# GS3DKViz_Arc_Plots

James Dalgleish

1/13/2021

```r
knitr::opts_chunk$set(echo = TRUE,message=F,error=F,warning=F)
knitr::opts_chunk$set(root.dir="./inst/extdata/")
library(Gviz)
library(GenomicInteractions)
library(magrittr)
library(GS3DKViz)
```

We'll start by loading a few libraries that will enable us to generate plots and loading the data.

For simple plots, we've created a convenient function for plotting.

```r
gs3dk_gint=readGint(system.file(
  "extdata","signficant_genes_genoscan_3D_knock.csv",
  package = "GS3DKViz"))
  filtered_gs3dk_gint=gs3dk_gint[(gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr:
  plotInteraction(gint = filtered_gs3dk_gint,
  chr=unique(janitor::clean_names(tibble::as_tibble(mcols(filtered_gs3dk_gint)))$chr),
  bounds_offset=1.5e4,
  main="Promoter-enhancer interactions, FEV1"
  )
```

Promoter–enhancer interactions, FEV1
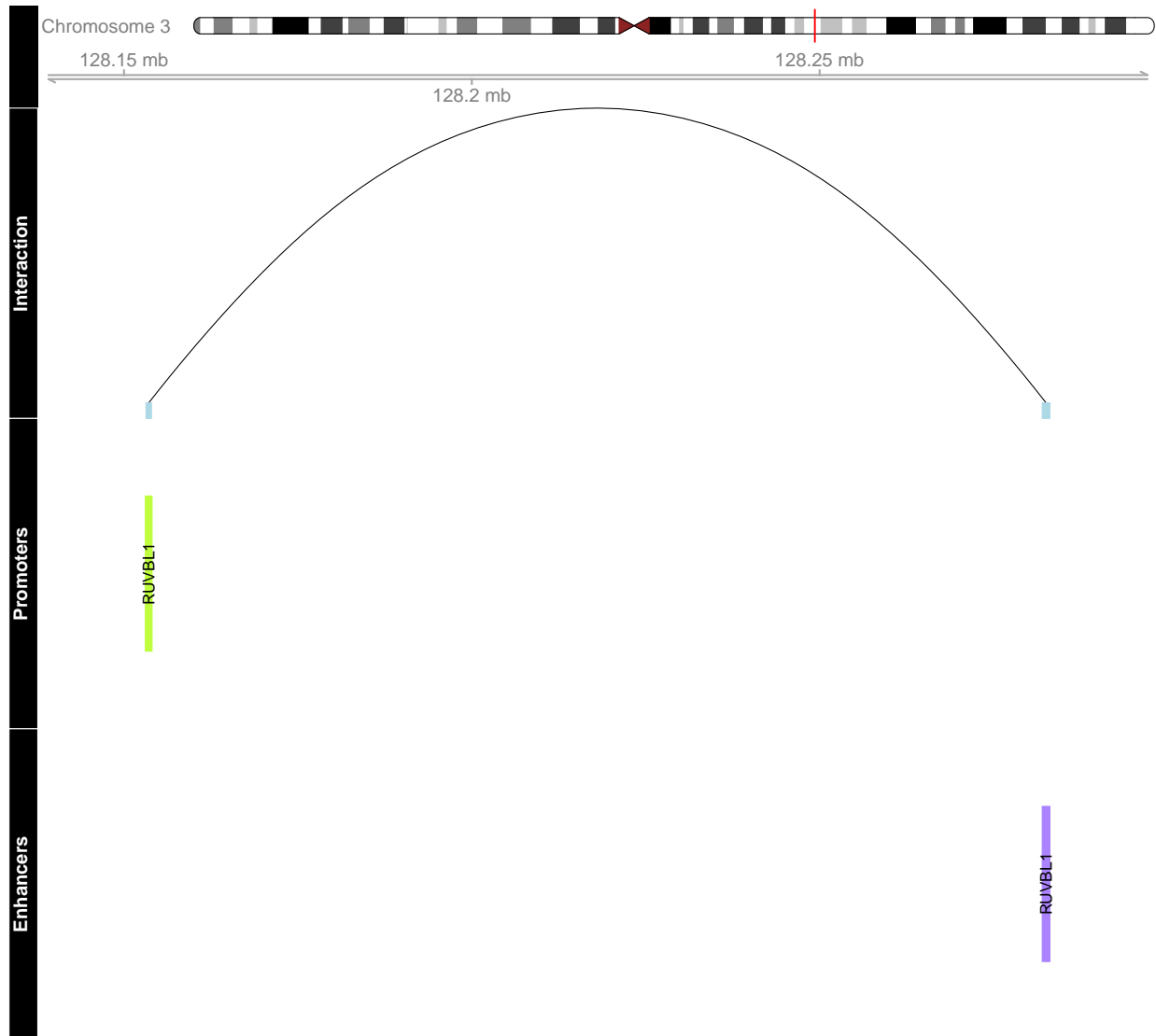

```
## NULL
```

A details plot is below, with unconnected promoters in orange. The code behind this plot will be explained towards the end of the tutorial.

```r
extra_prom_gr_tm=data.table::fread(system.file(
"extdata","more_promoters_topmed.csv",
  package = "GS3DKViz")
) %>% janitor::clean_names( ) %>%  GRanges()
GS3DKViz::plotInteractionDetail(gint = gs3dk_gint,
  chr="chr19", #unique(janitor::clean_names(tibble::as_tibble(mcols(gs3dk_gint)))$chr)
  bounds_offset=0,#1.5e4
  main="Promoter-enhancer interactions, AD",
  extra_prom_gr = extra_prom_gr_tm )
```

