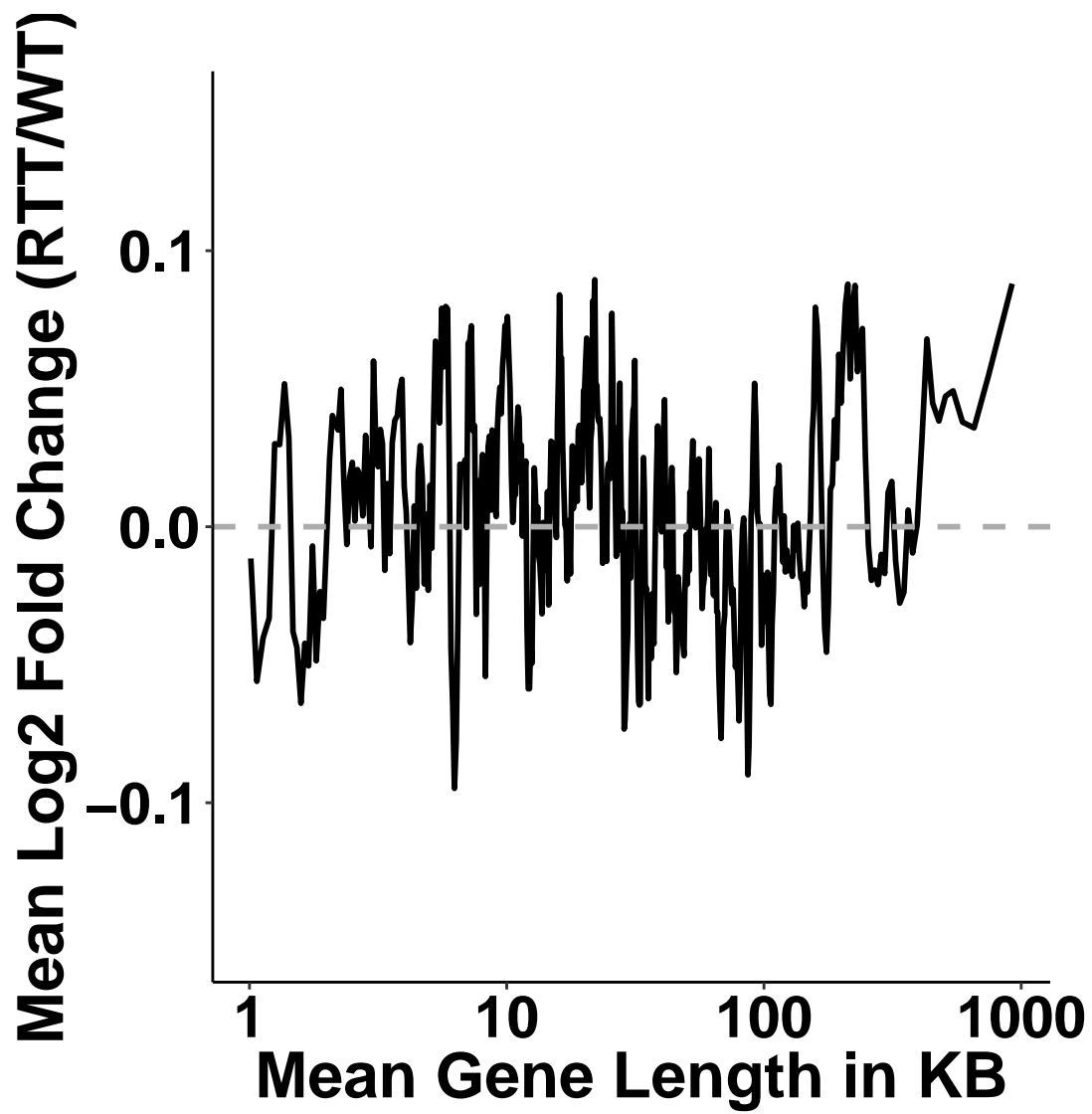
```
##
##
## Fig. 2(C) -- Lin's Dataset (RTT/WT)
##
##
##
## Frontal Cortex
```



```
##  
##  
## Frontal Cortex
```

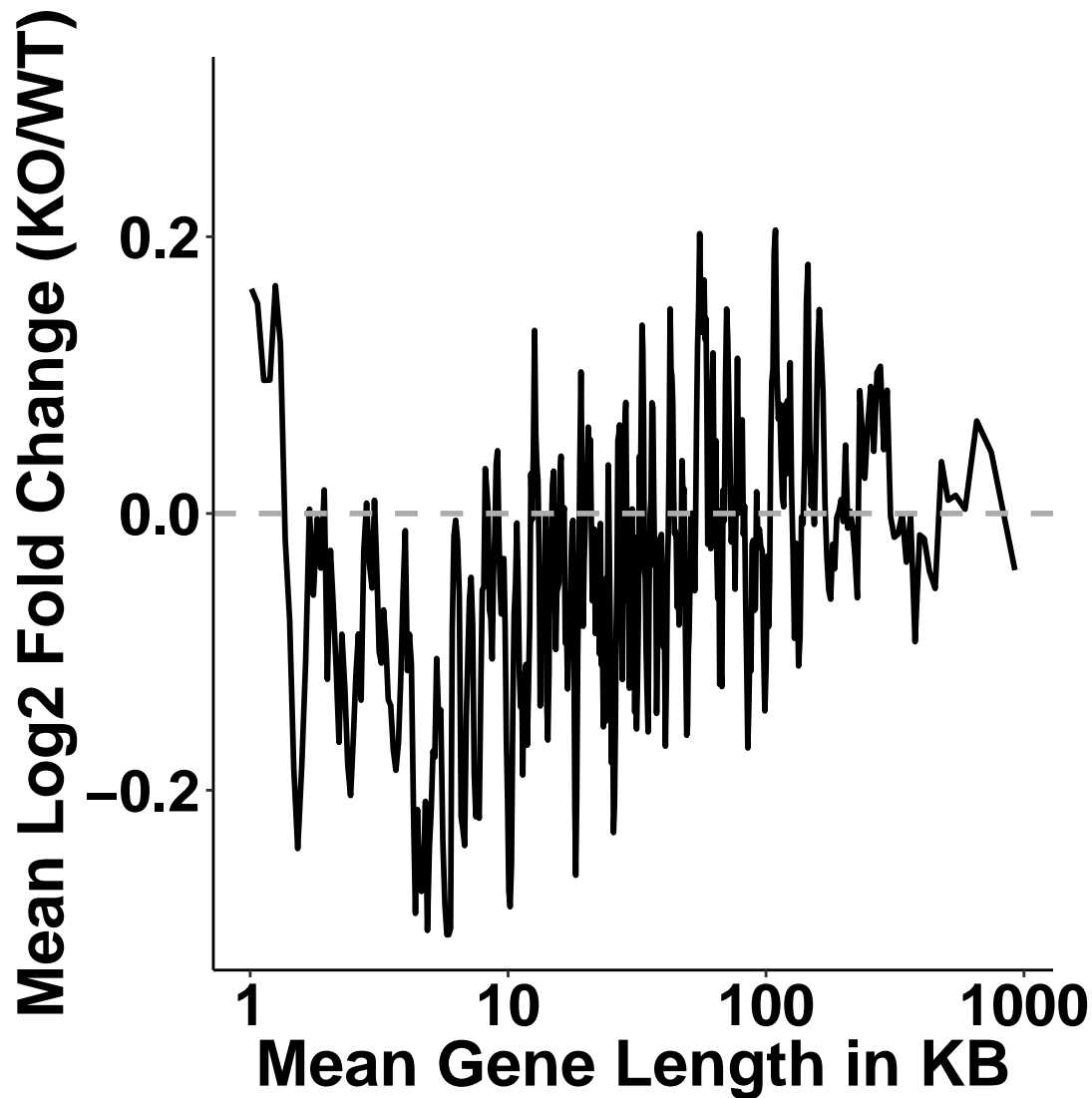


```
##
##
## Fig. 2(D) -- Lowry's Human RTT Dataset (RTT/WT)
##
##
##
## Frontal Cortex (18 years; Female)
```



##  
##  
## Frontal Cortex (1 year; Male)





Section 5: Figure 3 and Supplementary Figure 5

```
cat("\n\n Printing Fig. 3 \n\n")
```

```
##
```

```
##
```

```
## Printing Fig. 3
```

```
figure3()
```

```
##
```

```
##
```

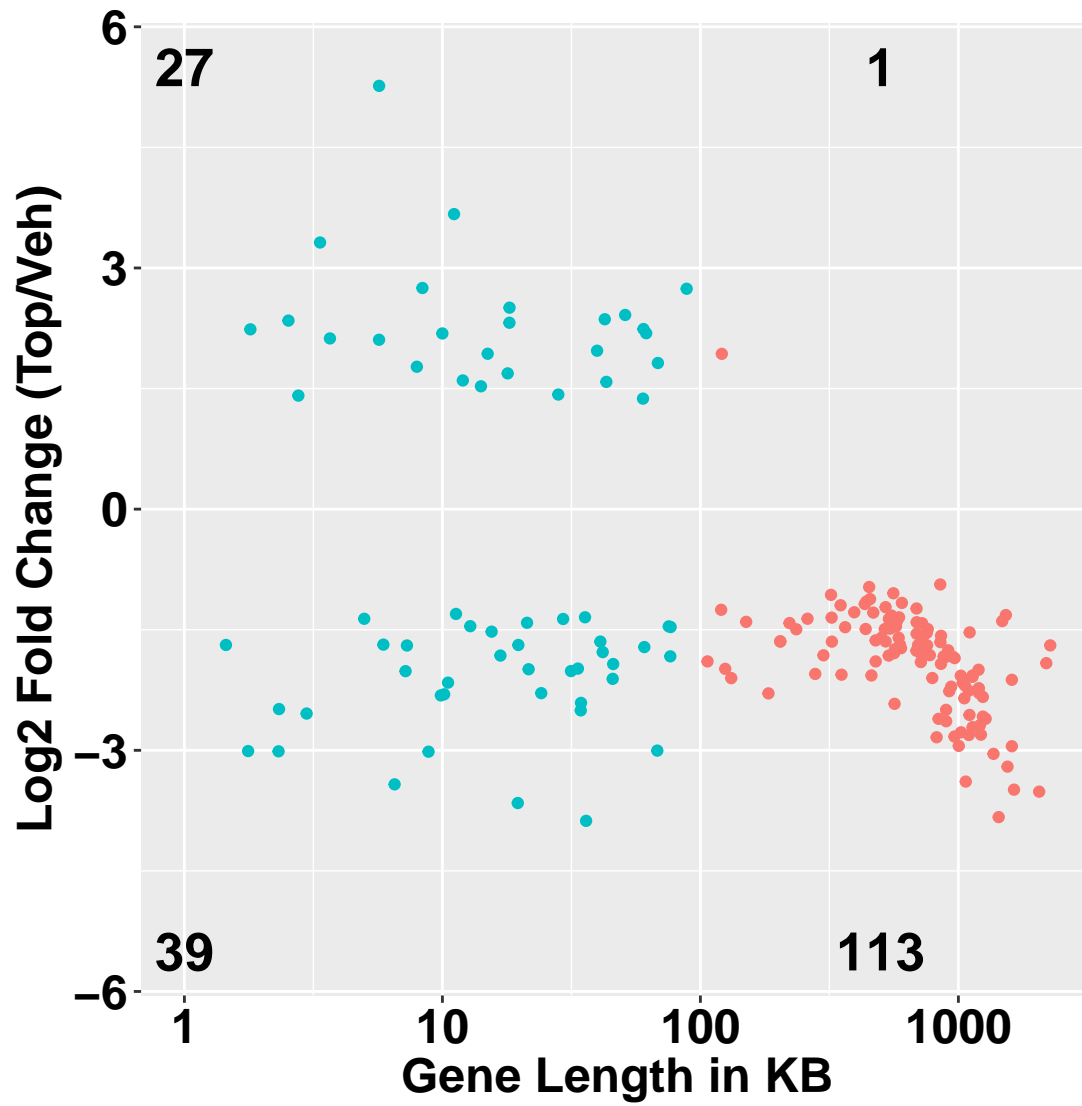
```
## Fig. 3(A) -- 300nM Topotecan Treatment RNA-seq Datasets
```

```
##
```

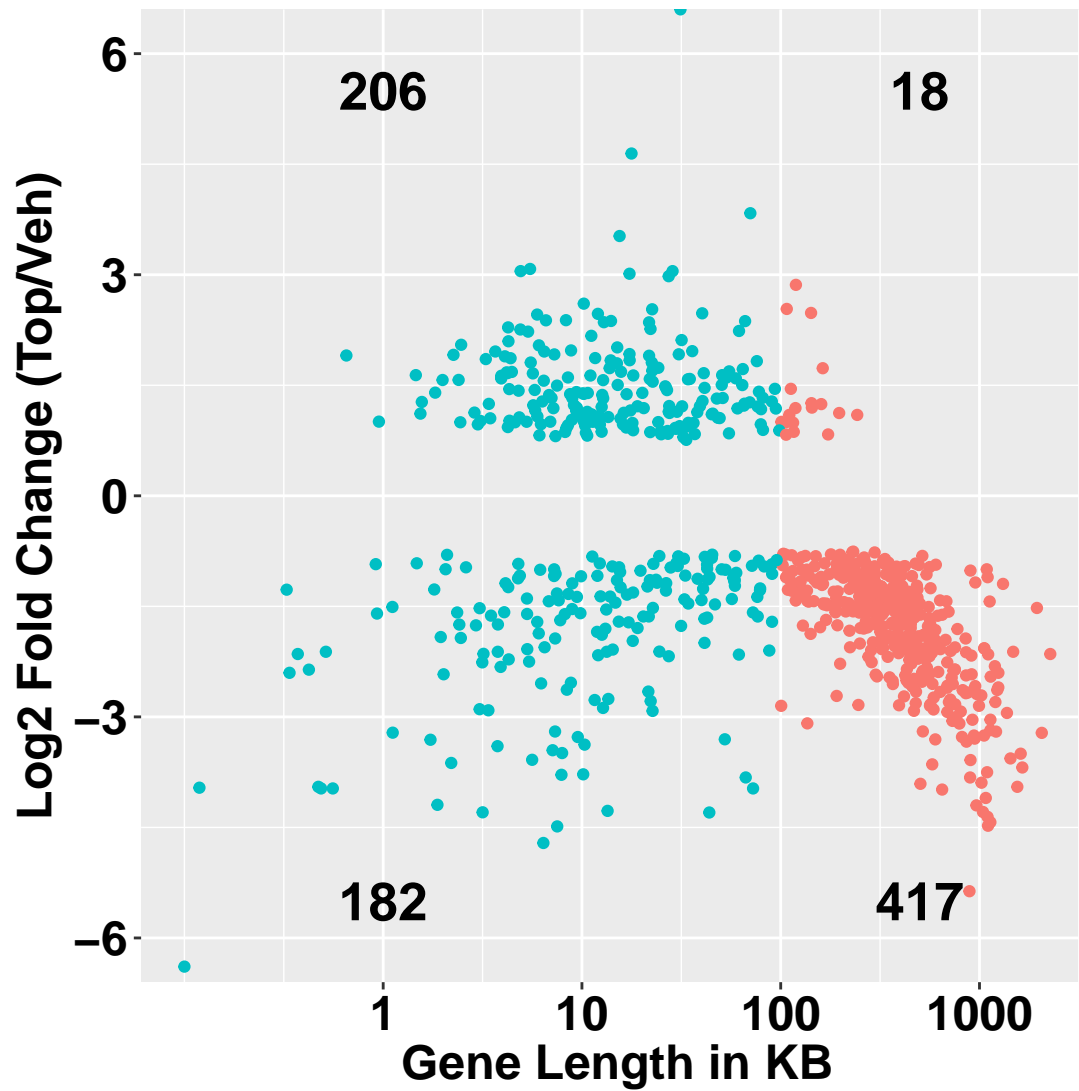
```
##
```

```
##
```

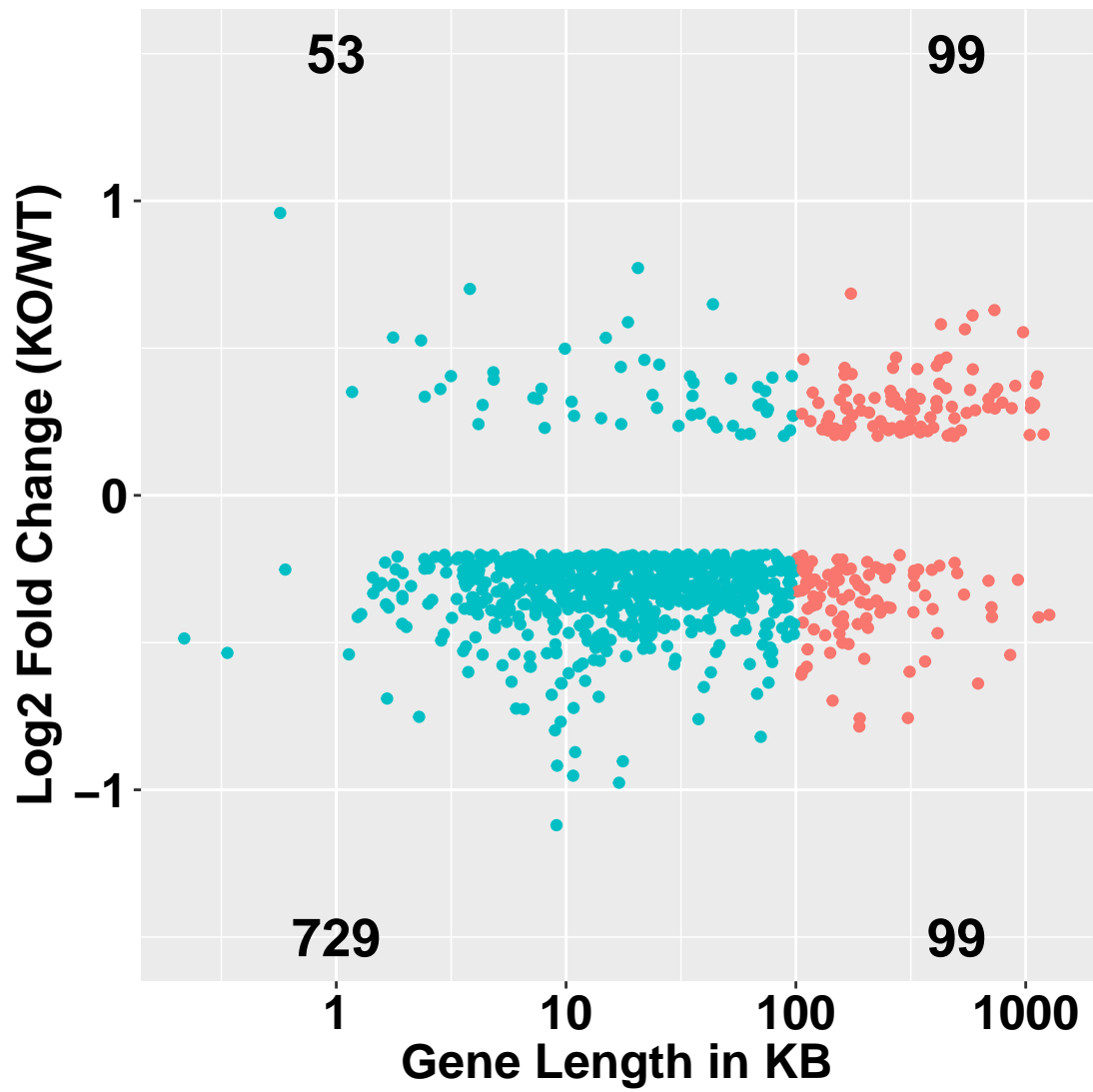
```
## Cortical Neurons (King et al.)
```



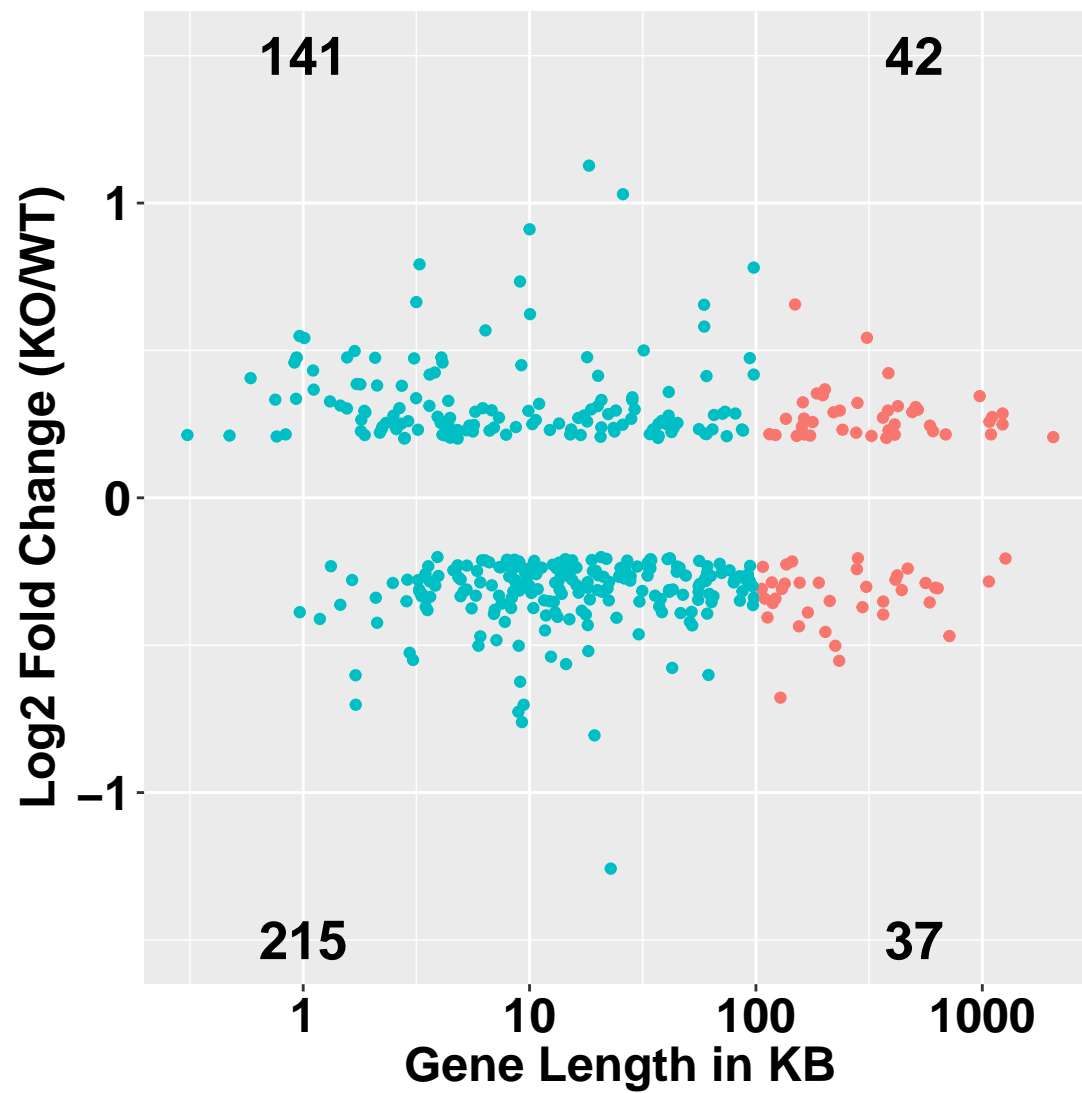
```
##
##
## Cortical Neurons (Mabb et al.)
```



