Gviz visualization: first plots

Iris Raes

29-4-2020

Installation of BiocManager and Gviz (= Bioconductor package)

```
if (!requireNamespace("BiocManager", quietly = TRUE)){
  install.packages("BiocManager")}
if (! requireNamespace("Gviz", quietly = TRUE)) {
  BiocManager::install("Gviz")
}
library(BiocManager)
```

Loading Gviz package

```
library(Gviz)
library(GenomicRanges)

setwd("C:/Users/irisr/Desktop/CSV")
```

Function to create a table with GRCh37 coordinates

```
## separate function requires tidyr package
library(tidyr)
Gviz_table <- function(filename) {
  X <- read.csv(file=filename, header = TRUE, sep=";")
  # GRCh37 coordinates in first assembly
  tablepart1 <- X[grep("^GRCh37", X$assembly1),]
  id1 <- tablepart1[1]
  Chr_1a <- tablepart1[["Chr_1"]]
  Chr_1a <- gsub("[A-Za-z]", "", Chr_1a)
  Chr_1a <- sub("", "chr", Chr_1a)
  assembly1 <- tablepart1[["assembly1"]]
  assembly1 <- gsub(".*?:", "", assembly1)
  assembly1 <- separate(data = as.data.frame(assembly1), col = assembly1,
                        into = c("start", "end"), sep = "-")
  GRCh37_assembly1 <- data.frame(id1,Chr_1a,assembly1)
  names(GRCh37_assembly1)[names(GRCh37_assembly1) == "assembly1"] <- "assembly"
  names(GRCh37_assembly1)[names(GRCh37_assembly1) == "Chr_1a"] <- "chr"
  # GRCh37 coordinates in second assembly
```

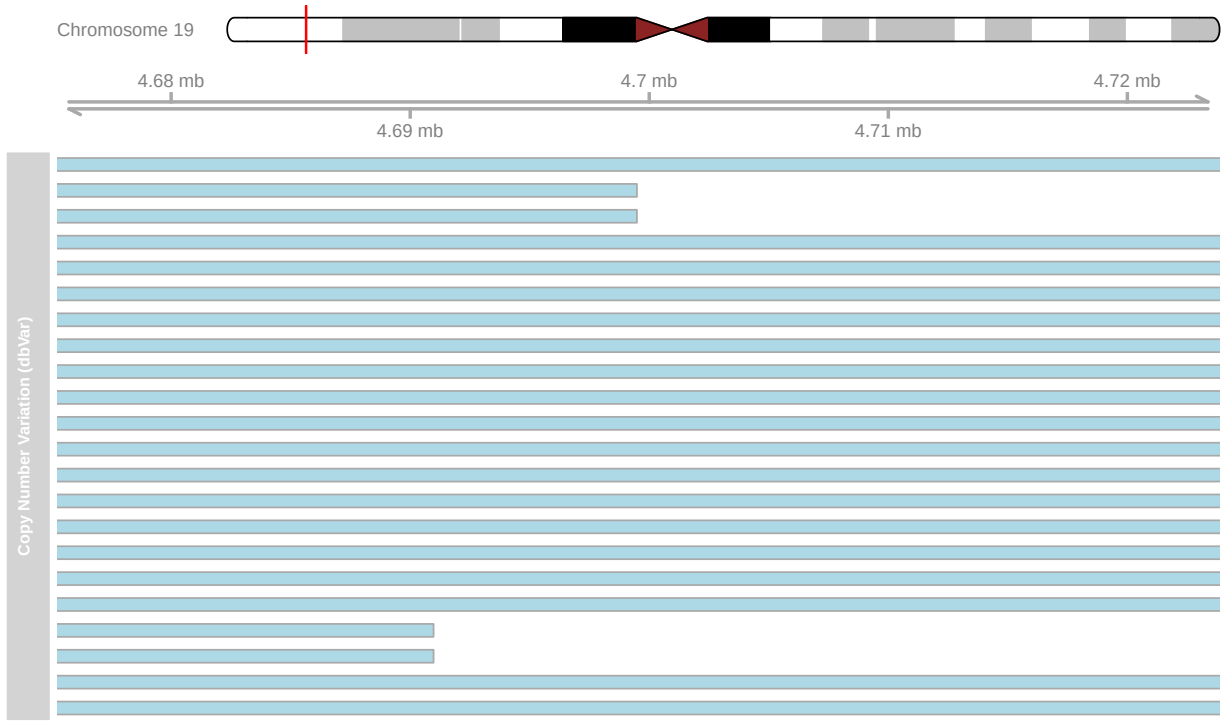
```
  tablepart2 <- X[grep("^GRCh37", X$assembly2),]
  id2 <- tablepart2[1]
  Chr_1b <- tablepart2[["Chr_2"]]
  Chr_1b <- gsub("[A-Za-z]", "", Chr_1b)
  Chr_1b <- sub("", "chr", Chr_1b)
  assembly2 <- tablepart2[["assembly2"]]
  assembly2 <- gsub(".*?:", "", assembly2)
  assembly2 <- separate(data = as.data.frame(assembly2), col = assembly2,
                        into = c("start", "end"), sep = "-")
  GRCh37_assembly2 <- data.frame(id2,Chr_1b,assembly2)
  names(GRCh37_assembly2)[names(GRCh37_assembly2) == "assembly2"] <- "assembly"
  names(GRCh37_assembly2)[names(GRCh37_assembly2) == "Chr_1b"] <- "chr"
  # Gviz table
  GRCh37_Gviz <- rbind(GRCh37_assembly1,GRCh37_assembly2)
}
```

