

Visualizing_genome

Paulyna Magana

19/05/2022

Libraries

```
library(Gviz)
library(rtracklayer)
library(GenomicRanges)
library(biomaRt)
```

```
session <- browserSession()
genome(session) <- "hg38"
trackNames(session)
```

| | | |
|----|---------------------|------------------------|
| ## | Base Position | Fix Patches |
| ## | "ruler" | "fixSeqLiftOverPsl" |
| ## | Alt Haplotypes | Assembly |
| ## | "altSeqLiftOverPsl" | "gold" |
| ## | Centromeres | Chromosome Band |
| ## | "centromeres" | "cytoBand" |
| ## | Clone Ends | Exome Probesets |
| ## | "cloneEndSuper" | "exomeProbesets" |
| ## | FISH Clones | Gap |
| ## | "fishClones" | "gap" |
| ## | GC Percent | GRC Contigs |
| ## | "gc5BaseBw" | "ctgPos2" |
| ## | GRC Incident | Hg19 Diff |
| ## | "grcIncidentDb" | "hg38ContigDiff" |
| ## | INSDC | LiftOver & ReMap |
| ## | "ucscToINSDC" | "liftHg19" |
| ## | LRG Regions | Mappability |
| ## | "lrg" | "mappability" |
| ## | RefSeq Acc | Restr Enzymes |
| ## | "ucscToRefSeq" | "cutters" |
| ## | Scaffolds | Short Match |
| ## | "scaffolds" | "oligoMatch" |
| ## | STS Markers | GENCODE V41 |
| ## | "stsMap" | "knownGene" |
| ## | NCBI RefSeq | All GENCODE |
| ## | "refSeqComposite" | "wgEncodeGencodeSuper" |
| ## | CCDS | CRISPR Targets |
| ## | "ccdsGene" | "crisprAllTargets" |

