Gviz_fig1

Nikita Verheyden

2023-04-17

setup

directory

```
setwd("D:/Krueger_Lab/Publications/miR181_paper/Figure1/Genome_tracks")
```

packages

```
library(Gviz)
```

```
## 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, setdiff, sort,
##     table, tapply, union, unique, unsplit, which.max, which.min
```

```
##
## Attaching package: 'S4Vectors'
```

```
## The following objects are masked from 'package:base':
##
##     expand.grid, I, unname
```

```
## Loading required package: IRanges
```

```
##
## Attaching package: 'IRanges'
```

```
## The following object is masked from 'package:grDevices':
##
##     windows
```

```
## Loading required package: GenomicRanges
## Loading required package: GenomeInfoDb
## Loading required package: grid
```

```
library(GenomicRanges)
library(rtracklayer)
library(dplyr)
```

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

```
library(plyranges)
```

```
##
## Attaching package: 'plyranges'
## The following objects are masked from 'package:dplyr':
##
##     between, n, n_distinct
## The following object is masked from 'package:IRanges':
##
##     slice
## The following object is masked from 'package:stats':
##
##     filter
```

```
library(BSgenome.Mmusculus.UCSC.mm10)
```

```
## Loading required package: BSgenome
## Loading required package: Biostrings
```

```
## Loading required package: XVector
##
## Attaching package: 'Biostrings'
## The following object is masked from 'package:grid':
##
##     pattern
## The following object is masked from 'package:base':
##
##     strsplit
```

Import data

```
# diferential results
diffres <- rtracklayer::import(con = "D:/Krueger_Lab/miReCLIP/Mirco/DifferentialBinding/BsDifferentialR

# non enriched non chimeric
IPKOminus <- import.bw("D:/Krueger_Lab/miReCLIP/Melina/pipe_output_22_02_14/pipe_output_22_02_14/non-chi
IPKOplus <- import.bw("D:/Krueger_Lab/miReCLIP/Melina/pipe_output_22_02_14/pipe_output_22_02_14/non-chim
IPWTminus <- import.bw("D:/Krueger_Lab/miReCLIP/Melina/pipe_output_22_02_14/pipe_output_22_02_14/non-chi
IPWTplus <- import.bw("D:/Krueger_Lab/miReCLIP/Melina/pipe_output_22_02_14/pipe_output_22_02_14/non-chim

IP181WTminus  <- import.bw("D:/Krueger_Lab/miReCLIP/Melina/pipe_output_22_02_14/pipe_output_22_02_14/nor
IP181WTplus <- import.bw("D:/Krueger_Lab/miReCLIP/Melina/pipe_output_22_02_14/pipe_output_22_02_14/non-



# enriched chimeric
IP181chimWTminus <- import.bw("D:/Krueger_Lab/miReCLIP/Melina/pipe_output_22_02_14/pipe_output_22_02_14/
IP181chimWTplus <- import.bw("D:/Krueger_Lab/miReCLIP/Melina/pipe_output_22_02_14/pipe_output_22_02_14/
```

setting regions

Use this panel to change the coordinates in all following plots

```
scoord = 80108008
ecoord = 80108895
tchr = "chr12"
clim = 45
```

Zfp36l1: chr12 80107716-80108070 Zfp36l2: chr17 84184290-84184345 Cblb: chr16:52171980-52172133 elmo1: chr13:20405350-20406237 srsf11: chr3:158,010,824-158,012,135 Aebp2: chr6:140652591-140652785 Zfp36L1: total: chr12:80108008-80108895 181bs:chr12:80108023-80108,242 otherbs: chr12:80108634-80108853

Tracks

Annotation track

```
# genome axis
gatrack <- GenomeAxisTrack()

# gene track----------------------------this is not working yet
# data(cpgIslands)
```

```
# class(cpgIslands)
# ## [1] "GRanges"
# ## attr(,"package")
# ## [1] "GenomicRanges"
# chr <- as.character(unique(seqnames(cpgIslands)))
# gen <- genome(cpgIslands)
# atrack <- AnnotationTrack(cpgIslands, name = "CpG")
#
# data(geneModels)
#
# gff <- import.gff3("D:/Krueger_Lab/Ribo_Profiling/run15112022M23/ref_genome/gencode.vM23.annotation.g
#
# gffloc <- as.data.frame(gff)
#
# gffloc <- gffloc[gffloc$seqnames==tchr,]
#
# grtrack <- GeneRegionTrack(gffloc, genome = "mm10", chromosome = tchr,
#                            name = "Gene")
#
# grtrack <- GeneRegionTrack(geneModels, genome = BSgenome.Mmusculus.UCSC.mm10, chromosome = tchr,
#                            name = "foo")
```

pureclip (merged)