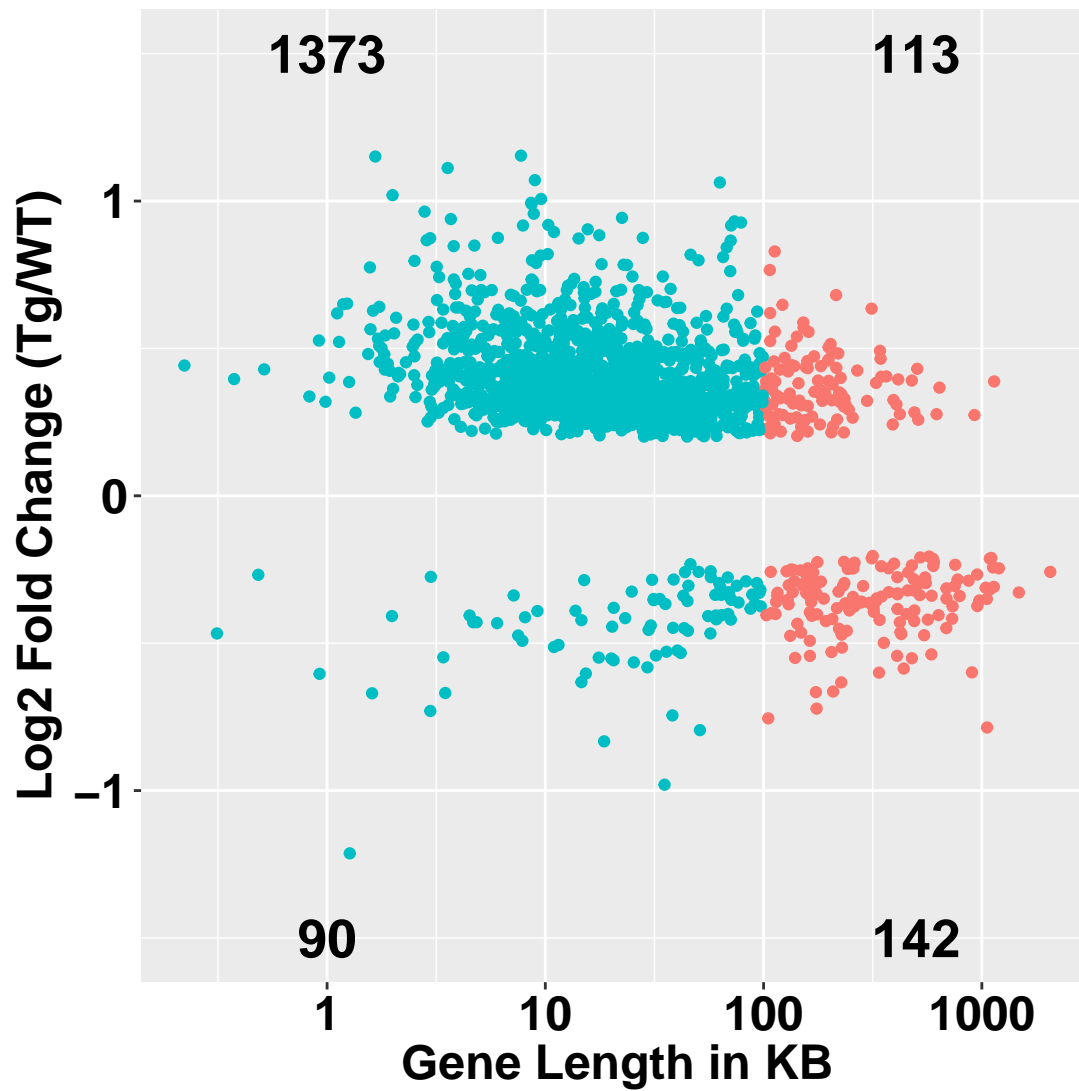
```
##
##
## Fig. 3(B) -- Mecp2 related Array (KO/WT)
##
##
##
## Hypothalamus (KO/WT)
```



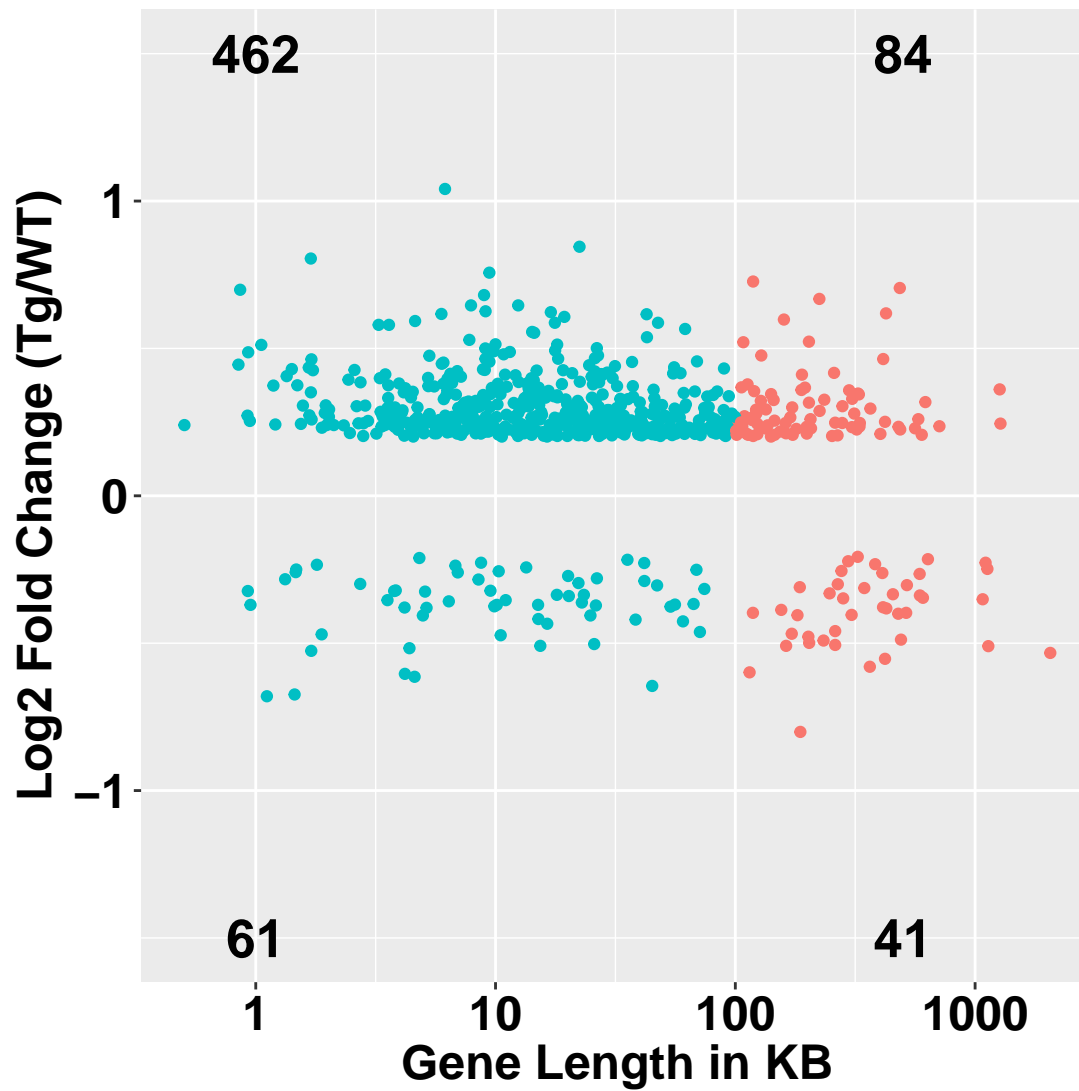
##  
##  
## Cerebellum (KO/WT)



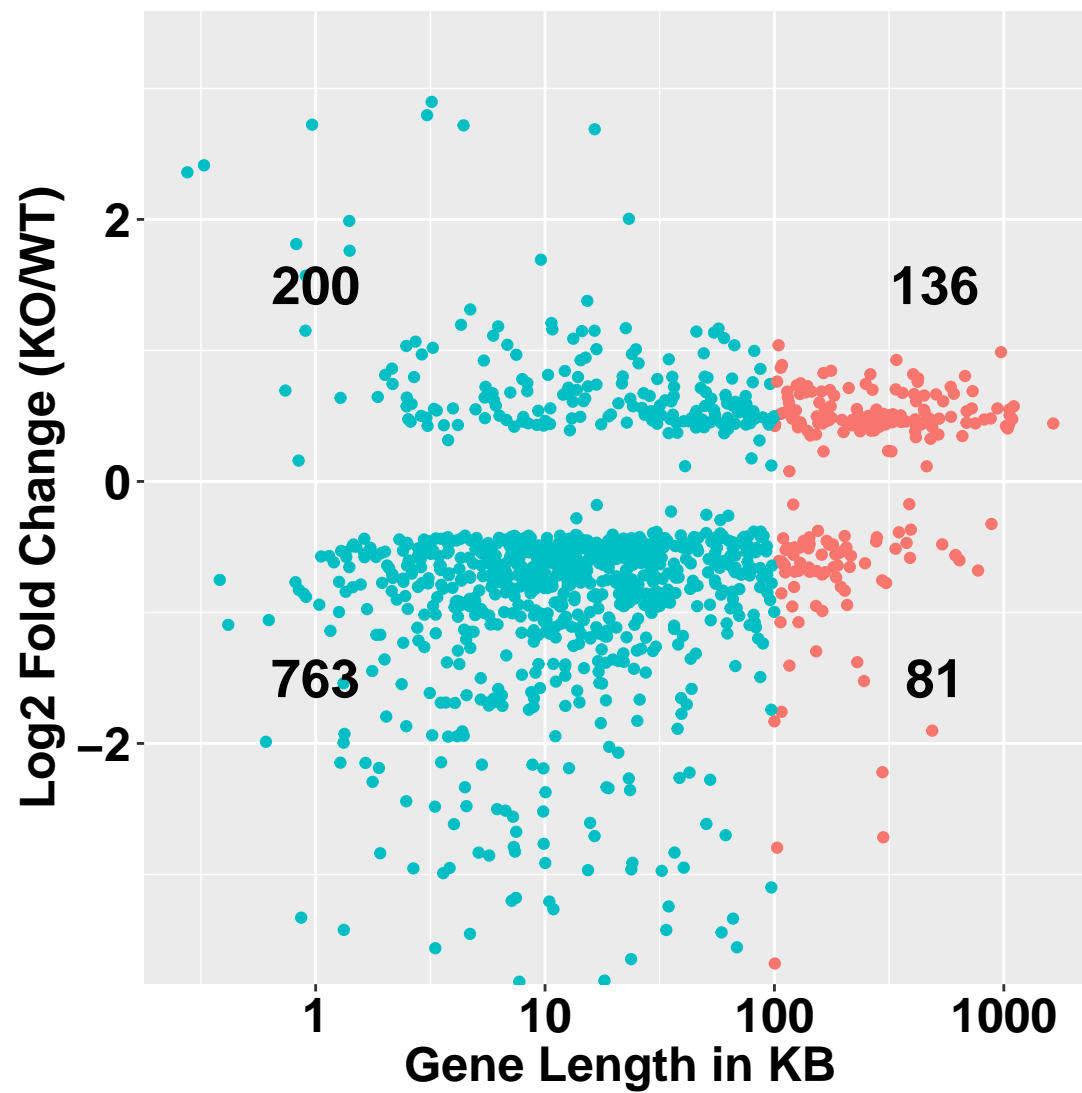
```
##
##
## Fig. 3(C) -- Mecp2 related Array (Tg/WT)
##
##
##
## Hypothalamus (Tg/WT)
```



##  
##  
## Cerebellum (Tg/WT)

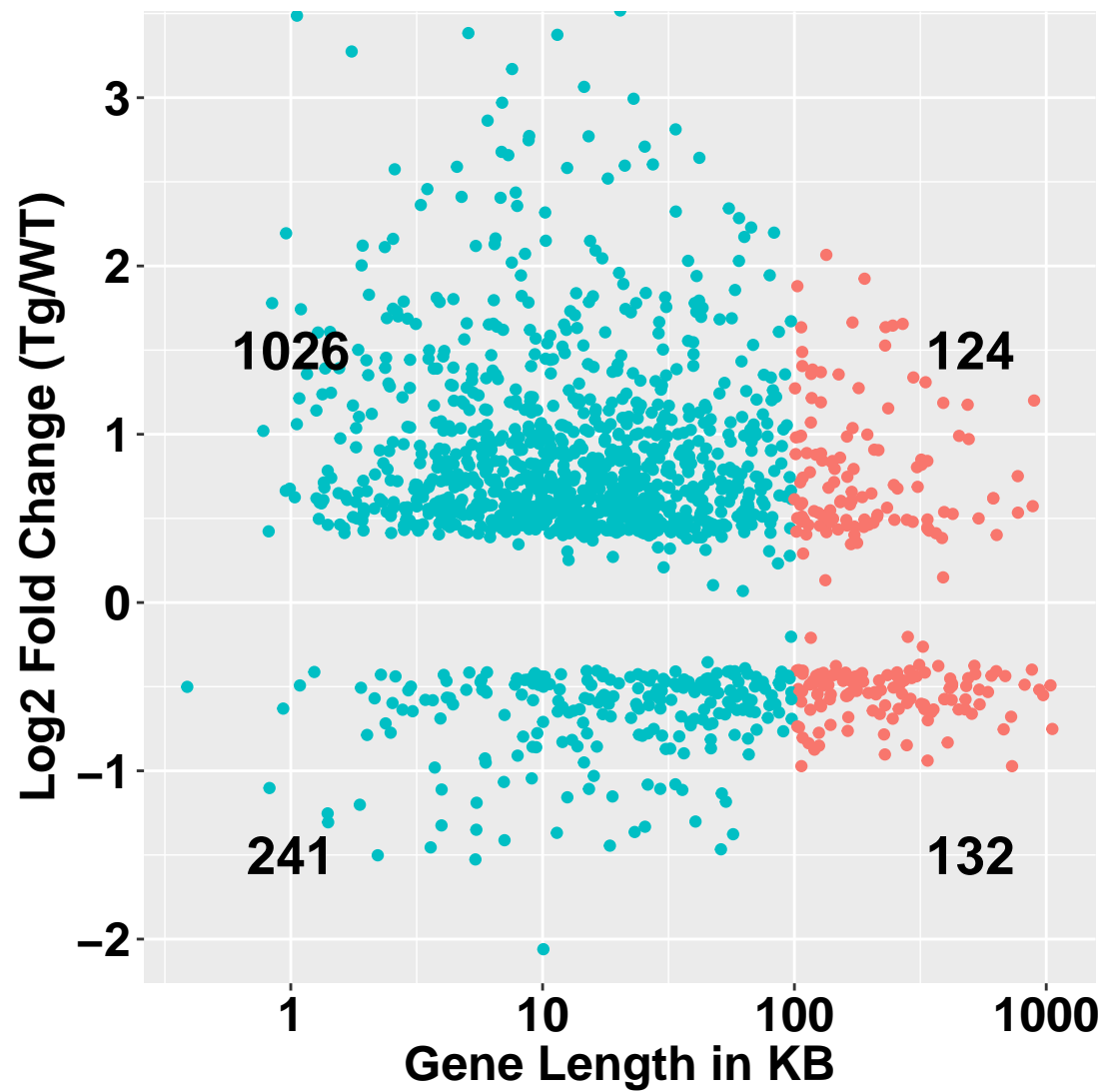


```
##
##
## Fig. 3(D) -- Mecp2 related RNA-seq Datasets
##
##
##
## Hypothalamus (KO/WT)
```



##  
##  
## Hypothalamus (Tg/WT)



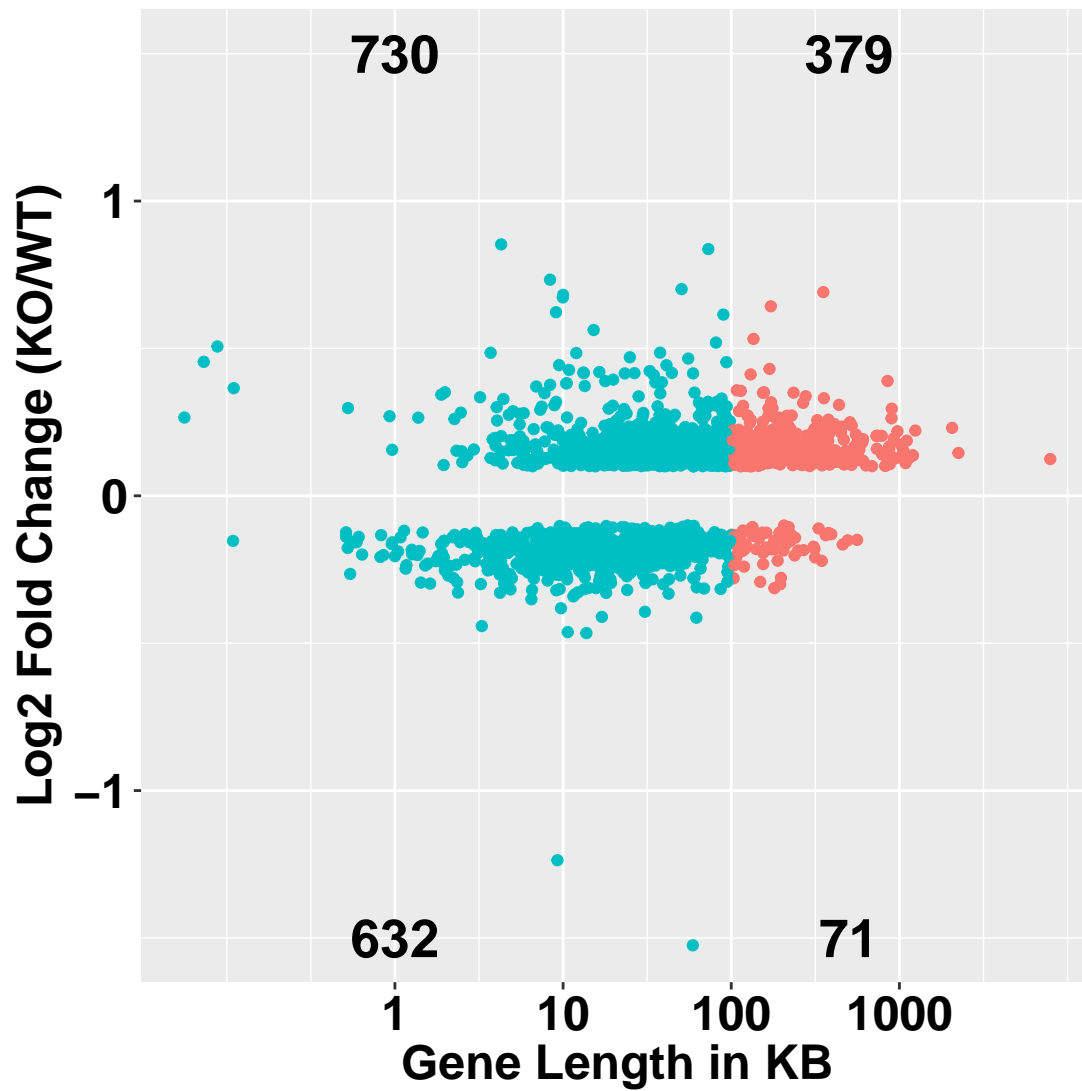


```
cat("\n\n Printing Supplementary Fig. 5 \n\n")
```

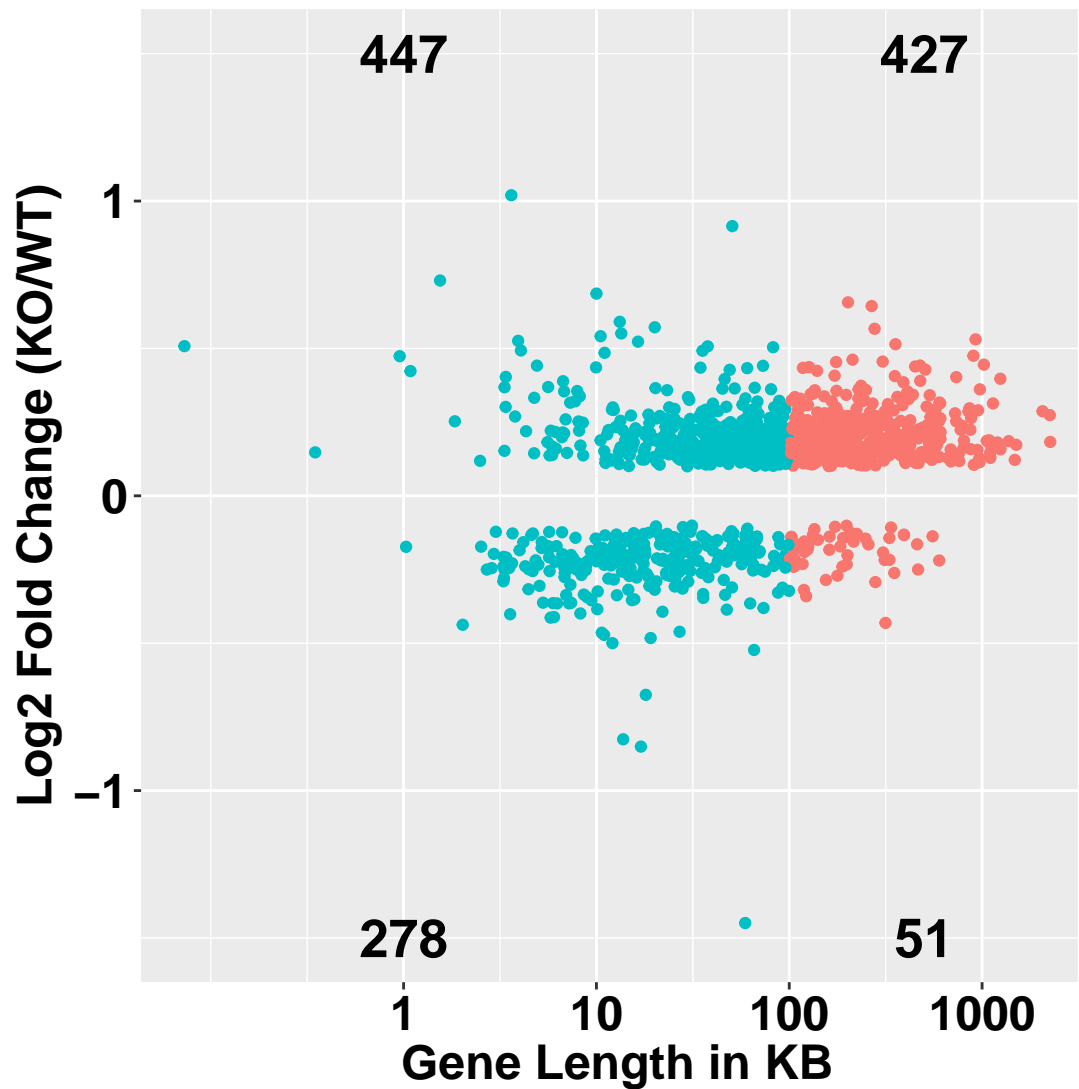
```
##
##
## Printing Supplementary Fig. 5
```

```
figureS5()
```

```
##
##
## Supplementary Fig. 5
##
##
## Hippocampus 4 weeks (KO/WT)
```



```
##  
##  
## Hippocampus 9 weeks (KO/WT)
```



Section 6: Figure 4 and Supplementary Fig. 6-9

```
cat("\n\n Printing Fig. 4 \n\n")
```

```
##
```

```
##
```