Converting tables to GRanges objects

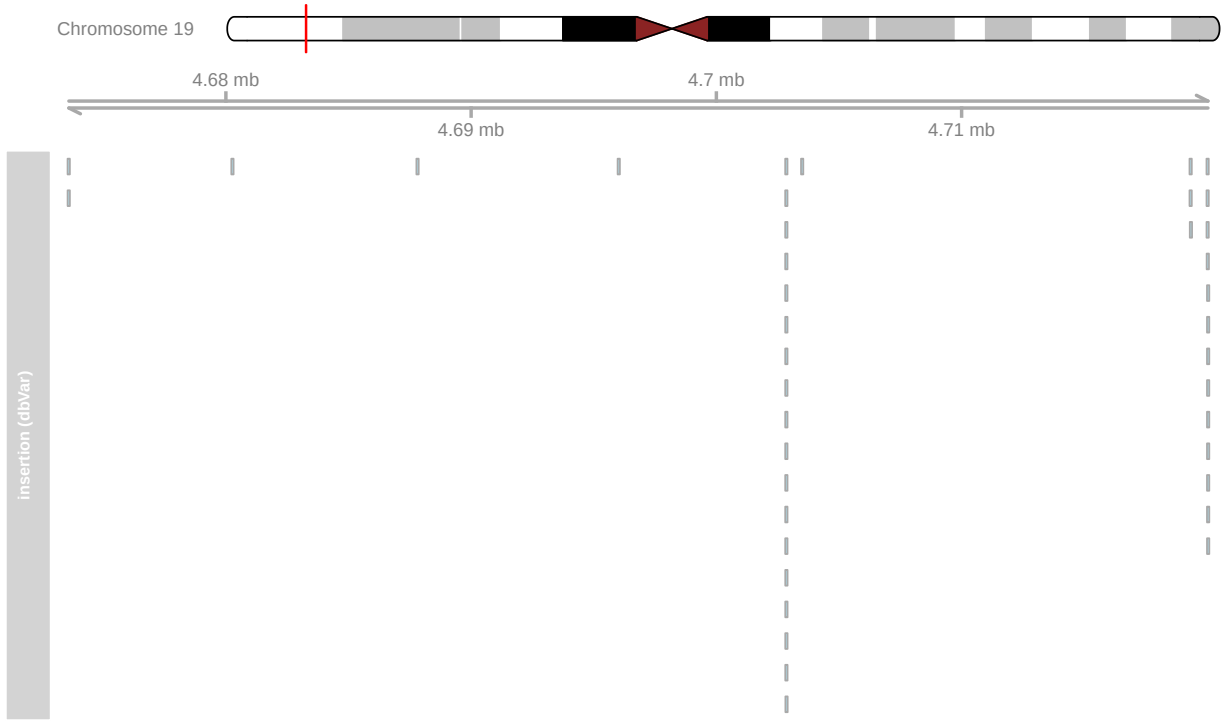
makeGRangesFromDataFrame function is used to create GRanges objects. Documentation: https://www.rdocumentation.org/packages/GenomicRanges/versions/1.24.1/topics/makeGRangesFromDataFrame

