

main

September 6, 2021

1 Generate circlized plot for schizophrenia analysis DEG results

```
[1]: library(biomaRt)
library(circlize)
library(tidyverse)
library(ComplexHeatmap)
```

```
=====
circlize version 0.4.13
CRAN page: https://cran.r-project.org/package=circlize
Github page: https://github.com/jokergoo/circlize
Documentation: https://jokergoo.github.io/circlize\_book/book/
```

If you use it in published research, please cite:
Gu, Z. circlize implements and enhances circular visualization
in R. Bioinformatics 2014.

This message can be suppressed by:
`suppressPackageStartupMessages(library(circlize))`

```
=====

Attaching packages                                tidyverse
1.3.1
```

```
ggplot2 3.3.5    purrr   0.3.4
tibble  3.1.4    dplyr   1.0.7
tidyr   1.1.3    stringr 1.4.0
readr   2.0.1    forcats 0.5.1
```

```
Conflicts
tidyverse_conflicts()
dplyr::filter() masks stats::filter()
dplyr::lag()    masks stats::lag()
dplyr::select() masks
biomaRt::select()
```

Loading required package: grid

```
=====
ComplexHeatmap version 2.6.2
Bioconductor page: http://bioconductor.org/packages/ComplexHeatmap/
Github page: https://github.com/jokergoo/ComplexHeatmap
Documentation: http://jokergoo.github.io/ComplexHeatmap-reference

If you use it in published research, please cite:
Gu, Z. Complex heatmaps reveal patterns and correlations in multidimensional
genomic data. Bioinformatics 2016.

This message can be suppressed by:
  suppressPackageStartupMessages(library(ComplexHeatmap))
=====
```

1.1 Prepare data

1.1.1 Get gene annotation

```
[2]: ensembl = useEnsembl(biomart="ensembl", dataset="hsapiens_gene_ensembl")
      biomart = getBM(attributes=c('ensembl_gene_id', 'chromosome_name',
                                   'start_position', 'end_position'), mart=ensembl)
      biomart %>% head(2)
```

A data.frame: 2 × 4		ensembl_gene_id <chr>	chromosome_name <chr>	start_position <int>	end_position <int>
1		ENSG00000210049	MT	577	647
2		ENSG00000211459	MT	648	1601

1.1.2 Get logFC for differential expression analysis

```
[3]: deg = data.table::fread("../_m/genes/diffExpr_szVctl_full.txt") %>%
      select(gencodeID, ensemblID, Symbol, logFC, "adj.P.Val")
      deg %>% head(2)
```

A data.table: 2 × 5		gencodeID <chr>	ensemblID <chr>	Symbol <chr>	logFC <dbl>	adj.P.Val <dbl>
		ENSG00000248587.7	ENSG00000248587	GDNF-AS1	0.8015019	1.387742e-26
		ENSG00000138944.7	ENSG00000138944	KIAA1644	0.5637328	1.707516e-24

1.1.3 Merge data and select regions

```
[4]: bed = deg %>%
      inner_join(biomart, by=c("ensemblID"="ensembl_gene_id")) %>%
      select(chromosome_name, start_position, end_position, logFC, "adj.P.Val")
      ↪ %>%
      mutate(chromosome_name=paste0('chr', chromosome_name))
```

```
bed_sz = bed %>% filter(logFC > 0, adj.P.Val < 0.05)
bed_ctl = bed %>% filter(logFC < 0, adj.P.Val < 0.05)
bed_nonsig = bed %>% filter(adj.P.Val > 0.05)
```

