

An Introduction to *Guitar* Package

Xiao Du

Modified: 26 April, 2019. Compiled: October 29, 2024

1 Quick Start with Guitar

This is a manual for Guitar package. The Guitar package is aimed for RNA landmark-guided transcriptomic analysis of RNA-related genomic features.

The Guitar package enables the comparison of multiple genomic features, which need to be stored in a name list. Please see the following example, which reads 1000 RNA m6A methylation sites into R for detection. Of course, in actual data analysis, features may come from multiple sets of resources.

```
library(Guitar)

## Loading required package: GenomicFeatures
## Loading required package: BiocGenerics
##
## Attaching package: 'BiocGenerics'
## The following objects are masked from 'package:stats':
##
##   IQR, mad, sd, var, xtabs
## The following objects are masked from 'package:base':
##
##   Filter, Find, Map, Position, Reduce,
##   anyDuplicated, aperm, append, as.data.frame,
##   basename, cbind, colnames, dirname, do.call,
##   duplicated, eval, evalq, get, grep, grepl,
##   intersect, is.unsorted, lapply, mapply, match,
##   mget, order, paste, pmax, pmax.int, pmin,
##   pmin.int, rank, rbind, rownames, sapply, saveRDS,
##   setdiff, table, tapply, union, unique, unsplit,
##   which.max, which.min
## Loading required package: S4Vectors
## Loading required package: stats4
##
## Attaching package: 'S4Vectors'
## The following object is masked from 'package:utils':
##
##   findMatches
## The following objects are masked from 'package:base':
##
##   I, expand.grid, unname
## Loading required package: IRanges
##
## Attaching package: 'IRanges'
```

```

## The following object is masked from 'package:grDevices':
##
## windows
## Loading required package: GenomeInfoDb
## Loading required package: GenomicRanges
## Loading required package: AnnotationDbi
## Loading required package: Biobase
## Welcome to Bioconductor
##
## Vignettes contain introductory material; view
## with 'browseVignettes()'. To cite Bioconductor,
## see 'citation("Biobase")', and for packages
## 'citation("pkgname")'.
## Loading required package: rtracklayer
## Loading required package: magrittr
##
## Attaching package: 'magrittr'
## The following object is masked from 'package:GenomicRanges':
##
## subtract
## Loading required package: ggplot2
## Loading required package: dplyr
##
## Attaching package: 'dplyr'
## The following object is masked from 'package:AnnotationDbi':
##
## select
## The following object is masked from 'package:Biobase':
##
## combine
## The following objects are masked from 'package:GenomicRanges':
##
## intersect, setdiff, union
## The following object is masked from 'package:GenomeInfoDb':
##
## intersect
## The following objects are masked from 'package:IRanges':
##
## collapse, desc, intersect, setdiff, slice, union
## The following objects are masked from 'package:S4Vectors':
##
## first, intersect, rename, setdiff, setequal,
## union
## The following objects are masked from 'package:BiocGenerics':
##
## combine, intersect, setdiff, union
## The following objects are masked from 'package:stats':
##
## filter, lag
## The following objects are masked from 'package:base':
##
## intersect, setdiff, setequal, union
##
## Attaching package: 'Guitar'

```

```
## The following object is masked from 'package:BiocGenerics':
##
##      normalize

# genomic features imported into named list
stBedFiles <- list(system.file("extdata", "m6A_mm10_exomePeak_1000peaks_bed12.bed",
                             package="Guitar"))
```

With the following script, we may generate the transcriptomic distribution of genomic features to be tested, and the result will be automatically saved into a PDF file under the working directory with prefix "example". With the `GuitarPlot` function, the gene annotation can be downloaded from internet automatically with a genome assembly number provided; however, this feature requires working internet and might take a longer time. The toy `Guitar` coordinates generated internally should never be re-used in other real data analysis.

```
count <- GuitarPlot(txGenomeVer = "mm10",
                   stBedFiles = stBedFiles,
                   miscOutFilePrefix = NA)
```

In a more efficient protocol, in order to re-use the gene annotation and *Guitar coordinates*, you will have to build `Guitar Coordinates` from a `txdb` object in a separate step. The `transcriptDb` contains the gene annotation information and can be obtained in a number of ways, e.g, download the complete gene annotation of species from UCSC automatically, which might takes a few minutes. In the following analysis, we load the `Txdb` object from a toy dataset provided with the `Guitar` package. Please note that this is only a very small part of the complete hg19 transcriptome, and the `Txdb` object provided with `Guitar` package should not be used in real data analysis. With a `Txdb` object that contains gene annotation information, we in the next build *Guitar coordinates*, which is essentially a bridge connects the transcriptomic landmarks and genomic coordinates.

```
txdb_file <- system.file("extdata", "mm10_toy.sqlite",
                        package="Guitar")
txdb <- loadDb(txdb_file)
guitarTxdb <- makeGuitarTxdb(txdb = txdb, txPrimaryOnly = FALSE)

## [1] "There are 2946 transcripts of 2946 genes in the genome."
## [1] "total 2946 transcripts extracted ..."
## [1] "total 2719 transcripts left after ambiguity filter ..."
## [1] "total 2719 transcripts left after check chromosome validity ..."
## [1] "total 1342 mRNAs left after component length filter ..."
## [1] "total 307 ncRNAs left after ncRNA length filter ..."
## [1] "generate components for all tx"
## [1] "generate components for mRNA"
## [1] "generate components for lncRNA"
## [1] "generate chiped transcriptome"
## [1] "generate coverage checking ranges for tx"
## [1] "generate coverage checking ranges for mrna"
## [1] "generate coverage checking ranges for ncRNA"

# Or use gff. file to generate guitarTxdb
# Or use getTxdb() to download TxDb from internet:
# txdb <- getTxdb(txGenomeVer="hg19")
# guitarTxdb <- makeGuitarTxdb(txdb)
```

