PepBed examples

November 18, 2017

importing bigbed files

```
library(PepBed)

# path to bigbed file
bigbedpath <- '/home/enrique/temp/pride_cluster_peptides_9606_Human.pogo.bb'

# convert bigbed to bed file (output bed file in the same directory)
bigbed2bed(inputFile = bigbedpath, compress = FALSE)

# getting basic information (output description file in the same directory)
getBigBedInfo(inputFile = bigbedpath)

# getting field names if available
fieldNames <- getBigBedFieldNames(inputFile = bigbedpath, only.names = TRUE)

print(fieldNames)
```

```
##  [1] "chrom"       "chromStart"  "chromEnd"    "name"        "score"
##  [6] "strand"      "thickStart"  "thickEnd"    "reserved"    "blockCount"
## [11] "blockSizes"  "chromStarts"
```

parsing Bed file

```
# path to bed file
bedpath <- '/home/enrique/temp/pride_cluster_peptides_9606_Human.bed'

# import bed file as  dataframe
df <- readBedFile(inputFile = bedpath)

# set column name to bed file
names(df) <- fieldNames

# convert dataframe to GRanges
granges_peptide <- buildGRangesFromData(data = df, chrColName = "chrom",
                                                   startColName = "chromStart",
                                                   endColName = "chromEnd")
```

some basic stats

```
# getting number of features(peptides) by chromosome
counts <- countsByChromosome(gr = granges_peptide)
print(counts)
```

```
##    chromosome countFeatures
## 1        chr1          9736
## 2        chr2          3688
## 3        chr3          5991
## 4        chr4          6191
## 5        chr5          1183
## 6        chr6          3822
## 7        chr7          2704
## 8        chr8          3524
## 9        chr9          6678
## 10      chr10          1176
## 11      chr11          5536
## 12      chr12          7490
## 13      chr13          2264
## 14      chr14           867
## 15      chr15          2236
## 16      chr16          5200
## 17      chr17          3505
## 18      chr18          4013
## 19      chr19          6463
## 20      chr20          4121
## 21      chr21          3064
## 22      chr22          3841
## 23       chrX            19
## 24       chrY          3138
## 25       chrM           131
```

```r
# removing duplicated entries from original granges_peptide
gr_unique <- getUniqueFeatures(granges_peptide, colFeatures = 'name')

# getting unique number of features(peptides) by chromosome
counts_unique <- countsByChromosome(gr = gr_unique)
print(counts_unique)
```

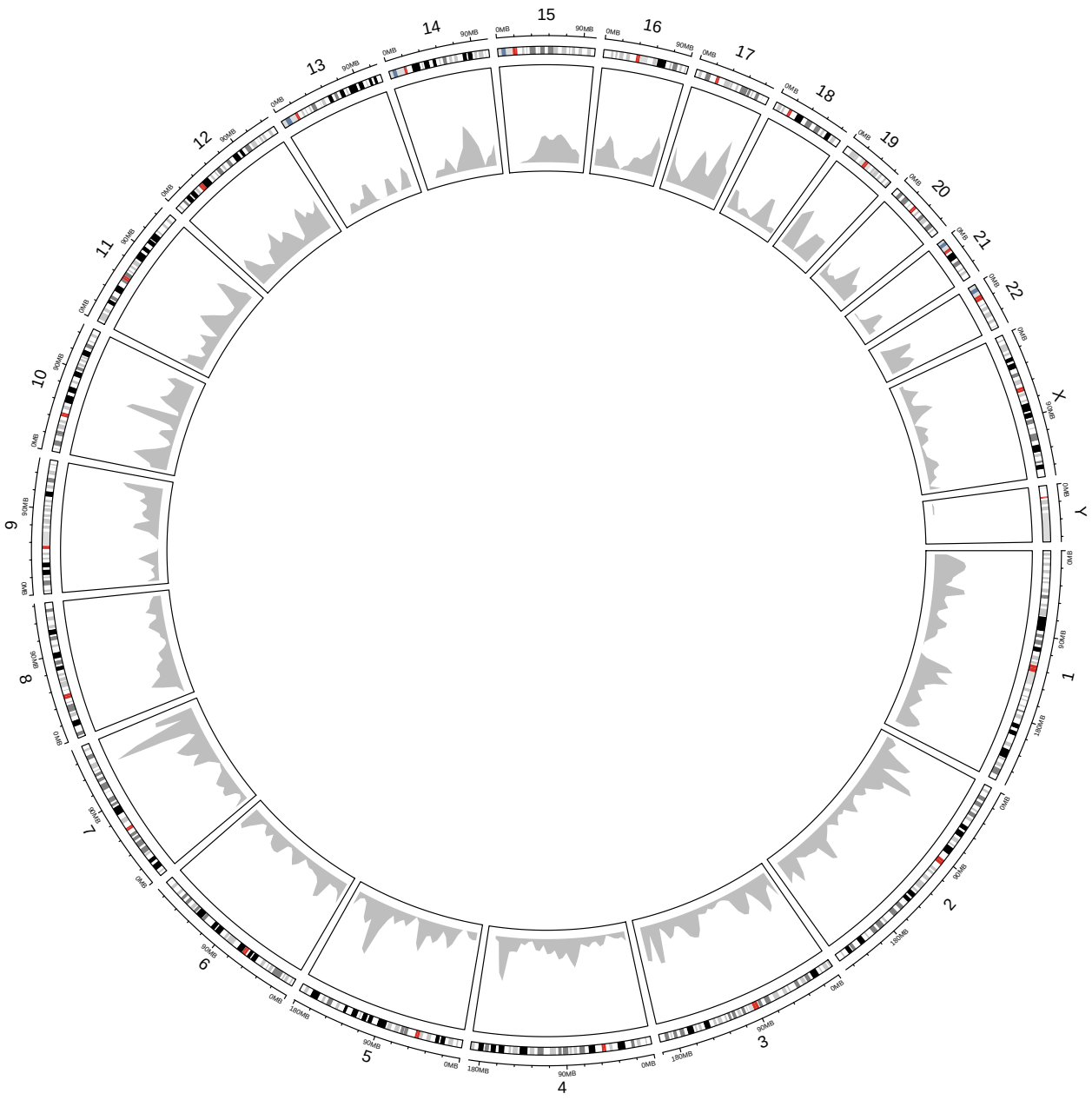
```
##    chromosome countFeatures
## 1        chr1          8801
## 2        chr2          3379
## 3        chr3          5341
## 4        chr4          4985
## 5        chr5          1031
## 6        chr6          3123
## 7        chr7          2336
## 8        chr8          2942
## 9        chr9          5491
## 10      chr10          1016
## 11      chr11          4666
## 12      chr12          6073
## 13      chr13          1935
## 14      chr14           666
## 15      chr15          1742
## 16      chr16          4622
## 17      chr17          3145
## 18      chr18          3385
## 19      chr19          3759
## 20      chr20          3338
## 21      chr21          2543
## 22      chr22          3224
## 23       chrX            19
## 24       chrY          2388
## 25       chrM             8
```

```r
## compute coverage of query (peptide evidences) on subject (transcripts) by crhomosome
data("protein_coding_transcript") # load protein coding transcript as GRanges object
coverage <- computeCoverageByChromosome(query = granges_peptide, subject = transcript)
print(coverage)
```