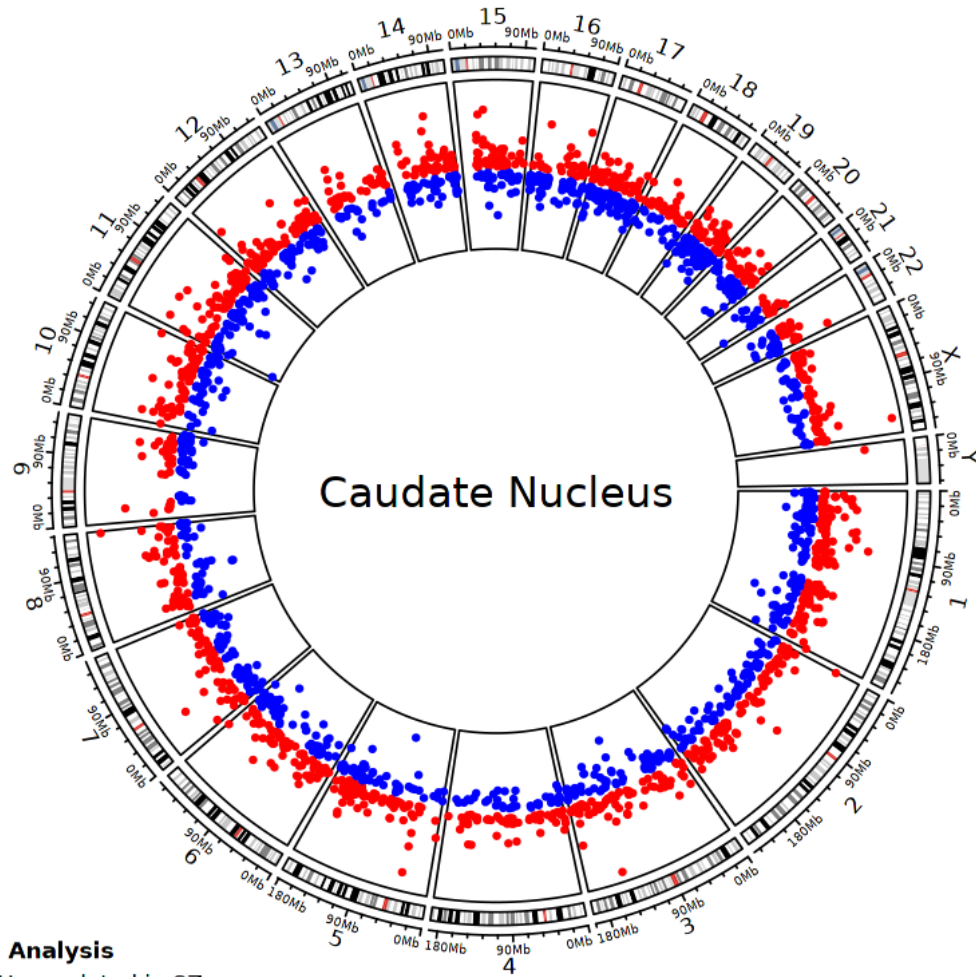
1.2 Circos Plot

```
[5]: plot_circos <- function(bed_list, colors, tissue_name){
  lgd_points = Legend(at=c("Upregulated in SZ", "Downregulated in SZ"),
    type="points", legend_gp=gpar(col = c("red", "blue")),
    title_position="topleft", title="DE Analysis",
    background="#FFFFFF")

  circos.clear() # clear plot if there is any
  circos.par("start.degree" = 0) # rotate 90 degrees
  # initialize with ideogram
  # use hg38, default is hg19
  circos.initializeWithIdeogram(species="hg38")
  # add logFC as a genomic track
  circos.genomicTrack(bed_list, track.height=inches_h(1),
    text(0, 0, tissue_name, cex = 1.5),
    panel.fun = function(region, value, ...) {
      i = getI(...)
      circos.genomicPoints(region, value, pch = 16,
        cex = 0.6, col = colors[i], ...)
    })
  draw(lgd_points, x=unit(5, "mm"), y=unit(5, "mm"), just=c("left", "bottom"))
}
```

```
[6]: tissue = 'Caudate Nucleus'
tissue_name = gsub(" ", "_", tissue)
```

```
[7]: plot_circos(list(bed_sz, bed_ctl), c("red", "blue"), tissue)
```



```
[8]: png(filename = paste0("significant_circos_plot_", tissue_name, ".png"))
plot_circos(list.bed_sz, bed_ctl), c("red", "blue"), tissue)
dev.off()

pdf(file = paste0("significant_circos_plot_", tissue_name, ".pdf"))
plot_circos(list.bed_sz, bed_ctl), c("red", "blue"), tissue)
dev.off()

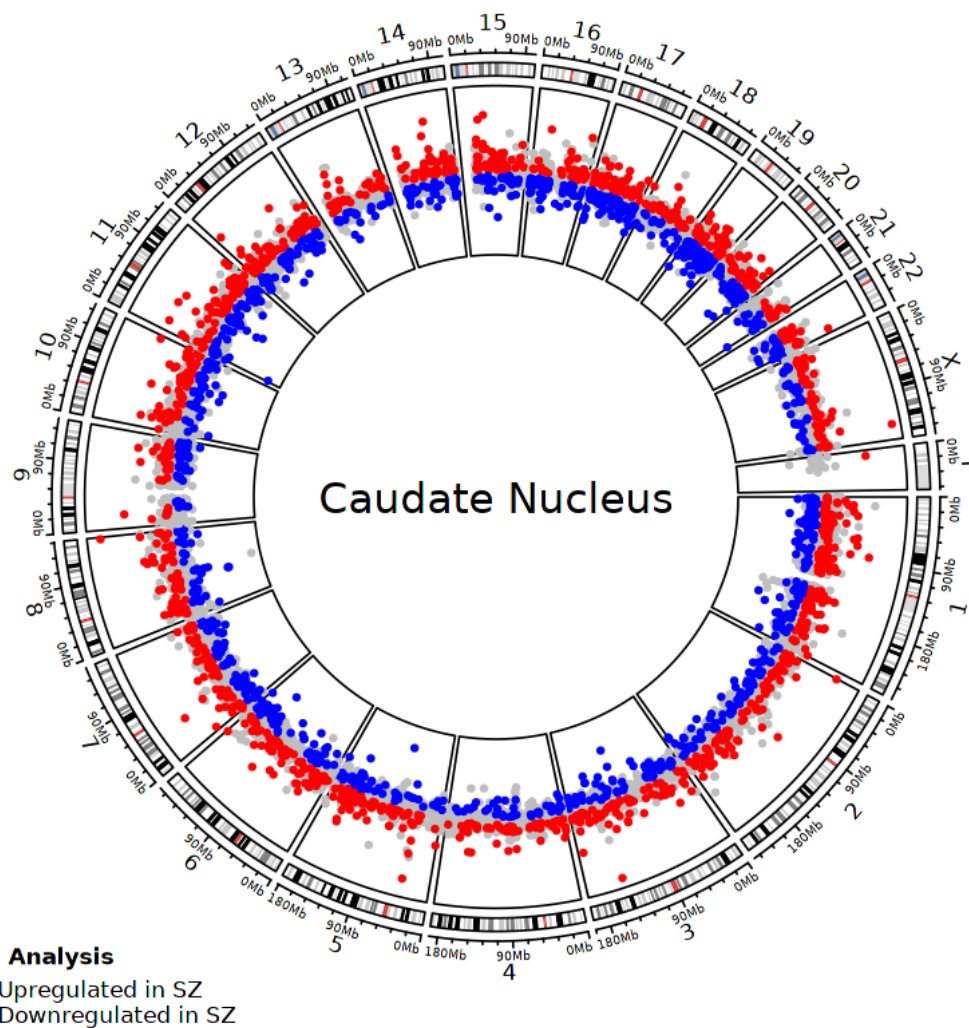
svg(filename = paste0("significant_circos_plot_", tissue_name, ".svg"))
plot_circos(list.bed_sz, bed_ctl), c("red", "blue"), tissue)
dev.off()
```