You may now generate the `Guitar` plot from the named list of genome-based features.

```

GuitarPlot(txTxdb = txdb,
            stBedFiles = stBedFiles,
            miscOutFilePrefix = "example")

## [1] "20241029231307"
## [1] "There are 2946 transcripts of 2946 genes in the genome."
## [1] "total 2946 transcripts extracted ..."
## [1] "total 2719 transcripts left after ambiguity filter ..."
## [1] "total 2719 transcripts left after check chromosome validity ..."
## [1] "total 1342 mRNAs left after component length filter ..."
## [1] "total 307 ncRNAs left after ncRNA length filter ..."
## [1] "generate components for all tx"
## [1] "generate components for mRNA"
## [1] "generate components for lncRNA"
## [1] "generate chiped transcriptome"
## [1] "generate coverage checking ranges for tx"
## [1] "generate coverage checking ranges for mrna"
## [1] "generate coverage checking ranges for ncRNA"
## [1] "20241029231321"
## [1] "import BED file F:/biocbuild/bbs-3.20-bioc/tmpdir/Rtmp0SedPo/Rinst179984dd0513c/Guitar/extdata/mRNA.bed"
## [1] "sample 10 points for Group1"
## [1] "start figure plotting for tx ..."

## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceDown' is discouraged.
## i Use 'confidenceDown' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceUp' is discouraged.
## i Use 'confidenceUp' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceDown' is discouraged.
## i Use 'confidenceDown' instead.
## Warning: Use of 'densityCI$confidenceUp' is discouraged.
## i Use 'confidenceUp' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'vline_pos$x1' is discouraged.
## i Use 'x1' instead.
## Warning: Use of 'vline_pos$y1' is discouraged.
## i Use 'y1' instead.
## Warning: Use of 'vline_pos$x2' is discouraged.
## i Use 'x2' instead.
## Warning: Use of 'vline_pos$y2' is discouraged.
## i Use 'y2' instead.

## [1] "start figure plotting for mrna ..."

## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.

```

```

## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceDown' is discouraged.
## i Use 'confidenceDown' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceUp' is discouraged.
## i Use 'confidenceUp' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceDown' is discouraged.
## i Use 'confidenceDown' instead.
## Warning: Use of 'densityCI$confidenceUp' is discouraged.
## i Use 'confidenceUp' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'vline_pos$x1' is discouraged.
## i Use 'x1' instead.
## Warning: Use of 'vline_pos$y1' is discouraged.
## i Use 'y1' instead.
## Warning: Use of 'vline_pos$x2' is discouraged.
## i Use 'x2' instead.
## Warning: Use of 'vline_pos$y2' is discouraged.
## i Use 'y2' instead.

## [1] "start figure plotting for ncrna ..."

## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceDown' is discouraged.
## i Use 'confidenceDown' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceUp' is discouraged.
## i Use 'confidenceUp' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceDown' is discouraged.
## i Use 'confidenceDown' instead.
## Warning: Use of 'densityCI$confidenceUp' is discouraged.
## i Use 'confidenceUp' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'vline_pos$x1' is discouraged.
## i Use 'x1' instead.
## Warning: Use of 'vline_pos$y1' is discouraged.
## i Use 'y1' instead.
## Warning: Use of 'vline_pos$x2' is discouraged.
## i Use 'x2' instead.
## Warning: Use of 'vline_pos$y2' is discouraged.
## i Use 'y2' instead.

```

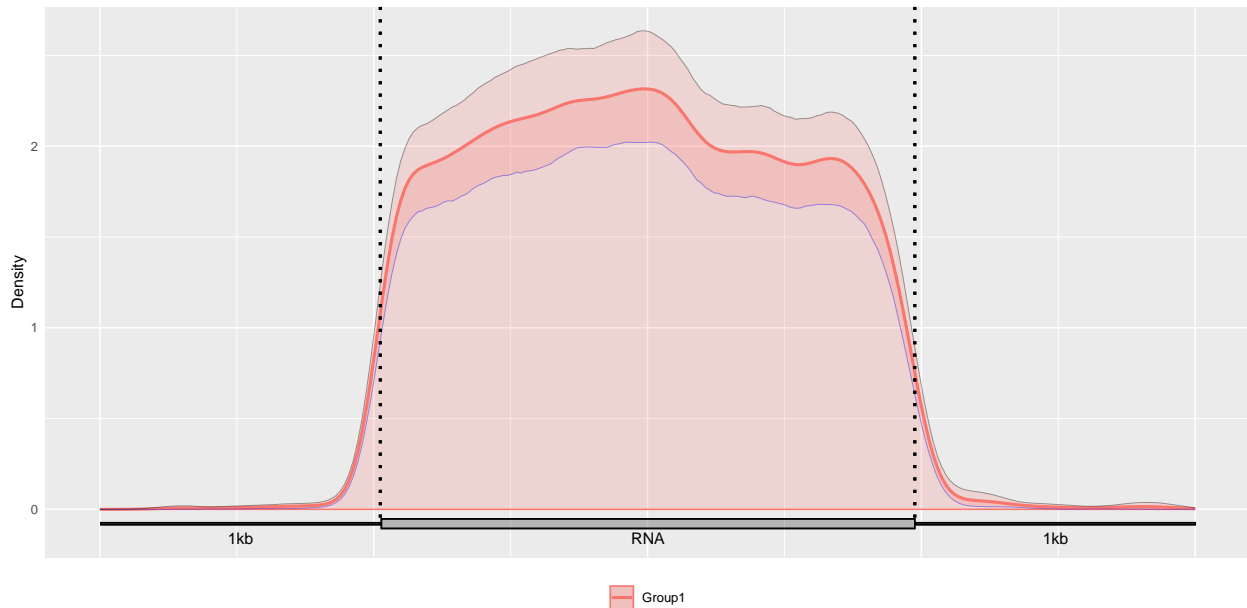
Alternatively, you may also optionally include the promoter DNA region and tail DNA region on the