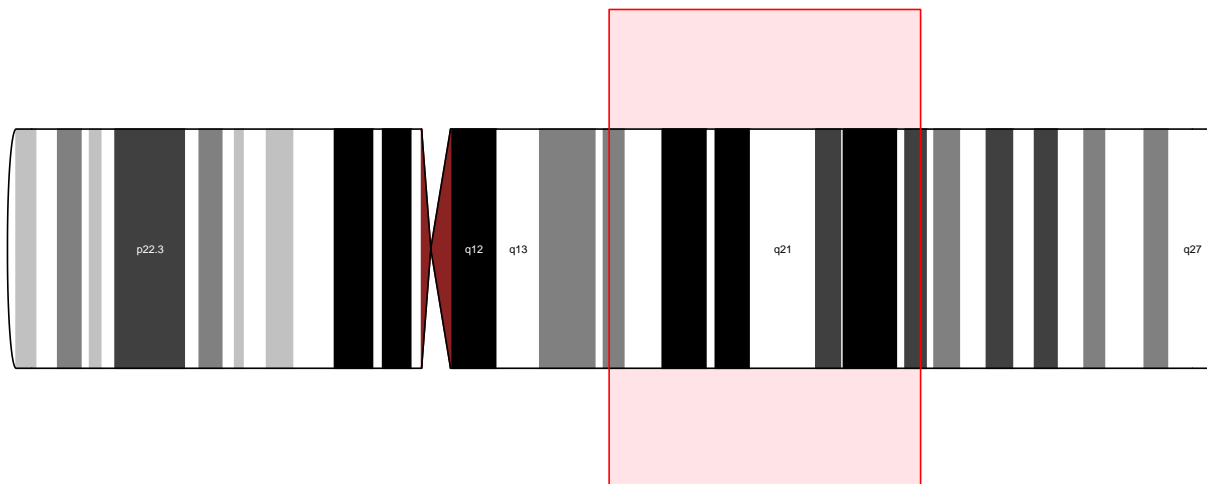
| | | |
|----|--------------------|----------------------|
| ## | HGNC | IKMC Genes Mapped |
| ## | "hgnc" | "hgIkmc" |
| ## | LRG Transcripts | MANE v1.0 |
| ## | "lrgTranscriptAli" | "mane" |
| ## | MGC Genes | Non-coding RNA |
| ## | "mgcFullMrna" | "nonCodingRNAs" |
| ## | Old UCSC Genes | ORFeome Clones |
| ## | "knownGeneOld12" | "orfeomeMrna" |
| ## | Other RefSeq | Pfam in GENCODE |
| ## | "xenoRefGene" | "ucscGenePfam" |
| ## | Prediction Archive | RetroGenes V9 |
| ## | "genePredArchive" | "ucscRetroAli9" |
| ## | TransMap V5 | UCSC Alt Events |
| ## | "transMapV5" | "knownAlt" |
| ## | UniProt | OMIM Alleles |
| ## | "uniprot" | "omimAvSnp" |
| ## | CADD | Cancer Gene Expr |
| ## | "caddSuper" | "cancerExpr" |
| ## | ClinGen | ClinGen CNVs |
| ## | "clinGenComp" | "iscaComposite" |
| ## | ClinVar Variants | Constraint scores |
| ## | "clinvar" | "constraintSuper" |
| ## | Coriell CNVs | COSMIC Regions |
| ## | "coriellDelDup" | "cosmicRegions" |
| ## | DECIPHER CNVs | DECIPHER SNVs |
| ## | "decipher" | "decipherSnvs" |
| ## | Development Delay | GenCC |
| ## | "cnvDevDelay" | "genCC" |
| ## | Gene Interactions | GeneReviews |
| ## | "interactions" | "geneReviews" |
| ## | GWAS Catalog | HGMD Variants |
| ## | "gwasCatalog" | "hgmd" |
| ## | LOVD Variants | OMIM Cyto Loci |
| ## | "lovdComp" | "omimLocation" |
| ## | OMIM Genes | Orphanet |
| ## | "omimGene2" | "orphadata" |
| ## | PanelApp | REVEL Scores |
| ## | "panelApp" | "revel" |
| ## | SNPedia | TCGA Pan-Cancer |
| ## | "snpedia" | "gdcCancer" |
| ## | UniProt Variants | Variants in Papers |
| ## | "spMut" | "varsInPubs" |
| ## | Rare Harmful Vars | Blood (PBMC) Hao |
| ## | "covidMuts" | "bloodHao" |
| ## | Colon Wang | Cortex Velmeshev |
| ## | "colonWang" | "cortexVelmeshev" |
| ## | Fetal Gene Atlas | Heart Cell Atlas |
| ## | "fetalGeneAtlas" | "heartCellAtlas" |
| ## | Ileum Wang | Kidney Stewart |
| ## | "ileumWang" | "kidneyStewart" |
| ## | Liver MacParland | Lung Travaglini |
| ## | "liverMacParland" | "lungTravaglini2020" |
| ## | Merged Cells | Muscle De Micheli |
| ## | "singleCellMerged" | "muscleDeMicheli" |