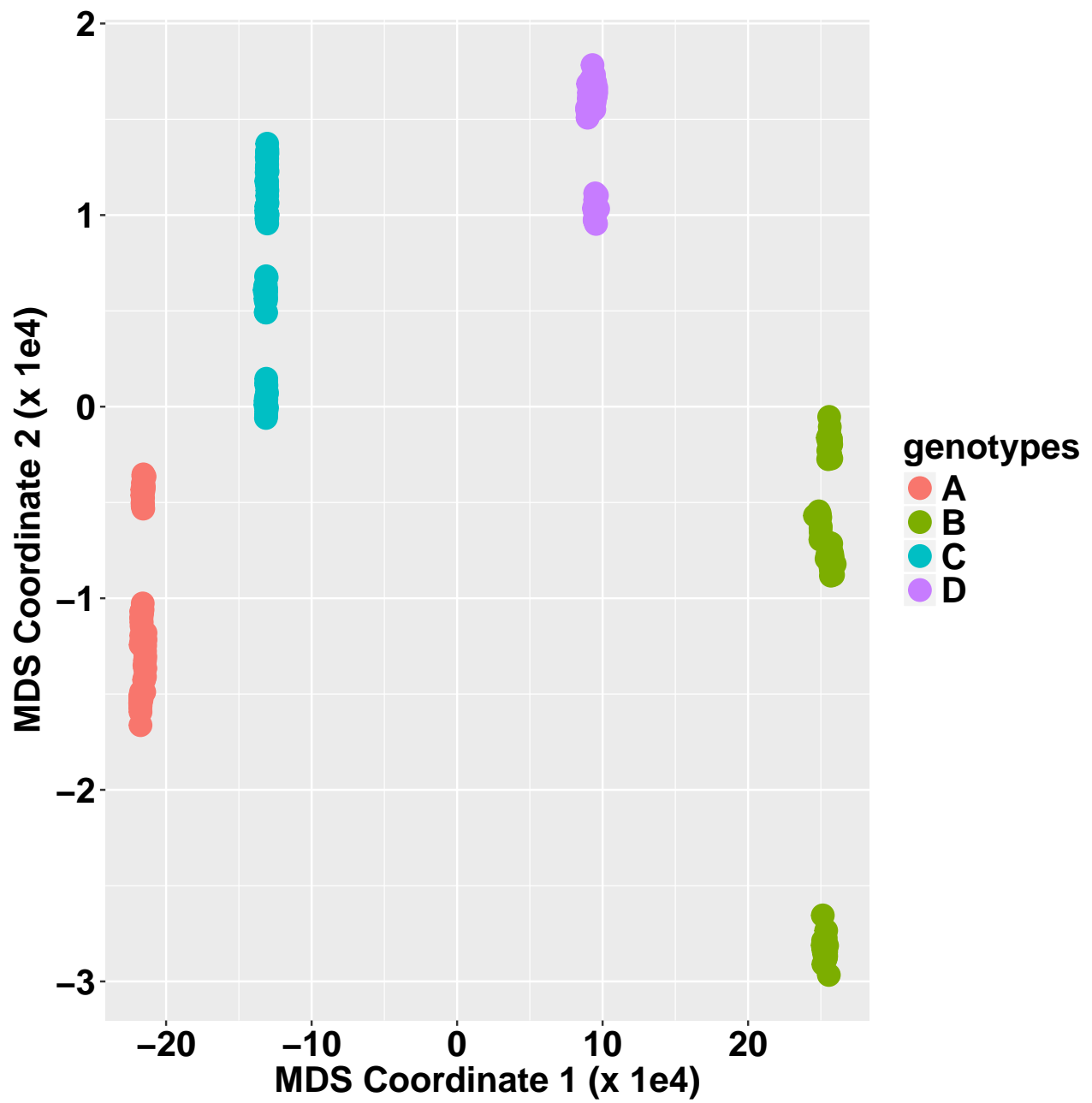
```
## Printing Fig. 4
```

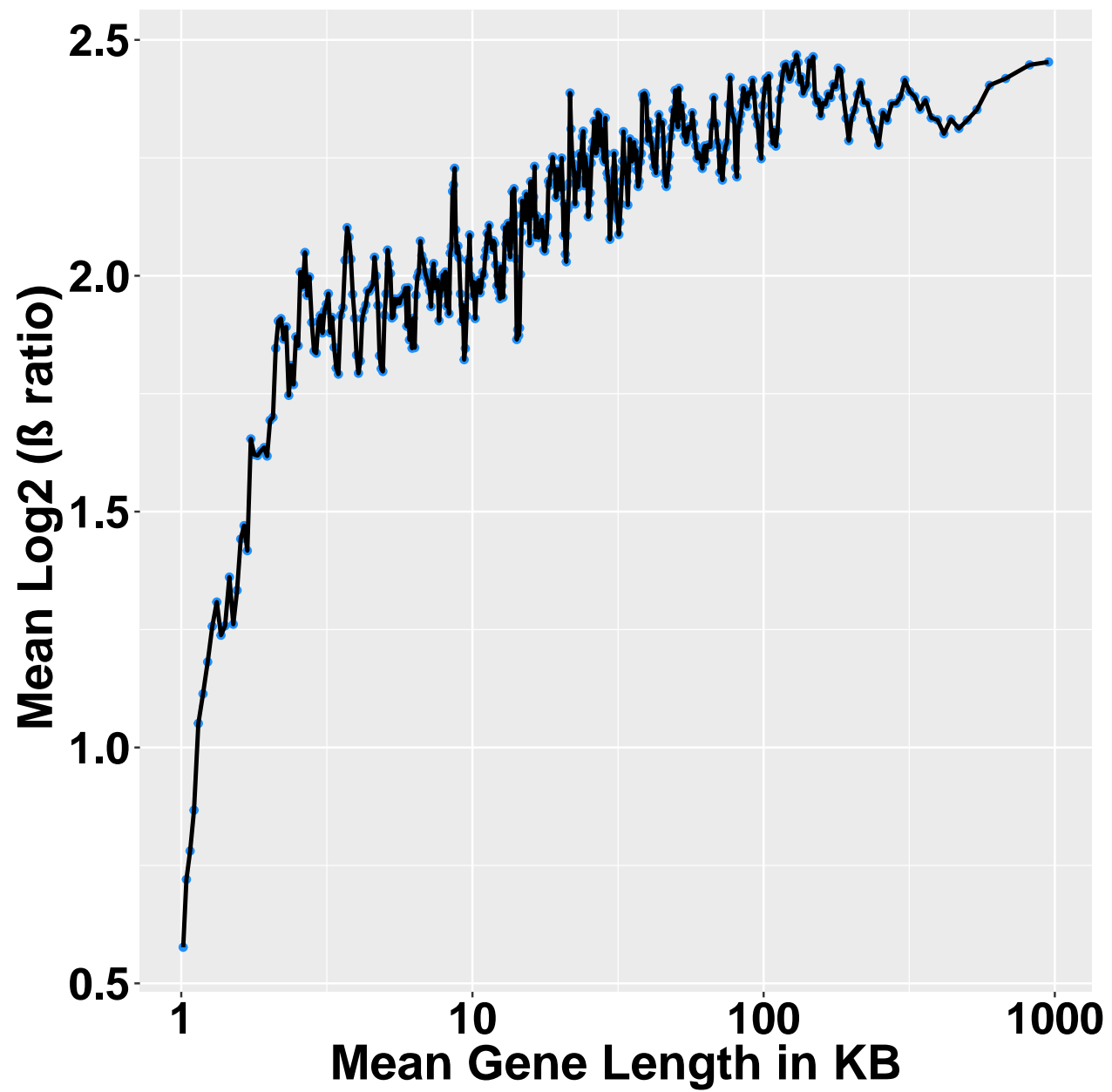
```
figure4()
```

```
##
```

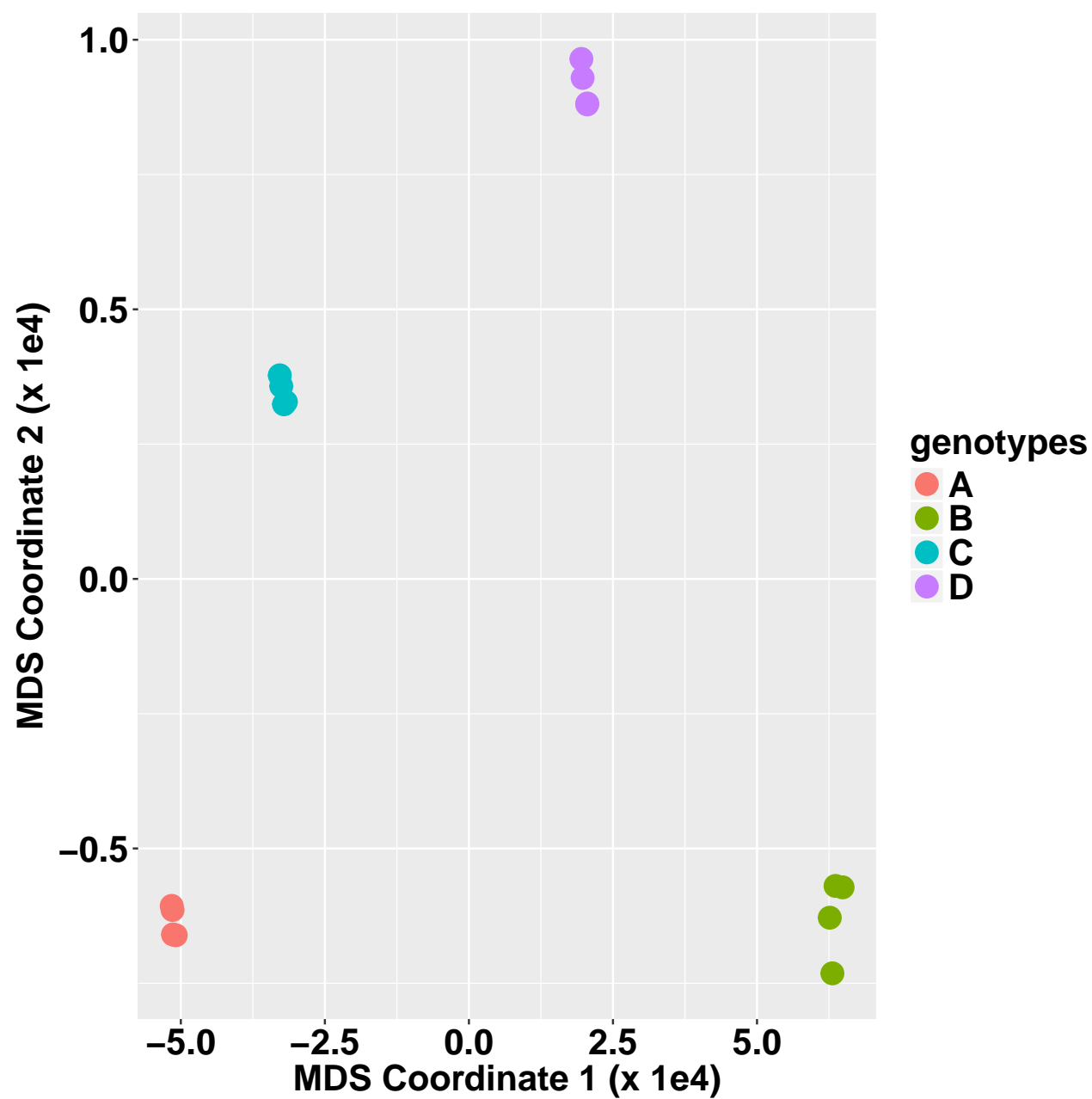
```
##
```

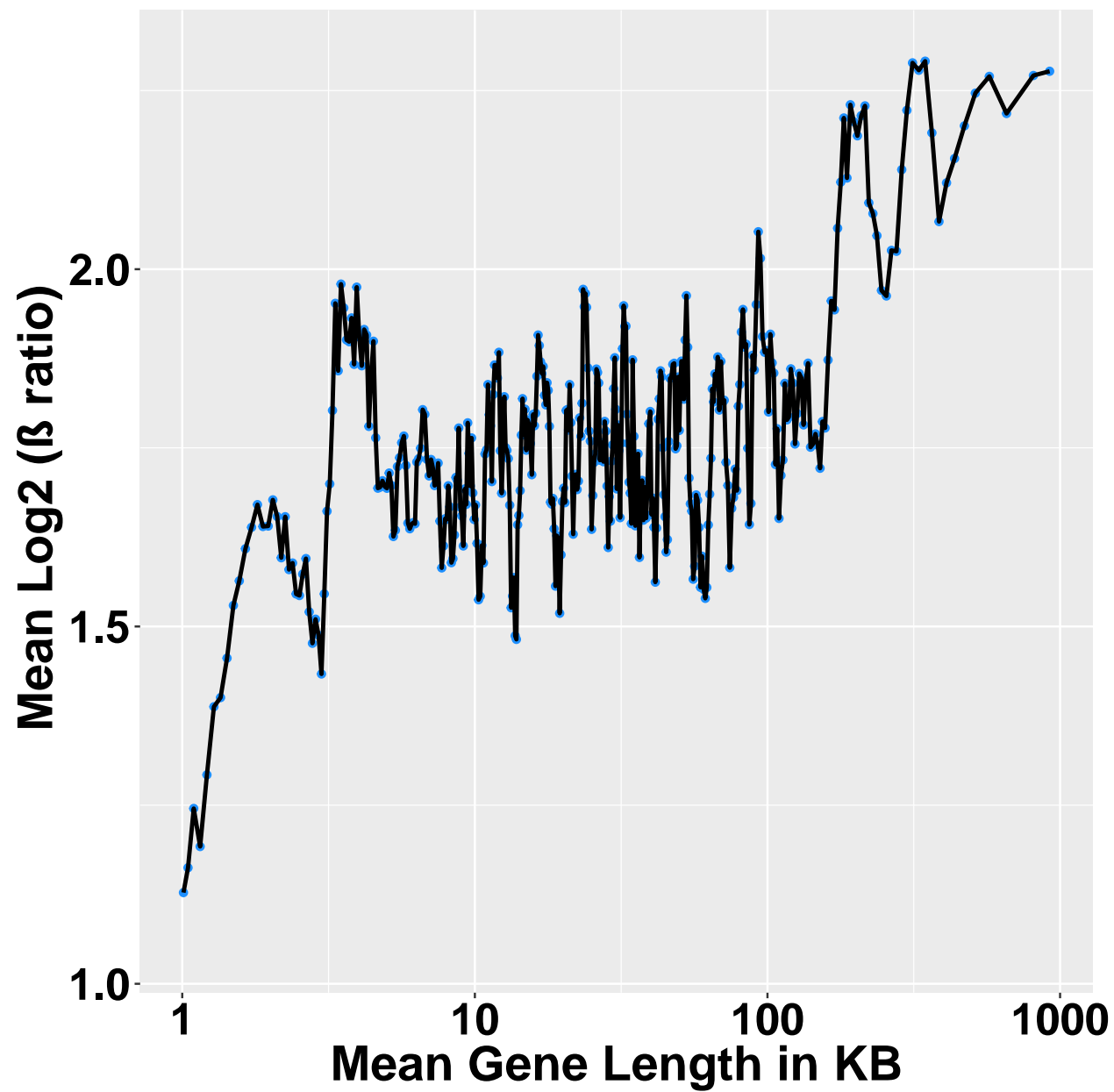
```
## Fig. 4(A) -- SEQC RNA-seq
```



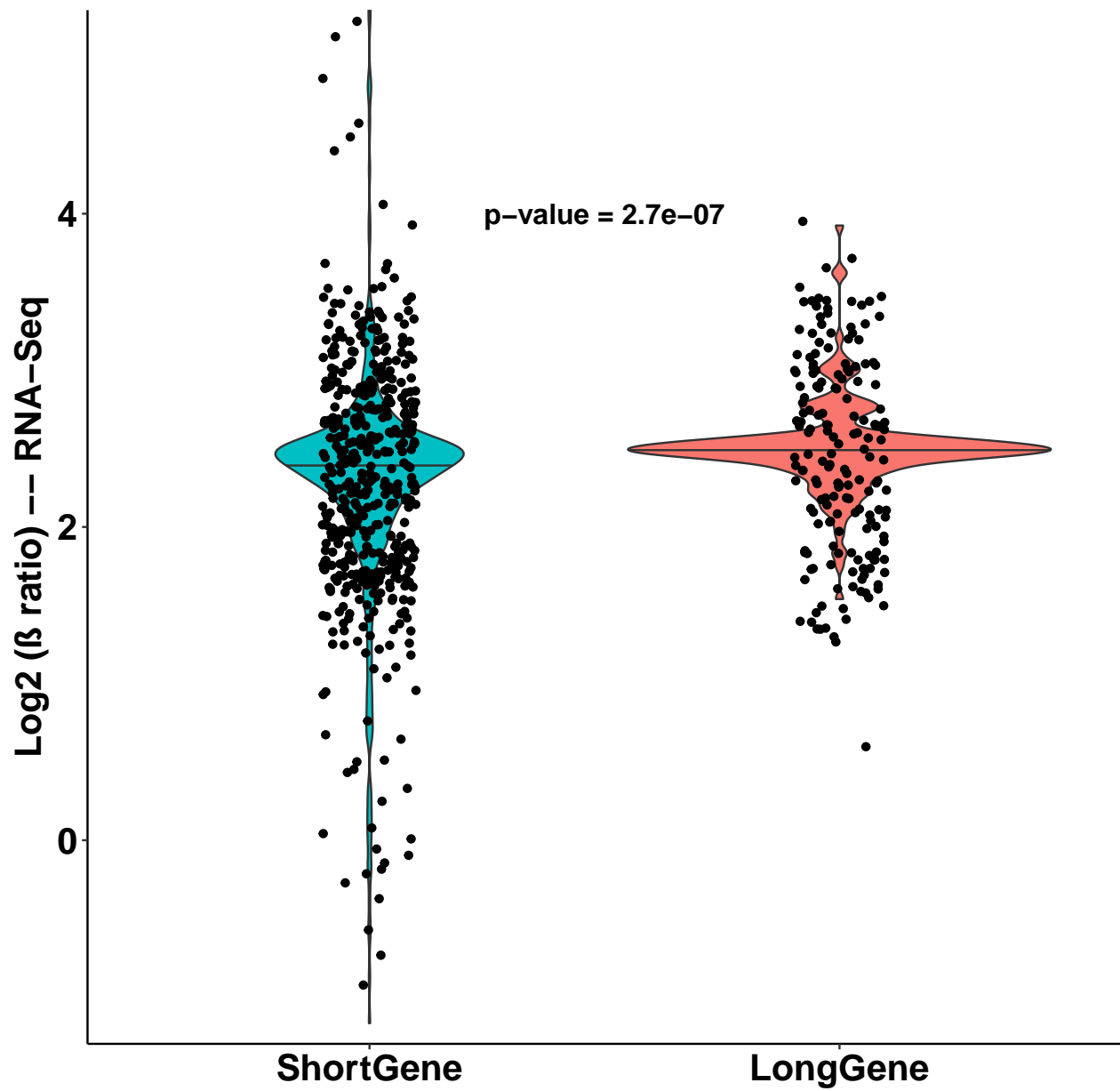


```
##  
##  
## Fig. 4(B) -- SEQC Array
```