png: 2

png: 2

png: 2

```
[9]: plot_circos(list.bed_nonsig, bed_sz, bed_ctl),  
      c("gray", "red", "blue"), tissue)
```



```
[10]: png(filename = paste0("allGenes_circos_plot_", tissue_name, ".png"))  
plot_circos(list.bed_nonsig, bed_sz, bed_ctl),  
           c('gray', 'red', 'blue'), tissue)  
dev.off()  
  
pdf(file = paste0("allGenes_circos_plot_", tissue_name, ".pdf"))
```

```

plot_circos(list.bed_nonsig, bed_sz, bed_ctl),
            c('gray', 'red', 'blue'), tissue)
dev.off()

svg(filename = paste0("allGenes_circos_plot_", tissue_name, ".svg"))
plot_circos(list.bed_nonsig, bed_sz, bed_ctl),
            c('gray', 'red', 'blue'), tissue)
dev.off()

```

png: 2

png: 2

png: 2

1.3 Reproducibility Information

```

[11]: Sys.time()
proc.time()
options(width = 120)
sessioninfo::session_info()

```

```
[1] "2021-09-06 08:41:54 EDT"
```

```

  user  system elapsed
51.516   0.743   59.749

```

Session info

```
setting  value
```

```
version  R version 4.0.3 (2020-10-10)
```

```
os       Arch Linux
```

```
system   x86_64, linux-gnu
```

```
ui       X11
```

```
language (EN)
```

```
collate  en_US.UTF-8
```

```
ctype    en_US.UTF-8
```

```
tz       America/New_York
```

```
date     2021-09-06
```