| | | |
|----|------------------------------|-------------------------|
| ## | Pancreas Baron | Placenta Vento-Tormo |
| ## | "pancreasBaron" | "placentaVentoTormo" |
| ## | Rectum Wang | Skin Sole-Boldo |
| ## | "rectumWang" | "skinSoleBoldo" |
| ## | Tabula Sapiens | Human ESTs |
| ## | "tabulaSapiens" | "est" |
| ## | Human mRNAs | Other ESTs |
| ## | "mrna" | "xenoEst" |
| ## | Other mRNAs | SIB Alt-Splicing |
| ## | "xenoMrna" | "sibTxGraph" |
| ## | Spliced ESTs | GTEX Gene V8 |
| ## | "intronEst" | "gtexGeneV8" |
| ## | GTEX RNA-Seq Coverage | Affy Archive |
| ## | "gtexCov" | "affyArchive" |
| ## | EPDnew Promoters | GNF Atlas 2 |
| ## | "epdNew" | "gnfAtlas2" |
| ## | GTEX Gene | GTEX Transcript |
| ## | "gtexGene" | "gtexTranscExpr" |
| ## | GWIPS-viz Riboseq | miRNA Tissue Atlas |
| ## | "gwipsvizRiboseq" | "miRnaAtlas" |
| ## | ENCODE cCREs | ENCODE Regulation |
| ## | "encodeCcreCombined" | "wgEncodeReg" |
| ## | CpG Islands | GeneHancer |
| ## | "cpgIslandSuper" | "geneHancer" |
| ## | GTEX cis-eQTLs | Hi-C and Micro-C |
| ## | "gtexEqtlHighConf" | "hicAndMicroC" |
| ## | JASPAR Transcription Factors | ORegAnno |
| ## | "jaspar" | "oreganno" |
| ## | RefSeq Func Elems | ReMap ChIP-seq |
| ## | "refSeqFuncElems" | "ReMap" |
| ## | Conservation | Cactus 241-way |
| ## | "cons100way" | "cons241way" |
| ## | Cons 30 Primates | Primate Chain/Net |
| ## | "cons30way" | "primateChainNet" |
| ## | Placental Chain/Net | Vertebrate Chain/Net |
| ## | "placentalChainNet" | "vertebrateChainNet" |
| ## | CHM13 alignments | Human Chain/Net |
| ## | "chm13LiftOver" | "humanChainNet" |
| ## | dbSNP 155 | dbSNP 153 |
| ## | "dbSnp155Composite" | "dbSnp153Composite" |
| ## | 1000G Archive | Array Probesets |
| ## | "tgpArchive" | "genotypeArrays" |
| ## | dbSNP Archive | dbVar Common Struct Var |
| ## | "dbSnpArchive" | "dbVarSv" |
| ## | DGV Struct Var | Genome In a Bottle |
| ## | "dgvPlus" | "giab" |
| ## | gnomAD Variants | Platinum Genomes |
| ## | "gnomadVariants" | "platinumGenomes" |
| ## | RepeatMasker | Interrupted Rpts |
| ## | "rmsk" | "nestedRepeats" |
| ## | Microsatellite | RepeatMasker Viz. |
| ## | "microsat" | "joinedRmsk" |
| ## | Segmental Dups | Self Chain |
| ## | "genomicSuperDups" | "chainSelf" |

```
##                               Simple Repeats           WM + SDust
##                               "simpleRepeat"             "windowmaskerSdust"
```

```
gtrack <- GenomeAxisTrack() #to add a genomic axis to the plot
ideoTrack <- IdeogramTrack(genome = "hg38", chromosome = "chr6")# ideogram

#plot
# showbandId - chromosome bands in the ideogram
# turn off the explicit plotting of the chromosome name by setting the showId display parameter to FALSE
plotTracks(ideoTrack, from = 85e6, to = 129e6, showId = FALSE,
            showBandId = TRUE, cex.bands = 0.5)
```



```
###Simple annotation tracks:
from <- 160115734
to <- 160160664

##Gene model type tracks:

ncbi_ref <- UscsTrack(genome="hg38", chromosome="chr6", track="NCBI RefSeq",
                      from=from, to=to, trackType="GeneRegionTrack",
                      rstarts="exonStarts", rends="exonEnds", gene="name",
                      symbol="name", transcript="name", strand="strand",
                      fill="#767676", stacking="squish", name="NCBI RefSeq",
                      showId=TRUE, geneSymbol=TRUE)

biomTrack <- BiomartGeneRegionTrack(genome="hg38", chromosome="chr6",
                                     start=from, end=to, name="Ensembl")

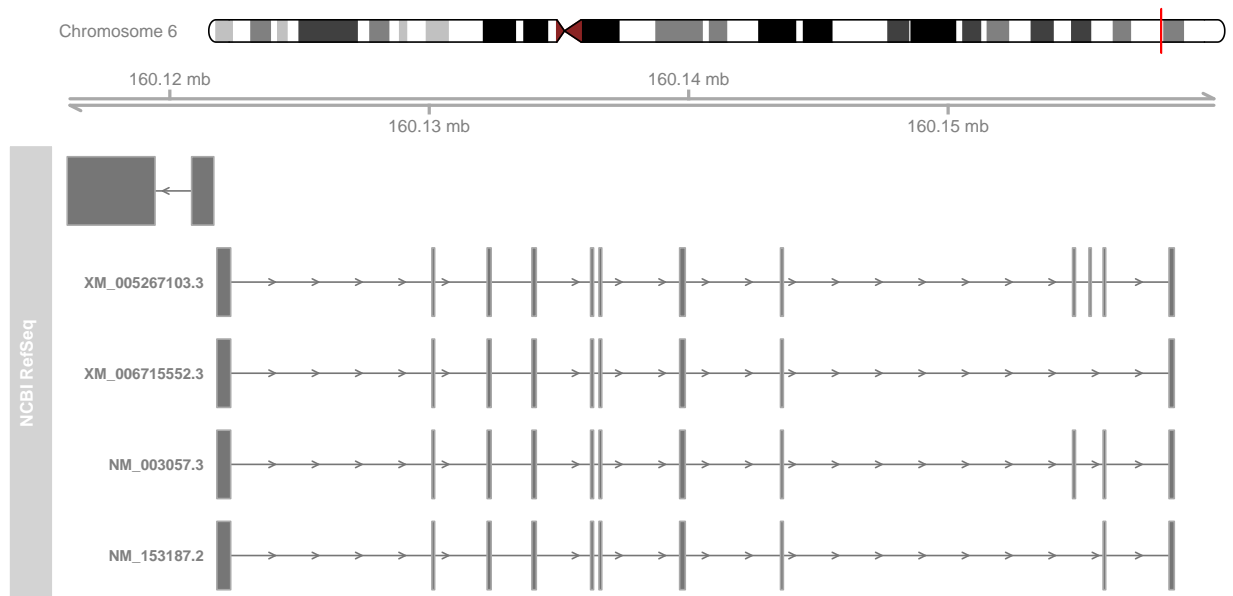
###Data tracks
```

```
###Other tracks and plotting:
```

```
axTrack <- GenomeAxisTrack()
idxTrack <- IdeogramTrack(genome="hg38", chromosome="chr6")
```

```
##PLOTING
```