5' and 3' side of a transcript in the plot with parameter headOrtail =TRUE.

```
GuitarPlot(txTxdb = txdb,
            stBedFiles = stBedFiles,
            headOrtail = TRUE)

## [1] "20241029231352"
## [1] "There are 2946 transcripts of 2946 genes in the genome."
## [1] "total 2946 transcripts extracted ..."
## [1] "total 2719 transcripts left after ambiguity filter ..."
## [1] "total 2719 transcripts left after check chromosome validity ..."
## [1] "total 1342 mRNAs left after component length filter ..."
## [1] "total 307 ncRNAs left after ncRNA length filter ..."
## [1] "generate components for all tx"
## [1] "generate components for mRNA"
## [1] "generate components for lncRNA"
## [1] "generate chiped transcriptome"
## [1] "generate coverage checking ranges for tx"
## [1] "generate coverage checking ranges for mrna"
## [1] "generate coverage checking ranges for ncRNA"
## [1] "20241029231406"
## [1] "import BED file F:/biocbuild/bbs-3.20-bioc/tmpdir/Rtmp0SedPo/Rinst179984dd0513c/Guitar/extdata/n"
## [1] "sample 10 points for Group1"
## [1] "start figure plotting for tx ..."

## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceDown' is discouraged.
## i Use 'confidenceDown' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceUp' is discouraged.
## i Use 'confidenceUp' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceDown' is discouraged.
## i Use 'confidenceDown' instead.
## Warning: Use of 'densityCI$confidenceUp' is discouraged.
## i Use 'confidenceUp' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'vline_pos$x1' is discouraged.
## i Use 'x1' instead.
## Warning: Use of 'vline_pos$y1' is discouraged.
## i Use 'y1' instead.
## Warning: Use of 'vline_pos$x2' is discouraged.
## i Use 'x2' instead.
## Warning: Use of 'vline_pos$y2' is discouraged.
## i Use 'y2' instead.
```



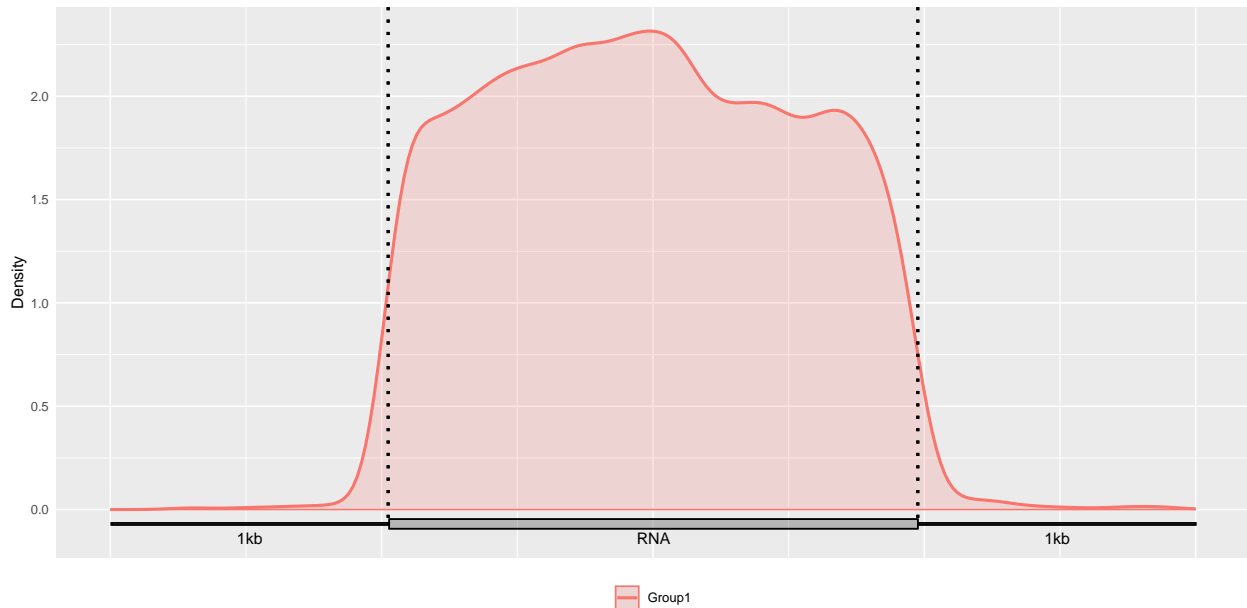
Alternatively, you may also optionally include the Confidence Interval for guitar plot with parameter `enableCI = FALSE`.

```
GuitarPlot(txTxdb = txdb,
            stBedFiles = stBedFiles,
            headOrtail = TRUE,
            enableCI = FALSE)

## [1] "20241029231425"
## [1] "There are 2946 transcripts of 2946 genes in the genome."
## [1] "total 2946 transcripts extracted ..."
## [1] "total 2719 transcripts left after ambiguity filter ..."
## [1] "total 2719 transcripts left after check chromosome validity ..."
## [1] "total 1342 mRNAs left after component length filter ..."
## [1] "total 307 ncRNAs left after ncRNA length filter ..."
## [1] "generate components for all tx"
## [1] "generate components for mRNA"
## [1] "generate components for lncRNA"
## [1] "generate chipped transcriptome"
## [1] "generate coverage checking ranges for tx"
## [1] "generate coverage checking ranges for mrna"
## [1] "generate coverage checking ranges for ncrna"
## [1] "20241029231438"
## [1] "import BED file F:/biocbuild/bbs-3.20-bioc/tmpdir/Rtmp0SedPo/Rinst179984dd0513c/Guitar/extdata/r"
## [1] "sample 10 points for Group1"
## [1] "start figure plotting for tx ..."

## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'vline_pos$x1' is discouraged.
## i Use 'x1' instead.
```

```
## Warning: Use of 'vline_pos$y1' is discouraged.
## i Use 'y1' instead.
## Warning: Use of 'vline_pos$x2' is discouraged.
## i Use 'x2' instead.
## Warning: Use of 'vline_pos$y2' is discouraged.
## i Use 'y2' instead.
```



2 Supported Data Format

Besides BED file, Guitar package also supports GRangesList and GRanges data structures. Please see the following examples.

```
# import different data formats into a named list object.
# These genomic features are using mm10 genome assembly
stBedFiles <- list(system.file("extdata", "m6A_mm10_exomePeak_1000peaks_bed12.bed",
                             package="Guitar"),
                  system.file("extdata", "m6A_mm10_exomePeak_1000peaks_bed6.bed",
                             package="Guitar"))

# Build Guitar Coordinates
txdb_file <- system.file("extdata", "mm10_toy.sqlite",
                        package="Guitar")
txdb <- loadDb(txdb_file)

# Guitar Plot
GuitarPlot(txTxdb = txdb,
           stBedFiles = stBedFiles,
           headOrtail = TRUE,
           enableCI = FALSE,
           mapFilterTranscript = TRUE,
           pltTxType = c("mrna"),
```



