

Draw Color Chromatin

Ricky Lim

March 1, 2013

1 Goal

- Draw five color chromatin states (Filion, et al., 2010) on Drosophila Genome

2 Load Dataset

2.1 Get Chromatin States

```
library(rtracklayer) #to import bed file into gRanges

## Warning: package 'rtracklayer' was built under R version 2.15.2
## Loading required package: GenomicRanges
## Warning: package 'GenomicRanges' was built under R version 2.15.2
## Loading required package: BiocGenerics
##
## Attaching package: 'BiocGenerics'
## The following object(s) are masked from 'package:stats':
##
##     xtabs
## The following object(s) are masked from 'package:base':
##
##     anyDuplicated, cbind, colnames, duplicated, eval, Filter,
##     Find, get, intersect, lapply, Map, mapply, mget, order, paste,
##     pmax, pmax.int, pmin, pmin.int, Position, rbind, Reduce,
##     rep.int, rownames, sapply, setdiff, table, tapply, union,
##     unique
## Loading required package: IRanges
## Warning: package 'IRanges' was built under R version 2.15.2
```

```

parent_dir = "/Users/RickyLim/Dropbox/BlackChromatin_project/"
setwd(parent_dir)
print(getwd())

## [1] "/Users/RickyLim/Dropbox/BlackChromatin_project"

kc_state <- import("RawDataSets/GSE22069_Drosophila_chromatin_domains.bed")
colnames(kc_state) <- "State"
kc_state <- as(kc_state, "GRanges")
print(kc_state)

## GRanges with 8428 ranges and 1 metadata column:
##           seqnames           ranges strand |           State
##           <Rle>             <IRanges> <Rle> | <character>
##      [1]   chr2L      [ 5302, 58754]    *   |         BLACK
##      [2]   chr2L      [58764, 65319]    *   |          BLUE
##      [3]   chr2L      [65317, 76559]    *   |           RED
##      [4]   chr2L      [76694, 84511]    *   |          BLUE
##      [5]   chr2L      [84651, 89599]    *   |           RED
##      [6]   chr2L      [89597, 94102]    *   |          BLUE
##      [7]   chr2L      [94100, 95516]    *   |           RED
##      [8]   chr2L      [95662, 96387]    *   |        YELLOW
##      [9]   chr2L      [96491, 99976]    *   |          GREEN
##      ...      ...      ...      ...      ...
## [8420]   chrX [22110550, 22116799]    *   |          GREEN
## [8421]   chrX [22116797, 22121916]    *   |         BLACK
## [8422]   chrX [22122147, 22132184]    *   |          GREEN
## [8423]   chrX [22132295, 22234481]    *   |         BLACK
## [8424]   chrX [22234529, 22247097]    *   |          BLUE
## [8425]   chrX [22247095, 22261952]    *   |           RED
## [8426]   chrX [22261950, 22264800]    *   |          BLUE
## [8427]   chrX [22264941, 22407410]    *   |         BLACK
## [8428]   chrX [22408460, 22422827]    *   |          GREEN
## ---
##      seqlengths:
##      chr2L chr2R chr3L chr3R  chr4  chrX
##      NA    NA    NA    NA    NA    NA

```

2.2 Get States on the Drosophila Genome

```

library(BSgenome.Dmelanogaster.UCSC.dm3) #get the seq.length of dros.Genome

## Loading required package: BSgenome
## Loading required package: Biostrings
## Warning: package 'Biostrings' was built under R version 2.15.2

chr.len = seqlengths(Dmelanogaster)
# exclude chromosomes with suffix '_' , 'M', 'Het', 'extra'.
chr.len = chr.len[grepl("_|M|U|Het|extra", names(chr.len), invert = T)]
# order the chromosomes
kc_state = keepSeqlevels(kc_state, names(chr.len))
seqlevels(kc_state) = names(chr.len)
seqlengths(kc_state) = (chr.len)
print(kc_state)

## GRanges with 8428 ranges and 1 metadata column:
##           seqnames           ranges strand |           State
##           <Rle>             <IRanges> <Rle> | <character>
##      [1]   chr2L      [ 5302, 58754]    *   |         BLACK
##      [2]   chr2L      [58764, 65319]    *   |          BLUE
##      [3]   chr2L      [65317, 76559]    *   |           RED
##      [4]   chr2L      [76694, 84511]    *   |          BLUE
##      [5]   chr2L      [84651, 89599]    *   |           RED
##      [6]   chr2L      [89597, 94102]    *   |          BLUE
##      [7]   chr2L      [94100, 95516]    *   |           RED
##      [8]   chr2L      [95662, 96387]    *   |        YELLOW
##      [9]   chr2L      [96491, 99976]    *   |          GREEN
##      ...      ...      ...      ...      ...
## [8420]   chrX [22110550, 22116799]    *   |          GREEN
## [8421]   chrX [22116797, 22121916]    *   |         BLACK
## [8422]   chrX [22122147, 22132184]    *   |          GREEN
## [8423]   chrX [22132295, 22234481]    *   |         BLACK
## [8424]   chrX [22234529, 22247097]    *   |          BLUE
## [8425]   chrX [22247095, 22261952]    *   |           RED
## [8426]   chrX [22261950, 22264800]    *   |          BLUE
## [8427]   chrX [22264941, 22407410]    *   |         BLACK
## [8428]   chrX [22408460, 22422827]    *   |          GREEN
##      ---
##      seqlengths:
##      chr2L   chr2R   chr3L   chr3R   chr4   chrX
## 23011544 21146708 24543557 27905053 1351857 22422827