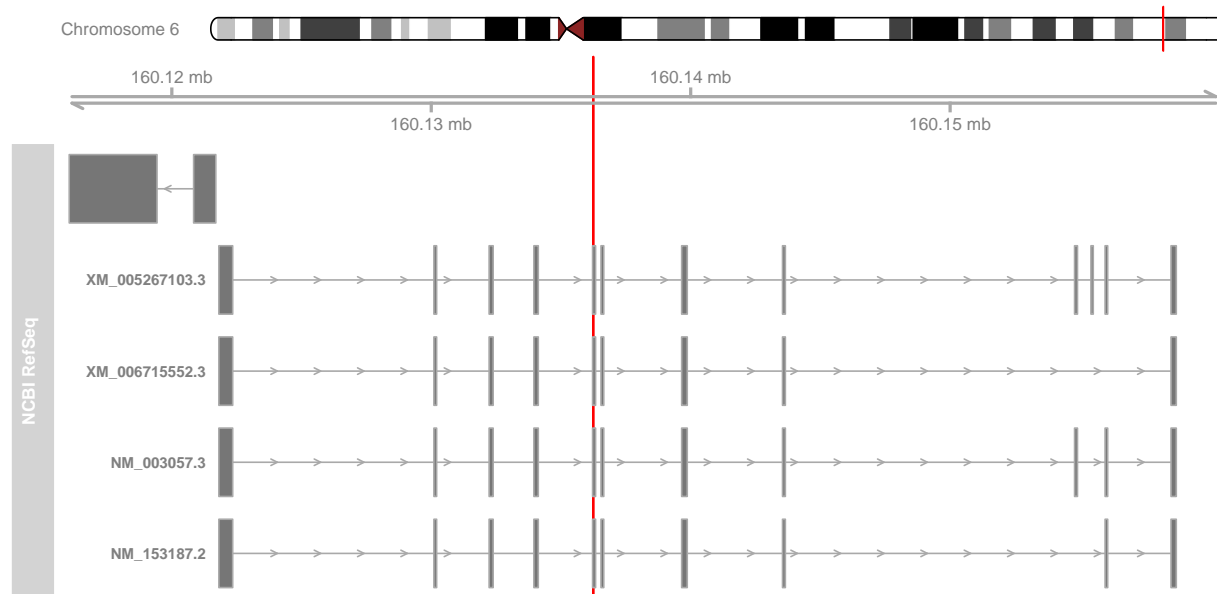
```
plotTracks(list(idxTrack, axTrack, ncbi_ref), from=from, to=to, showTitle=TRUE, col.line=NULL,
           collapseTranscripts="longest", transcriptAnnotation = "symbol")
```



```
ht <- HighlightTrack(trackList = list(axTrack, ncbi_ref),
                    start = c(160136234, 160136250), width = 7,
                    chromosome = 6)
```

```
lim <- c(from, to)
```

```
plotTracks(list(idxTrack, ht), from = lim[1], to = lim[2])
```