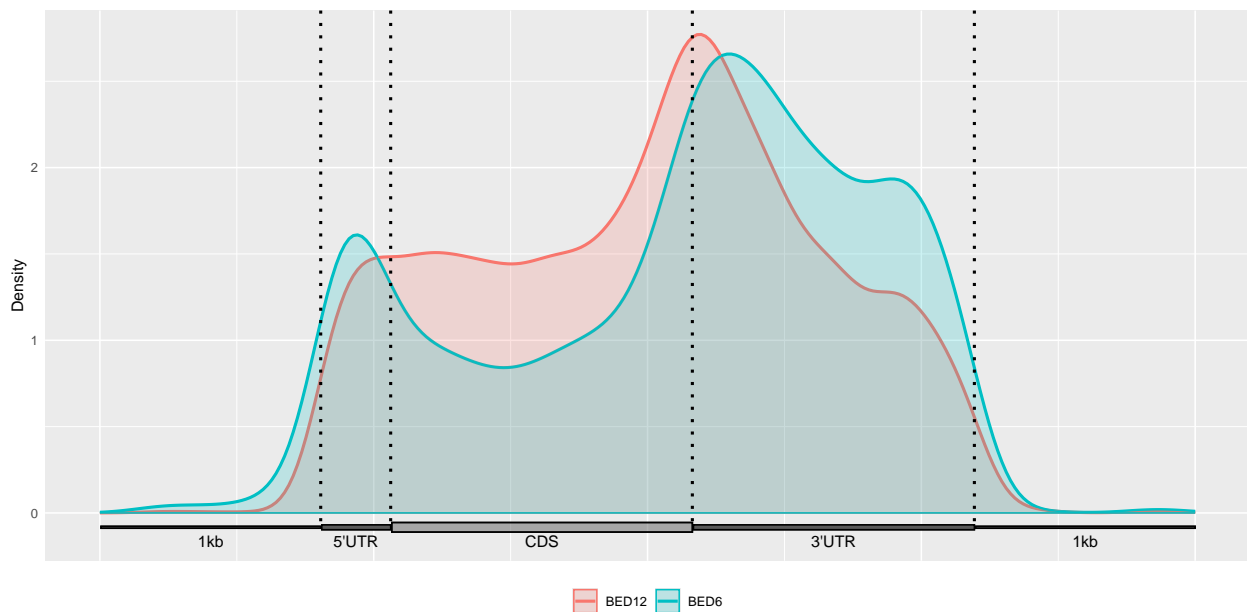
```

stGroupName = c("BED12", "BED6"))

## [1] "20241029231442"
## [1] "There are 2946 transcripts of 2946 genes in the genome."
## [1] "total 2946 transcripts extracted ..."
## [1] "total 2719 transcripts left after ambiguity filter ..."
## [1] "total 2719 transcripts left after check chromosome validity ..."
## [1] "total 1342 mRNAs left after component length filter ..."
## [1] "total 307 ncRNAs left after ncRNA length filter ..."
## [1] "generate components for mRNA"
## [1] "generate chiped transcriptome"
## [1] "generate coverage checking ranges for mrna"
## [1] "20241029231453"
## [1] "import BED file F:/biocbuild/bbs-3.20-bioc/tmpdir/Rtmp0SedPo/Rinst179984dd0513c/Guitar/extdata/n
## [1] "import BED file F:/biocbuild/bbs-3.20-bioc/tmpdir/Rtmp0SedPo/Rinst179984dd0513c/Guitar/extdata/n
## [1] "sample 10 points for BED12"
## [1] "sample 10 points for BED6"
## [1] "start figure plotting for mrna ..."

## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'vline_pos$x1' is discouraged.
## i Use 'x1' instead.
## Warning: Use of 'vline_pos$y1' is discouraged.
## i Use 'y1' instead.
## Warning: Use of 'vline_pos$x2' is discouraged.
## i Use 'x2' instead.
## Warning: Use of 'vline_pos$y2' is discouraged.
## i Use 'y2' instead.

```



3 Processing of sampling sites information

We can select parameters for site sampling.

```
stGRangeLists = vector("list", length(stBedFiles))
sitesPoints <- list()
for (i in seq_len(length(stBedFiles))) {
  stGRangeLists[[i]] <- blocks(import(stBedFiles[[i]]))
}
for (i in seq_len(length(stGRangeLists))) {
  sitesPoints[[i]] <- samplePoints(stGRangeLists[i],
                                   stSampleNum = 10,
                                   stAmbiguity = 5,
                                   pltTxType = c("mrna"),
                                   stSampleModle = "Equidistance",
                                   mapFilterTranscript = FALSE,
                                   guitarTxdb = guitarTxdb)
}
```

4 Guitar Coordinates - Transcriptomic Landmarks Projected on Genome

The guitarTxdb object contains the genome-projected transcriptome coordinates, which can be valuable for evaluating transcriptomic information related applications, such as checking the quality of MeRIP-Seq data. The Guitar coordinates are essentially the genomic projection of standardized transcript-based coordinates, making a viable bridge between the landmarks on transcript and genome-based coordinates.