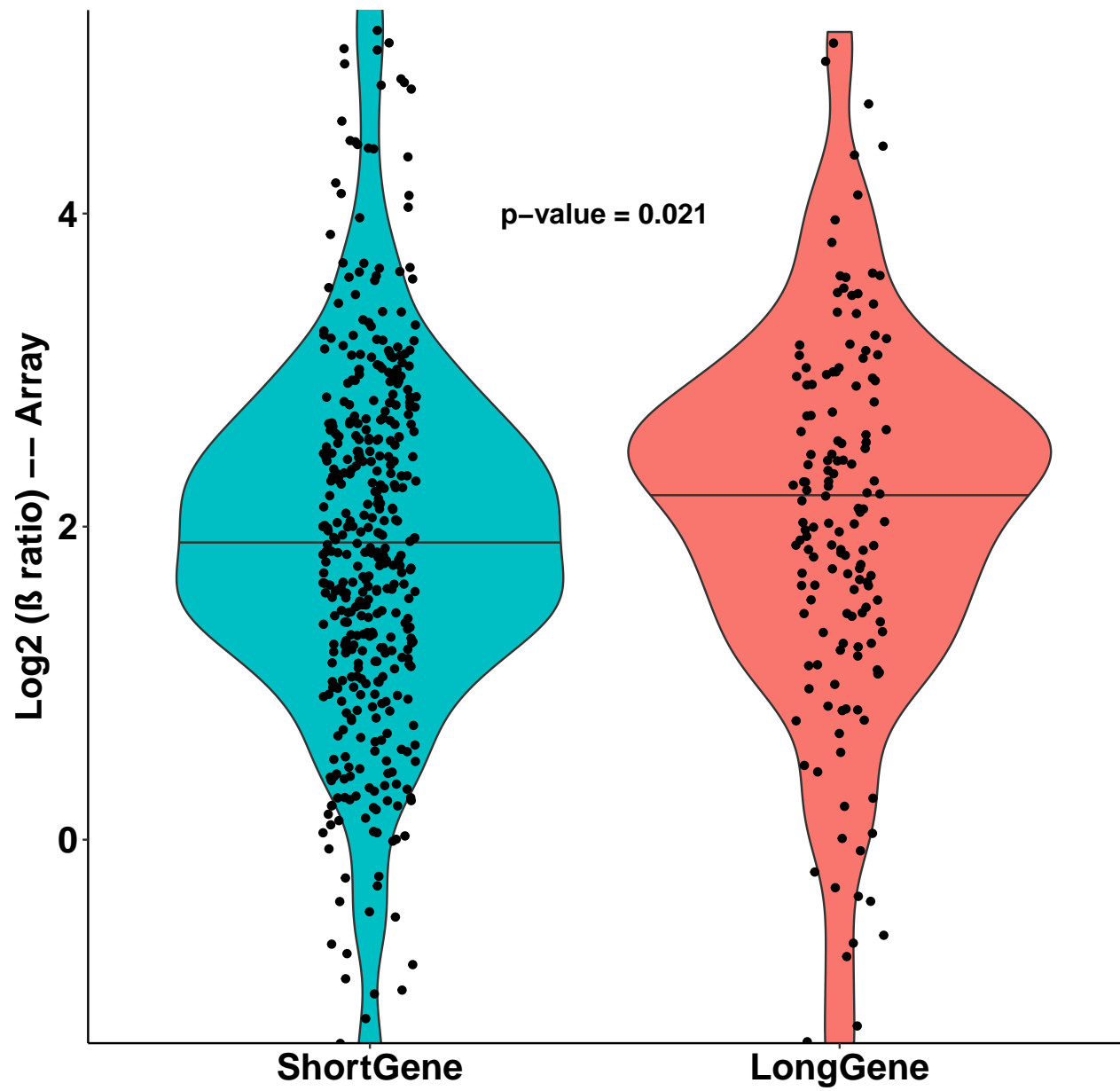


##  
##  
## Fig. 4(E) -- RNA-seq

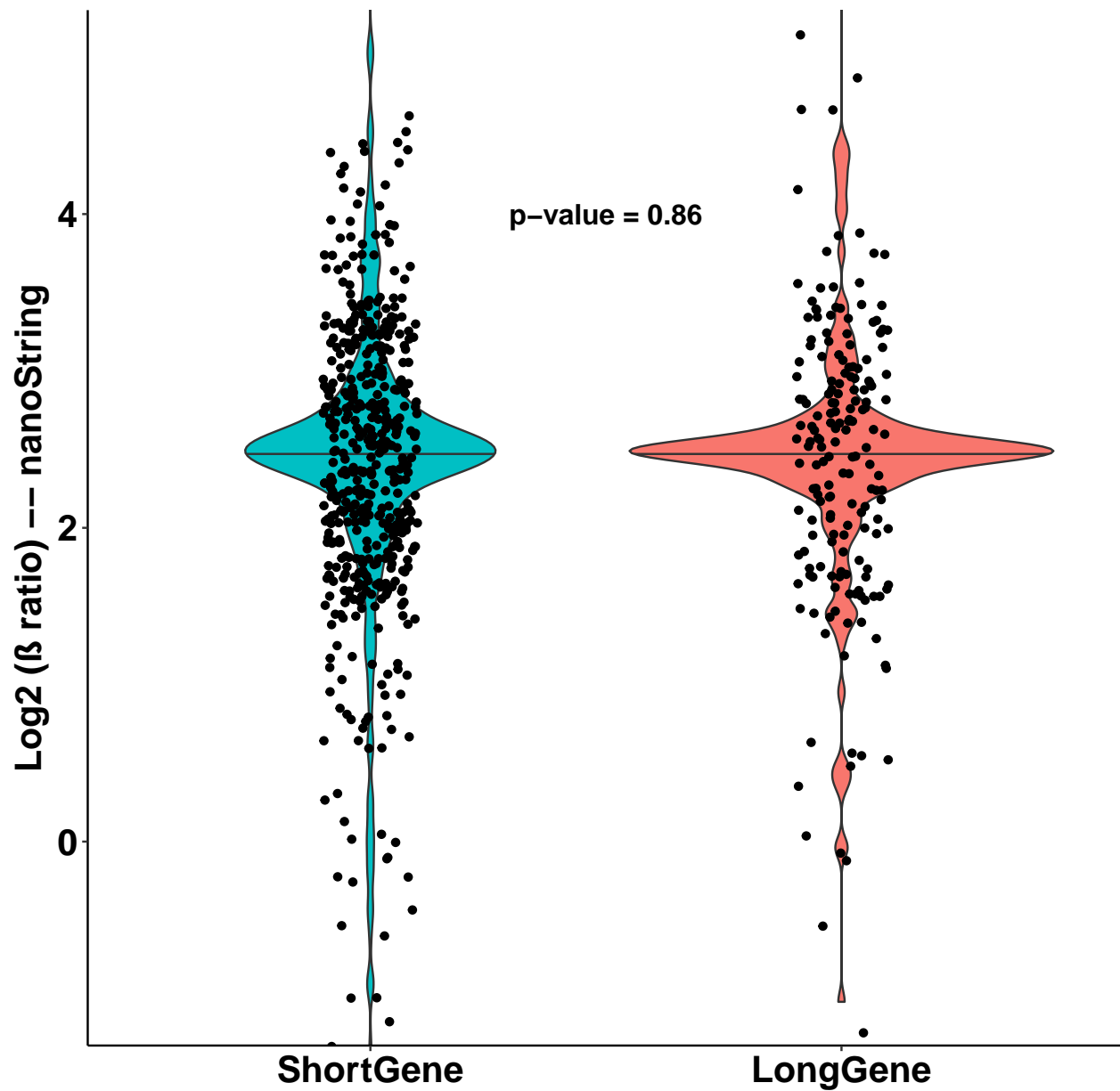


##  
 ##  
 ## Fig. 4(F) -- Microarray





##  
##  
## Fig. 4(F) -- Nanostring

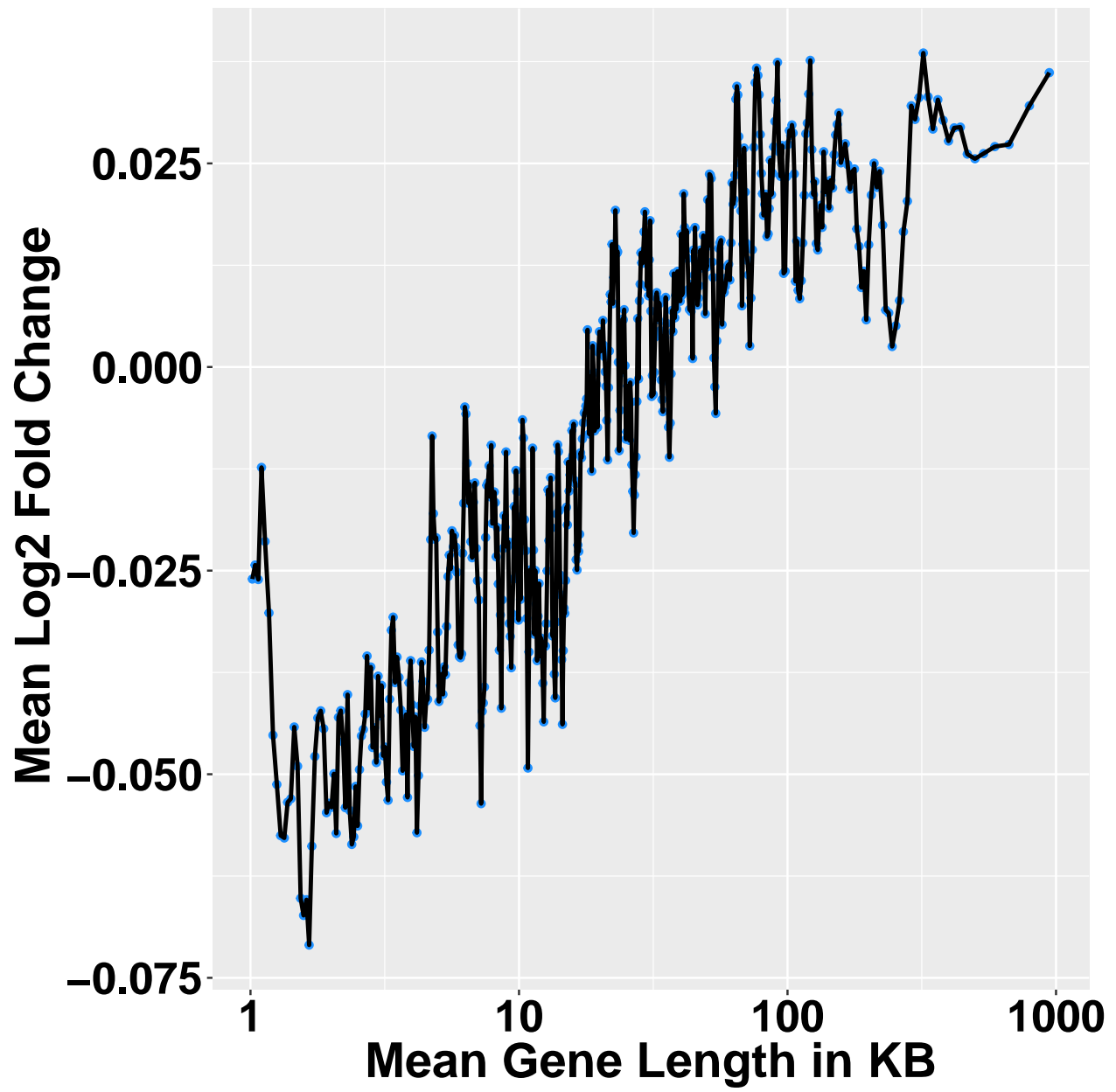


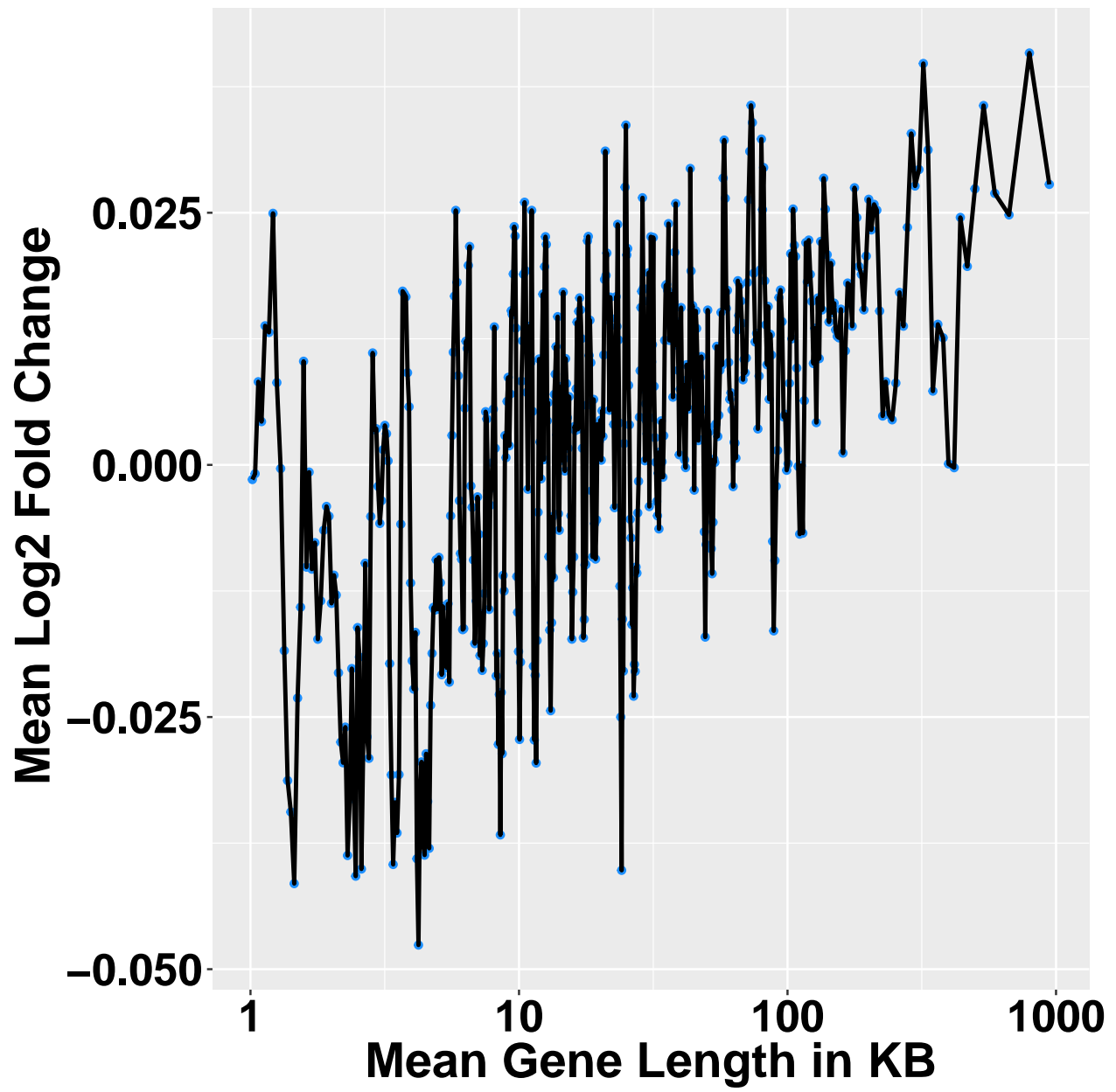
```
cat("\n\n Printing Supplementary Fig. 6 \n\n")
```

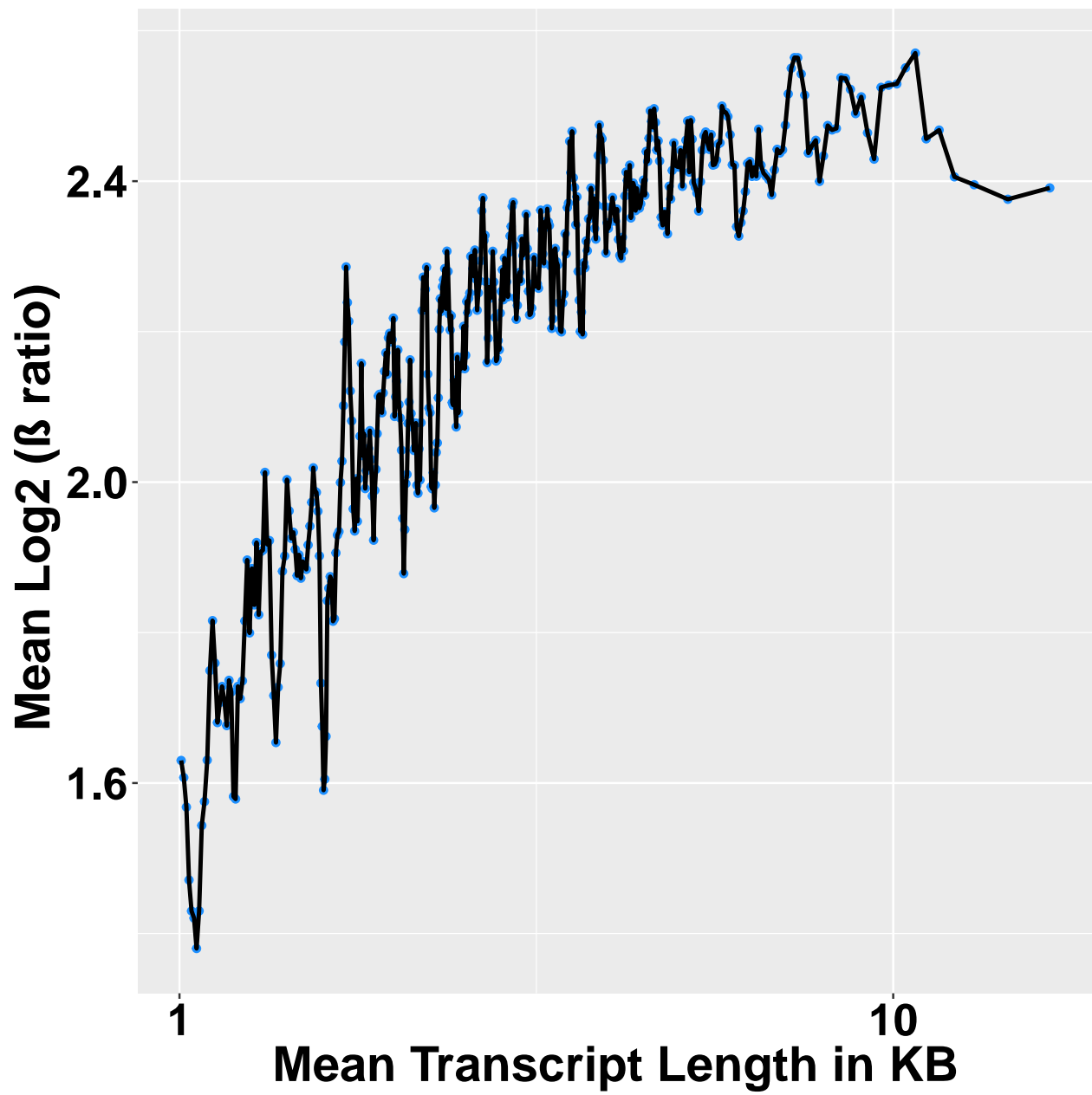
```
##
##
## Printing Supplementary Fig. 6
```

```
figureS6()
```

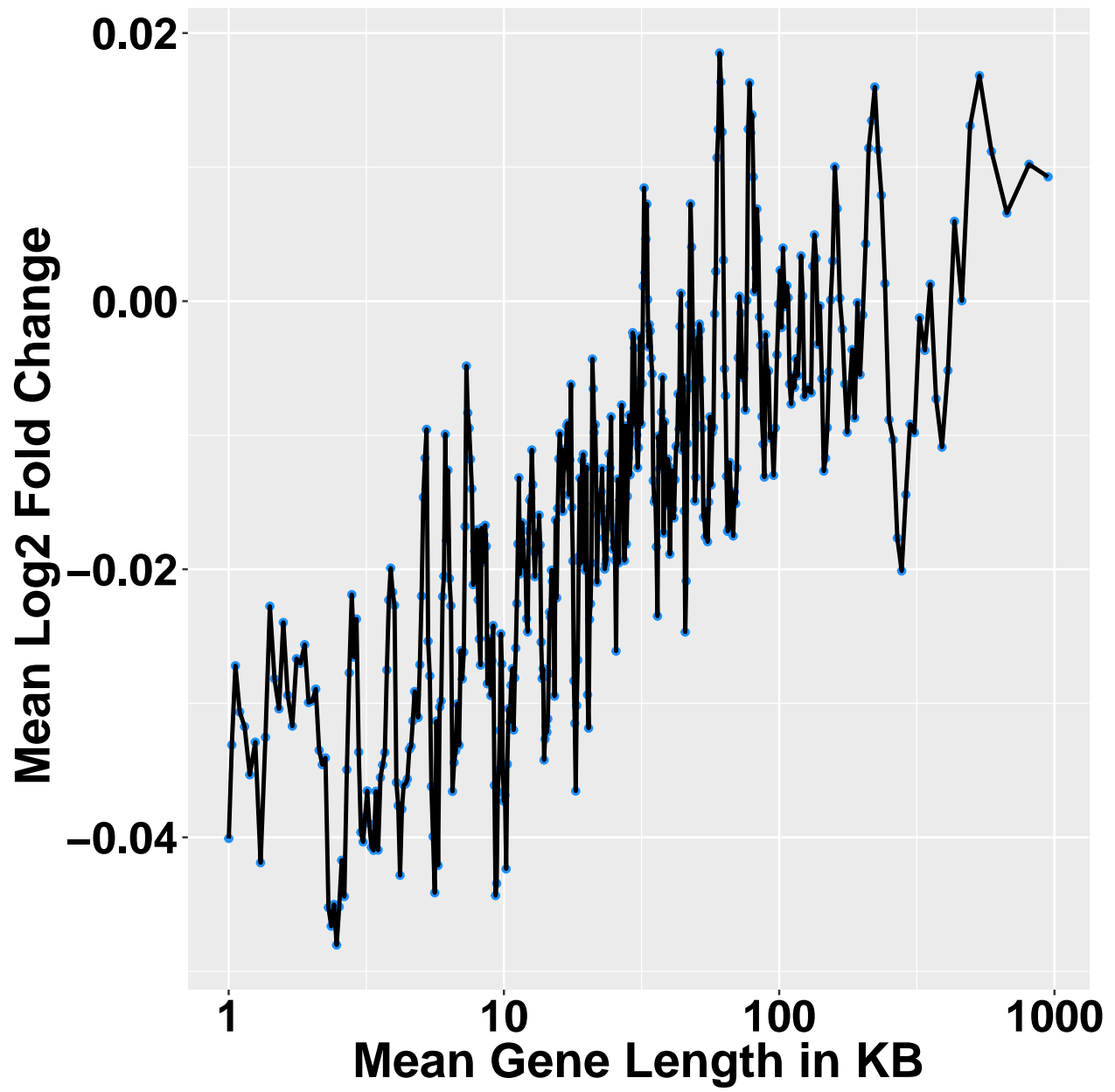
```
##
##
## Supplementary Fig. 6(A) -- SEQ RNA-seq
```

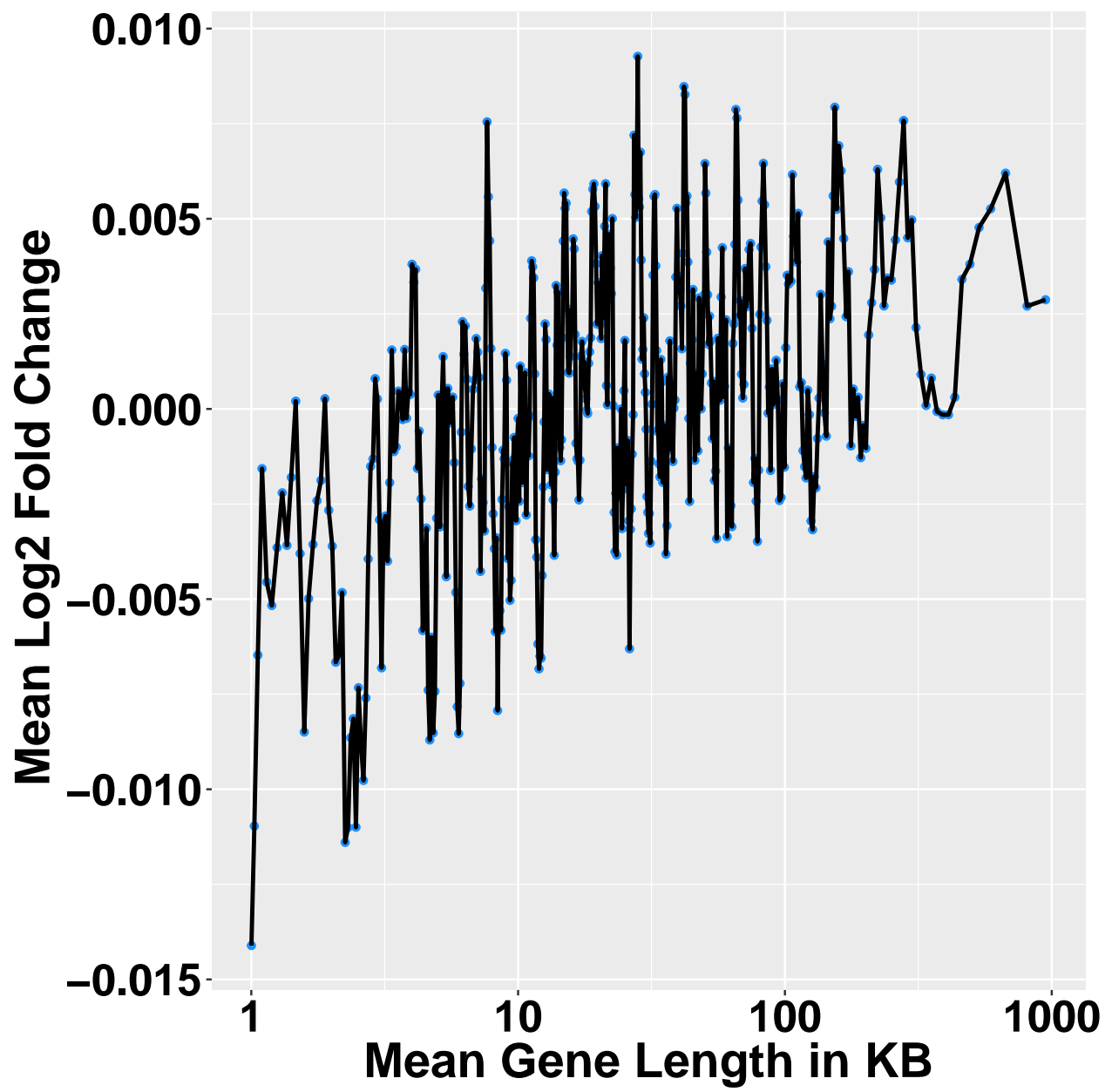


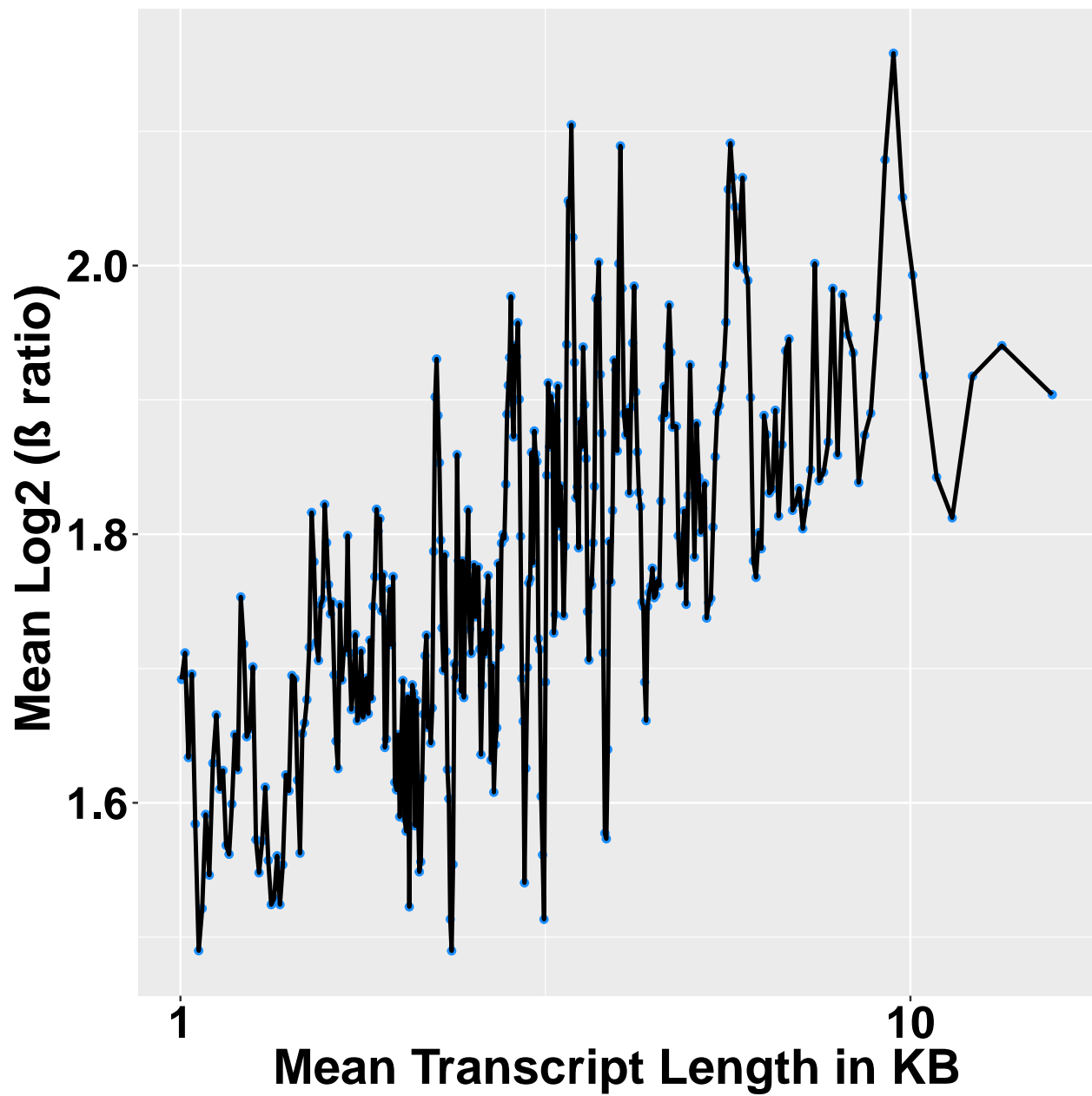




```
##  
##  
## Supplementary Fig. 6(B) -- SEQC Array
```