SLC22A5

###Simple annotation tracks:

```
from <- 132366800
```

```
to <- 132396017
```

##Gene model type tracks:

```
ncbi_ref <- UcsTrack(genome="hg38", chromosome="chr5", track="NCBI RefSeq",
  from=from, to=to, trackType="GeneRegionTrack",
  rstarts="exonStarts", rends="exonEnds", gene="name",
  symbol="name", transcript="name", strand="strand",
  fill="#767676", stacking="squish", name="NCBI RefSeq",
  showId=TRUE, geneSymbol=TRUE)
```

```
biomTrack <- BiomartGeneRegionTrack(genome="hg38", chromosome="chr5",
  start=from, end=to, name="Ensembl")
```

###Data tracks

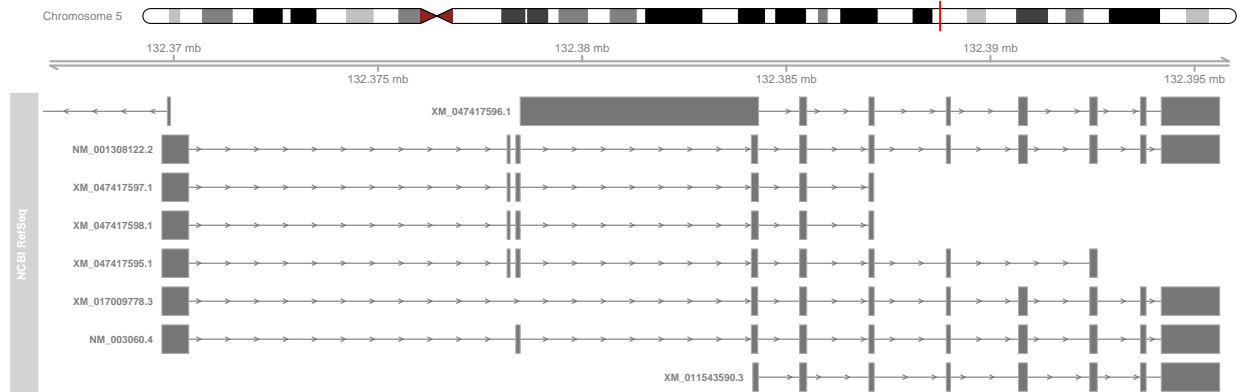
###Other tracks and plotting:

```
axTrack <- GenomeAxisTrack()
```

```
idxTrack <- IdeogramTrack(genome="hg38", chromosome="chr5")
```

##PLOTING

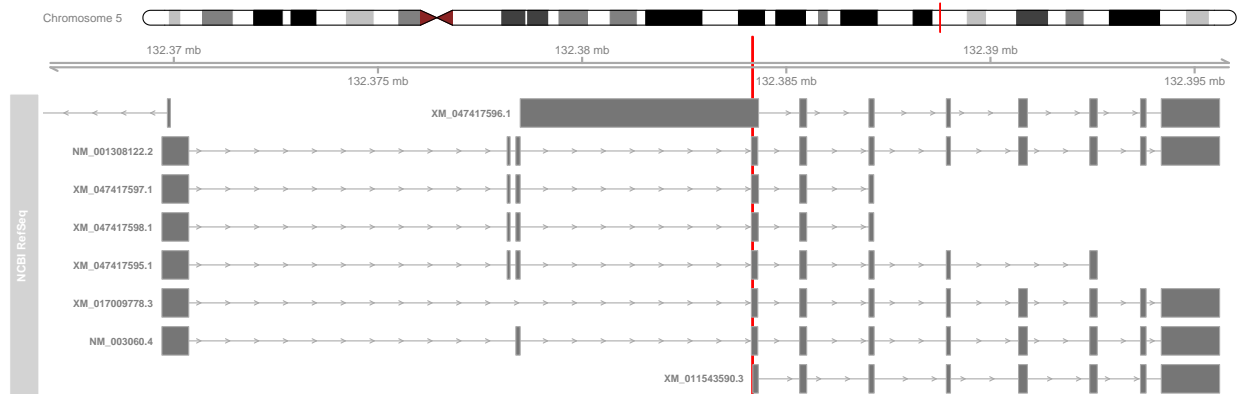
```
plotTracks(list(idxTrack, axTrack, ncbi_ref), from=from, to=to, showTitle=TRUE, col.line=NULL,
collapseTranscripts="longest", transcriptAnnotation = "symbol")
```



```
ht <- HighlightTrack(trackList = list(axTrack, ncbi_ref),
start = c(132384160, 132384180), width = 7,
chromosome = 5)
```

```
lim <- c(from, to)
```

```
plotTracks(list(idxTrack, ht), from = lim[1], to = lim[2])
```