It is based on the *txdb* object input, extracts the transcript information in *txdb*, selects the transcripts that match the parameters according to the component parameters set by the user, and saves according to the transcript type (tx, mrna, ncRNA).

```
guitarTxdb <- makeGuitarTxdb(txdb = txdb,
                             txAmbiguity = 5,
                             txMrnaComponentProp = c(0.1,0.15,0.6,0.05,0.1),
                             txLncrnaComponentProp = c(0.2,0.6,0.2),
                             pltTxType = c("tx","mrna","ncrna"),
                             txPrimaryOnly = FALSE)

## [1] "There are 2946 transcripts of 2946 genes in the genome."
## [1] "total 2946 transcripts extracted ..."
## [1] "total 2719 transcripts left after ambiguity filter ..."
## [1] "total 2719 transcripts left after check chromosome validity ..."
## [1] "total 1342 mRNAs left after component length filter ..."
## [1] "total 307 ncRNAs left after ncRNA length filter ..."
## [1] "generate components for all tx"
## [1] "generate components for mRNA"
## [1] "generate components for lncRNA"
## [1] "generate chiped transcriptome"
## [1] "generate coverage checking ranges for tx"
## [1] "generate coverage checking ranges for mrna"
## [1] "generate coverage checking ranges for ncRNA"
```

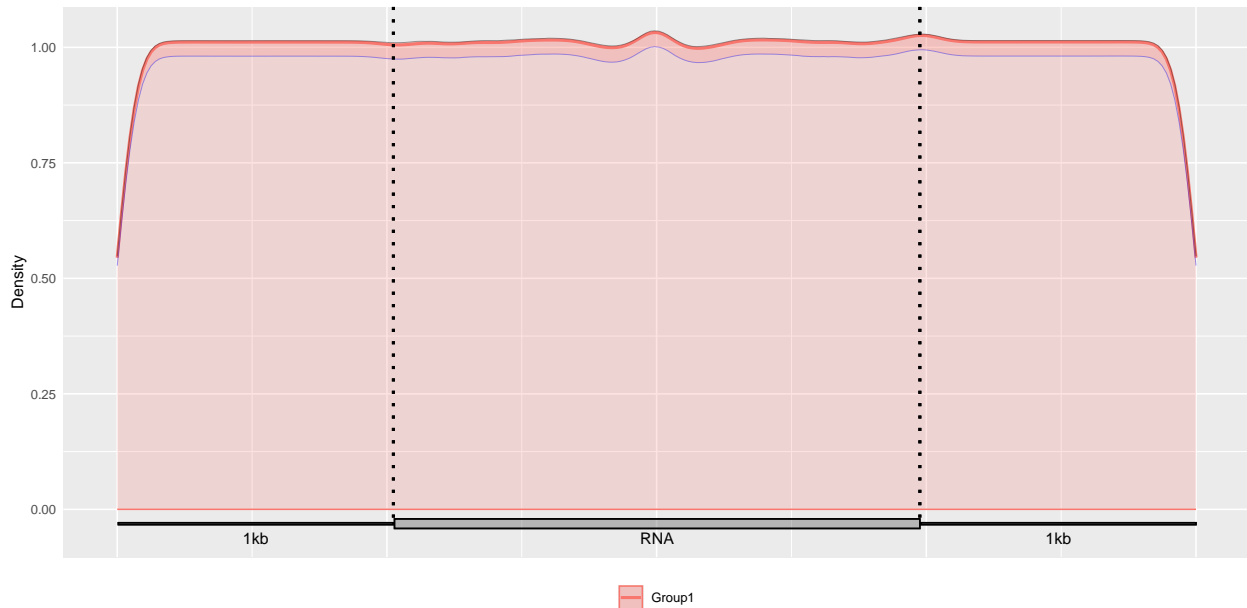
5 Check the Overlapping between Different Components

We can also check the distribution of the Guitar coordinates built.

```
gcl <- list(guitarTxdb$tx$tx)
GuitarPlot(txTxdb = txdb,
           stGRangeLists = gcl,
           stSampleNum = 200,
           enableCI = TRUE,
           pltTxType = c("tx"),
           txPrimaryOnly = FALSE
           )

## [1] "20241029231511"
## [1] "There are 2946 transcripts of 2946 genes in the genome."
## [1] "total 2946 transcripts extracted ..."
## [1] "total 2719 transcripts left after ambiguity filter ..."
## [1] "total 2719 transcripts left after check chromosome validity ..."
## [1] "total 1342 mRNAs left after component length filter ..."
## [1] "total 307 ncRNAs left after ncRNA length filter ..."
## [1] "generate components for all tx"
## [1] "generate chiped transcriptome"
## [1] "generate coverage checking ranges for tx"
## [1] "20241029231527"
## [1] "sample 200 points for Group1"
## [1] "start figure plotting for tx ..."

## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceDown' is discouraged.
## i Use 'confidenceDown' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceUp' is discouraged.
## i Use 'confidenceUp' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'densityCI$confidenceDown' is discouraged.
## i Use 'confidenceDown' instead.
## Warning: Use of 'densityCI$confidenceUp' is discouraged.
## i Use 'confidenceUp' instead.
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'vline_pos$x1' is discouraged.
## i Use 'x1' instead.
## Warning: Use of 'vline_pos$y1' is discouraged.
## i Use 'y1' instead.
## Warning: Use of 'vline_pos$x2' is discouraged.
## i Use 'x2' instead.
## Warning: Use of 'vline_pos$y2' is discouraged.
## i Use 'y2' instead.
```