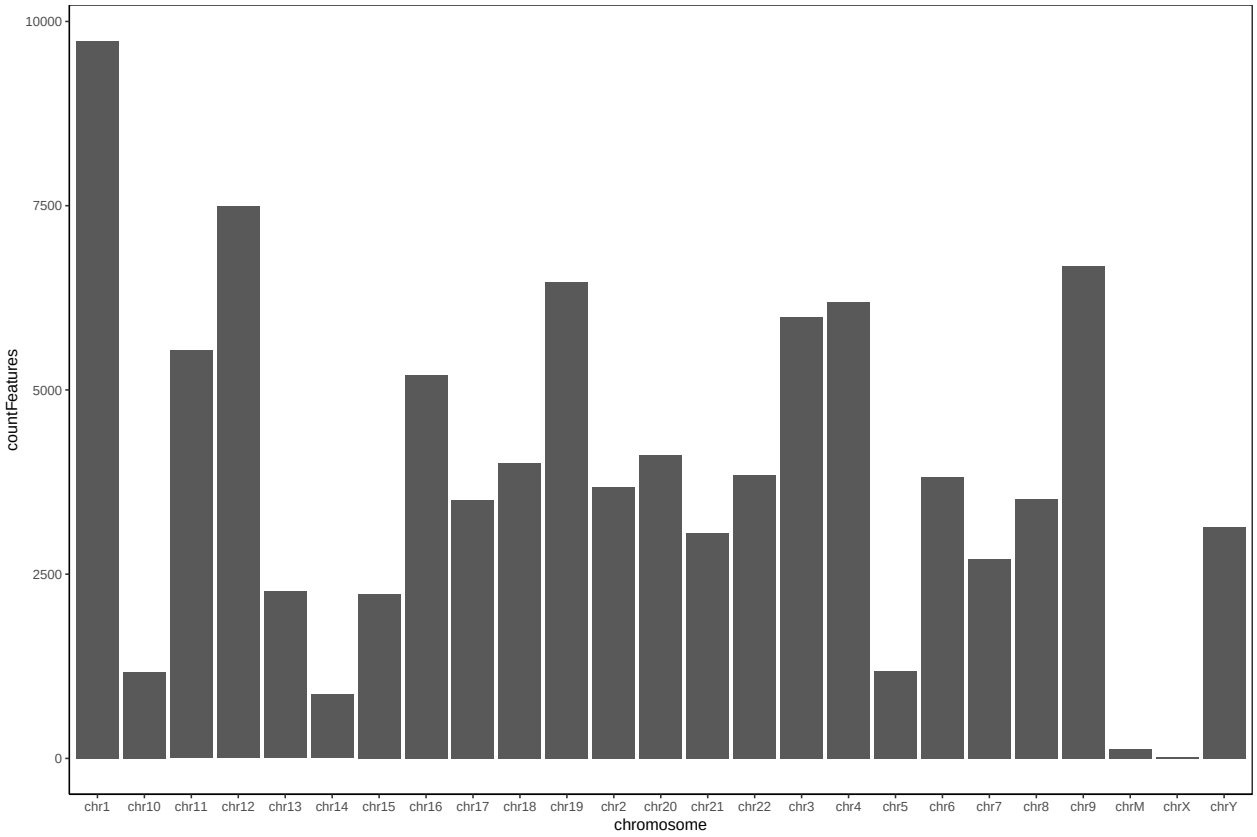
```
##    Chromosome Coverage
## 1        chr1   50.781
## 2        chr2   50.613
## 3        chr3   46.526
## 4        chr4   51.239
## 5        chr5   47.049
## 6        chr6   51.290
## 7        chr7   58.682
## 8        chr8   52.876
## 9        chr9   48.021
## 10      chr10   44.897
## 11      chr11   46.371
## 12      chr12   43.799
## 13      chr13   31.403
## 14      chr14   49.515
## 15      chr15   46.962
## 16      chr16   47.596
## 17      chr17   51.952
## 18      chr18   51.426
## 19      chr19   48.749
## 20      chr20   49.875
## 21      chr21   34.799
## 22      chr22   47.774
## 23       chrX   42.346
## 24       chrY   22.274
## 25       chrM    5.882
```