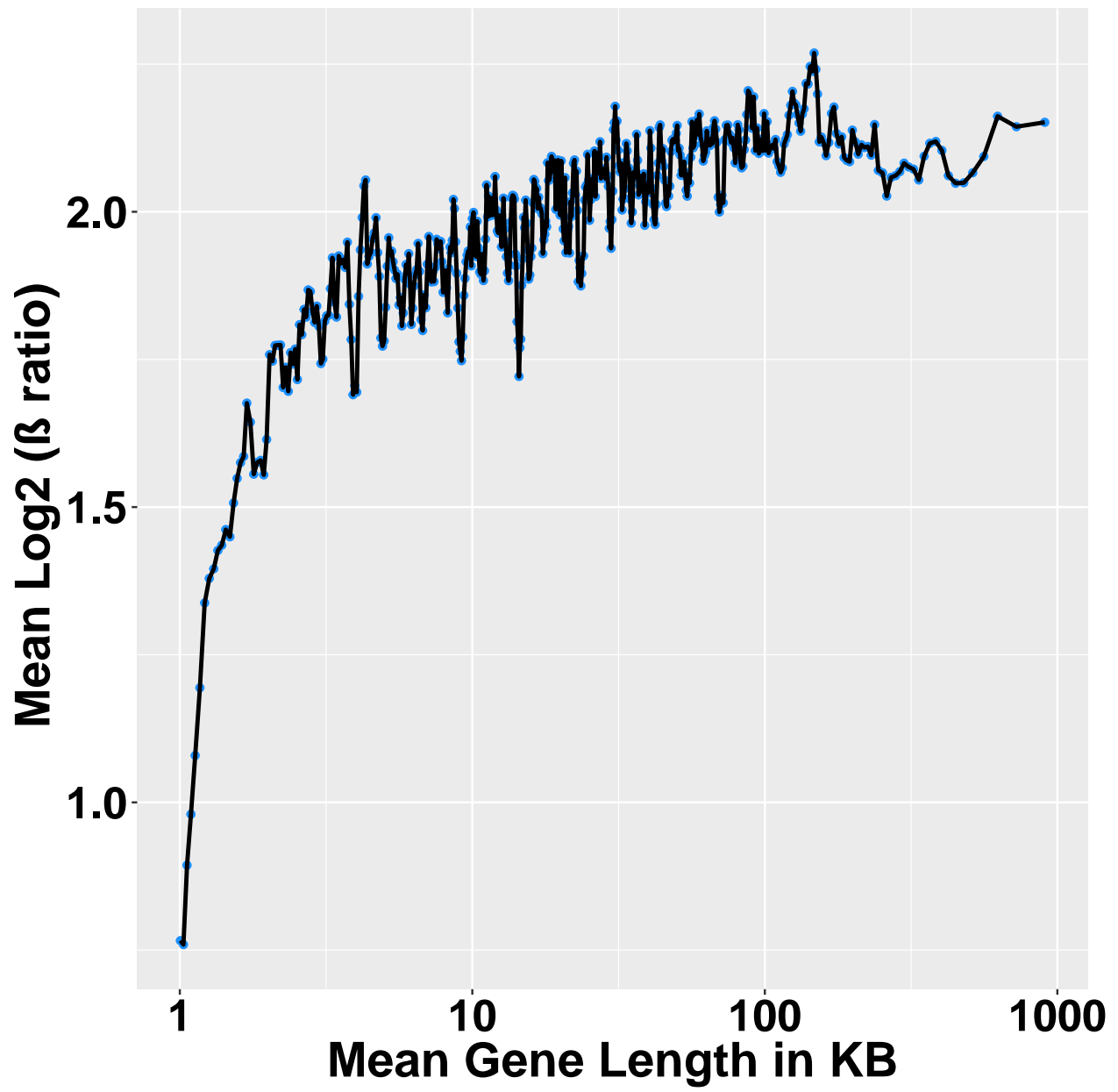
```
cat("\n\n Printing Supplementary Fig. 7 \n\n")
```

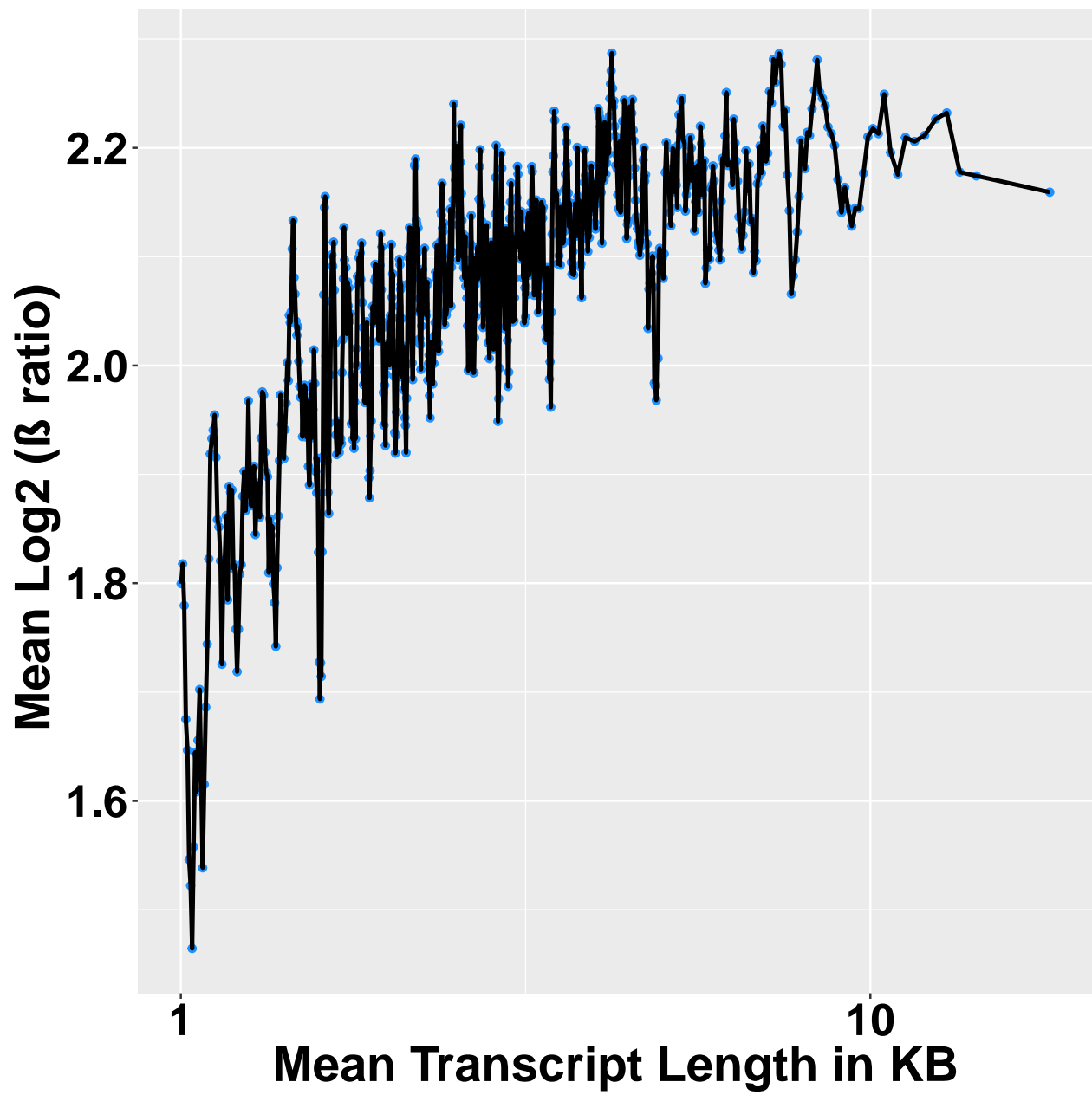
```
##
##
## Printing Supplementary Fig. 7
```

```
figureS7()
```

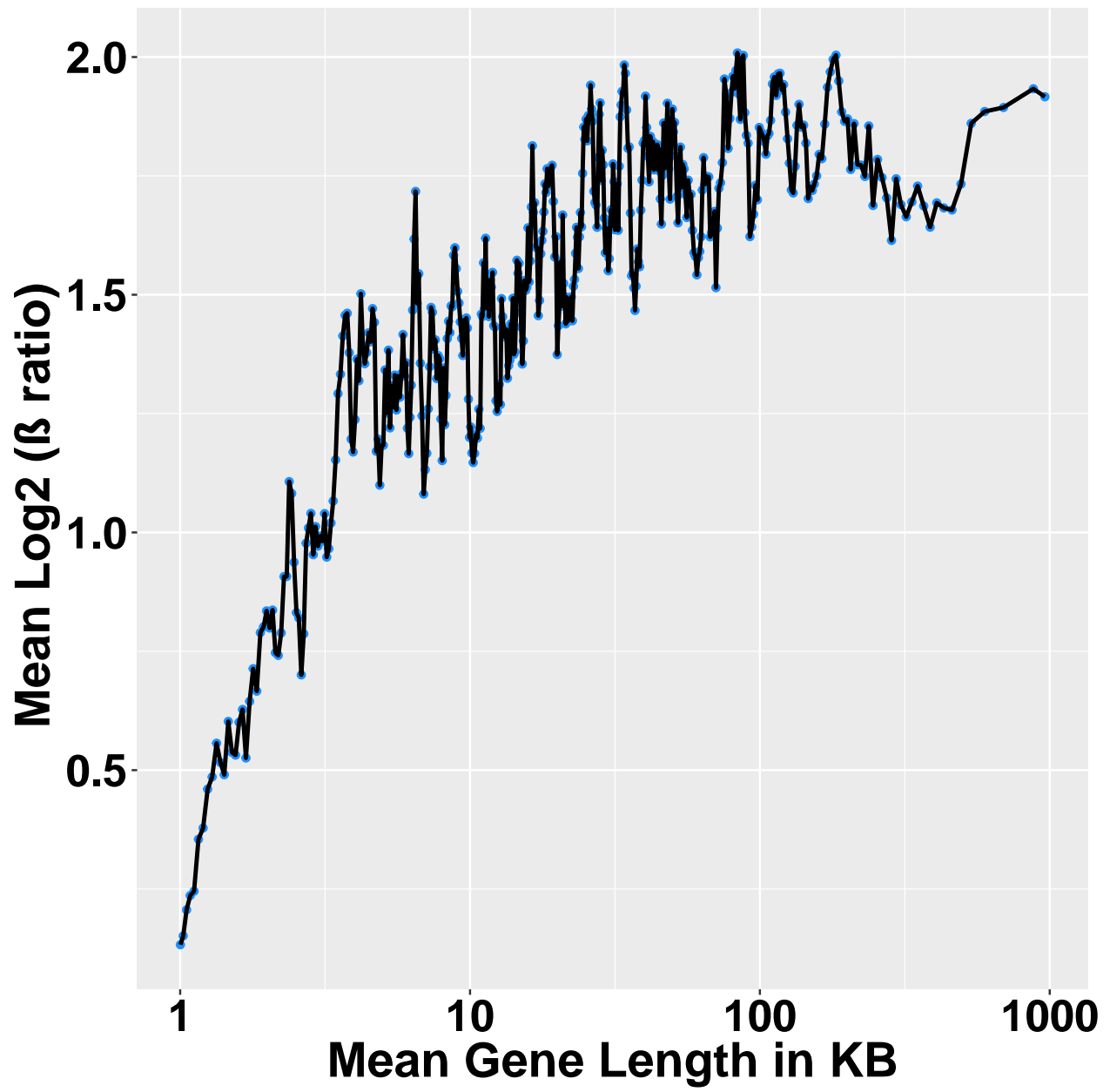
```
##
##
## Supplementary Fig. 7(A) -- Total Count
```

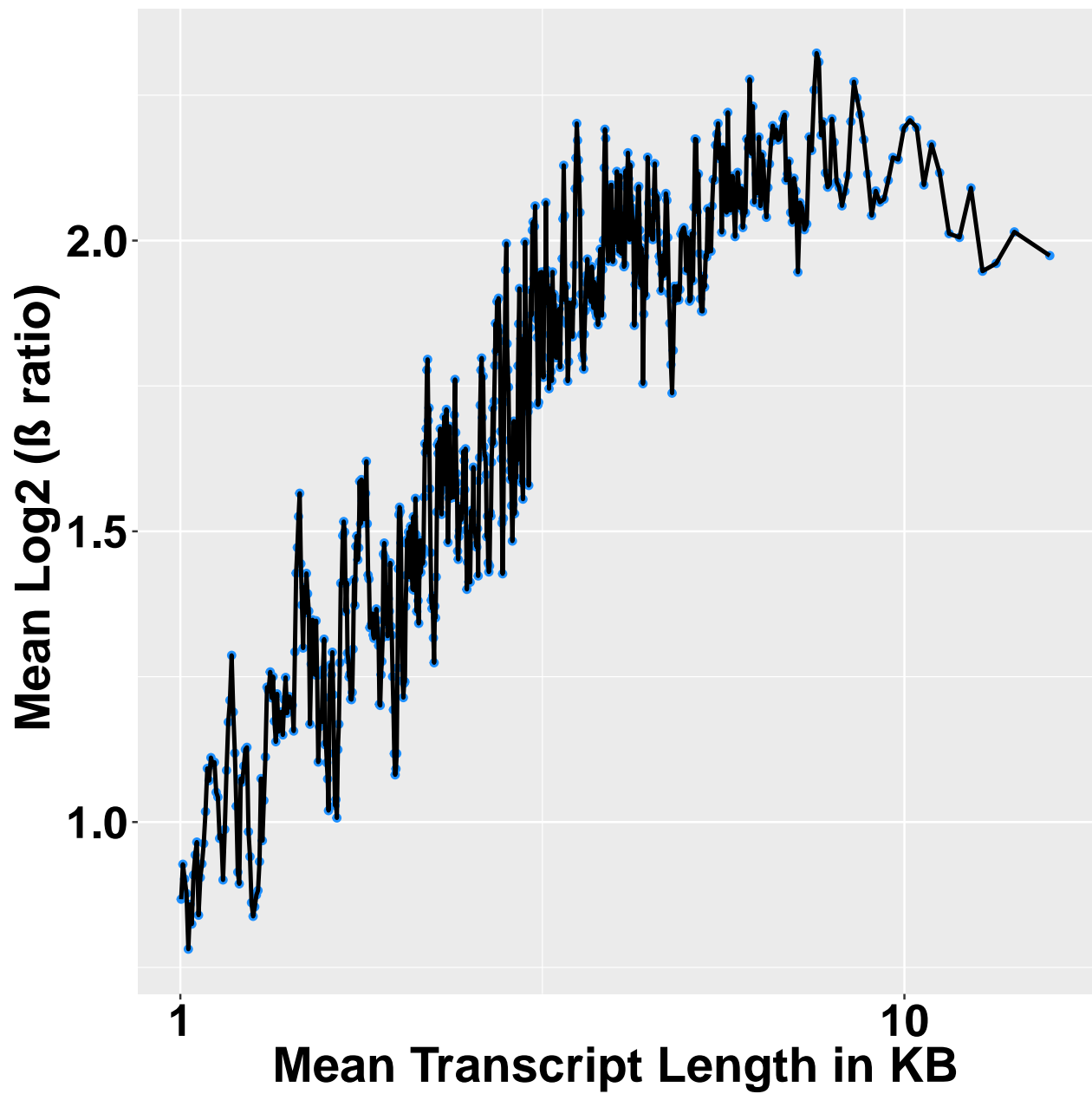






```
##  
##  
## Supplementary Fig. 7(B) -- TMM (edgeR)
```





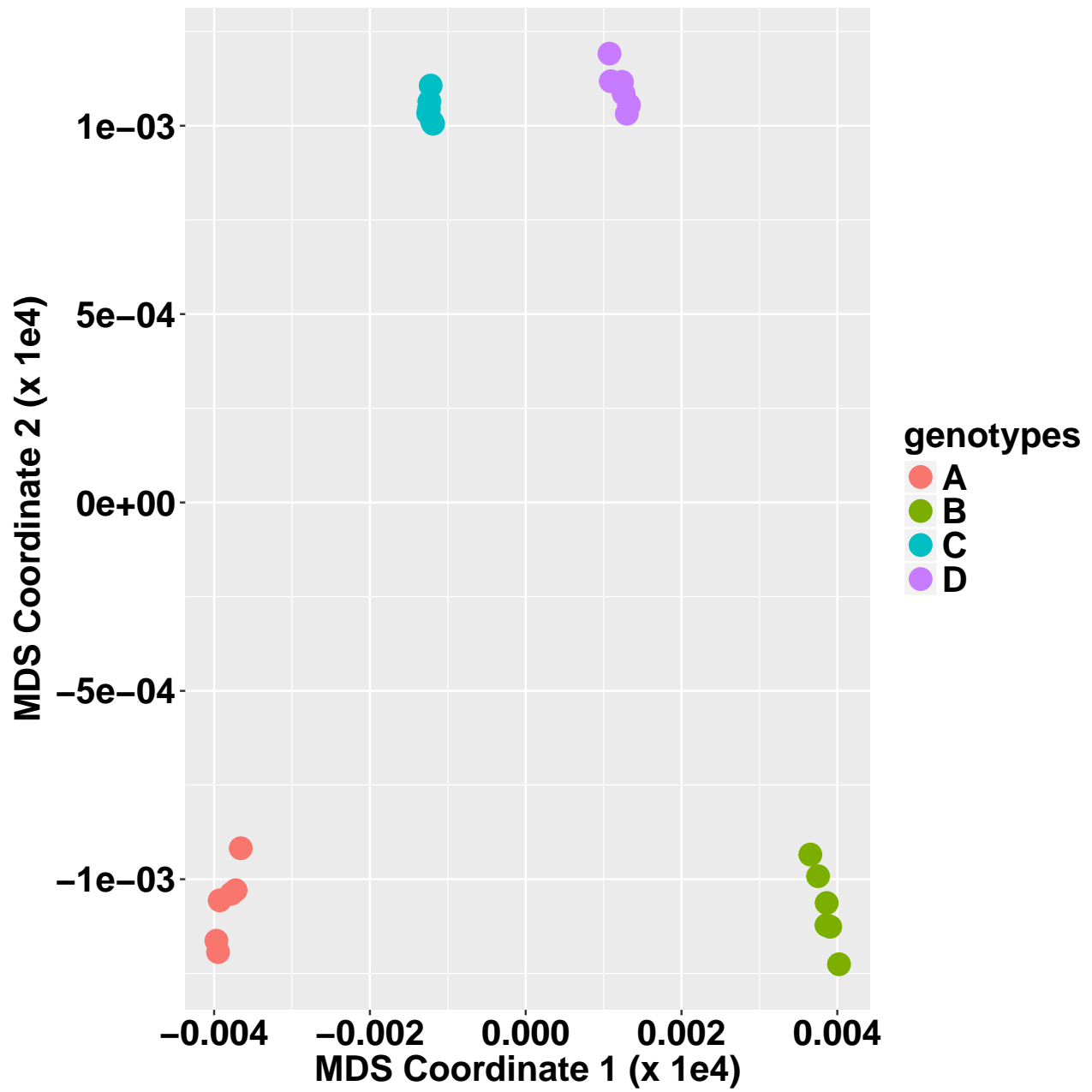
```
cat("\n\n Printing Supplementary Fig. 8 \n\n")
```