Promoter–enhancer interactions, AD
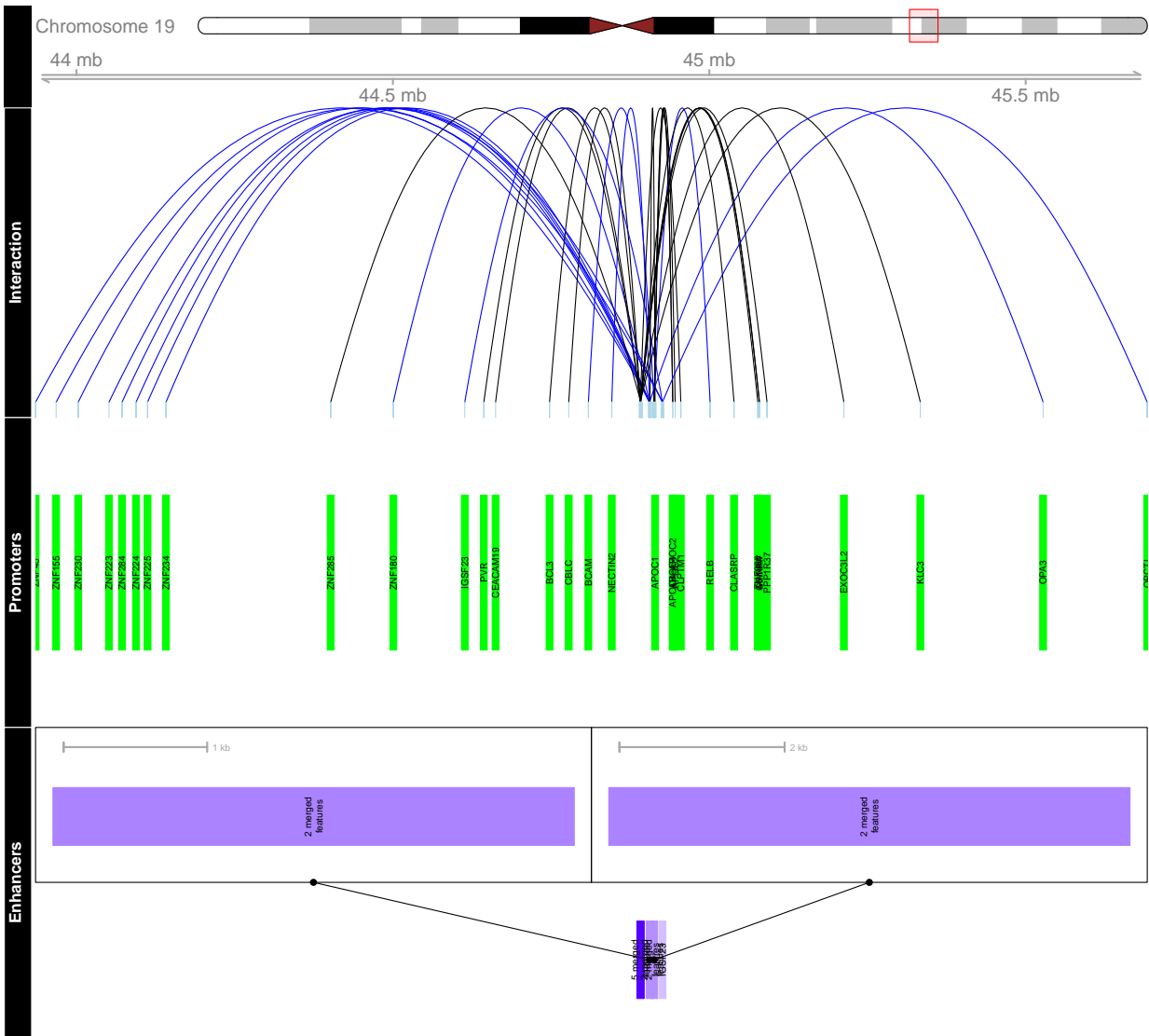
NULL

I have also included a custom function to convert data frames to GenomicInteractions objects, but in any way that you create such an object is fine. The counts column will contain the gene-enhancer score (from either ABC or GeneCards). I implement a scaling algorithm here so that the larger scaled GeneHancer does not have vastly larger arc height. GenomicInteractions Objects just require 6 columns in the following order: chr1-chromosome of the first range start1-start positions of the first range end1-end positions of the first range chr2-Chromosomes for the second set of ranges start1-start position of the second range start2-start position of the second range Thus, we'll need to split our data into start and end positions for both promoters (range 1, which can be accessed by the standard anchorOne function), and enhancers (range 2, which can be accessed with the anchorTwo function). I will provide the detailed code here on how our results were generated, but if desired, one can skip to the next section and just pull in our sample data. If executing this code, be sure to install data.table and tidyverse.

Essentially, GenomicInteractions object will work to reproduce the plots, so long as the counts have a