plotting

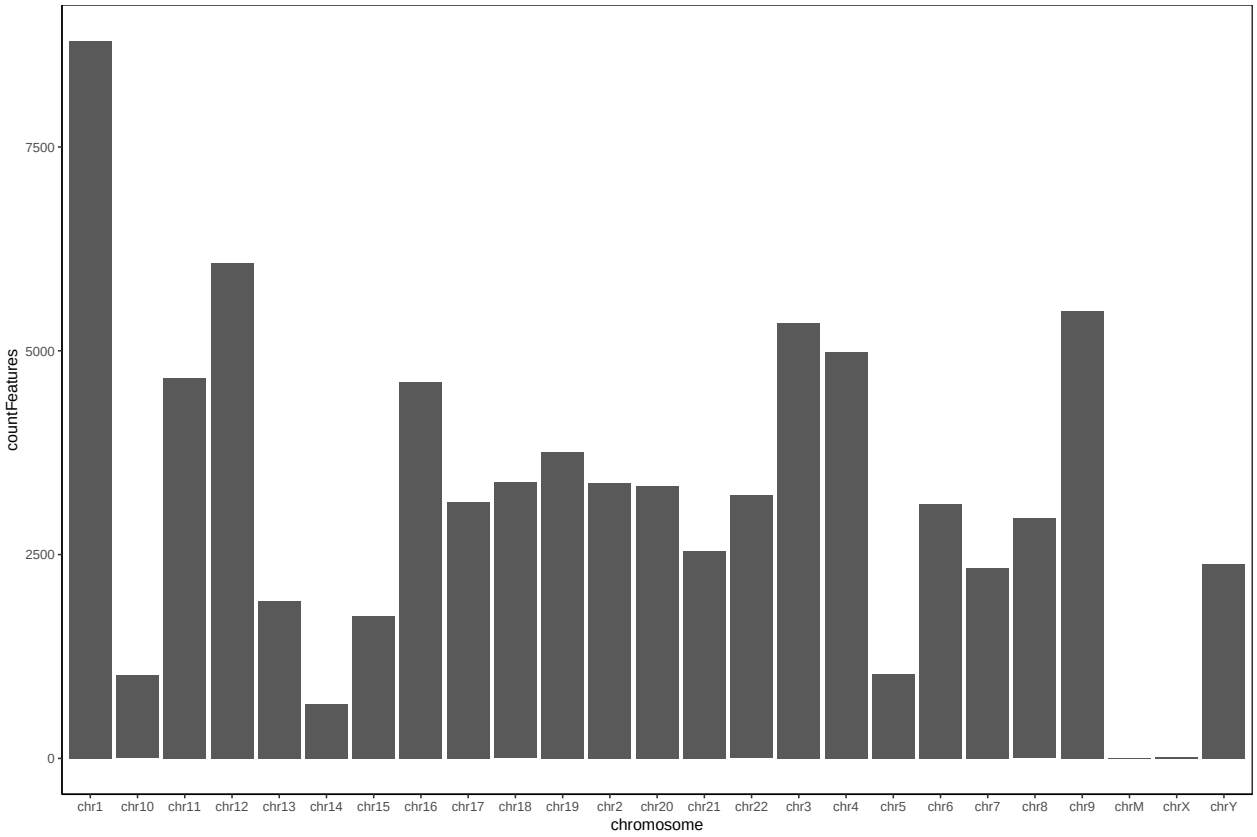
- The distribution of features (peptides) by chromosome



- barplot with counts by chromosome



- barplot with unique counts by chromosome



- barplot with coverage by chromosome