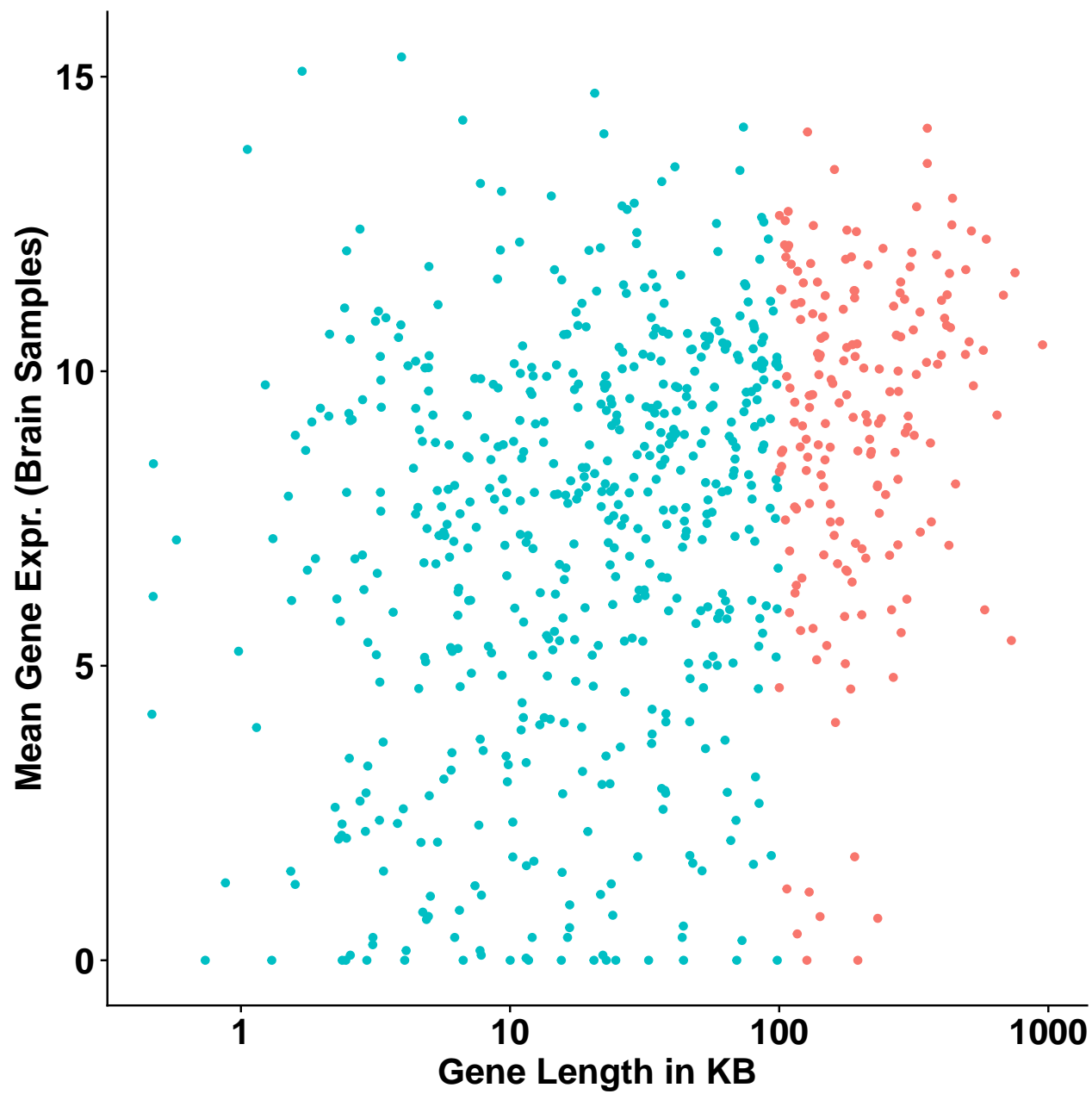
```
##
```

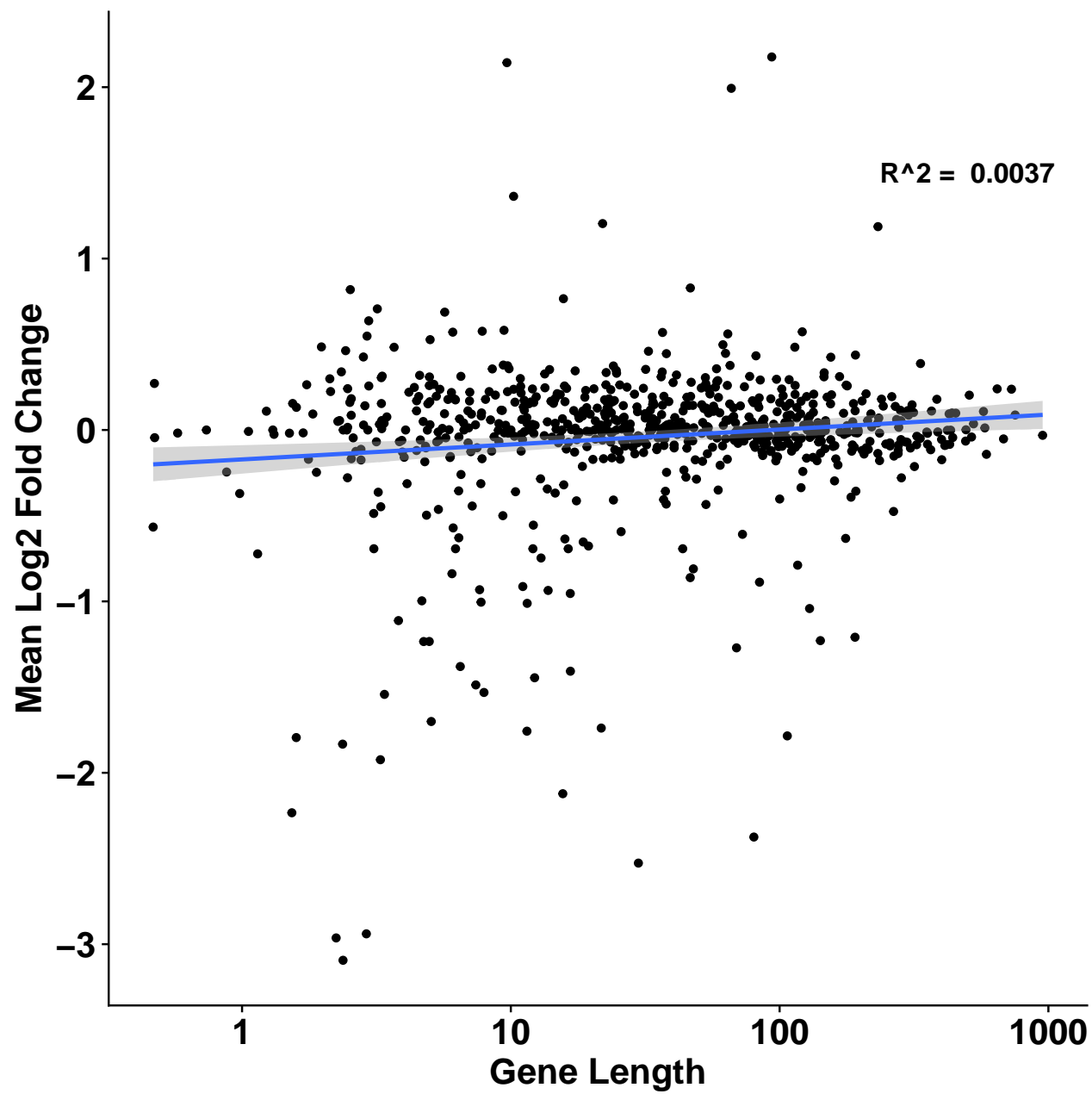
```
##
```

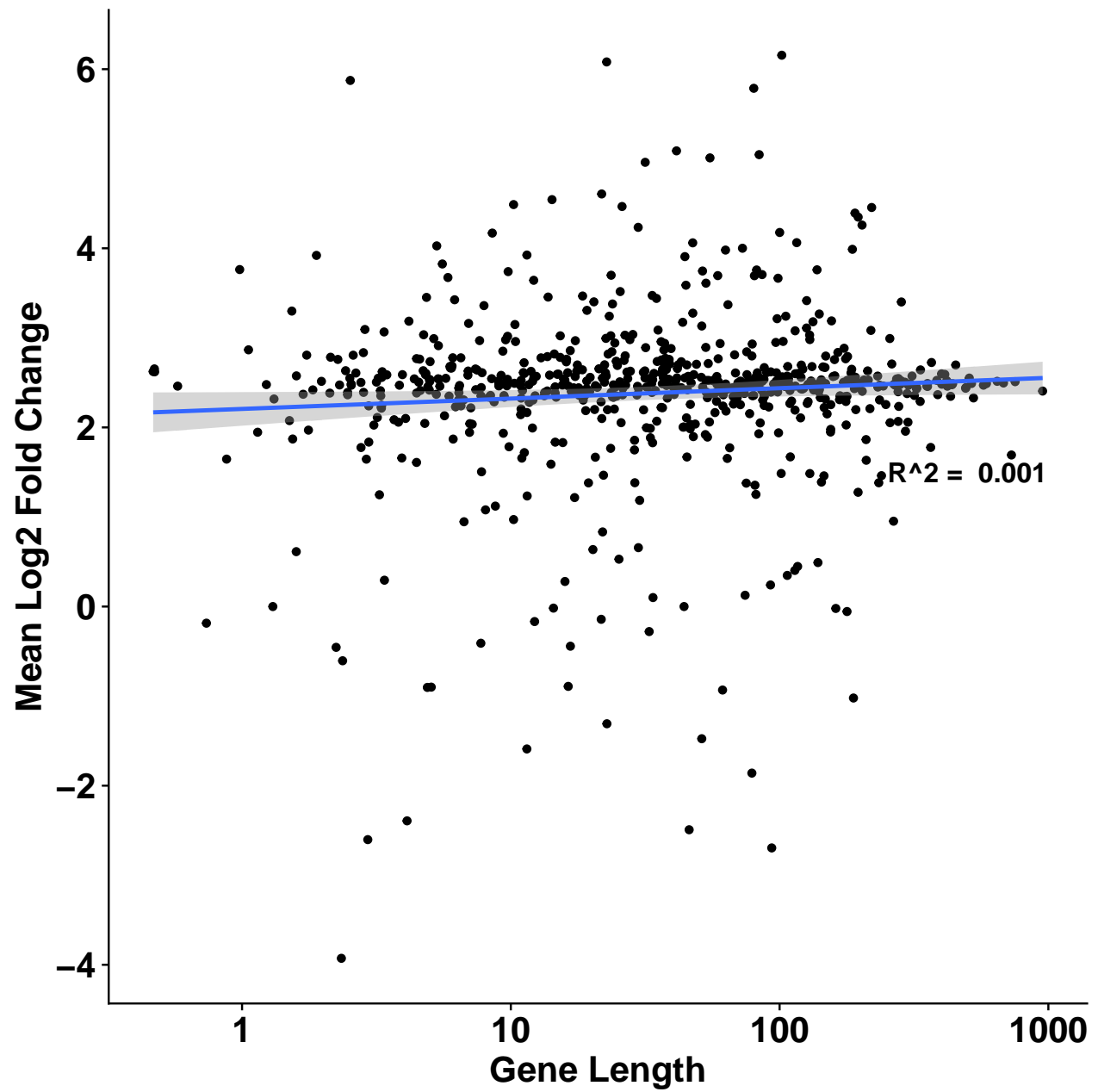
```
## Printing Supplementary Fig. 8
```

```
figureS8()
```









```
cat("\n\n Printing Supplementary Fig. 9 \n\n")
```

```
##
```

```
##
```

```
## Printing Supplementary Fig. 9
```

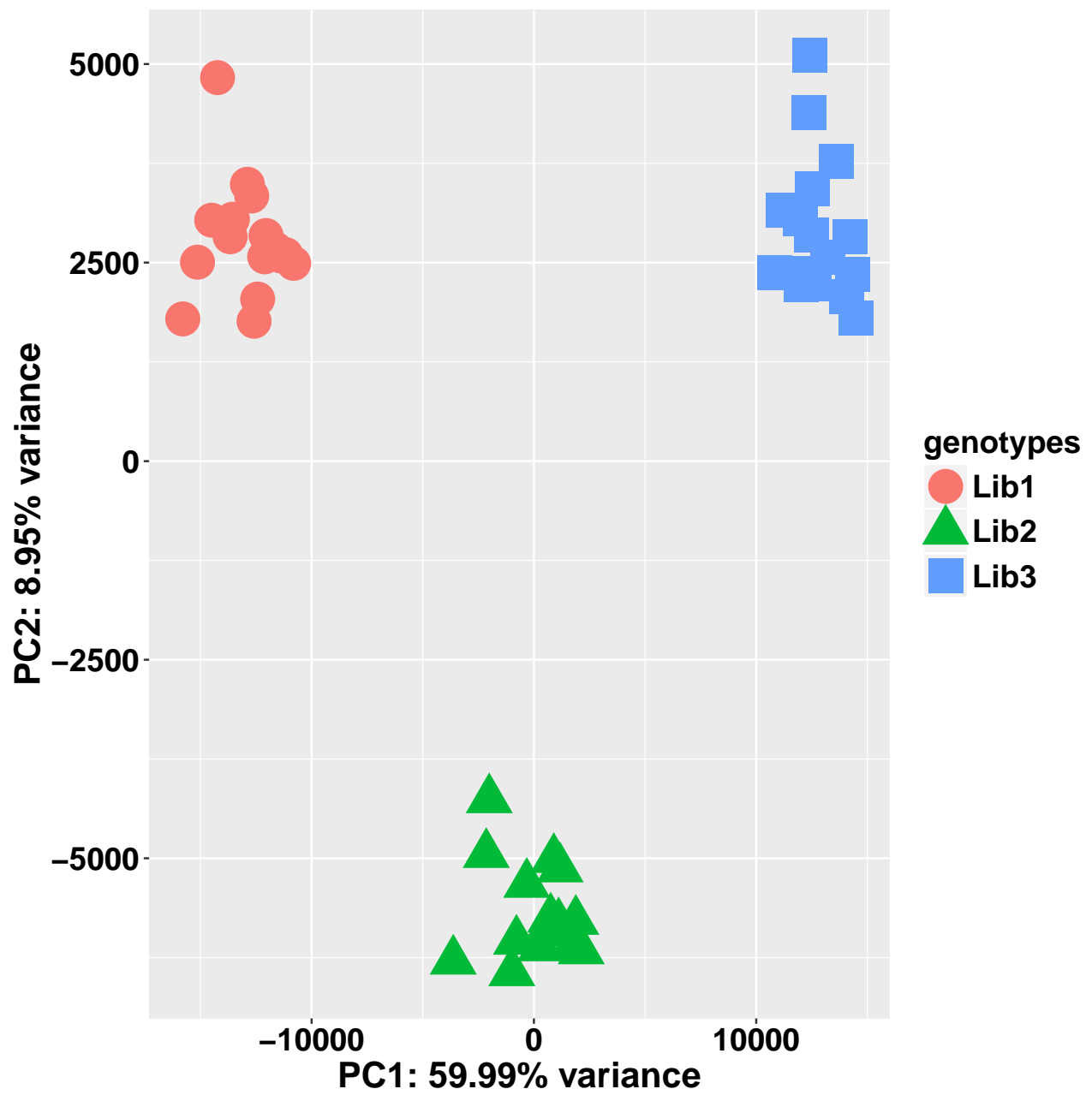
```
figureS9()
```

```
##
```

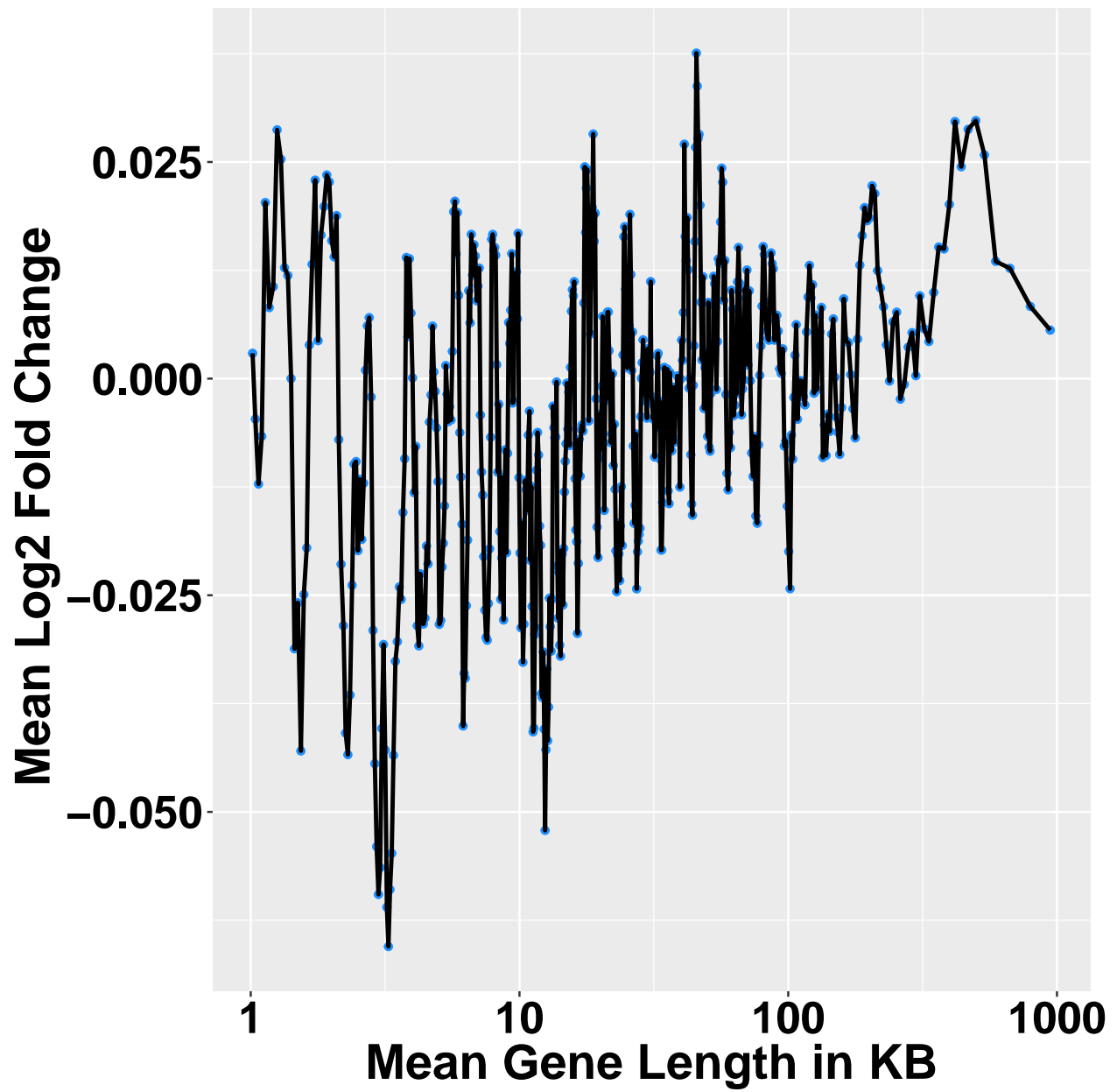
```
##
```

```
## Supplementary Fig. 9(A) -- Technical Brain Replicates
```

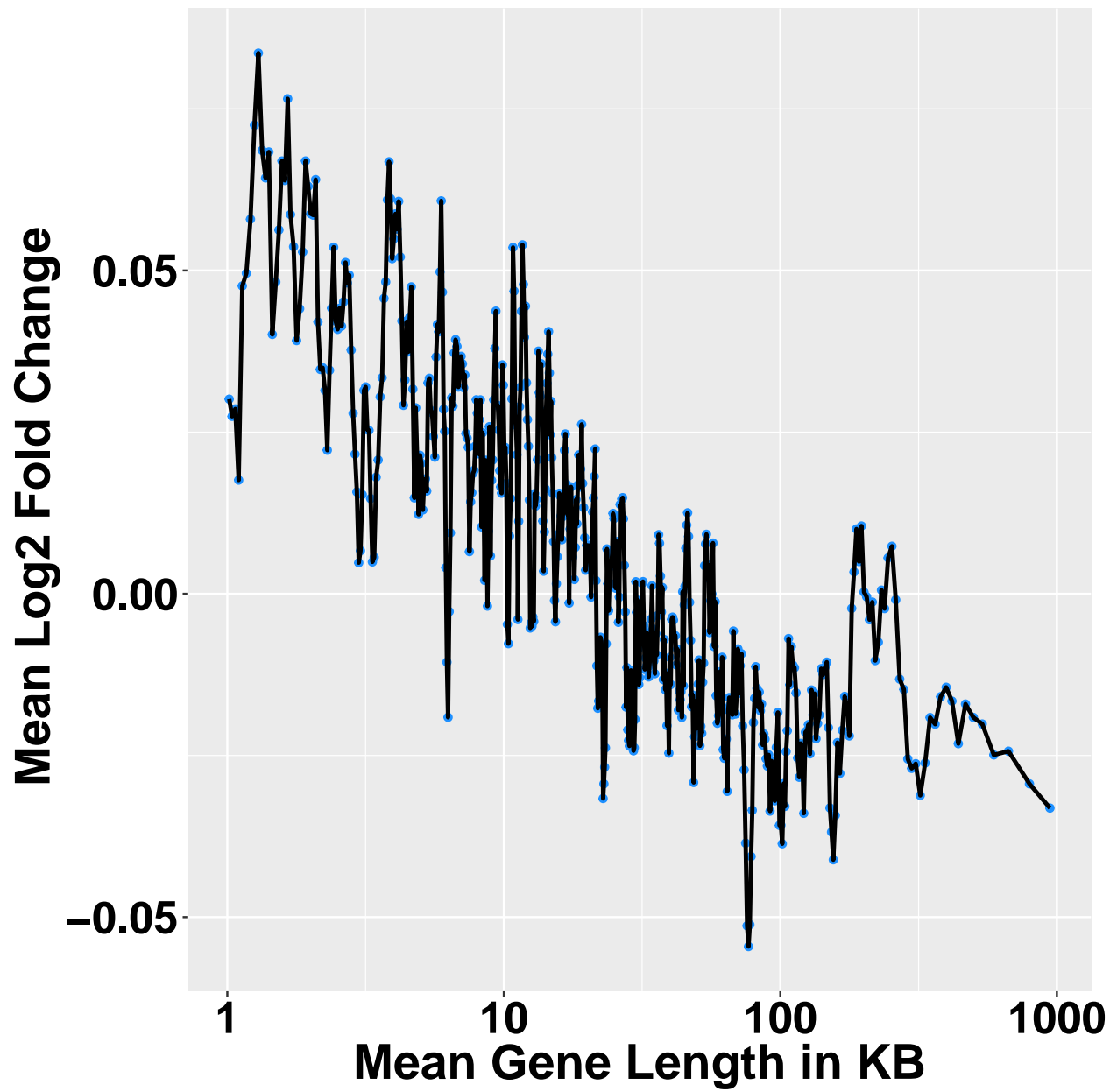




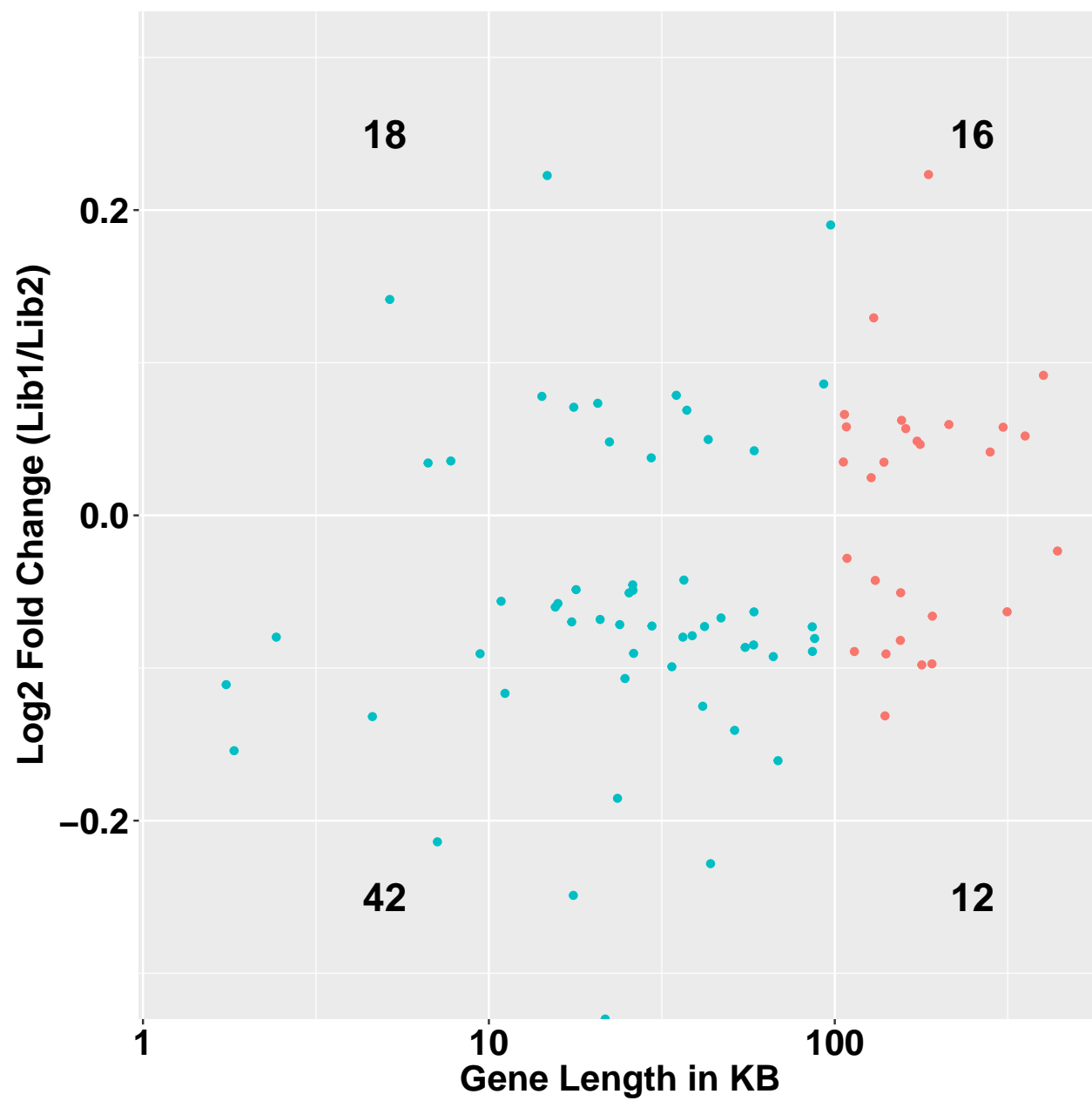
##  
 ##  
 ## Supplementary Fig. 9(B) -- Comparison of Lib1/Lib2

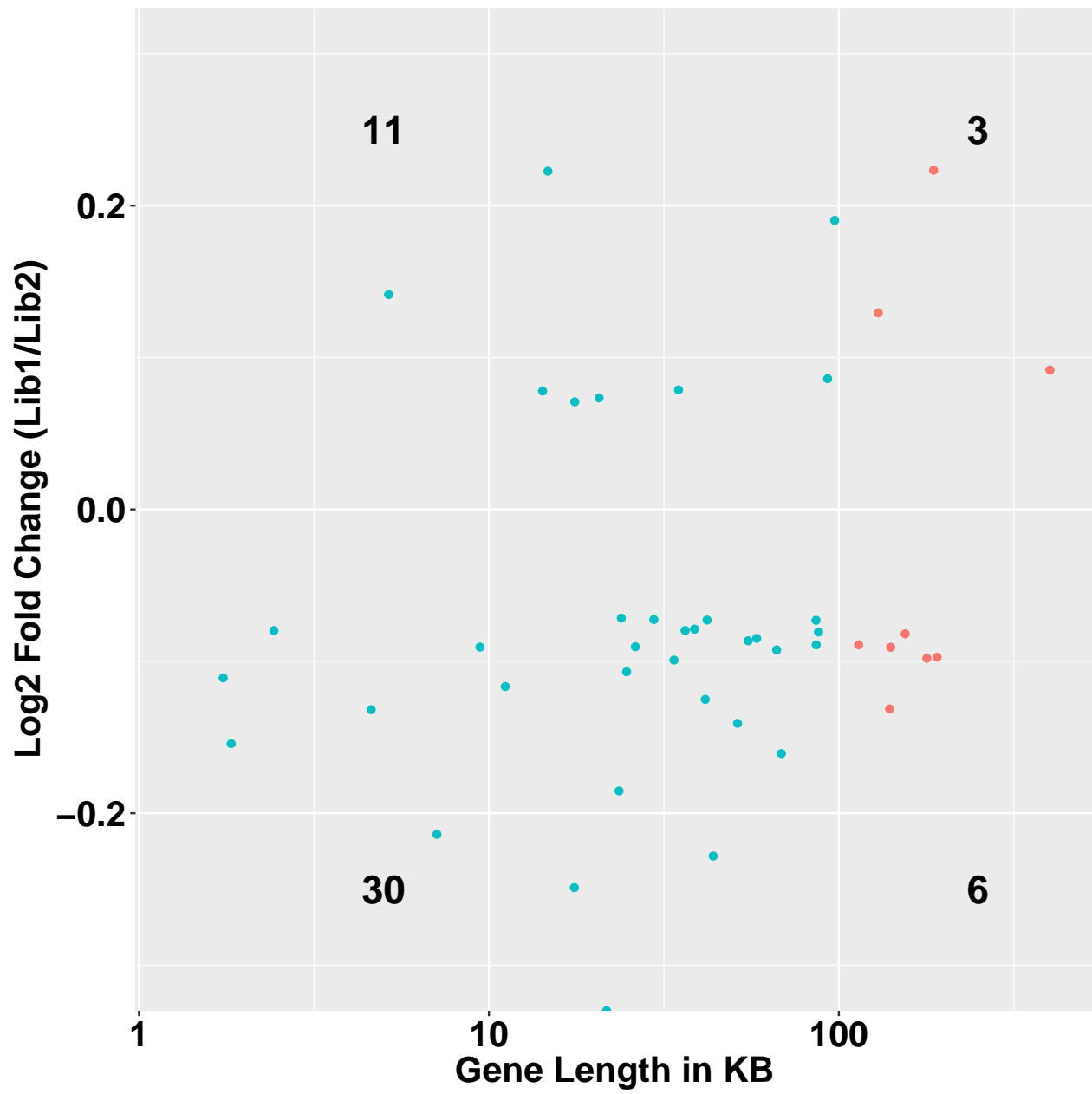


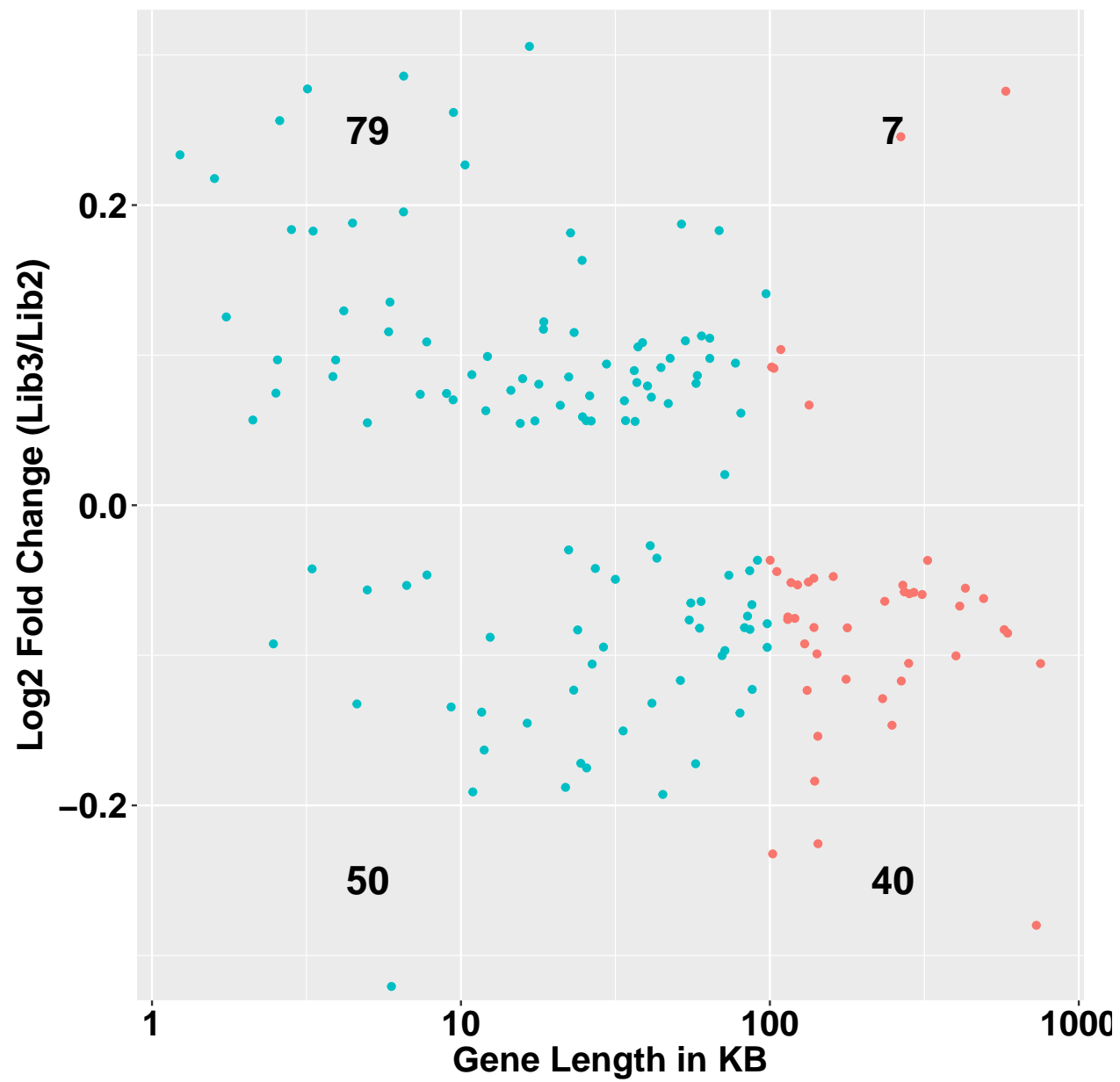
```
##
##
## Supplementary Fig. 9(C) -- Comparison of Lib3/Lib1
```