Plots for CNVs, insertions, STRs and ClinVar variants

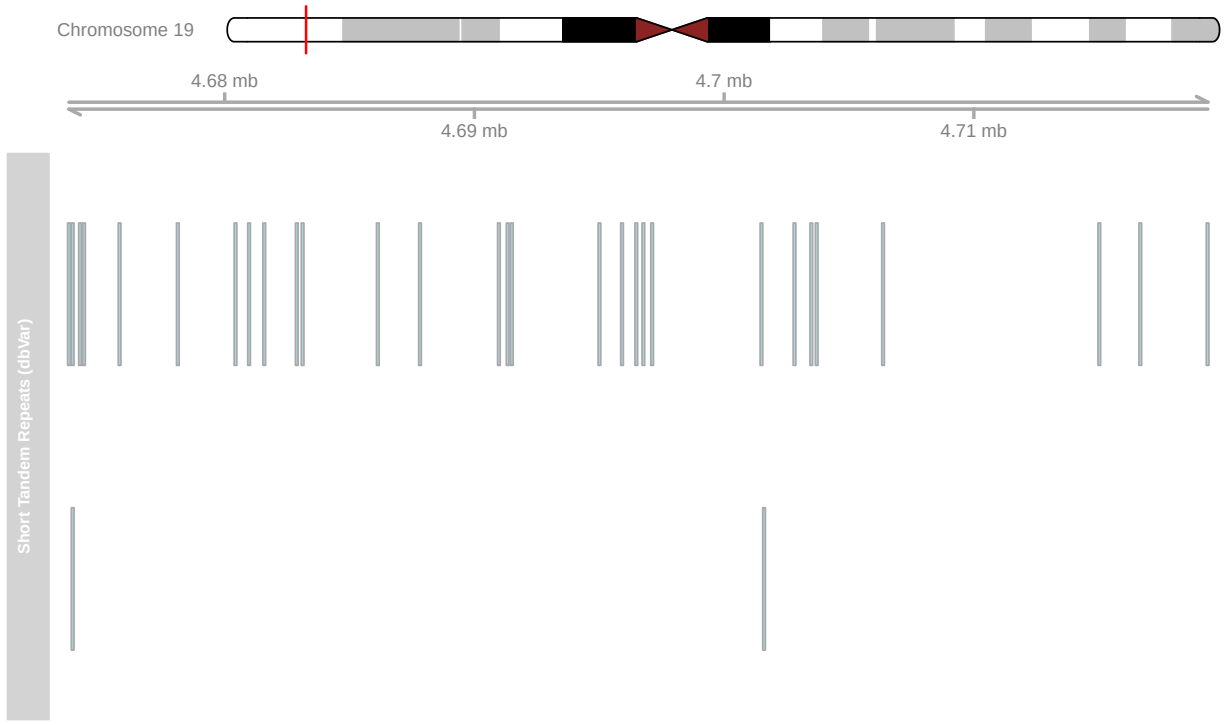
```
# Creating annotation track
gtrack <- GenomeAxisTrack()
atrack1 <- AnnotationTrack(CNVdbVar_GR, name = "Copy Number Variation (dbVar)")
# Ideogram -> start-end = [4675239-4723855]
chr1 <- as.character(unique(seqnames(CNVdbVar_GR)))
itrack1 <- IdeogramTrack(genome = "hg19", chromosome = chr1)
plotTracks(list(itrack1,gtrack, atrack1),from = 4675239, to = 4723855)
```