```

3 Let's Draw it

```
library(ggbio)

## Warning: package 'ggbio' was built under R version 2.15.2
## Loading required package: ggplot2
## Need specific help about ggbio? try mailing
## the maintainer or visit http://tengfei.github.com/ggbio/
##
## Attaching package: 'ggbio'
## The following object(s) are masked from 'package:ggplot2':
##
##   geom_bar, geom_rect, geom_segment, stat_bin, stat_identity,
##   xlim

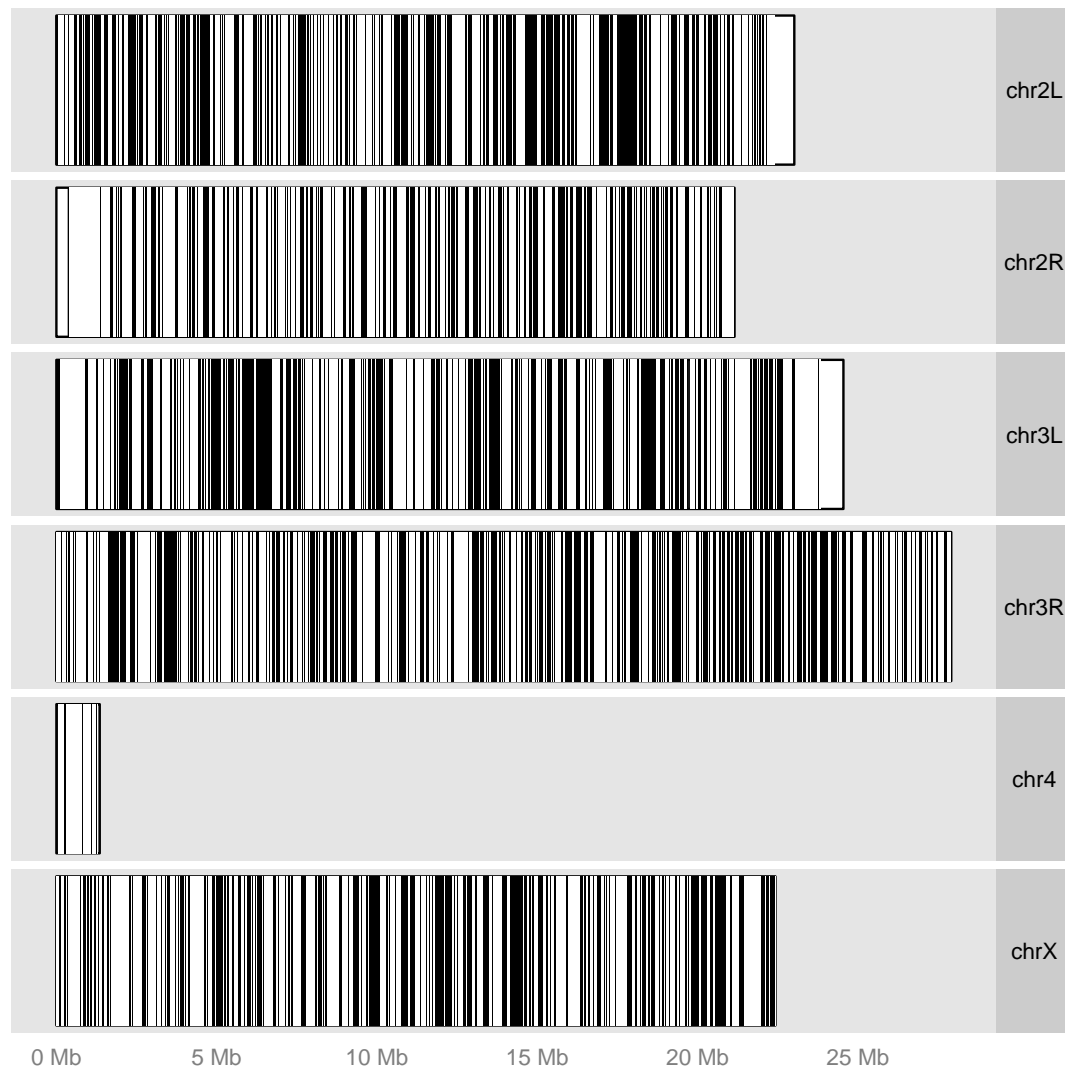
chrom.col <- c("black", "blue", "seagreen3", "red", "gold2")
black.col <- c("black", "white", "white", "white", "white")
p <- autoplot(kc_state, layout = "karyogram", aes(fill = State))

## Scale for 'x' is already present. Adding another scale for 'x', which will
## replace the existing scale.

p + scale_fill_manual(values = black.col) + opts(legend.position = "none")

## 'opts' is deprecated. Use 'theme' instead. (Deprecated; last used in version
## 0.9.1)

## Object of class "ggbio"
```