meaningful measure of interaction score and all the interactions are on the same chromosome.

```r
#read in data
gs3dk_gint=data.table::fread(system.file(
"extdata","signficant_genes_genoscan_3D_knock.csv",
  package = "GS3DKViz")) %>%
  #automatically clean column names, converting to standard format.
janitor::clean_names() %>%
  #split the gene position into start and end
  tidyr::separate(gene_position,c("genestart","geneend"),  "-") %>%
    #split the promoter position into start and end
  tidyr::separate(promoter_position,c("promoterstart","promoterend"),  "-") %>%
      #split the enhancer position into start and end
  tidyr::separate(best_enhancer_position,c("enhstart","enhend"),  "-") %>%
  #convert the promoter starts to numeric, assign replacements for missing data.
  dplyr::mutate(promoterstart=as.numeric(promoterstart),chr1=chr,chr2=chr,
               #replace missing promoter starts with TSS
  promoterstart=ifelse(is.na(promoterstart),genestart,promoterstart),
  #replace missing promoter ends with TSS
  promoterend=ifelse(is.na(promoterend),genestart,promoterend)) %>%
  #create metadata column that determines whether an interaction is from
  #the ABC model or from GeneHancer.
  dplyr::mutate(enhancer_type=ifelse(grepl(pattern="GH",x=enhancer_id),"GH","ABC")) %>%
  #rearrange data columns for GenomicInteractions Objects
  dplyr::select(chr1,promoterstart,promoterend,chr2,enhstart,enhend,dplyr::everything()) %>%
  #rename columns to standard GenomicInteractions format.
  dplyr::rename(start1=promoterstart,end1=promoterend,start2=enhstart,end2=enhend) %>%
  #remove chromosomes from second range (ranges should only be positions).
  dplyr::mutate(start2=gsub("chr19: ","",start2)) %>%
  #create the GI.
  makeGenomicInteractionsFromDataFrame() %>%
  #remove arcs where the promoter and enhancers overlap.
  .[which(calculateDistances(.)!=0),]
#with the GenomicInteractions object created, we'll normalize the interaction score
#and store it in the counts column for plotting.
mcols(gs3dk_gint)=gs3dk_gint %>% tibble::as_tibble() %>% janitor::clean_names() %>%
  dplyr::group_by(enhancer_type) %>% dplyr::mutate(interaction_score=interaction_score/mean(interaction_
mcols(gs3dk_gint)$counts<-ifelse(!is.na(mcols(gs3dk_gint)$interaction_score),
                                 mcols(gs3dk_gint)$interaction_score,1)

#save the object

save(list = c("gs3dk_gint"),file = "gs3dk_sample_data.rda",compress = "xz",
     compression_level = 9)
```

Now with the data loaded, we can create tracks with Gviz and plot them. First we'll load the data, extract promoter and enhancer granges objects We'll use PSMA4 to get the promoter, but let the bounds be defined by the min and max ranges in the available data.

```r
load(system.file(
"extdata","gs3dk_sample_data.rda",
  package = "GS3DKViz"))
#extract promoter and enhancer data from GenomicInteractions object.
promoter_gr=anchorOneWithMetadata(gs3dk_gint)
enhancer_gr=anchorTwoWithMetadata(gs3dk_gint)
geneSymbol="PSMA4"
```

```r
#set the chromosome containing the PSMA4 gene to be the plotted chromosome.
selected_chromosome=as.character(gs3dk_gint %>% as.data.frame() %>%
  janitor::clean_names() %>% dplyr::filter(gene_id==geneSymbol))
```

We will now construct the interaction, promoter, and enhancer tracks.

```r
interaction_track <- InteractionTrack(gs3dk_gint, name = "Interaction", chromosome = "chr15")

bounds=c(gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FEV1/FVC")
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FEV1/FVC") %>% dply
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FEV1/FVC") %>% dply
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FEV1/FVC") %>% dply
promoterTrack <- AnnotationTrack(promoter_gr, genome="hg38", name="Promoters",
                           id=mcols(promoter_gr)$gene_id, featureAnnotation="id")
enhancerTrack <- AnnotationTrack(enhancer_gr, genome="hg38", name="Enhancers",
                           id=mcols(enhancer_gr)$gene_id, featureAnnotation="id")
feature(enhancerTrack)<-mcols(gs3dk_gint)$enhancer_type

displayPars(promoterTrack) <- list(fill = "olivedrab1", col = NA,
                                  fontcolor.feature = "black", fontsize=8,
                                  just.group="below",rotation=90,rotation.group=90,rotation.item=90)
displayPars(enhancerTrack) <- list(fill = "mediumpurple1", col = NA,
                                  fontcolor.feature = "black", fontsize=10,
                                  just.group="below",rotation.item=90,
                                  collapse=T,mergeGroups=T,showOverplotting=T,groupAnnotation="group",
displayPars(interaction_track) <- list(fill = "deepskyblue", col = NA,
                                  fontcolor.feature = "black", fontsize=8,
                                  just.group="below",plot.anchors=T,plot.outside=T,col.outside="lightbl
                                  interaction.dimension="height",
                                  col.interactions="black",
                                  plot.trans=T,
                                           fontsize.legend=200
                                  )
itrack <- IdeogramTrack(genome = "hg38", chromosome =as.character(gs3dk_gint %>% as.data.frame() %>% jan
gtrack <- GenomeAxisTrack()
displayPars(enhancerTrack)=list(rotation.item=0)
plotTracks(list(itrack,gtrack,interaction_track,promoterTrack,enhancerTrack),
          chromosome=as.character(gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr:
          from = (min(bounds))-1.5e4,
          to = (max(bounds))+1.5e4,
          type = c("b"),
          showSampleNames = TRUE,
          cex.sampleNames = 0.6,
          cex.main=3,
          fontsize=18,fontsize.item=12,collapse=T,min.width=5,mergeGroups=T
          ,stacking="squish"
          ,main="Promoter-enhancer interactions, FEV1/FVC",
          background.title = "black"
)
```