Packages

package	* version	date	lib	source
AnnotationDbi	1.52.0	2020-10-27	[1]	Bioconductor
askpass	1.1	2019-01-13	[1]	CRAN (R 4.0.2)
assertthat	0.2.1	2019-03-21	[1]	CRAN (R 4.0.2)
backports	1.2.1	2020-12-09	[1]	CRAN (R 4.0.2)
base64enc	0.1-3	2015-07-28	[1]	CRAN (R 4.0.2)
Biobase	2.50.0	2020-10-27	[1]	Bioconductor
BiocFileCache	1.14.0	2020-10-27	[1]	Bioconductor
BiocGenerics	0.36.1	2021-04-16	[1]	Bioconductor
biomaRt	* 2.46.3	2021-02-09	[1]	Bioconductor

bit	4.0.4	2020-08-04	[1]	CRAN	(R 4.0.2)
bit64	4.0.5	2020-08-30	[1]	CRAN	(R 4.0.2)
blob	1.2.2	2021-07-23	[1]	CRAN	(R 4.0.3)
broom	0.7.9	2021-07-27	[1]	CRAN	(R 4.0.3)
cachem	1.0.6	2021-08-19	[1]	CRAN	(R 4.0.3)
Cairo	1.5-12.2	2020-07-07	[1]	CRAN	(R 4.0.2)
cellranger	1.1.0	2016-07-27	[1]	CRAN	(R 4.0.2)
circlize	* 0.4.13	2021-06-09	[1]	CRAN	(R 4.0.3)
cli	3.0.1	2021-07-17	[1]	CRAN	(R 4.0.3)
clue	0.3-59	2021-04-16	[1]	CRAN	(R 4.0.3)
cluster	2.1.0	2019-06-19	[2]	CRAN	(R 4.0.3)
colorspace	2.0-2	2021-06-24	[1]	CRAN	(R 4.0.3)
ComplexHeatmap	* 2.6.2	2020-11-12	[1]	Bioconductor	
crayon	1.4.1	2021-02-08	[1]	CRAN	(R 4.0.3)
curl	4.3.2	2021-06-23	[1]	CRAN	(R 4.0.3)
data.table	1.14.0	2021-02-21	[1]	CRAN	(R 4.0.3)
DBI	1.1.1	2021-01-15	[1]	CRAN	(R 4.0.2)
dbplyr	2.1.1	2021-04-06	[1]	CRAN	(R 4.0.3)
digest	0.6.27	2020-10-24	[1]	CRAN	(R 4.0.2)
dplyr	* 1.0.7	2021-06-18	[1]	CRAN	(R 4.0.3)
ellipsis	0.3.2	2021-04-29	[1]	CRAN	(R 4.0.3)
evaluate	0.14	2019-05-28	[1]	CRAN	(R 4.0.2)
fansi	0.5.0	2021-05-25	[1]	CRAN	(R 4.0.3)
fastmap	1.1.0	2021-01-25	[1]	CRAN	(R 4.0.2)
forcats	* 0.5.1	2021-01-27	[1]	CRAN	(R 4.0.2)
fs	1.5.0	2020-07-31	[1]	CRAN	(R 4.0.2)
generics	0.1.0	2020-10-31	[1]	CRAN	(R 4.0.2)
GetoptLong	1.0.5	2020-12-15	[1]	CRAN	(R 4.0.2)
ggplot2	* 3.3.5	2021-06-25	[1]	CRAN	(R 4.0.3)
GlobalOptions	0.1.2	2020-06-10	[1]	CRAN	(R 4.0.2)
glue	1.4.2	2020-08-27	[1]	CRAN	(R 4.0.2)
gtable	0.3.0	2019-03-25	[1]	CRAN	(R 4.0.2)
haven	2.4.3	2021-08-04	[1]	CRAN	(R 4.0.3)
hms	1.1.0	2021-05-17	[1]	CRAN	(R 4.0.3)
htmltools	0.5.2	2021-08-25	[1]	CRAN	(R 4.0.3)
httr	1.4.2	2020-07-20	[1]	CRAN	(R 4.0.2)
IRanges	2.24.1	2020-12-12	[1]	Bioconductor	
IRdisplay	1.0	2021-01-20	[1]	CRAN	(R 4.0.2)
IRkernel	1.2	2021-05-11	[1]	CRAN	(R 4.0.3)
jsonlite	1.7.2	2020-12-09	[1]	CRAN	(R 4.0.2)
lifecycle	1.0.0	2021-02-15	[1]	CRAN	(R 4.0.3)
lubridate	1.7.10	2021-02-26	[1]	CRAN	(R 4.0.3)
magrittr	2.0.1	2020-11-17	[1]	CRAN	(R 4.0.2)
matrixStats	0.60.1	2021-08-23	[1]	CRAN	(R 4.0.3)
memoise	2.0.0	2021-01-26	[1]	CRAN	(R 4.0.2)
modelr	0.1.8	2020-05-19	[1]	CRAN	(R 4.0.2)
munsell	0.5.0	2018-06-12	[1]	CRAN	(R 4.0.2)
openssl	1.4.5	2021-09-02	[1]	CRAN	(R 4.0.3)