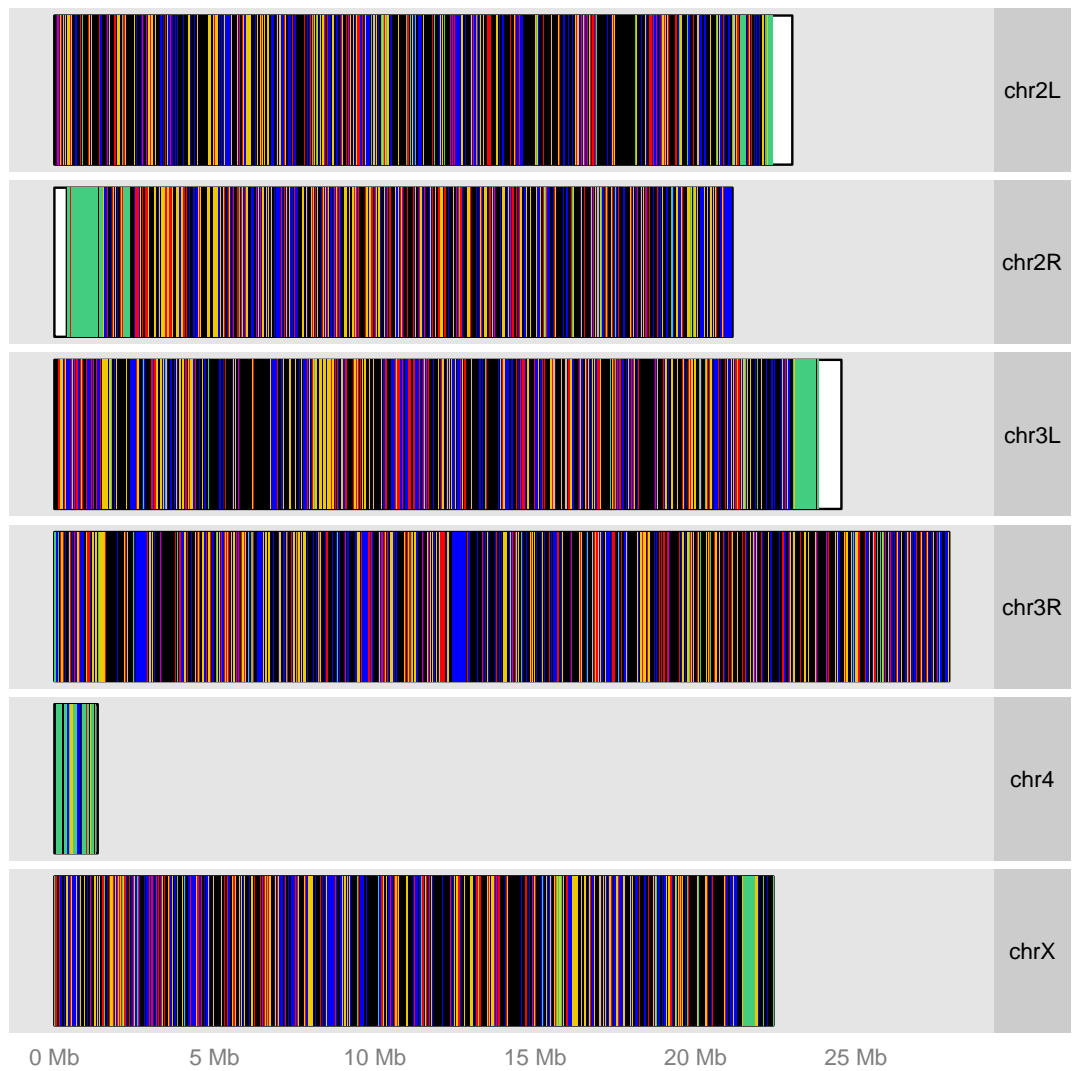


```
## NULL

p + scale_fill_manual(values = chrom.col) + opts(legend.position = "none")

## 'opts' is deprecated. Use 'theme' instead. (Deprecated; last used in version
0.9.1)

## Object of class "ggbio"
```