Always just activate or inactivate the strand that actually contains data

```
# -

#KO minus
KOm <- as.data.frame(IPKOminus)
KOm <- KOm[KOm$seqnames == tchr & KOm$start >= scoord & KOm$end <= ecoord,]
KOm$seqnames <- as.character(KOm$seqnames)
KOm <- makeGRangesFromDataFrame(KOm, keep.extra.columns = T)

KOmT <- DataTrack(KOm, name="KO-", ylim = c(0,clim))


#WT minus
WTm <- as.data.frame(IPWTminus)
WTm <- WTm[WTm$seqnames == tchr & WTm$start >= scoord & WTm$end <= ecoord,]
WTm$seqnames <- as.character(WTm$seqnames)
WTm <- makeGRangesFromDataFrame(WTm, keep.extra.columns = T)

WTmT <- DataTrack(WTm, name="WT-", ylim = c(0,clim))

#181 WT minus
WT181m <- as.data.frame(IP181WTminus)
WT181m <- WT181m[WT181m$seqnames == tchr & WT181m$start >= scoord & WT181m$end <= ecoord,]
WT181m$seqnames <- as.character(WT181m$seqnames)
WT181m <- makeGRangesFromDataFrame(WT181m, keep.extra.columns = T)

WT181mT <- DataTrack(WT181m, name="WT181-", ylim = c(0,clim))
```

```
# +

# #KO plus
# KOp <- as.data.frame(IPKOplus)
# KOp <- KOp[KOp$seqnames == tchr & KOp$start >= scoord & KOp$end <= ecoord,]
# KOp$seqnames <- as.character(KOp$seqnames)
# KOp <- makeGRangesFromDataFrame(KOp, keep.extra.columns = T)
#
# KOpT <- DataTrack(KOp, name="KO+", ylim = c(0,clim))
#
# #WT plus
# WTp <- as.data.frame(IPWTplus)
# WTp <- WTp[WTp$seqnames == tchr & WTp$start >= scoord & WTp$end <= ecoord,]
# WTp$seqnames <- as.character(WTp$seqnames)
# WTp <- makeGRangesFromDataFrame(WTp, keep.extra.columns = T)
#
# WTpT <- DataTrack(WTp, name="WT+", ylim = c(0,clim))

# #181 WT plus
# WT181p <- as.data.frame(IP181WTplus)
# WT181p <- WT181p[WT181p$seqnames == tchr & WT181p$start >= scoord & WT181p$end <= ecoord,]
# WT181p$seqnames <- as.character(WT181p$seqnames)
# WT181p <- makeGRangesFromDataFrame(WT181p, keep.extra.columns = T)
#
# WT181pT <- DataTrack(WT181p, name="WT181+", ylim = c(0,clim))
```

chimeric reads

```
# minus
chim181m <- as.data.frame(IP181chimWTminus)
chim181m <- chim181m[chim181m$seqnames == tchr & chim181m$start >= scoord & chim181m$end <= ecoord,]
chim181m$seqnames <- as.character(chim181m$seqnames)
chim181m <- makeGRangesFromDataFrame(chim181m, keep.extra.columns = T)

chim181mT <- DataTrack(chim181m, name="WT181rich-", ylim = c(0,clim))

# #plus
# chim181p <- as.data.frame(IP181chimWTplus)
# chim181p <- chim181p[chim181p$seqnames == tchr & chim181p$start >= scoord & chim181p$end <= ecoord,]
# chim181p$seqnames <- as.character(chim181p$seqnames)
# chim181p <- makeGRangesFromDataFrame(chim181p, keep.extra.columns = T)
#
# chim181pT <- DataTrack(chim181p, name="WT181rich-", ylim = c(0,clim))
```