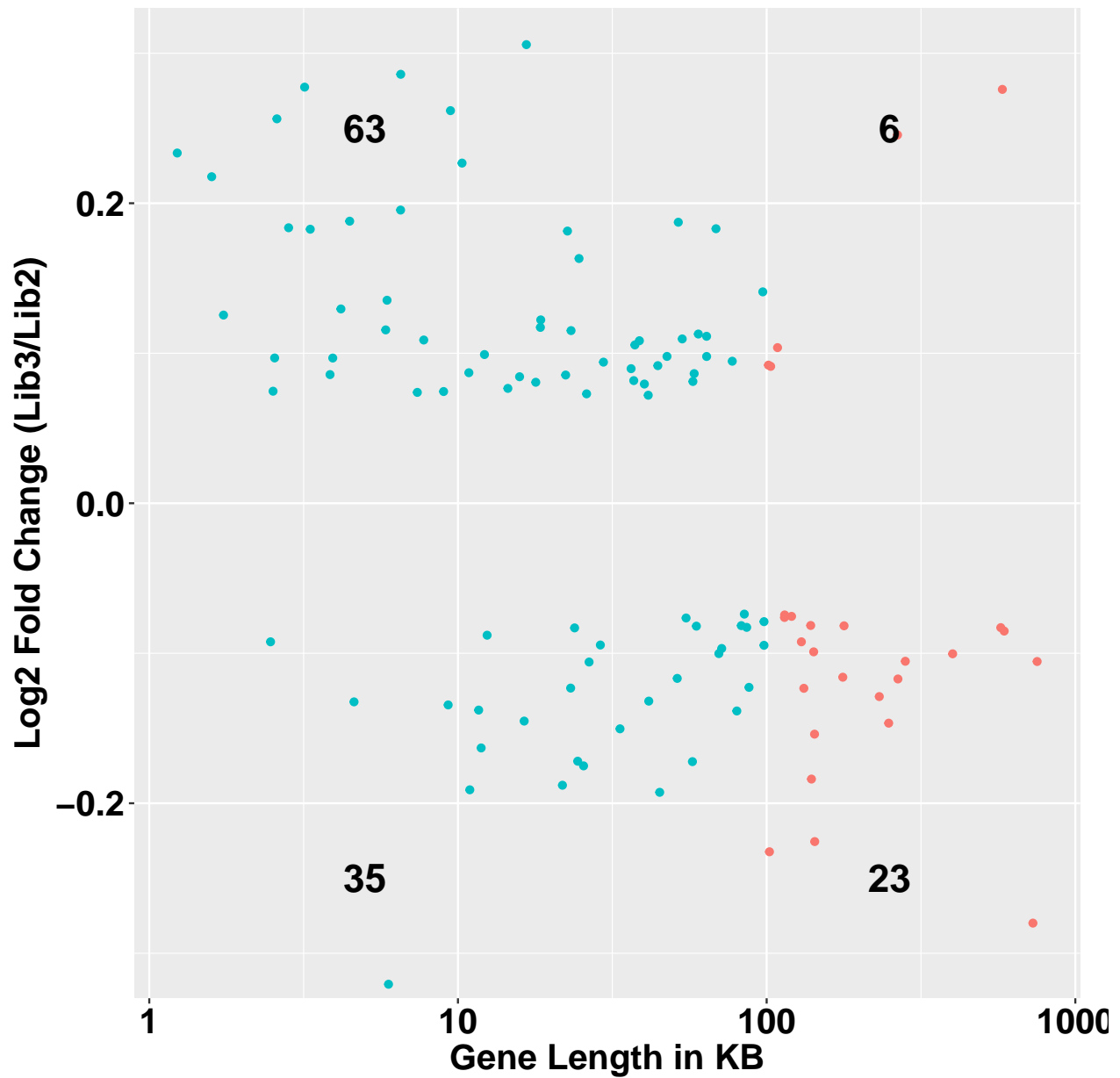


```
##  
##  
## Supplementary Fig. 9(D) -- DEGs Comparison of Lib1/Lib2
```









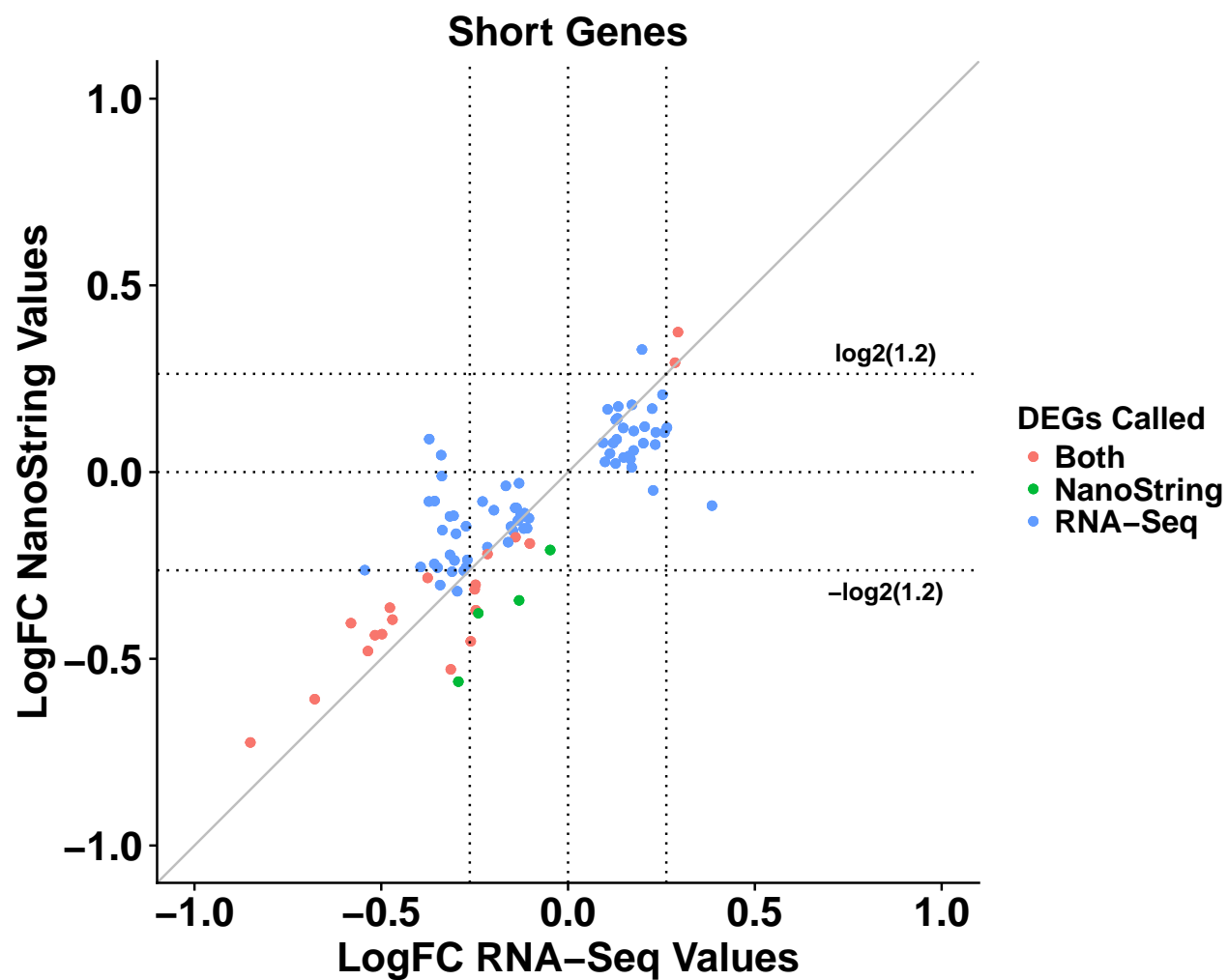
Section 7: Figure 5 and Supplementary Fig. 10

```
cat("\n\n Printing Fig. 5 \n\n")
```

```
##
##
## Printing Fig. 5
```

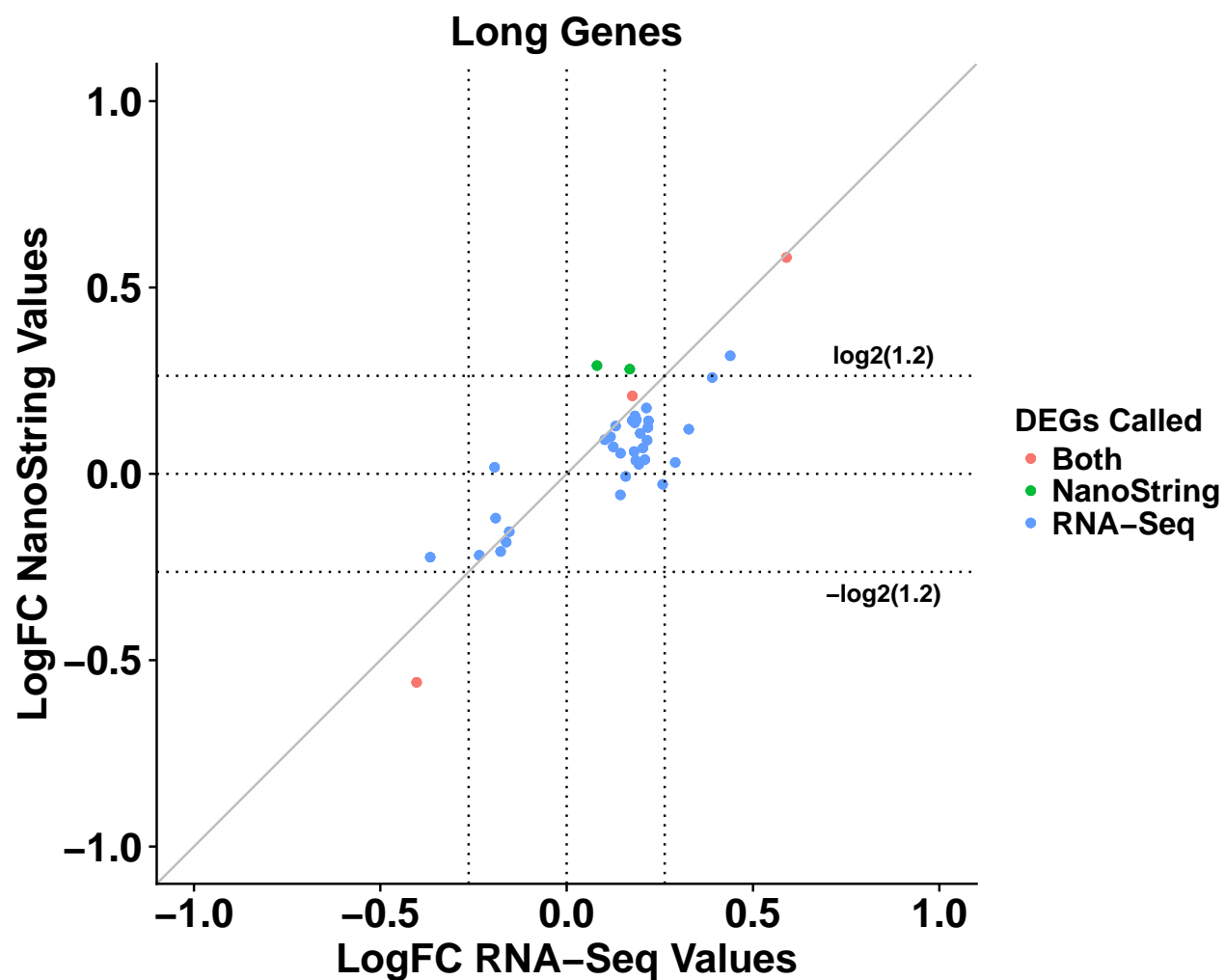
```
figure5()
```

```
##
##
## Fig. 5(A) -- Short Genes
```

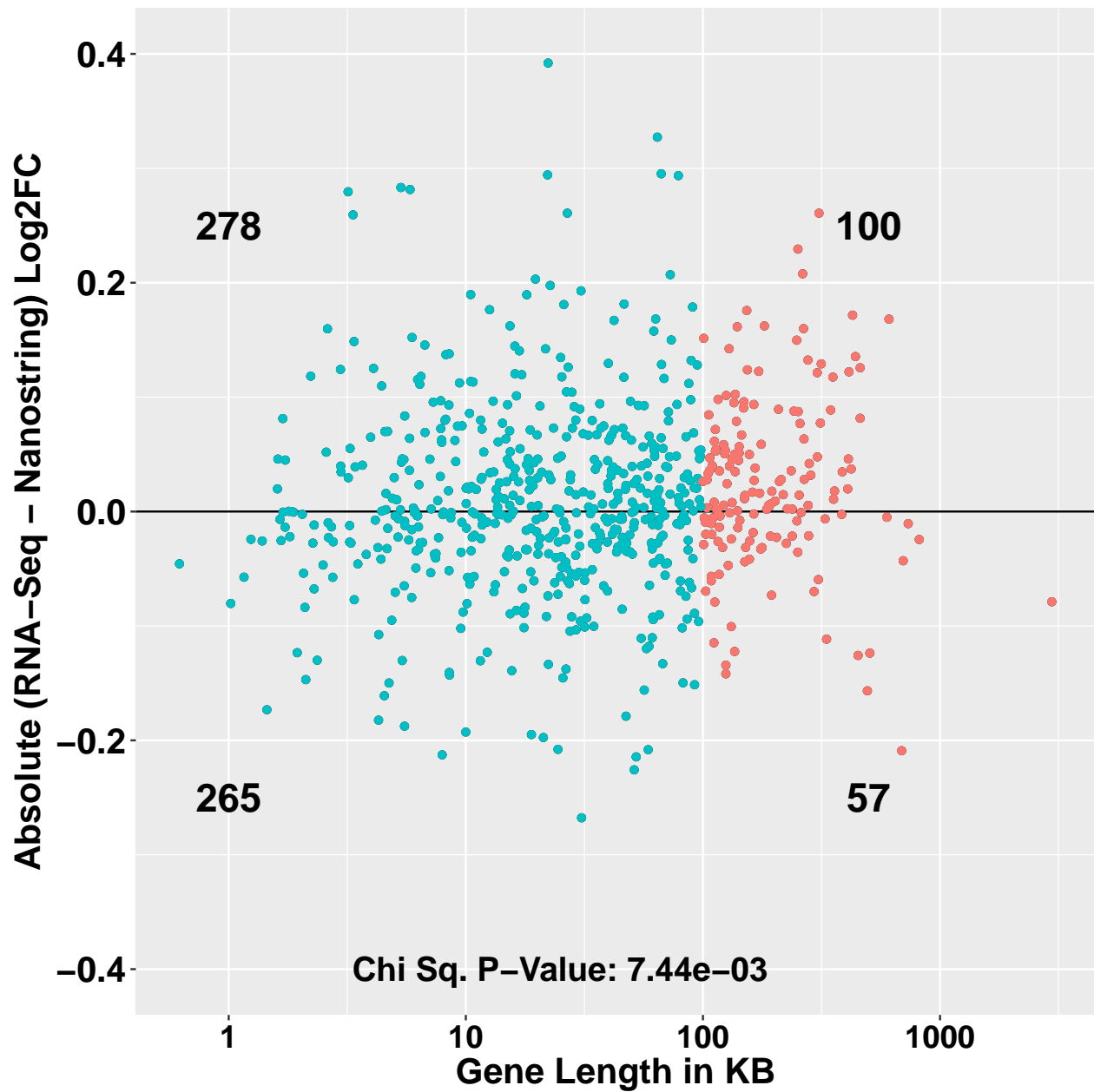


##  
 ##  
 ## Fig. 5(B) -- Long Genes





##  
 ##  
 ## Fig. 5(C)

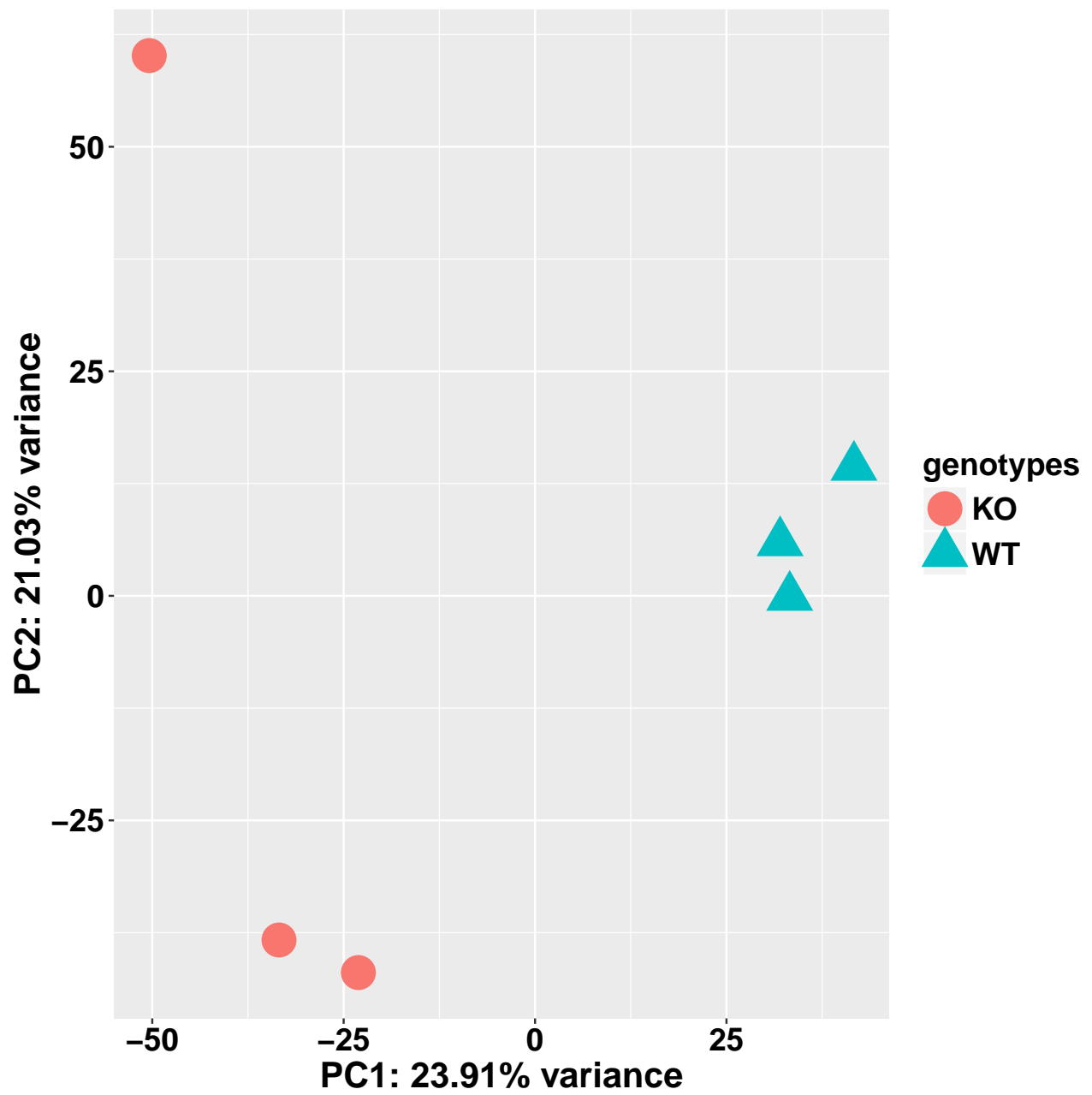


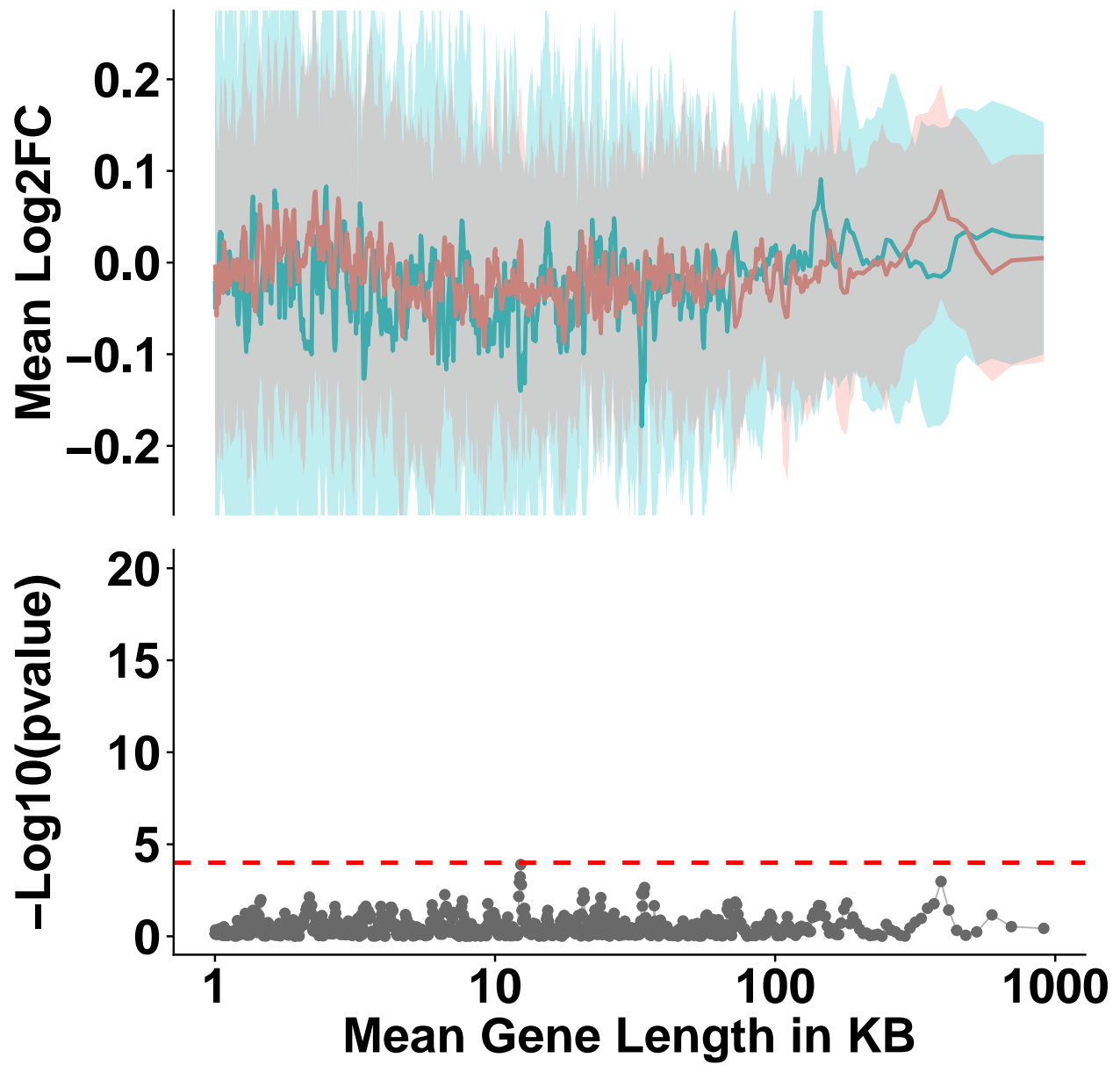
```
cat("\n\n Printing Supplementary Fig. 10 \n\n")
```