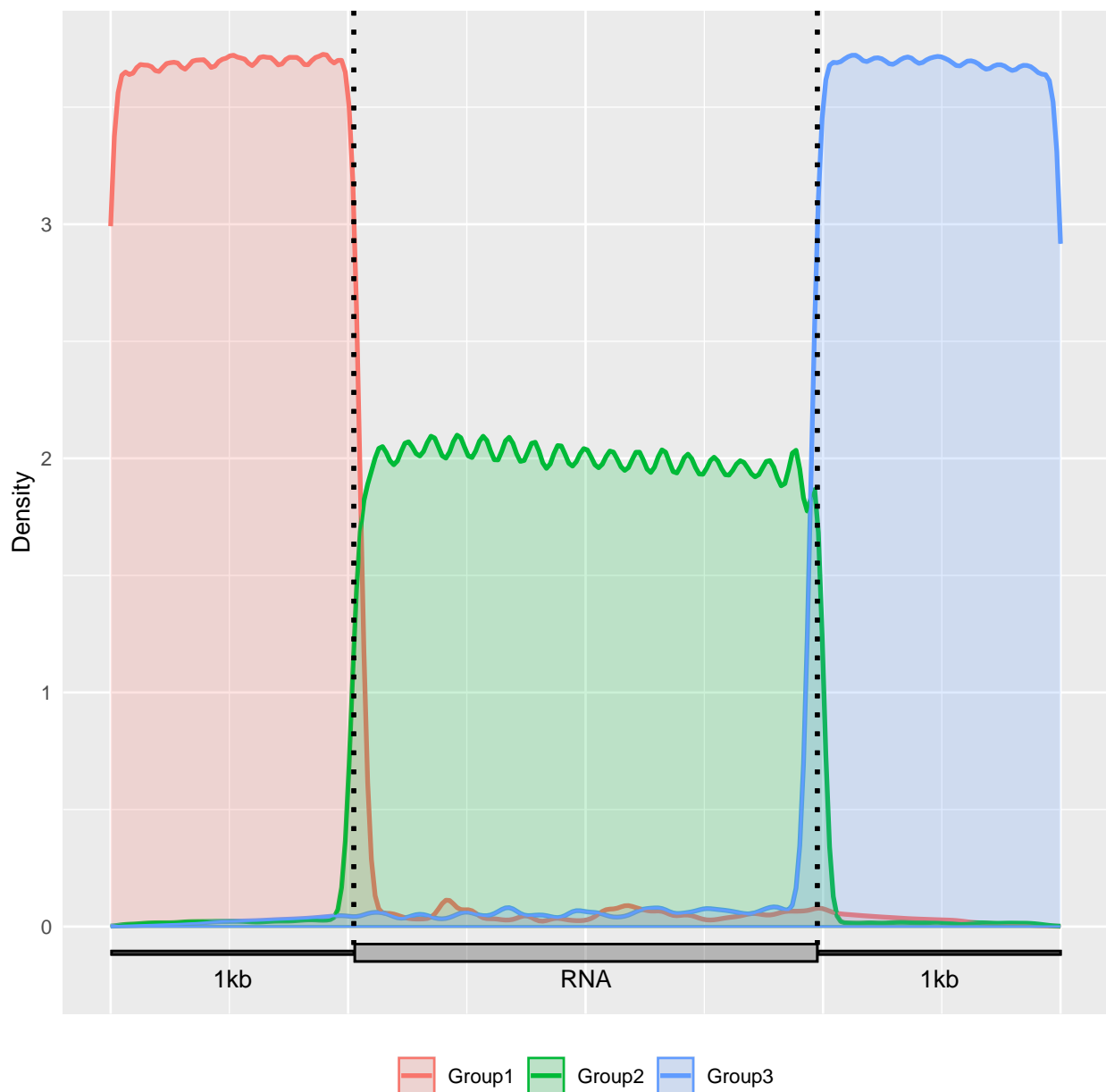
Alternatively, we can extract the RNA components, check the distribution of tx components in the transcriptome

```
GuitarCoords <- guitarTxdb$tx$txComponentGRange
type <- paste(mcols(GuitarCoords)$componentType, mcols(GuitarCoords)$txType)
key <- unique(type)
landmark <- list(1,2,3,4,5,6,7,8,9,10,11)
names(landmark) <- key
for (i in 1:length(key)) {
  landmark[[i]] <- GuitarCoords[type==key[i]]
}
GuitarPlot(txTxdb = txdb ,
            stGRangeLists = landmark[1:3],
            pltTxType = c("tx"),
            enableCI = FALSE
)

## [1] "20241029232300"
## [1] "There are 2946 transcripts of 2946 genes in the genome."
## [1] "total 2946 transcripts extracted ..."
## [1] "total 2719 transcripts left after ambiguity filter ..."
## [1] "total 2719 transcripts left after check chromosome validity ..."
## [1] "total 1342 mRNAs left after component length filter ..."
## [1] "total 307 ncRNAs left after ncRNA length filter ..."
## [1] "generate components for all tx"
## [1] "generate chiped transcriptome"
## [1] "generate coverage checking ranges for tx"
## [1] "20241029232317"
## [1] "sample 10 points for Group1"
## [1] "sample 10 points for Group2"
## [1] "sample 10 points for Group3"
## [1] "start figure plotting for tx ..."

## Warning: Use of 'densityCI$x' is discouraged.
```

```
## i Use 'x' instead.
## Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'vline_pos$x1' is discouraged.
## i Use 'x1' instead.
## Warning: Use of 'vline_pos$y1' is discouraged.
## i Use 'y1' instead.
## Warning: Use of 'vline_pos$x2' is discouraged.
## i Use 'x2' instead.
## Warning: Use of 'vline_pos$y2' is discouraged.
## i Use 'y2' instead.
```