differential binding

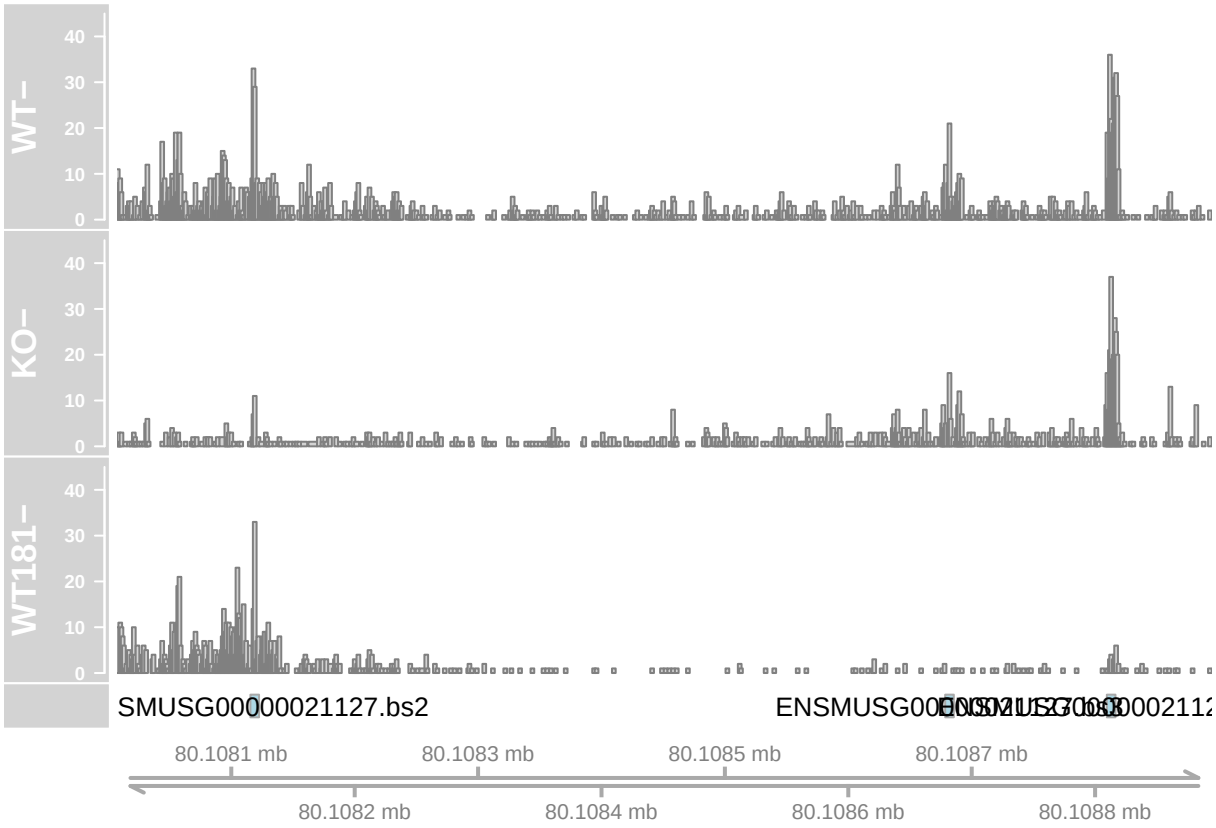
```
diffresf <- diffres %>% filter(seqnames == tchr)

diffresT <- AnnotationTrack(diffresf, name = "dBS", id=diffresf$name)
```

plot

```
# # +
# plotTracks(list(WTpT, KOpT, WT181pT, chim181pT, diffresT,  gatrack), from = scoord, to= ecoord, type=
#            fontsize.feature=10, fontcolor.feature = "black")

# -
plotTracks(list(WTmT, KOmT,WT181mT, diffresT,  gatrack), from = scoord, to= ecoord, type="histogram", f
           fontsize.feature=10, fontcolor.feature = "black")
```