NULL

4 MetaInfo

```
## \begin{itemize}\raggedright
##   \item R version 2.15.1 (2012-06-22), \verb|x86_64-apple-darwin9.8.0|
##   \item Locale: \verb|en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8|
##   \item Base packages: base, datasets, graphics, grDevices,
##     methods, stats, utils
##   \item Other packages: BiocGenerics~0.4.0, Biostrings~2.26.3,
##     BSgenome~1.26.1, BSgenome.Dmelanogaster.UCSC.dm3~1.3.19,
##     GenomicRanges~1.10.6, ggbio~1.6.6, ggplot2~0.9.3,
##     IRanges~1.16.6, knitr~1.1, rtracklayer~1.18.2
##   \item Loaded via a namespace (and not attached):
##     AnnotationDbi~1.20.3, Biobase~2.18.0, biomaRt~2.14.0,
##     biovizBase~1.6.2, bitops~1.0-4.2, cluster~1.14.3,
##     colorspace~1.2-1, DBI~0.2-5, dichromat~2.0-0, digest~0.6.3,
##     evaluate~0.4.3, formatR~0.7, GenomicFeatures~1.10.1,
##     grid~2.15.1, gridExtra~0.9.1, gtable~0.1.2, Hmisc~3.10-1,
##     labeling~0.1, lattice~0.20-13, MASS~7.3-23, munsell~0.4,
##     parallel~2.15.1, plyr~1.8, proto~0.3-10, RColorBrewer~1.0-5,
##     RCurl~1.95-3, reshape2~1.2.2, Rsamtools~1.10.2,
##     RSQLite~0.11.2, scales~0.2.3, stats4~2.15.1, stringr~0.6.2,
##     tools~2.15.1, VariantAnnotation~1.4.9, XML~3.95-0.1,
##     zlibbioc~1.4.0
## \end{itemize}
```

- R version 2.15.1 (2012-06-22), x86_64-apple-darwin9.8.0
- Locale: en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
- Base packages: base, datasets, graphics, grDevices, methods, stats, utils
- Other packages: BiocGenerics 0.4.0, Biostrings 2.26.3, BSgenome 1.26.1, BSgenome.Dmelanogaster.UCSC.dm3 1.3.19, GenomicRanges 1.10.6, ggbio 1.6.6, ggplot2 0.9.3, IRanges 1.16.6, rtracklayer 1.18.2
- Loaded via a namespace (and not attached): AnnotationDbi 1.20.3, Biobase 2.18.0, biomaRt 2.14.0, biovizBase 1.6.2, bitops 1.0-4.2, cluster 1.14.3, colorspace 1.2-1, DBI 0.2-5, dichromat 2.0-0, digest 0.6.3, evaluate 0.4.3, formatR 0.7, GenomicFeatures 1.10.1, grid 2.15.1, gridExtra 0.9.1, gtable 0.1.2, Hmisc 3.10-1, knitr 1.1, labeling 0.1, lattice 0.20-13, MASS 7.3-23, munsell 0.4, parallel 2.15.1, plyr 1.8, proto 0.3-10, RColorBrewer 1.0-5, RCurl 1.95-3, reshape2 1.2.2, Rsamtools 1.10.2, RSQLite 0.11.2, scales 0.2.3, stats4 2.15.1, stringr 0.6.2, tools 2.15.1, VariantAnnotation 1.4.9, XML 3.95-0.1, zlibbioc 1.4.0