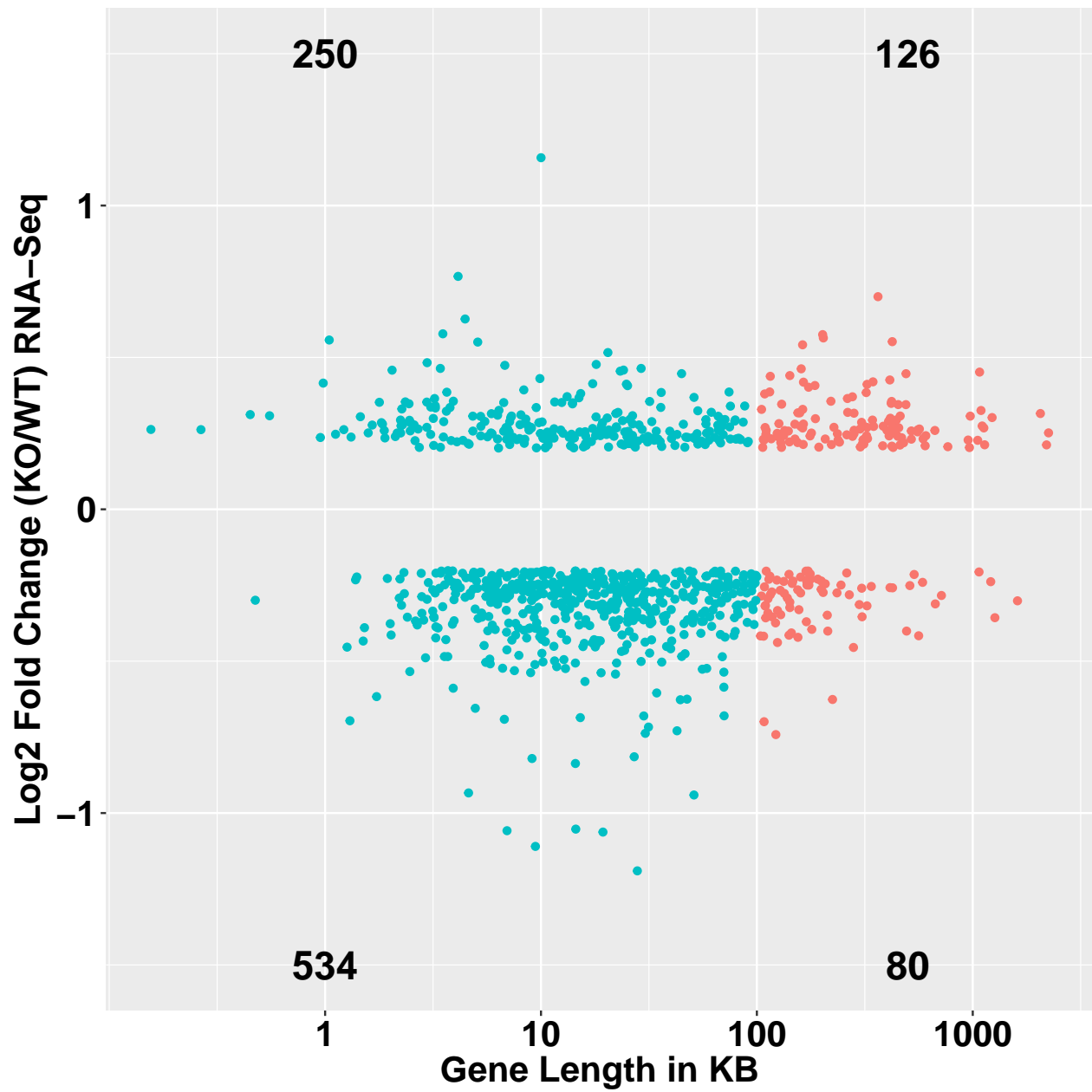
```
##
##
## Printing Supplementary Fig. 10
```

```
figureS10()
```

```
##
##
## Supplementary Fig. 10(A) -- Mecp2 Cerebellum RNA-seq KO/WT Dataset (Whole Genome)
```



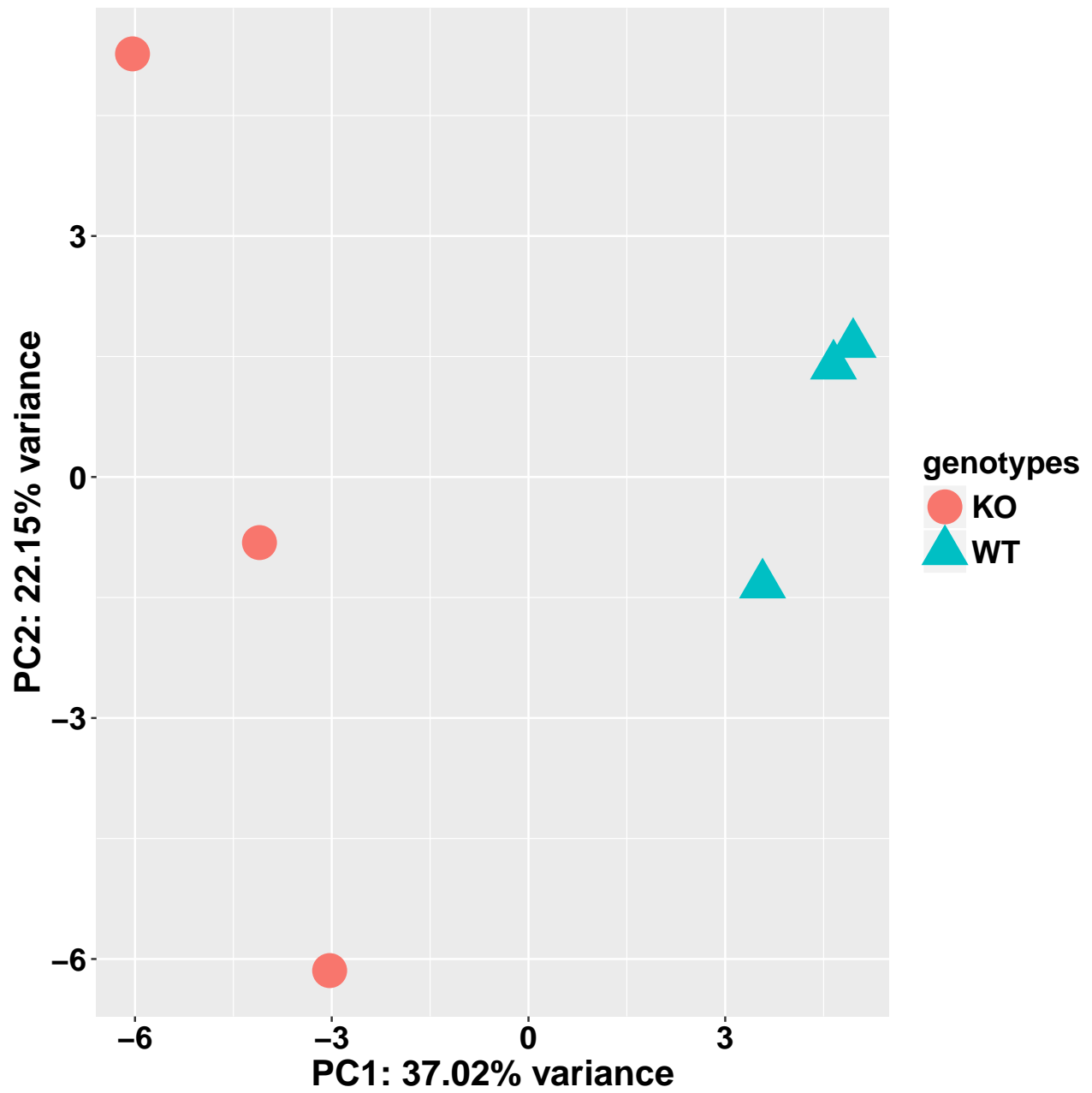


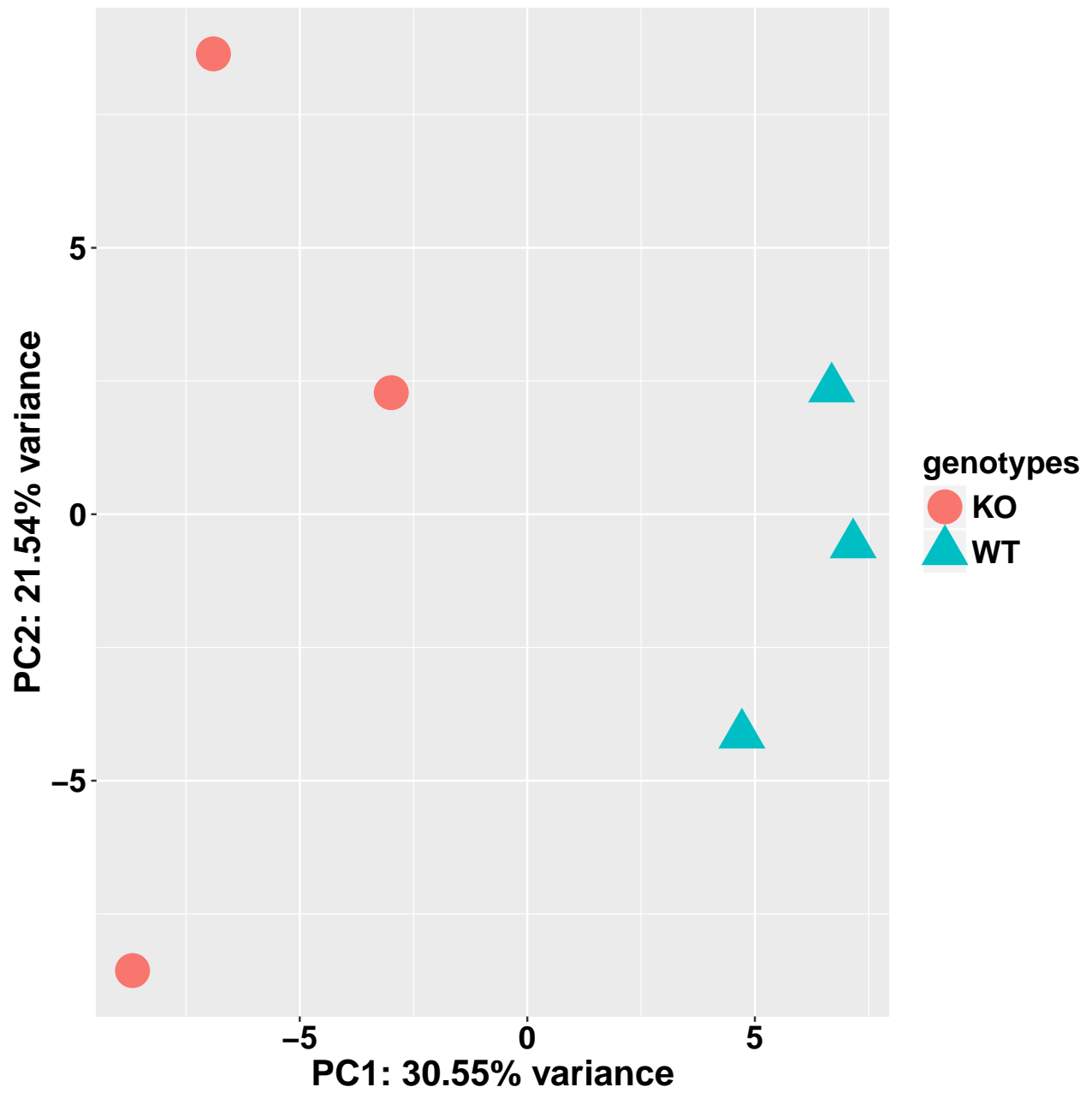


##

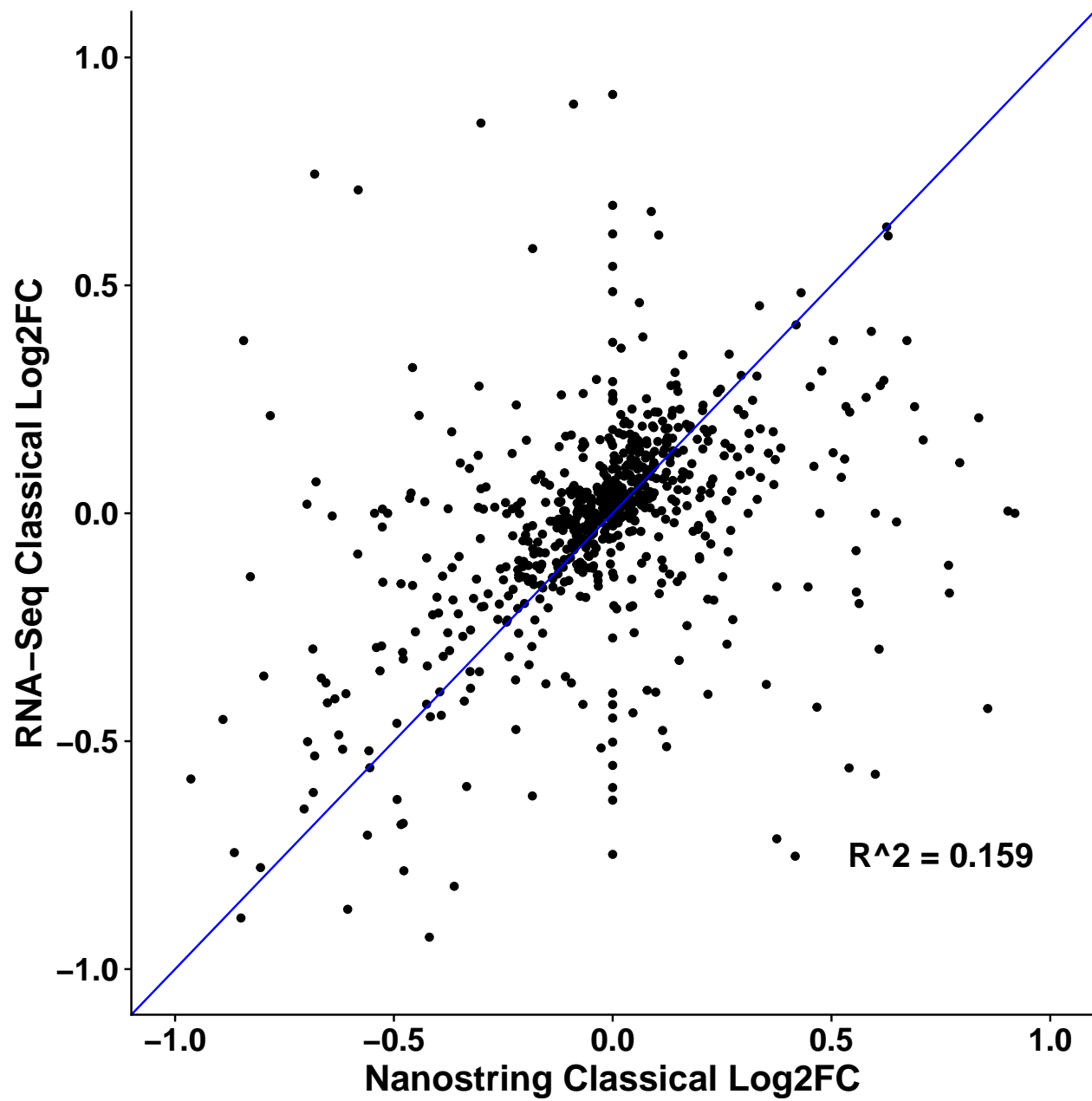
##

## Supplementary Fig. 10(B) -- 750 common genes between RNA-seq and Nanostring

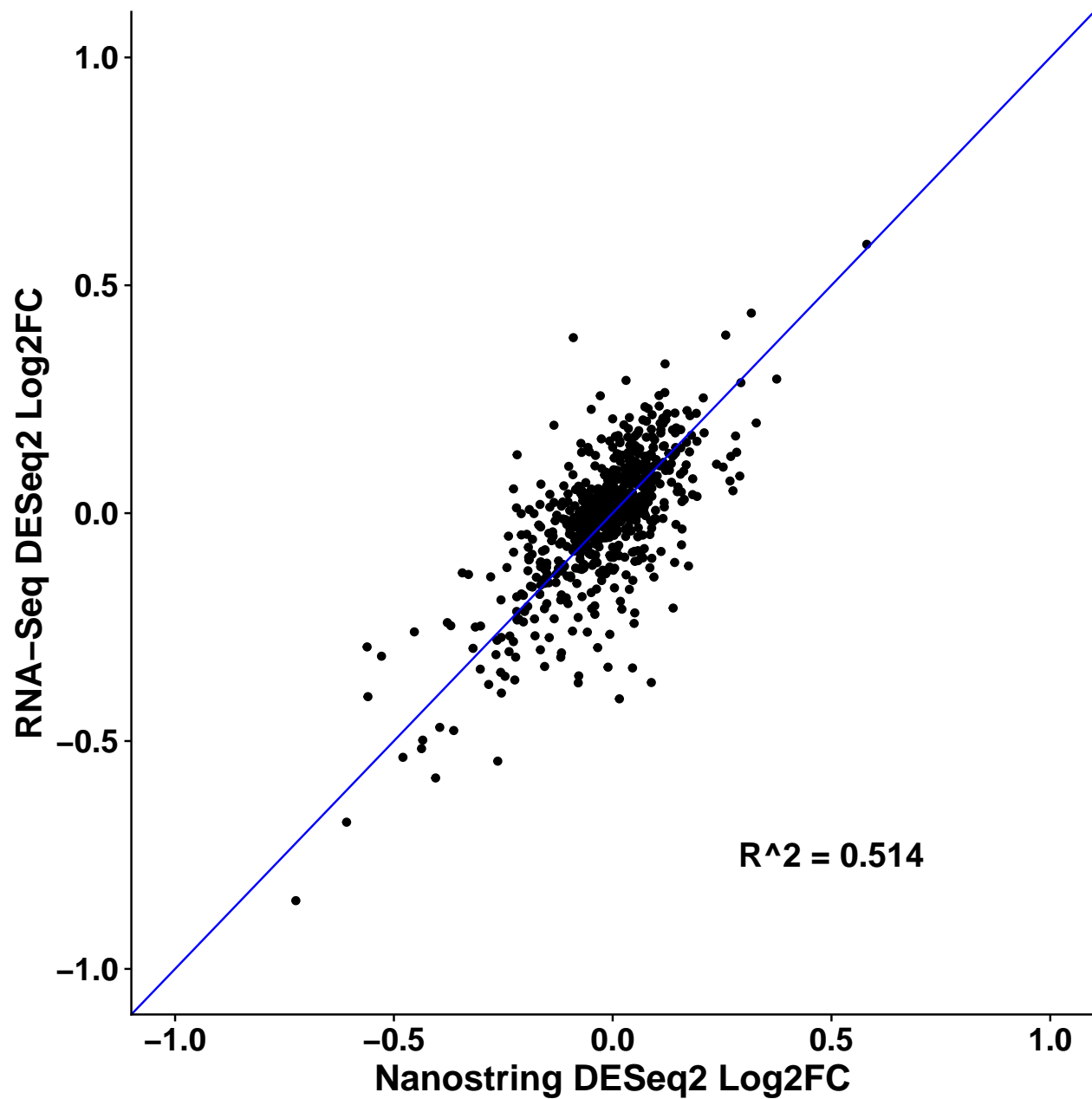




##  
##  
## Supplementary Fig. 10(C)







## Section 8: Session info

```
sessionInfo()
```

```
## R version 3.4.3 (2017-11-30)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS Sierra 10.12.6
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRlapack.dylib
```

```

##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats4      parallel  stats      graphics  grDevices  utils      datasets
## [8] methods     base
##
## other attached packages:
## [1] pd.hg.u95av2_3.12.0      DBI_0.8
## [3] RSQLite_2.1.0            tibble_1.4.2
## [5] scales_0.5.0             reshape2_1.4.3
## [7] rafalib_1.0.0            oligo_1.40.2
## [9] Biostings_2.44.2         XVector_0.16.0
## [11] oligoClasses_1.38.0      NanoStringNorm_1.2.1
## [13] vsn_3.44.0               gdata_2.18.0
## [15] gridExtra_2.3            GEOquery_2.42.0
## [17] edgeR_3.18.1             limma_3.32.10
## [19] dplyr_0.7.4              DESeq2_1.16.1
## [21] SummarizedExperiment_1.6.5 DelayedArray_0.2.7
## [23] matrixStats_0.53.1       GenomicRanges_1.28.6
## [25] GenomeInfoDb_1.12.3      cowplot_0.9.2
## [27] ggplot2_2.2.1.9000       annotate_1.54.0
## [29] XML_3.98-1.10            AnnotationDbi_1.38.2
## [31] IRanges_2.10.5           S4Vectors_0.14.7
## [33] affy_1.54.0              Biobase_2.36.2
## [35] BiocGenerics_0.22.1
##
## loaded via a namespace (and not attached):
## [1] bitops_1.0-6             bit64_0.9-7
## [3] RColorBrewer_1.1-2       httr_1.3.1
## [5] rprojroot_1.3-2         tools_3.4.3
## [7] backports_1.1.2         R6_2.2.2
## [9] affyio_1.46.0           rpart_4.1-13
## [11] Hmisc_4.1-1             lazyeval_0.2.1
## [13] colorspace_1.3-2        nnet_7.3-12
## [15] bit_1.1-12              compiler_3.4.3
## [17] preprocessCore_1.38.1    htmlTable_1.11.2
## [19] labeling_0.3            checkmate_1.8.5
## [21] genefilter_1.58.1       stringr_1.3.0
## [23] digest_0.6.15           foreign_0.8-69
## [25] rmarkdown_1.9           base64enc_0.1-3
## [27] pkgconfig_2.0.1         htmltools_0.3.6
## [29] htmlwidgets_1.0         rlang_0.2.0
## [31] rstudioapi_0.7          BiocInstaller_1.26.1
## [33] bindr_0.1.1             BiocParallel_1.10.1
## [35] gtools_3.5.0            acepack_1.4.1
## [37] RCurl_1.95-4.10         magrittr_1.5
## [39] GenomeInfoDbData_0.99.0 Formula_1.2-2
## [41] Matrix_1.2-14           Rcpp_0.12.16
## [43] munsell_0.4.3           stringi_1.1.7
## [45] yaml_2.1.18             zlibbioc_1.22.0
## [47] plyr_1.8.4              affxparser_1.48.0
## [49] grid_3.4.3              blob_1.1.1

```

```
## [51] lattice_0.20-35      splines_3.4.3
## [53] locfit_1.5-9.1       knitr_1.20
## [55] pillar_1.2.1         codetools_0.2-15
## [57] geneplotter_1.54.0   glue_1.2.0
## [59] evaluate_0.10.1      latticeExtra_0.6-28
## [61] data.table_1.10.4-3  foreach_1.4.4
## [63] gtable_0.2.0         assertthat_0.2.0
## [65] xtable_1.8-2         ff_2.2-13
## [67] survival_2.42-3      iterators_1.0.9
## [69] memoise_1.1.0        bindrcpp_0.2.2
## [71] cluster_2.0.7-1
```