SLC22A4

```
###Simple annotation tracks:
```

```
from <- 132290928
```

```
to <- 132345343
```

```
##Gene model type tracks:
```

```
ncbi_ref <- UcscTrack(genome="hg38", chromosome="chr5", track="NCBI RefSeq",
from=from, to=to, trackType="GeneRegionTrack",
rstarts="exonStarts", rends="exonEnds", gene="name",
```

```

symbol="name", transcript="name", strand="strand",
fill="#767676", stacking="squish", name="NCBI RefSeq",
showId=TRUE, geneSymbol=TRUE)

```

```

biomTrack <- BiomartGeneRegionTrack(genome="hg38", chromosome="chr5",
                                   start=from, end=to, name="Ensembl")

```

###Data tracks

###Other tracks and plotting:

```

axTrack <- GenomeAxisTrack()

```

```

idxTrack <- IdeogramTrack(genome="hg38", chromosome="chr5")

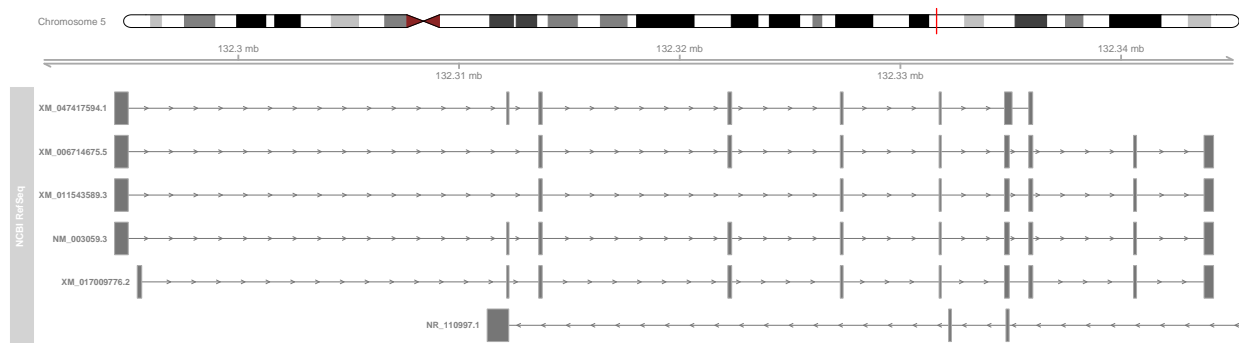
```

##PLOTING

```

plotTracks(list(idxTrack, axTrack, ncbi_ref), from=from, to=to, showTitle=TRUE, col.line=NULL,
            collapseTranscripts="longest", transcriptAnnotation = "symbol")

```



```

ht <- HighlightTrack(trackList = list(axTrack, ncbi_ref),
                    start = c(132313626, 132313650), width = 7,
                    chromosome = 5)

```

```

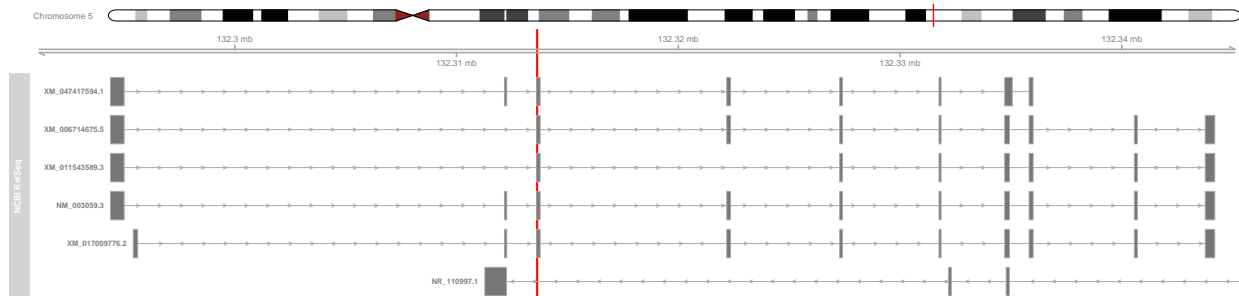
lim <- c(from, to)

```

```

plotTracks(list(idxTrack, ht), from = lim[1], to = lim[2])