Promoter–enhancer interactions, FEV1/FVC
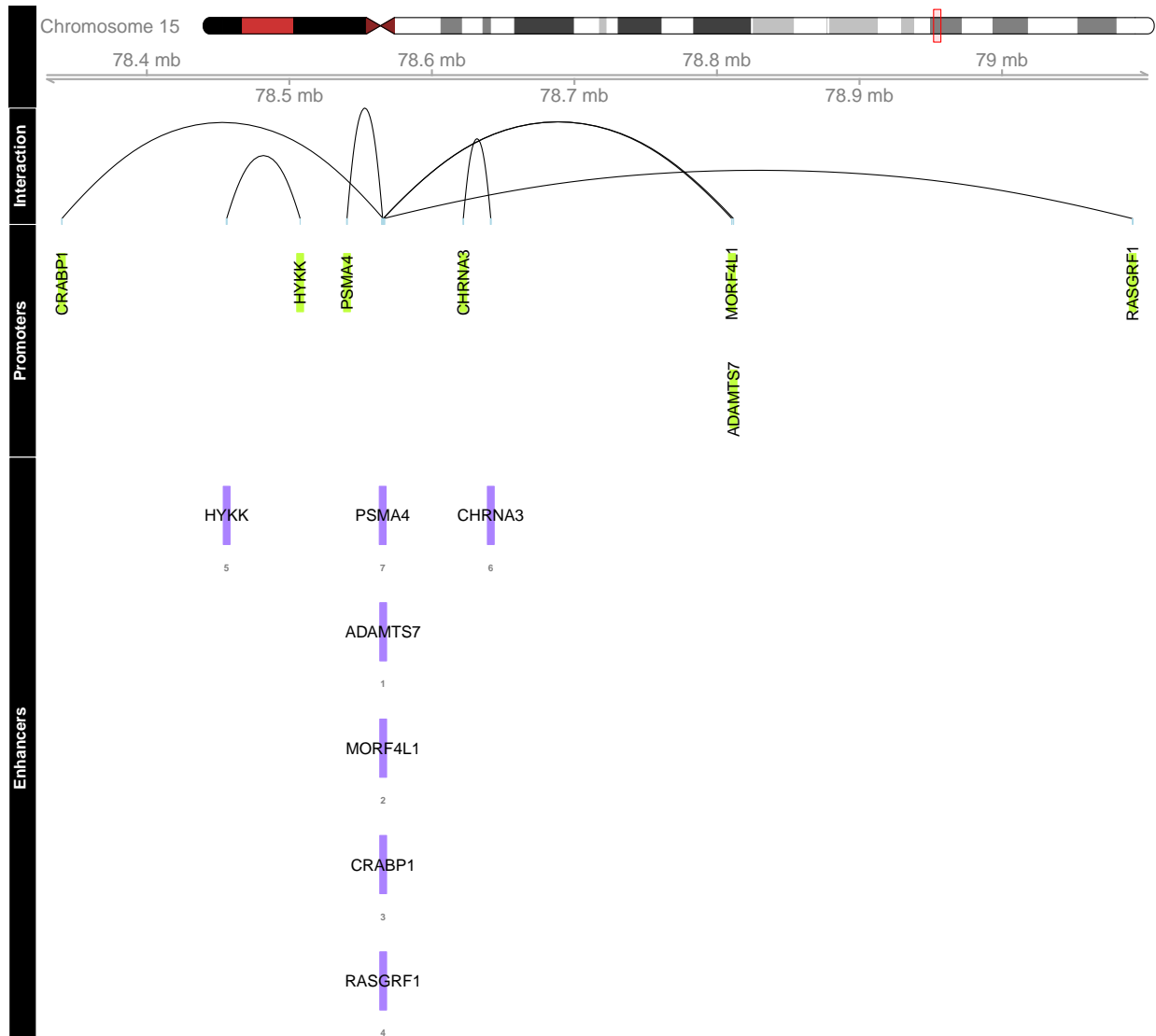
Now, we'll add more promoters from a separate csv file that do not have enhancer links in a separate track (which will be vertically combined later). Track properties are copied from the previous promoter track and color is changed to dark orange. We previously did the color change in Illustrator and combined the two tracks vertically using a separate plot that used stacking=dense. We exported to postscript from this point and embedded the fonts

```
extra_prom_gr_tm=data.table::fread(system.file(
"extdata","more_promoters_topmed.csv",
  package = "GS3DKViz")
) %>% janitor::clean_names( ) %>%  GRanges()
extraPromoterTrackTm <- AnnotationTrack(extra_prom_gr_tm, genome="hg38", name=" Promoters",
                        id=mcols(extra_prom_gr_tm)$gene_id,  featureAnnotation="id")
displayPars(extraPromoterTrackTm)=displayPars(promoterTrack)
displayPars(extraPromoterTrackTm) <- list(fill = "darkorange")
#postscript("fev1-fvc-psma4-dense-extra-prom.eps",width=27,height=9,family="sans")
```

```r
plotTracks(list(itrack,gtrack,interaction_track,promoterTrack,extraPromoterTrackTm,enhancerTrack),
        chromosome=as.character(gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr:
        from = (min(bounds))-1.5e4,
        to = (max(bounds))+1.5e4,
        type = c("b"),
        showSampleNames = TRUE,
        cex.sampleNames = 0.6,
        cex.main=3,
        fontsize=18,fontsize.item=12,collapse=T,min.width=5,mergeGroups=T
        ,stacking="squish" #comment out to restore stacking
        ,main="Promoter-enhancer interactions, FEV1/FVC",
        background.title = "black"
)
```