Check the distribution of mRNA components in the transcriptome

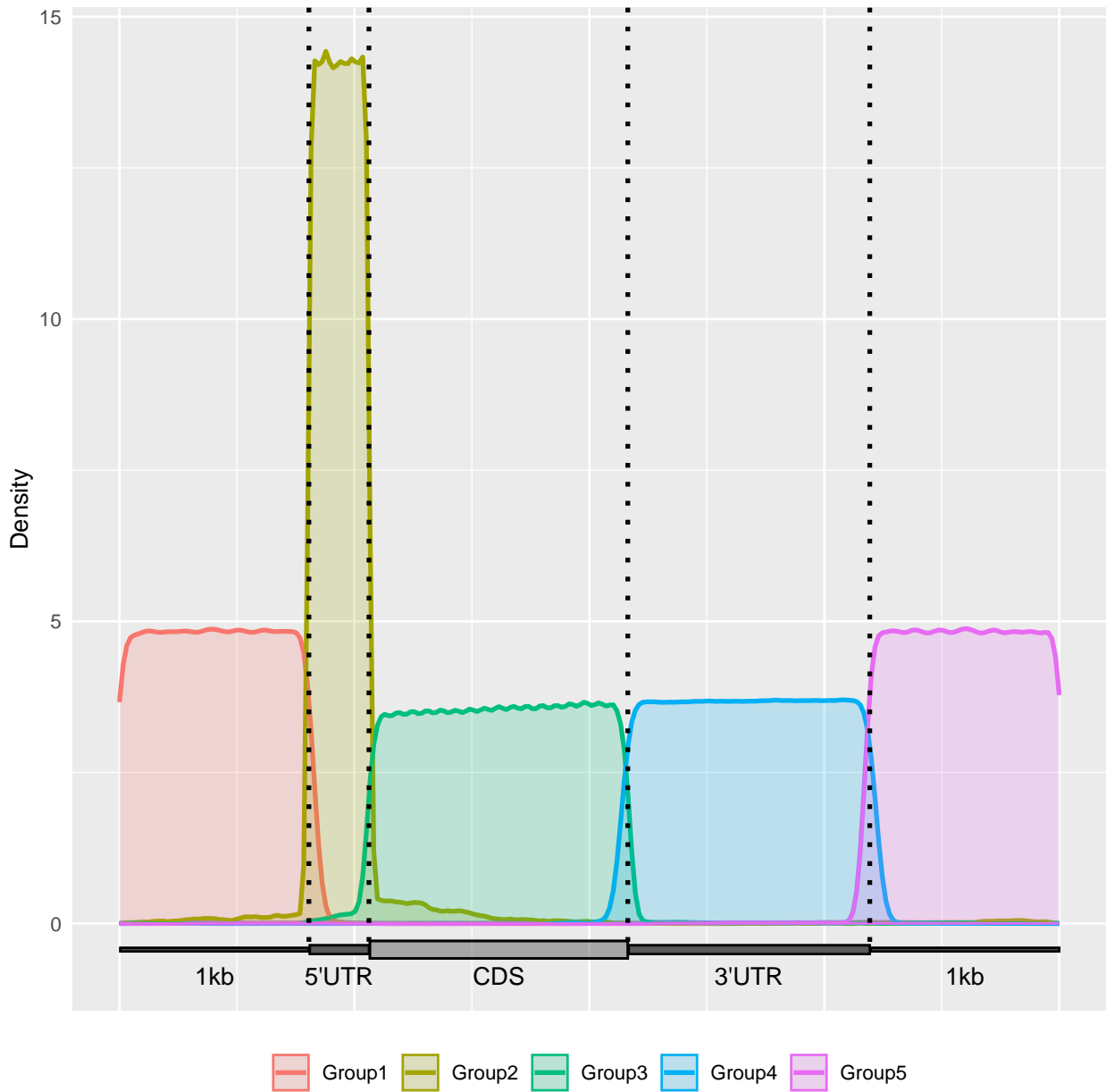
```

GuitarPlot(txTxdb = txdb ,
            stGRangeLists = landmark[4:8],
            pltTxType = c("mrna"),
            enableCI = FALSE
)

## [1] "20241029232338"
## [1] "There are 2946 transcripts of 2946 genes in the genome."
## [1] "total 2946 transcripts extracted ..."
## [1] "total 2719 transcripts left after ambiguity filter ..."
## [1] "total 2719 transcripts left after check chromosome validity ..."
## [1] "total 1342 mRNAs left after component length filter ..."
## [1] "total 307 ncRNAs left after ncRNA length filter ..."
## [1] "generate components for mRNA"
## [1] "generate chiped transcriptome"
## [1] "generate coverage checking ranges for mrna"
## [1] "20241029232351"
## [1] "sample 10 points for Group1"
## [1] "sample 10 points for Group2"
## [1] "sample 10 points for Group3"
## [1] "sample 10 points for Group4"
## [1] "sample 10 points for Group5"
## [1] "start figure plotting for mrna ..."

## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'vline_pos$x1' is discouraged.
## i Use 'x1' instead.
## Warning: Use of 'vline_pos$y1' is discouraged.
## i Use 'y1' instead.
## Warning: Use of 'vline_pos$x2' is discouraged.
## i Use 'x2' instead.
## Warning: Use of 'vline_pos$y2' is discouraged.
## i Use 'y2' instead.

