

Visualizing_genome

Paulyna Magana

19/05/2022

Libraries

```
library(Gviz)
library(GenomicRanges)
```

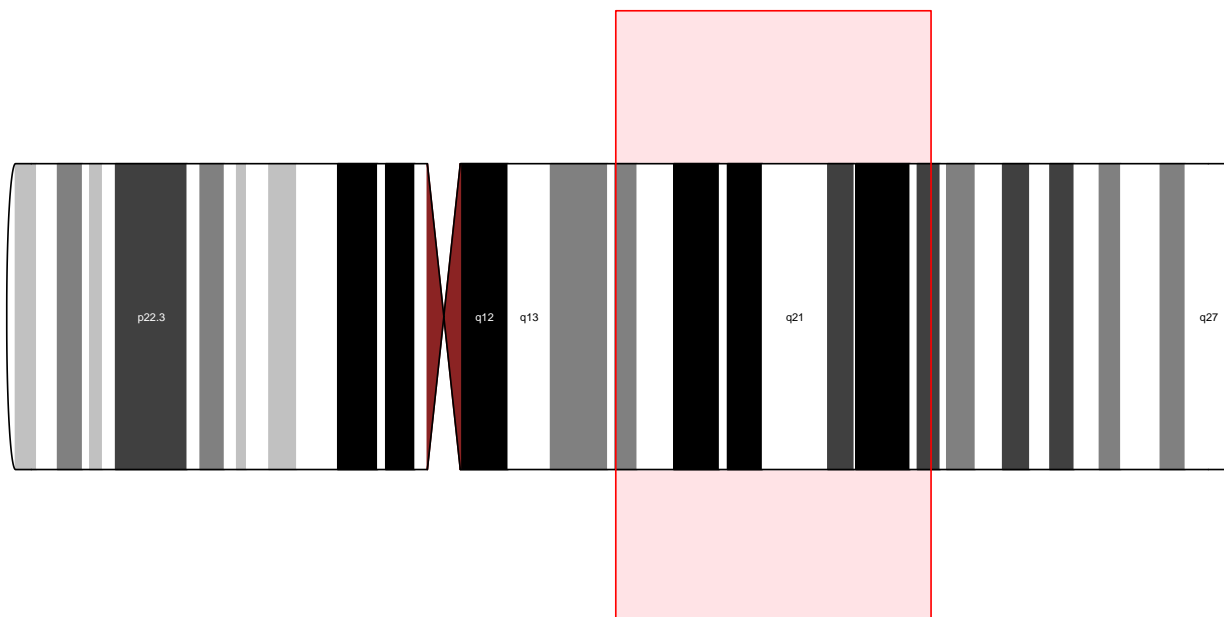
```
data(cpgIslands)
class(cpgIslands)
```

```
## [1] "GRanges"
## attr(,"package")
## [1] "GenomicRanges"
```

```
chr <- as.character(unique(seqnames(cpgIslands)))
gen <- genome(cpgIslands) #plot the single CpG islands annotation track.
```

```
atrack <- AnnotationTrack(cpgIslands, name= "CpG") #plot cpgislands
gtrack <- GenomeAxisTrack() #to add a genomic axis to the plot
ideoTrack <- IdeogramTrack(genome = "hg19", 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
cpgIslands <- UcsTrack(genome="hg38", chromosome="chr6",
                      track="CpG Islands", from=from, to=to,
                      trackType="AnnotationTrack", start="chromStart", end="chromEnd",
                      id="name", shape="box", fill="#006400", name="CpG Islands")

snpLocations <- UcsTrack(genome="hg38", chromosome="chr6",
                        track="All SNPs(144)", from=from, to=to, trackType="AnnotationTrack",
                        start="chromStart", end="chromEnd", id="name", feature="func",
                        strand="strand", shape="box", stacking="dense", fill="black",
                        name="SNPs")

oreganno <- UcsTrack(genome="hg38", chromosome="chr6",
                    track="ORegAnno", from=from, to=to,
                    trackType="AnnotationTrack", start="chromStart", end="chromEnd",
                    id="name", shape="box", fill="#CE796B", col="transparent", name="Regulatory Ele

##Gene model type tracks:

#knownGenes <- UcsTrack(genome="hg38", chromosome="chr6",
#                      # track="knownGene", from=from, to=to, trackType="GeneRegionTrack",
#                      # rstarts="exonStarts", rends="exonEnds", gene="name", symbol="name",
#                      # transcript="name", strand="strand", fill="#8282d2", name="UCSC Genes")

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