make seperate plots for different binding sites

session info

```
sessionInfo()
```

```
## R version 4.2.3 (2023-03-15 ucrt)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 19045)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=German_Germany.utf8  LC_CTYPE=German_Germany.utf8
## [3] LC_MONETARY=German_Germany.utf8 LC_NUMERIC=C
```

```
## [5] LC_TIME=German_Germany.utf8
##
## attached base packages:
## [1] grid      stats4    stats     graphics  grDevices utils     datasets
## [8] methods   base
##
## other attached packages:
##  [1] BSgenome.Mmusculus.UCSC.mm10_1.4.3 BSgenome_1.66.3
##  [3] Biostrings_2.66.0                  XVector_0.38.0
##  [5] plyranges_1.18.0                   dplyr_1.1.1
##  [7] rtracklayer_1.58.0                 Gviz_1.42.1
##  [9] GenomicRanges_1.50.2               GenomeInfoDb_1.34.9
## [11] IRanges_2.32.0                     S4Vectors_0.36.2
## [13] BiocGenerics_0.44.0
##
## loaded via a namespace (and not attached):
##   [1] ProtGenerics_1.30.0         bitops_1.0-7
##   [3] matrixStats_0.63.0          bit64_4.0.5
##   [5] filelock_1.0.2              RColorBrewer_1.1-3
##   [7] progress_1.2.2              httr_1.4.5
##   [9] tools_4.2.3                 backports_1.4.1
##  [11] utf8_1.2.3                  R6_2.5.1
##  [13] rpart_4.1.19                lazyeval_0.2.2
##  [15] Hmisc_5.0-1                 DBI_1.1.3
##  [17] colorspace_2.1-0            nnet_7.3-18
##  [19] tidyselect_1.2.0            gridExtra_2.3
##  [21] prettyunits_1.1.1           bit_4.0.5
##  [23] curl_5.0.0                  compiler_4.2.3
##  [25] cli_3.6.0                   Biobase_2.58.0
##  [27] htmlTable_2.4.1             xml2_1.3.3
##  [29] DelayedArray_0.23.2         scales_1.2.1
##  [31] checkmate_2.1.0             rappdirs_0.3.3
##  [33] stringr_1.5.0               digest_0.6.31
##  [35] Rsamtools_2.14.0            foreign_0.8-84
##  [37] rmarkdown_2.21              jpeg_0.1-10
##  [39] dichromat_2.0-0.1           base64enc_0.1-3
##  [41] pkgconfig_2.0.3             htmltools_0.5.4
##  [43] MatrixGenerics_1.10.0       highr_0.10
##  [45] ensembldb_2.22.0            dbplyr_2.3.2
##  [47] fastmap_1.1.1               htmlwidgets_1.6.2
##  [49] rlang_1.1.0                 rstudioapi_0.14
##  [51] RSQLite_2.3.1               BiocIO_1.8.0
##  [53] generics_0.1.3              BiocParallel_1.32.6
##  [55] VariantAnnotation_1.44.1    RCurl_1.98-1.12
##  [57] magrittr_2.0.3              GenomeInfoDbData_1.2.9
##  [59] Formula_1.2-5               interp_1.1-4
##  [61] Matrix_1.5-3                Rcpp_1.0.10
##  [63] munsell_0.5.0               fansi_1.0.4
##  [65] lifecycle_1.0.3             stringi_1.7.12
##  [67] yaml_2.3.7                  SummarizedExperiment_1.28.0
##  [69] zlibbioc_1.44.0             BiocFileCache_2.6.1
##  [71] blob_1.2.4                  parallel_4.2.3
##  [73] crayon_1.5.2                deldir_1.0-6
##  [75] lattice_0.20-45             GenomicFeatures_1.50.4
```

```
##  [77] hms_1.1.3                  KEGGREST_1.38.0
##  [79] knitr_1.42                 pillar_1.9.0
##  [81] rjson_0.2.21               codetools_0.2-19
##  [83] biomaRt_2.54.1             XML_3.99-0.14
##  [85] glue_1.6.2                 evaluate_0.20
##  [87] latticeExtra_0.6-30        biovizBase_1.46.0
##  [89] data.table_1.14.8          png_0.1-8
##  [91] vctrs_0.6.1                gtable_0.3.3
##  [93] cachem_1.0.7               ggplot2_3.4.2
##  [95] xfun_0.37                  AnnotationFilter_1.22.0
##  [97] restfulr_0.0.15            tibble_3.2.1
##  [99] GenomicAlignments_1.34.1   AnnotationDbi_1.60.2
## [101] memoise_2.0.1              cluster_2.1.4
```