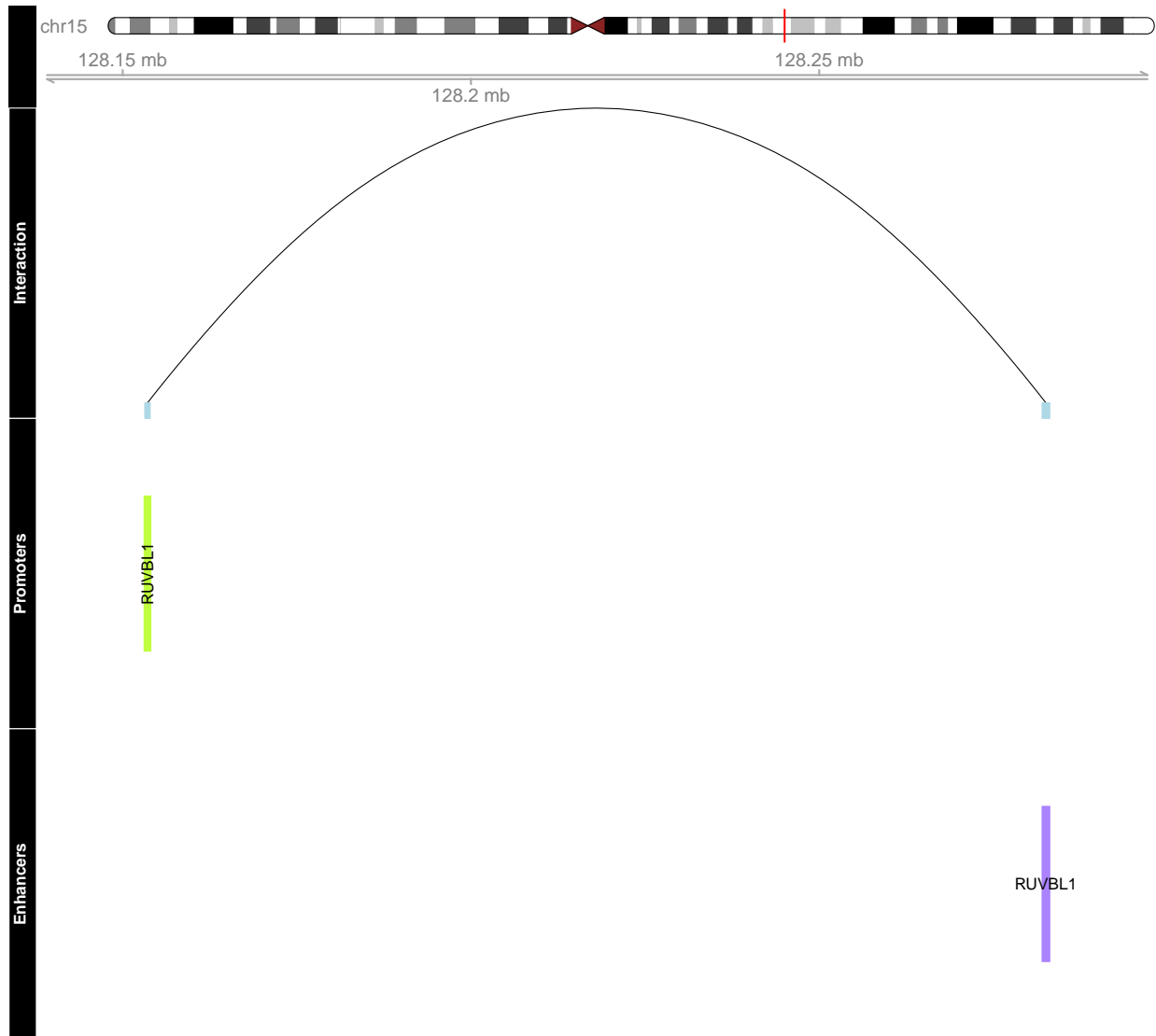# Promoter–enhancer interactions, FEV1/FVC
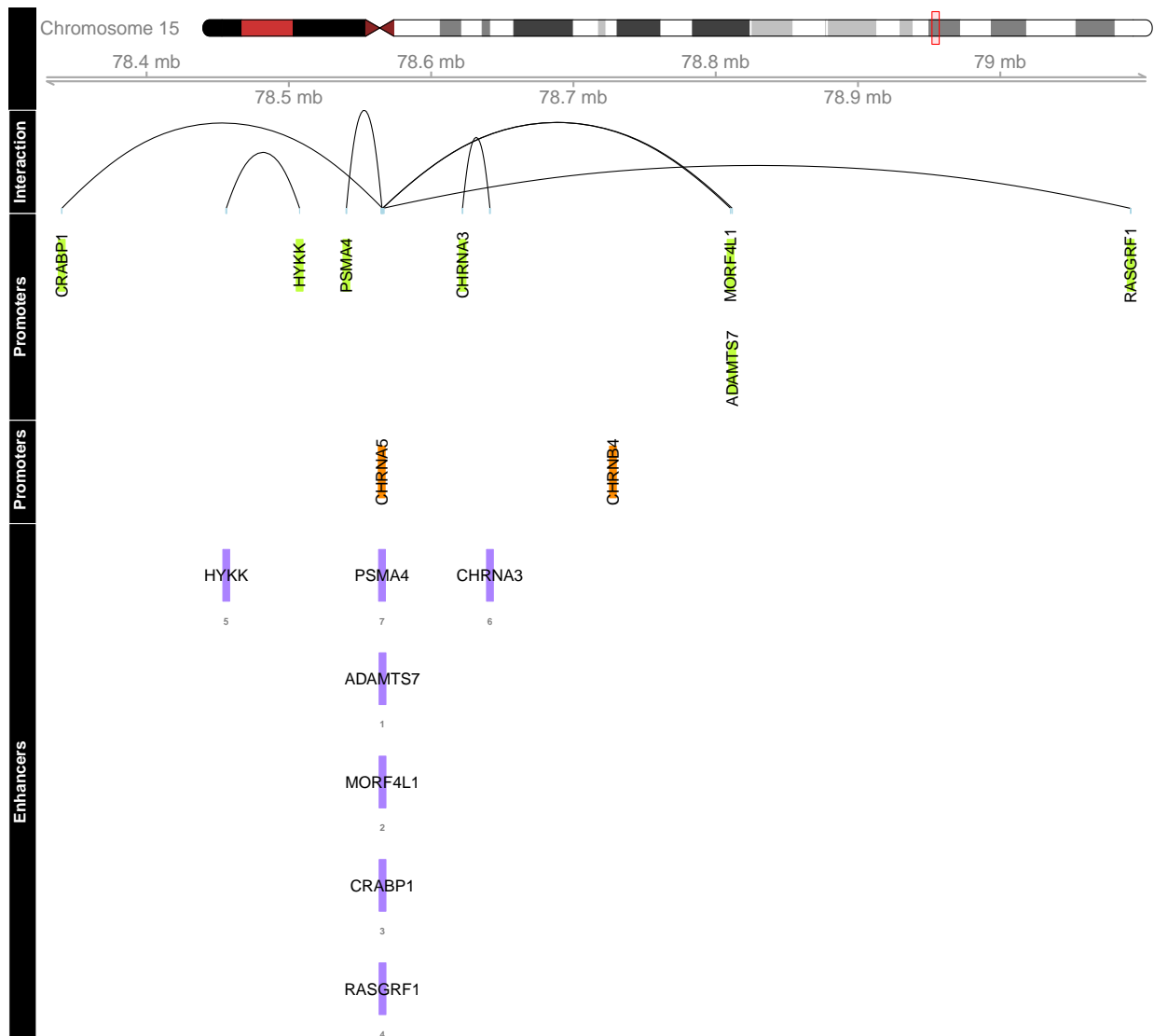
```
#dev.off()
```