```



R Markdown

```
## R version 4.1.1 (2021-08-10)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 18363)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_United Kingdom.1252
## [2] LC_CTYPE=English_United Kingdom.1252
## [3] LC_MONETARY=English_United Kingdom.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United Kingdom.1252
##
## attached base packages:
## [1] grid      stats4    stats      graphics  grDevices  utils      datasets
## [8] methods   base
##
## other attached packages:
## [1] biomaRt_2.50.3      rtracklayer_1.54.0  Gviz_1.36.2
## [4] GenomicRanges_1.46.1 GenomeInfoDb_1.30.1 IRanges_2.28.0
## [7] S4Vectors_0.32.4    BiocGenerics_0.40.0
##
## loaded via a namespace (and not attached):
## [1] ProtGenerics_1.26.0      bitops_1.0-7
## [3] matrixStats_0.62.0      bit64_4.0.5
## [5] filelock_1.0.2          RColorBrewer_1.1-3
## [7] progress_1.2.2          httptr_1.4.4
## [9] backports_1.4.1         tools_4.1.1
## [11] utf8_1.2.2              R6_2.5.1
## [13] rpart_4.1-15            lazyeval_0.2.2
## [15] Hmisc_4.7-1            DBI_1.1.3
## [17] colorspace_2.0-3       nnet_7.3-16
## [19] withr_2.5.0            gridExtra_2.3
## [21] tidyselect_1.2.0       prettyunits_1.1.1
## [23] bit_4.0.4              curl_4.3.3
## [25] compiler_4.1.1         cli_3.4.1
## [27] Biobase_2.54.0         htmlTable_2.4.1
## [29] xml2_1.3.3             DelayedArray_0.20.0
## [31] checkmate_2.1.0        scales_1.2.1
## [33] rappdirs_0.3.3         stringr_1.4.1
## [35] digest_0.6.29          Rsamtools_2.10.0
## [37] foreign_0.8-81         rmarkdown_2.17
## [39] XVector_0.34.0         dichromat_2.0-0.1
## [41] base64enc_0.1-3        jpeg_0.1-9
## [43] pkgconfig_2.0.3        htmltools_0.5.3
## [45] MatrixGenerics_1.6.0   highr_0.9
## [47] ensemblDb_2.18.4       dbplyr_2.2.1
## [49] fastmap_1.1.0          BSgenome_1.62.0
## [51] htmlwidgets_1.5.4      rlang_1.0.6
## [53] rstudioapi_0.14        RSQLite_2.2.17
## [55] BiocIO_1.4.0           generics_0.1.3
## [57] BiocParallel_1.28.3    dplyr_1.0.10
```

```

## [59] VariantAnnotation_1.40.0    RCurl_1.98-1.8
## [61] magrittr_2.0.3              GenomeInfoDbData_1.2.7
## [63] Formula_1.2-4               interp_1.1-3
## [65] Matrix_1.5-1                Rcpp_1.0.9
## [67] munsell_0.5.0               fansi_1.0.3
## [69] lifecycle_1.0.3             stringi_1.7.6
## [71] yaml_2.3.5                  SummarizedExperiment_1.24.0
## [73] zlibbioc_1.40.0             BiocFileCache_2.2.1
## [75] blob_1.2.3                  parallel_4.1.1
## [77] crayon_1.5.2                deldir_1.0-6
## [79] lattice_0.20-44             Biostrings_2.62.0
## [81] splines_4.1.1               GenomicFeatures_1.46.5
## [83] hms_1.1.2                   KEGGREST_1.34.0
## [85] knitr_1.40                   pillar_1.8.1
## [87] rjson_0.2.21                codetools_0.2-18
## [89] XML_3.99-0.10               glue_1.6.2
## [91] evaluate_0.17               biovizBase_1.42.0
## [93] latticeExtra_0.6-30         data.table_1.14.2
## [95] png_0.1-7                   vctrs_0.4.2
## [97] purrr_0.3.5                 gtable_0.3.1
## [99] assertthat_0.2.1            cachem_1.0.6
## [101] ggplot2_3.3.6               xfun_0.33
## [103] AnnotationFilter_1.18.0     restfulr_0.0.15
## [105] survival_3.2-11             tibble_3.1.8
## [107] GenomicAlignments_1.30.0    AnnotationDbi_1.56.2
## [109] memoise_2.0.1               cluster_2.1.2
## [111] ellipsis_0.3.2

```