```

```

#ensGenes <- UcsTrack(genome="hg38", chromosome="chr6",
  # track="All GENCODE V25", from=from, to=to, trackType="GeneRegionTrack",
  # rstarts="exonStarts", rends="exonEnds", gene="name",
  # symbol="name2", transcript="name", strand="strand",
  # fill="#960000", name="Ensembl Genes")

###Data tracks

conservation <- UcsTrack(genome="hg38", chromosome="chr6",
  track="Conservation", table="phastCons100way", from=from, to=to,
  trackType="DataTrack", start="start", end="end", data="score",
  type="hist", window="auto", col.histogram="darkblue",
  fill.histogram="darkblue", ylim=c(-3.7, 4), name="Conservation")

clinvar <- UcsTrack(genome="hg38", chromosome="chr6",
  track="clinvar", table="clinvarCnv", from=from, to=to, trackType="DataTrack",
  start="start", end="end", data="score",
  type="p", stacking="dense", bol="lightgrey",
  name="ClinVar")

gcContent <- UcsTrack(genome="hg38", chromosome="chr6",
  track="GC Percent", table="gc5Base", from=from, to=to,
  trackType="DataTrack", start="start", end="end", data="score",
  type="hist", window=-1, windowSize=1500, fill.histogram="black",
  col.histogram="black", ylim=c(30, 70), name="GC Percent")

EPD <- UcsTrack(genome="hg38", chromosome="chr6",
  track="EPDnew", table="epdNewPromoter", from=from, to=to,
  trackType="DataTrack", start="start", end="end", data="score",
  type="p", stacking="dense", fill="black", name="EPD new")

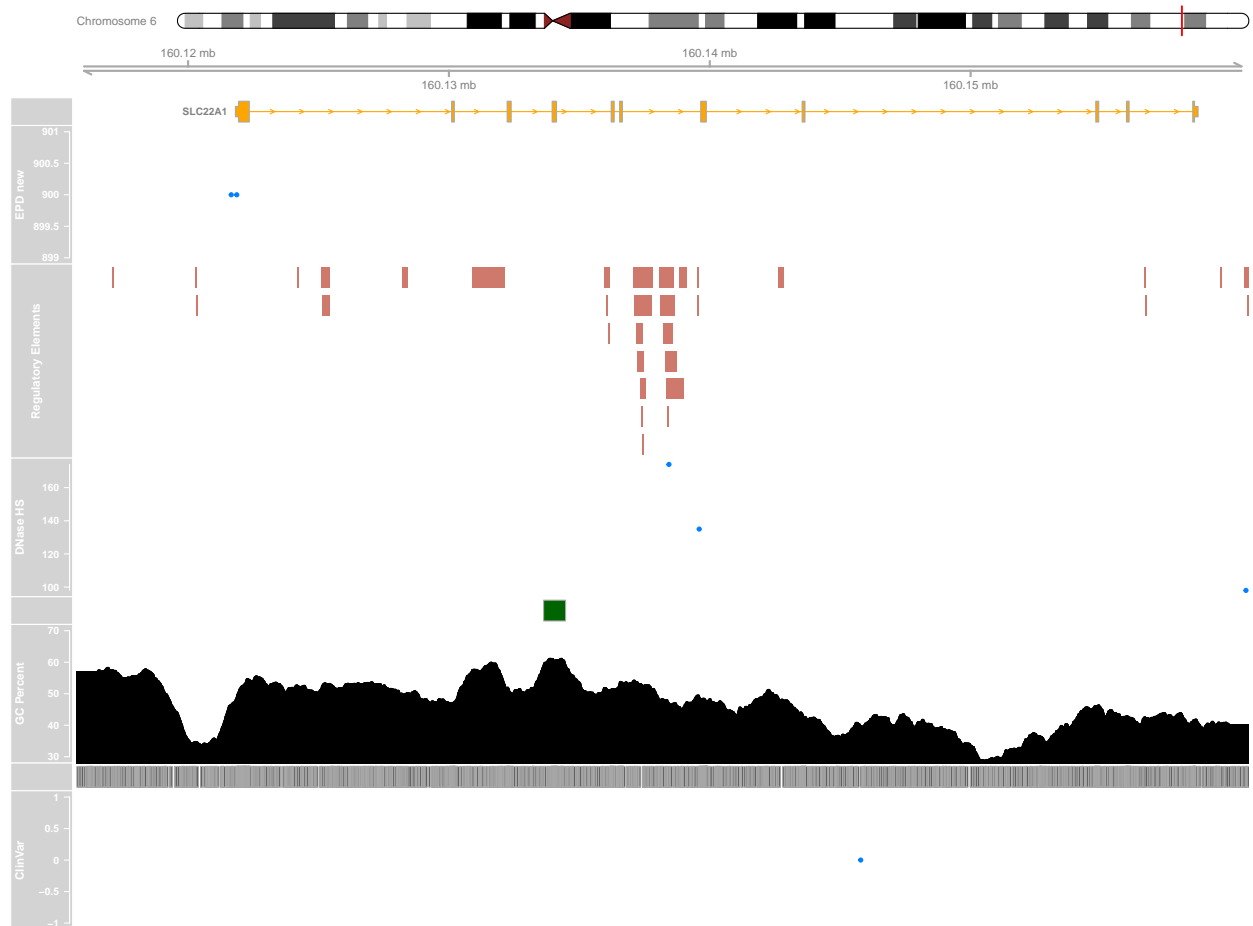
DNase <- UcsTrack(genome="hg38", chromosome="chr6",
  track="DNase HS", table="wgEncodeRegDnaseUwA549Peak", from=from, to=to,
  trackType="DataTrack", start="start", end="end", data="score",
  type="p", stacking="dense", bol="lightgrey", name="DNase HS")

###Other tracks and plotting:

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

##PLOTING
plotTracks(list(idxTrack, axTrack, biomTrack, EPD, oreganno, DNase,
  cpgIslands, gcContent, snpLocations, clinvar), from=from, to=to, showTitle=TRUE, col.lin
  collapseTranscripts="longest", transcriptAnnotation = "symbol")