Now we'll do FEV1 and FVC by changing the bounds and the chromosome to those matching detected FEV1 and FVC promoters.

```
interaction_track <- InteractionTrack(gs3dk_gint, name = "Interaction", chromosome = as.character(gs3dk
displayPars(interaction_track)=list(col.interactions="black")
bounds=c(gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FEV1") %>%
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FEV1") %>% dplyr::pu
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FEV1") %>% dplyr::pu
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FEV1") %>% dplyr::pu

plotTracks(list(itrack,gtrack,interaction_track,promoterTrack,enhancerTrack),
           chromosome=as.character(gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr:
           from = (min(bounds))-1.5e4,
           to = (max(bounds))+1.5e4,
           type = c("b"),
           showSampleNames = TRUE,
           cex.sampleNames = 0.6,
           cex.main=3,
           fontsize=18,fontsize.item=12,collapse=T,min.width=5,mergeGroups=T
           ,stacking="dense" #comment out to restore stacking
           ,main="Promoter-enhancer interactions, FEV1",
           background.title = "black"
)
```

Promoter–enhancer interactions, FEV1


```
interaction_track <- InteractionTrack(gs3dk_gint, name = "Interaction", chromosome = as.character(gs3dk
displayPars(interaction_track)=list(col.interactions="black")
bounds=c(gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FVC") %>% c
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FVC") %>% dplyr::pul
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FVC") %>% dplyr::pul
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FVC") %>% dplyr::pul
plotTracks(list(itrack,gtrack,interaction_track,promoterTrack,enhancerTrack),
           chromosome=as.character(gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr:
           from = (min(bounds))-1.5e4,
           to = (max(bounds))+1.5e4,
           type = c("b"),
           showSampleNames = TRUE,
           cex.sampleNames = 0.6,
           cex.main=3,
```

```
                fontsize=18,fontsize.item=12,collapse=T,min.width=5,mergeGroups=T
                ,stacking="dense" #comment out to restore stacking
                ,main="Promoter-enhancer interactions, FVC",
                background.title = "black"
)
```