```



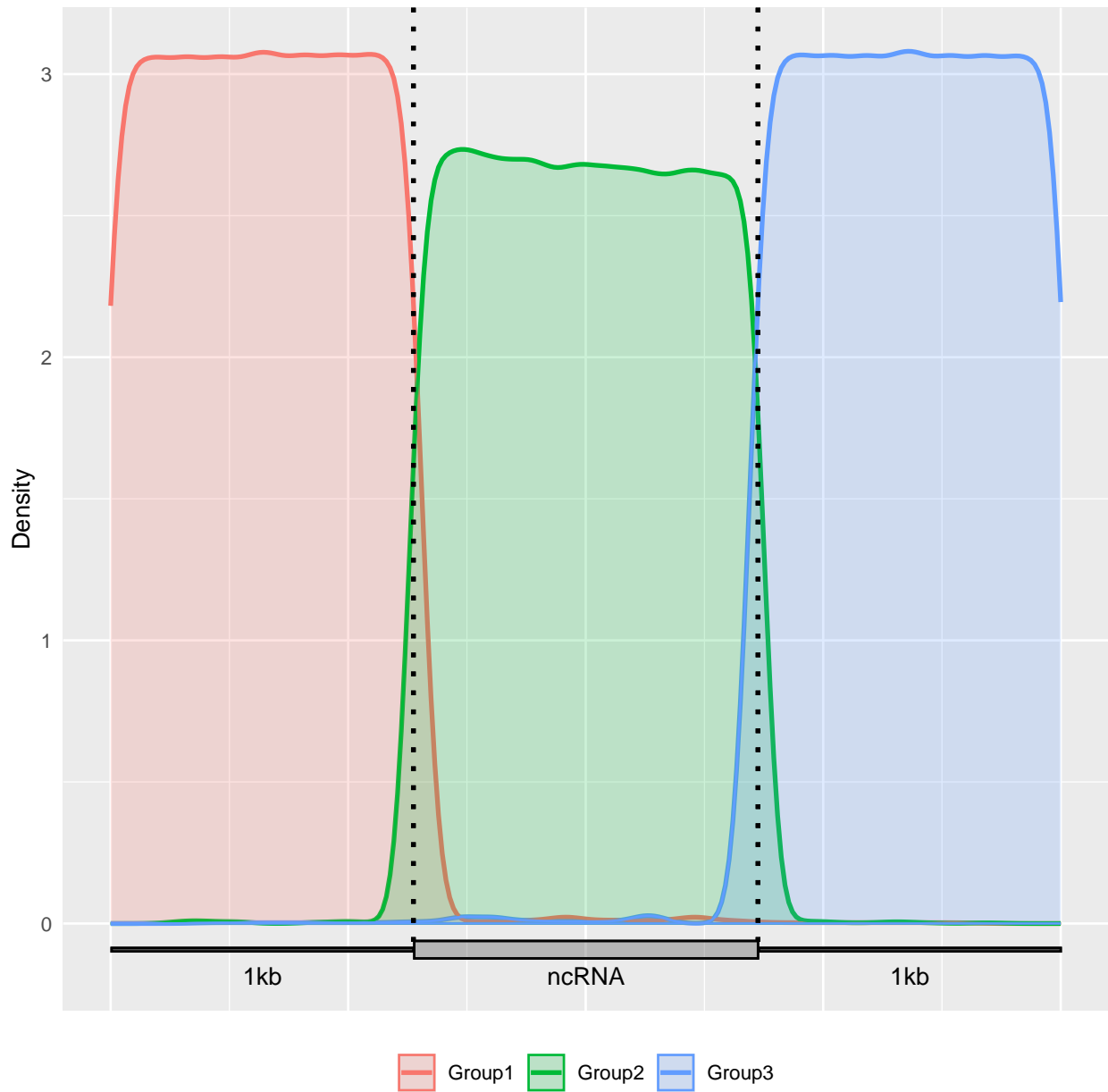
Check the distribution of lncRNA components in the transcriptome

```
GuitarPlot(txTxdb = txdb ,
            stGRangeLists = landmark[9:11],
            pltTxType = c("ncrna"),
            enableCI = FALSE
)

## [1] "20241029232406"
## [1] "There are 2946 transcripts of 2946 genes in the genome."
## [1] "total 2946 transcripts extracted ..."
## [1] "total 2719 transcripts left after ambiguity filter ..."
## [1] "total 2719 transcripts left after check chromosome validity ..."
## [1] "total 1342 mRNAs left after component length filter ..."
## [1] "total 307 ncRNAs left after ncRNA length filter ..."
```

```
## [1] "generate components for lncRNA"
## [1] "generate chiped transcriptome"
## [1] "generate coverage checking ranges for ncrna"
## [1] "20241029232428"
## [1] "sample 10 points for Group1"
## [1] "sample 10 points for Group2"
## [1] "sample 10 points for Group3"
## [1] "start figure plotting for ncrna ..."
```

```
## Warning: Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Use of 'densityCI$x' is discouraged.
## i Use 'x' instead.
## Warning: Use of 'vline_pos$x1' is discouraged.
## i Use 'x1' instead.
## Warning: Use of 'vline_pos$y1' is discouraged.
## i Use 'y1' instead.
## Warning: Use of 'vline_pos$x2' is discouraged.
## i Use 'x2' instead.
## Warning: Use of 'vline_pos$y2' is discouraged.
## i Use 'y2' instead.
```

6 Session Information

```
sessionInfo()

## R version 4.4.1 (2024-06-14 ucrt)
## Platform: x86_64-w64-mingw32/x64
## Running under: Windows Server 2022 x64 (build 20348)
##
## Matrix products: default
##
##
```

```

## locale:
## [1] LC_COLLATE=C
## [2] LC_CTYPE=English_United States.utf8
## [3] LC_MONETARY=English_United States.utf8
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.utf8
##
## time zone: America/New_York
## tzcode source: internal
##
## attached base packages:
## [1] stats4      stats      graphics  grDevices utils
## [6] datasets    methods    base
##
## other attached packages:
## [1] Guitar_2.22.0      dplyr_1.1.4
## [3] ggplot2_3.5.1      magrittr_2.0.3
## [5] rtracklayer_1.66.0 GenomicFeatures_1.58.0
## [7] AnnotationDbi_1.68.0 Biobase_2.66.0
## [9] GenomicRanges_1.58.0 GenomeInfoDb_1.42.0
## [11] IRanges_2.40.0      S4Vectors_0.44.0
## [13] BiocGenerics_0.52.0 knitr_1.48
##
## loaded via a namespace (and not attached):
## [1] KEGGREST_1.46.0      SummarizedExperiment_1.36.0
## [3] gtable_0.3.6         rjson_0.2.23
## [5] xfun_0.48            lattice_0.22-6
## [7] generics_0.1.3       vctrs_0.6.5
## [9] tools_4.4.1          bitops_1.0-9
## [11] curl_5.2.3           parallel_4.4.1
## [13] tibble_3.2.1         fansi_1.0.6
## [15] RSQLite_2.3.7        highr_0.11
## [17] blob_1.2.4           pkgconfig_2.0.3
## [19] Matrix_1.7-1         lifecycle_1.0.4
## [21] GenomeInfoDbData_1.2.13 farver_2.1.2
## [23] compiler_4.4.1       textshaping_0.4.0
## [25] Rsamtools_2.22.0     Biostrings_2.74.0
## [27] munsell_0.5.1        codetools_0.2-20
## [29] RCurl_1.98-1.16      yaml_2.3.10
## [31] pillar_1.9.0         crayon_1.5.3
## [33] BiocParallel_1.40.0  cachem_1.1.0
## [35] DelayedArray_0.32.0  abind_1.4-8
## [37] tidyselect_1.2.1     restfulr_0.0.15
## [39] labeling_0.4.3       fastmap_1.2.0
## [41] grid_4.4.1           colorspace_2.1-1
## [43] cli_3.6.3            SparseArray_1.6.0
## [45] S4Arrays_1.6.0       XML_3.99-0.17
## [47] utf8_1.2.4           withr_3.0.2
## [49] UCSC.utils_1.2.0     scales_1.3.0
## [51] bit64_4.5.2          XVector_0.46.0
## [53] httr_1.4.7           matrixStats_1.4.1
## [55] bit_4.5.0            ragg_1.3.3
## [57] png_0.1-8            memoise_2.0.1
## [59] evaluate_1.0.1       BiocIO_1.16.0

```

```
## [61] rlang_1.1.4          glue_1.8.0
## [63] DBI_1.2.3            jsonlite_1.8.9
## [65] R6_2.5.1             systemfonts_1.1.0
## [67] MatrixGenerics_1.18.0 GenomicAlignments_1.42.0
## [69] zlibbioc_1.52.0
```