```



```
## Add a highlight
```

```
h1= 160120029
```

```
h2 = 160123742
```

```
chromo = 6
```

```
library(rtracklayer)
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 V39
##	"stsMap"	"knownGene"
##	NCBI RefSeq	All GENCODE
##	"refSeqComposite"	"wgEncodeGencodeSuper"
##	CCDS	CRISPR Targets
##	"ccdsGene"	"crisprAllTargets"
##	IKMC Genes Mapped	LRG Transcripts
##	"hgIkmc"	"lrgTranscriptAli"
##	MANE v1.0	MGC Genes
##	"mane"	"mgcFullMrna"
##	Non-coding RNA	Old UCSC Genes
##	"nonCodingRNAs"	"knownGeneOld12"
##	ORFeome Clones	Other RefSeq
##	"orfeomeMrna"	"xenoRefGene"
##	Pfam in GENCODE	Prediction Archive
##	"ucscGenePfam"	"genePredArchive"
##	RetroGenes V9	TransMap V5
##	"ucscRetroAli9"	"transMapV5"
##	UCSC Alt Events	UniProt
##	"knownAlt"	"uniprot"
##	OMIM Alleles	CADD
##	"omimAvSnp"	"caddSuper"
##	Cancer Gene Expr	ClinGen
##	"cancerExpr"	"clinGenComp"
##	ClinGen CNVs	ClinVar Variants
##	"iscaComposite"	"clinvar"
##	Coriell CNVs	COSMIC Regions
##	"coriellDelDup"	"cosmicRegions"
##	DECIPHER CNVs	DECIPHER SNVs
##	"decipher"	"decipherSnvs"
##	Development Delay	Gene Interactions
##	"cnvDevDelay"	"interactions"
##	GeneReviews	GWAS Catalog
##	"geneReviews"	"gwasCatalog"
##	HGMD Variants	LOVD Variants
##	"hgmd"	"lovdComp"
##	OMIM Cyto Loci	OMIM Genes
##	"omimLocation"	"omimGene2"
##	Orphanet	REVEL Scores
##	"orphadata"	"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 153	1000G Archive
##	"dbSnp153Composite"	"tgpArchive"
##	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"
```

```
query <- ucscTableQuery(session, "clinvar",
                        GRangesForUCSCGenome("hg38", "chr6",
                                              IRanges(from,to)))
tableNames(query)
```

```
## [1] "clinvarMain"      "clinvarCnv"      "clinvarSubLolly"
```