Promoter–enhancer interactions, FVC

Another example using alternative colors for the interactions:

```
annotateInteractions(GIObject = gs3dk_gint,annotations =  list(promoter=promoter_gr, enhancer=enhancer_g

knockoff_genes=mcols(gs3dk_gint) %>% tibble::as_tibble() %>% dplyr::filter(identified_by_knockoff_or_not

gs3dk_gint@regions@elementMetadata$node.class=(mcols(regions(gs3dk_gint)) %>% tibble::as_tibble() %>% dp

interaction_track <- InteractionTrack(gs3dk_gint, name = "Interaction", chromosome = as.character(gs3dk_
```
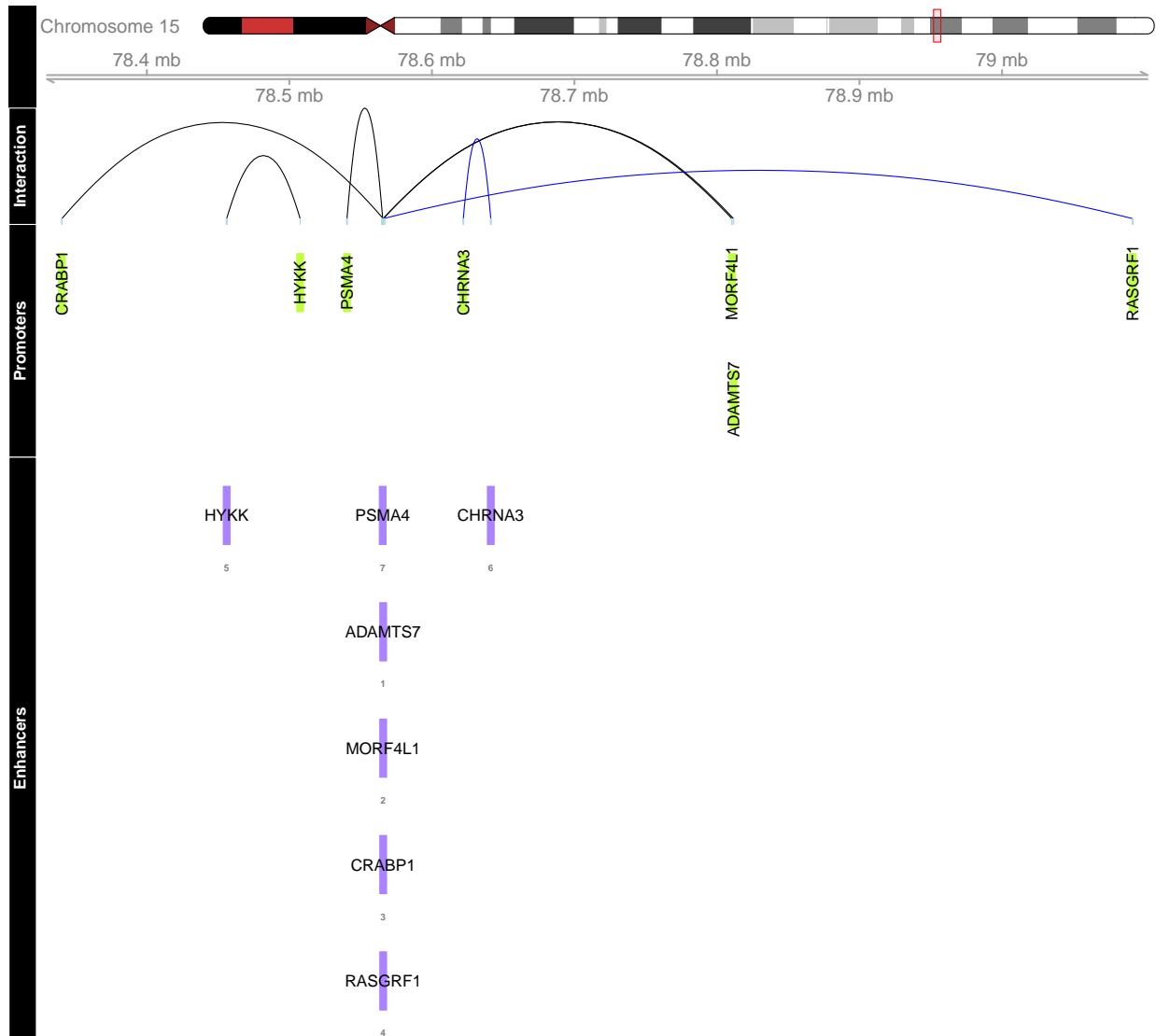
```r
displayPars(interaction_track)=list(col.interactions="black",col.interaction.types=c('knockoff_detected
bounds=c(gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FEV1/FVC")
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FEV1/FVC") %>% dply
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FEV1/FVC") %>% dply
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="FEV1/FVC") %>% dply
plotTracks(list(itrack,gtrack,interaction_track,promoterTrack,enhancerTrack),
          chromosome=as.character(gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr:
          from = (min(bounds))-1.5e4,
          to = (max(bounds))+1.5e4,
          type = c("b"),
          showSampleNames = TRUE,
          cex.sampleNames = 0.6,
          cex.main=3,
          fontsize=18,fontsize.item=12,collapse=T,min.width=5,mergeGroups=T
          ,stacking="squish" #comment out to restore stacking
          ,main="Promoter-enhancer interactions, FEV1/FVC",
          background.title = "black"
)
```

Promoter–enhancer interactions, FEV1/FVC

Lastly, we'll do a details plot, which requires a details function and selection function. It's a bit complex, but the code is below.

```
annotateInteractions(GIObject = gs3dk_gint,annotations =  list(promoter=promoter_gr, enhancer=enhancer_g
knockoff_genes=mcols(gs3dk_gint) %>% tibble::as_tibble() %>% dplyr::filter(identified_by_knockoff_or_no
gs3dk_gint@regions@elementMetadata$node.class=(mcols(regions(gs3dk_gint)) %>% tibble::as_tibble() %>% dp
interaction_track <- InteractionTrack(gs3dk_gint, name = "Interaction", chromosome = as.character(gs3dk_
displayPars(interaction_track)=list(col.interactions="black",col.interaction.types=c('knockoff_detected-
bounds=c(gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="AD") %>% dp
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="AD") %>% dplyr::pul
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="AD") %>% dplyr::pul
gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr::filter(trait=="AD") %>% dplyr::pul
selFun <- function(identifier, start, end, track, GdObject, ...){
    gcount <- table(group(GdObject))
```

```r
    pxRange <- Gviz:::.pxResolution(min.width = 50, coord = "x")
    return((end - start) < pxRange && gcount[identifier] == 1 && (distanceToNearest(GdObject[group(GdOb
              )
}
detFun <- function(identifier, GdObject.original, ...){
  plotTracks(list(GenomeAxisTrack(scale = 0.3, size = 0.2, cex = 0.7),
        GdObject.original[group(GdObject.original) == identifier]),          add = TRUE, showTitle = FALS
}
deTrackEnh2 <- AnnotationTrack(name = "Enhancers",enhancer_gr, fun = detFun,
                              selectFun = selFun,
                              groupDetails = TRUE, details.size = 0.5,
                              detailsConnector.cex = 0.5,
                              detailsConnector.lty = "dotted",
                              shape = c("smallArrow", "arrow"),
                              groupAnnotation = "group",
                              id=mcols(enhancer_gr)$gene_id,
                              group=subjectHits(findOverlaps(enhancer_gr, reduce(enhancer_gr))),featureAnn
                  )
displayPars(deTrackEnh2) <- list(fill = "mediumpurple1", col = NA,
                                  fontcolor.feature = "black", fontsize=10,
                                  just.group="below",rotation.item=90,
                                  collapse=T,mergeGroups=T,showOverplotting=T,groupAnnotation="group",
                               groupDetails=T,detailsConnector.lty="solid",
                               detailsConnector.col="black",
                               detailsBorder.col="black",
                               detailsBorder.lty="solid",showId=F
                               ,min.distance=5,min.width=5,title="Enhancers")
extra_prom_gr=data.table::fread(system.file("extdata","more_promoters.csv",
  package = "GS3DKViz")) %>% janitor::clean_names( ) %>%  dplyr::filter(!(gene_id %in%  janitor::clean_n
itrack <- IdeogramTrack(genome = "hg38", chromosome =unique(as.character(gs3dk_gint %>% as.data.frame()
combined_prom_gr=c(promoter_gr,extra_prom_gr)
CombinedPromoterTrack <- AnnotationTrack(combined_prom_gr, genome="hg38", name="Promoters",
                            id=mcols(combined_prom_gr)$gene_id,  featureAnnotation="id",groupAnnotation
displayPars(CombinedPromoterTrack)=list(fill = "olivedrab1", col = NA,
                                  fontcolor.feature = "black", fontsize=8,                    just.group=
                                  min.width=1,min.distance=5)
feature(CombinedPromoterTrack)=c(rep("connected",length(promoter_gr)),rep("unconnected",length(extra_pr
plotTracks(list(itrack,gtrack,interaction_track,CombinedPromoterTrack,deTrackEnh2),chromosome=as.charac
           from = (min(bounds)-1.5e4),
           to = (max(bounds)+1.5e4),
           type = c("b"),
           showSampleNames = TRUE,
           cex.sampleNames = 0.6,
           cex.main=3,
           fontsize=18,fontsize.item=7,collapse=T,
           mergeGroups=T
           ,stacking="dense"
           ,main="Promoter-enhancer interactions, AD",
           background.title = "black",
           unconnected="orange",connected="green"
)
```