pbmZMQ	0.3-5	2021-02-10	[1]	CRAN	(R 4.0.3)
pillar	1.6.2	2021-07-29	[1]	CRAN	(R 4.0.3)
pkgconfig	2.0.3	2019-09-22	[1]	CRAN	(R 4.0.2)
png	0.1-7	2013-12-03	[1]	CRAN	(R 4.0.2)
prettyunits	1.1.1	2020-01-24	[1]	CRAN	(R 4.0.2)
progress	1.2.2	2019-05-16	[1]	CRAN	(R 4.0.2)
purrr	* 0.3.4	2020-04-17	[1]	CRAN	(R 4.0.2)
R6	2.5.1	2021-08-19	[1]	CRAN	(R 4.0.3)
rappdirs	0.3.3	2021-01-31	[1]	CRAN	(R 4.0.2)
RColorBrewer	1.1-2	2014-12-07	[1]	CRAN	(R 4.0.2)
Rcpp	1.0.7	2021-07-07	[1]	CRAN	(R 4.0.3)
readr	* 2.0.1	2021-08-10	[1]	CRAN	(R 4.0.3)
readxl	1.3.1	2019-03-13	[1]	CRAN	(R 4.0.2)
repr	1.1.3	2021-01-21	[1]	CRAN	(R 4.0.2)
reprex	2.0.1	2021-08-05	[1]	CRAN	(R 4.0.3)
rjson	0.2.20	2018-06-08	[1]	CRAN	(R 4.0.2)
rlang	0.4.11	2021-04-30	[1]	CRAN	(R 4.0.3)
RSQLite	2.2.8	2021-08-21	[1]	CRAN	(R 4.0.3)
rstudioapi	0.13	2020-11-12	[1]	CRAN	(R 4.0.2)
rvest	1.0.1	2021-07-26	[1]	CRAN	(R 4.0.3)
S4Vectors	0.28.1	2020-12-09	[1]	Bioconductor	
scales	1.1.1	2020-05-11	[1]	CRAN	(R 4.0.2)
sessioninfo	1.1.1	2018-11-05	[1]	CRAN	(R 4.0.2)
shape	1.4.6	2021-05-19	[1]	CRAN	(R 4.0.3)
stringi	1.7.4	2021-08-25	[1]	CRAN	(R 4.0.3)
stringr	* 1.4.0	2019-02-10	[1]	CRAN	(R 4.0.2)
tibble	* 3.1.4	2021-08-25	[1]	CRAN	(R 4.0.3)
tidyr	* 1.1.3	2021-03-03	[1]	CRAN	(R 4.0.3)
tidyselect	1.1.1	2021-04-30	[1]	CRAN	(R 4.0.3)
tidyverse	* 1.3.1	2021-04-15	[1]	CRAN	(R 4.0.3)
tzdb	0.1.2	2021-07-20	[1]	CRAN	(R 4.0.3)
utf8	1.2.2	2021-07-24	[1]	CRAN	(R 4.0.3)
uuid	0.1-4	2020-02-26	[1]	CRAN	(R 4.0.2)
vctrs	0.3.8	2021-04-29	[1]	CRAN	(R 4.0.3)
withr	2.4.2	2021-04-18	[1]	CRAN	(R 4.0.3)
XML	3.99-0.7	2021-08-17	[1]	CRAN	(R 4.0.3)
xml2	1.3.2	2020-04-23	[1]	CRAN	(R 4.0.2)

[1] /home/jbenja13/R/x86_64-pc-linux-gnu-library/4.0

[2] /usr/lib/R/library