R Markdown

```
## R version 4.0.5 (2021-03-31)
## Platform: x86_64-apple-darwin17.0 (64-bit)
## Running under: macOS Big Sur 10.16
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRblas.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_GB.UTF-8/en_GB.UTF-8/en_GB.UTF-8/C/en_GB.UTF-8/en_GB.UTF-8
##
## attached base packages:
## [1] grid      parallel  stats4    stats      graphics  grDevices  utils
## [8] datasets  methods   base
##
## other attached packages:
## [1] rtracklayer_1.50.0  Gviz_1.34.1          GenomicRanges_1.42.0
## [4] GenomeInfoDb_1.26.7 IRanges_2.24.1       S4Vectors_0.28.1
## [7] BiocGenerics_0.36.1
##
## loaded via a namespace (and not attached):
## [1] ProtGenerics_1.22.0      bitops_1.0-7
## [3] matrixStats_0.62.0      bit64_4.0.5
## [5] RColorBrewer_1.1-3      progress_1.2.2
## [7] httr_1.4.3              backports_1.4.1
## [9] tools_4.0.5             utf8_1.2.2
## [11] R6_2.5.1                rpart_4.1.16
## [13] lazyeval_0.2.2          Hmisc_4.6-0
## [15] DBI_1.1.2               colorspace_2.0-3
## [17] nnet_7.3-17             gridExtra_2.3
## [19] tidyselect_1.1.2        prettyunits_1.1.1
## [21] bit_4.0.4              curl_4.3.2
```

## [23] compiler_4.0.5	cli_3.3.0
## [25] Biobase_2.50.0	htmlTable_2.4.0
## [27] xml2_1.3.3	DelayedArray_0.16.3
## [29] checkmate_2.1.0	scales_1.2.0
## [31] askpass_1.1	rappdirs_0.3.3
## [33] stringr_1.4.0	digest_0.6.29
## [35] Rsamtools_2.6.0	foreign_0.8-82
## [37] rmarkdown_2.14	XVector_0.30.0
## [39] dichromat_2.0-0.1	base64enc_0.1-3
## [41] jpeg_0.1-9	pkgconfig_2.0.3
## [43] htmltools_0.5.2	MatrixGenerics_1.2.1
## [45] highr_0.9	ensembl_2.14.1
## [47] dbplyr_2.1.1	fastmap_1.1.0
## [49] BSgenome_1.58.0	htmlwidgets_1.5.4
## [51] rlang_1.0.2	rstudioapi_0.13
## [53] RSQLite_2.2.14	generics_0.1.2
## [55] BiocParallel_1.24.1	dplyr_1.0.9
## [57] VariantAnnotation_1.36.0	RCurl_1.98-1.6
## [59] magrittr_2.0.3	GenomeInfoDbData_1.2.4
## [61] Formula_1.2-4	Matrix_1.4-1
## [63] Rcpp_1.0.8.3	munsell_0.5.0
## [65] fansi_1.0.3	lifecycle_1.0.1
## [67] stringi_1.7.6	yaml_2.3.5
## [69] SummarizedExperiment_1.20.0	zlibbioc_1.36.0
## [71] BiocFileCache_1.14.0	blob_1.2.3
## [73] crayon_1.5.1	lattice_0.20-45
## [75] Biostrings_2.58.0	splines_4.0.5
## [77] GenomicFeatures_1.42.3	hms_1.1.1
## [79] knitr_1.39	pillar_1.7.0
## [81] biomaRt_2.46.3	XML_3.99-0.9
## [83] glue_1.6.2	evaluate_0.15
## [85] biovizBase_1.38.0	latticeExtra_0.6-29
## [87] data.table_1.14.2	vctrs_0.4.1
## [89] png_0.1-7	gtable_0.3.0
## [91] openssl_2.0.1	purrr_0.3.4
## [93] assertthat_0.2.1	cachem_1.0.6
## [95] ggplot2_3.3.6	xfun_0.31
## [97] AnnotationFilter_1.14.0	survival_3.3-1
## [99] tibble_3.1.7	GenomicAlignments_1.26.0
## [101] AnnotationDbi_1.52.0	memoise_2.0.1
## [103] cluster_2.1.3	ellipsis_0.3.2