Promoter–enhancer interactions, AD
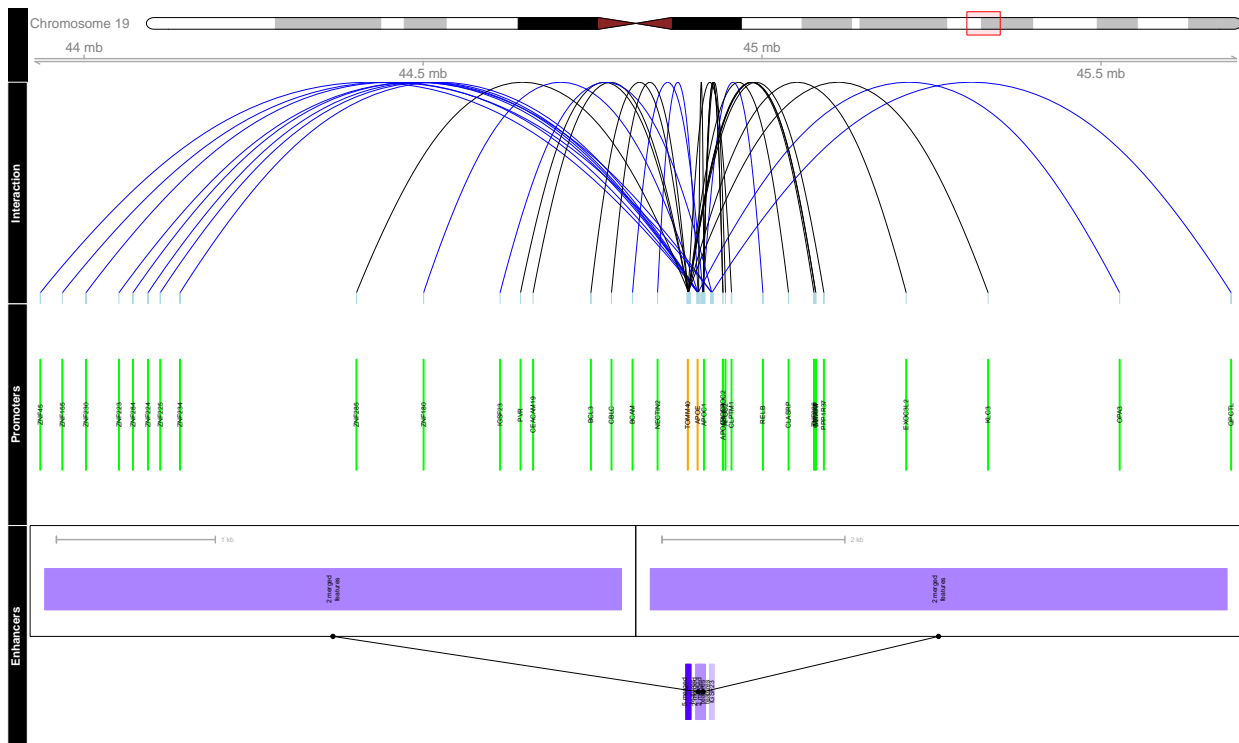
```
extra_prom_gr=data.table::fread(system.file("extdata","more_promoters.csv",
  package = "GS3DKViz")) %>% janitor::clean_names( ) %>%  dplyr::filter(!(gene_id %in%  janitor::clean_
```

To export it to EPS, for further edits, do the following:

```
filename=gsub(":","_",gsub(" ","_",paste(date(),"AD-dense-details-extra_color_arcs.eps")))
postscript(filename,width=50,height=500,family="sans")
plotTracks(list(itrack,gtrack,interaction_track,promoterTrack,extraPromoterTrack,deTrackEnh2),
           chromosome=as.character(gs3dk_gint %>% as.data.frame() %>% janitor::clean_names() %>% dplyr:
           from = (min(bounds)-1.5e4),
           to = (max(bounds)+1.5e4),
           type = c("b"),
           showSampleNames = TRUE,
           cex.sampleNames = 0.6,
           cex.main=3,
           fontsize=18,fontsize.item=7,collapse=T,
           #min.width=1,min.distance=5,
           mergeGroups=T
           ,stacking="dense" #comment out to restore stacking
           ,main="Promoter-enhancer interactions, AD",
           background.title = "black"
)
dev.off()
extrafont::embed_fonts(filename,outfile=paste0(filename,"-embed.eps"))
```