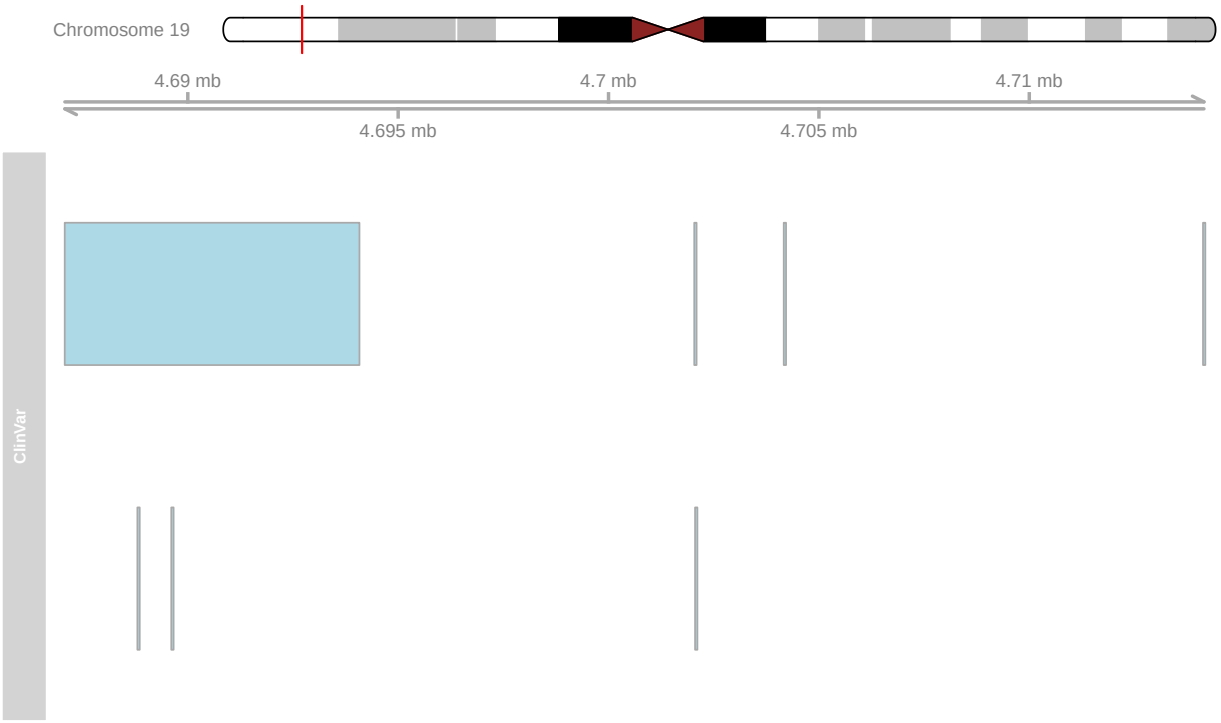

```
gtrack <- GenomeAxisTrack()
atrack2 <- AnnotationTrack(insertiondbVar_GR, name = "insertion (dbVar)")
chr2 <- as.character(unique(seqnames(insertiondbVar_GR)))
itrack2 <- IdeogramTrack(genome = "hg19", chromosome = chr2)
plotTracks(list(itrack2, gtrack, atrack2))
```



```
gtrack <- GenomeAxisTrack()
atrack3 <- AnnotationTrack(STRdbVar_GR, name = "Short Tandem Repeats (dbVar)")
chr3 <- as.character(unique(seqnames(STRdbVar_GR)))
itrack3 <- IdeogramTrack(genome = "hg19", chromosome = chr3)
plotTracks(list(itrack3, gtrack, atrack3))
```




```
gtrack <- GenomeAxisTrack()
atrack4 <- AnnotationTrack(ClinVar_GR, name = "ClinVar")
chr4 <- as.character(unique(seqnames(STRdbVar_GR)))
itrack4 <- IdeogramTrack(genome = "hg19", chromosome = chr4)
plotTracks(list(itrack4, gtrack, atrack4))
```


Chromosome 19
4.69 mb
4.7 mb
4.71 mb
4.695 mb
4.705 mb
ClinVar