SushiPresentation

Erika Bueno

December 1, 2018

```
#--------------------------INSTALLING SUSHI FROM BIOCONDUCTOR-----------
------------
#Loading the package, sushi
#Use the BiocManager package to install and manage packages from the Bioc
onductor project for the statistical analysis and comprehension of high-t
hroughput genomic data.
if (!requireNamespace("BiocManager", quietly = TRUE))
    install.packages("BiocManager")
BiocManager::install("Sushi", version = "3.8")
```

```
## Bioconductor version 3.8 (BiocManager 1.30.4), R 3.5.1 (2018-07-02)
```

```
## Installing package(s) 'Sushi'
```

```
## package 'Sushi' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\ebuen\AppData\Local\Temp\RtmpIRlkMN\downloaded_packages
```

```
## installation path not writeable, unable to update packages: foreign,
##   lattice, MASS, Matrix, mgcv, survival
```

```
## Update old packages: 'ggtree'
```

```
#optional installing via devtools
library("devtools")
install_github("dphansti/Sushi")
```

```
## Downloading GitHub repo dphansti/Sushi@master
```

```
## Skipping 6 packages ahead of CRAN: AnnotationDbi, Biobase, BiocGeneric
s, biomaRt, IRanges, S4Vectors
```

```
##


   checking for file 'C:\Users\ebuen\AppData\Local\Temp\RtmpIRlkMN\remote
s56284d6d4025\dphansti-Sushi-f53dd94/DESCRIPTION' ...

   checking for file 'C:\Users\ebuen\AppData\Local\Temp\RtmpIRlkMN\remote
s56284d6d4025\dphansti-Sushi-f53dd94/DESCRIPTION' ...

v  checking for file 'C:\Users\ebuen\AppData\Local\Temp\RtmpIRlkMN\remote
s56284d6d4025\dphansti-Sushi-f53dd94/DESCRIPTION'
##



-  preparing 'Sushi': (1.7s)
##

v  checking DESCRIPTION meta-information
##



-  checking for LF line-endings in source and make files and shell script
s
##



-  checking for empty or unneeded directories
##



-  looking to see if a 'data/datalist' file should be added
##



-  building 'Sushi_1.7.1.tar.gz' (860ms)
##
```

```
##
```

```
## Installing package into 'C:/Users/ebuen/OneDrive/Documents/R/win-libra
ry/3.5'
## (as 'lib' is unspecified)
```

```
library(Sushi) #loading sushi
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
## Loading required package: biomaRt
```

```
##what is sushi? Sushi is an R package for visualizing genomic data

##What formats does it use? BED, BEDGRAPH, BEDPE, interaction matrix

##what type of plots does it make? sequencing tracks, various chromatin i
nteractions plots from HiC data, manhattan plots for GWAS, transcript str
uctures from RNA SEQ data, gene structures, and gene density plots

##What other features does it have? allows you to define regions of plots
 to highlight and zoom into in subsequent graphs for a more detailed/more
 resolution visualization. Uses the layout function to create multipanele
d plots for publications

#Let's look at the available datasets in Sushi

Sushi_data = data(package = 'Sushi') #access the datasets with function d
ata
Sushi_data
data(list = Sushi_data$results[,3])
Sushi_data$results[,3] #viewing the datasets available
```

```
##  [1] "Sushi_5C.bedpe"                    "Sushi_ChIAPET_pol2.bedpe"

##  [3] "Sushi_ChIPExo_CTCF.bedgraph"       "Sushi_ChIPSeq_CTCF.bedgraph"

##  [5] "Sushi_ChIPSeq_pol2.bed"            "Sushi_ChIPSeq_pol2.bedgraph"

##  [7] "Sushi_ChIPSeq_severalfactors.bed" "Sushi_DNaseI.bedgraph"

##  [9] "Sushi_GWAS.bed"                    "Sushi_HiC.matrix"

## [11] "Sushi_RNASeq_K562.bedgraph"        "Sushi_genes.bed"

## [13] "Sushi_hg18_genome"                 "Sushi_transcripts.bed"
```

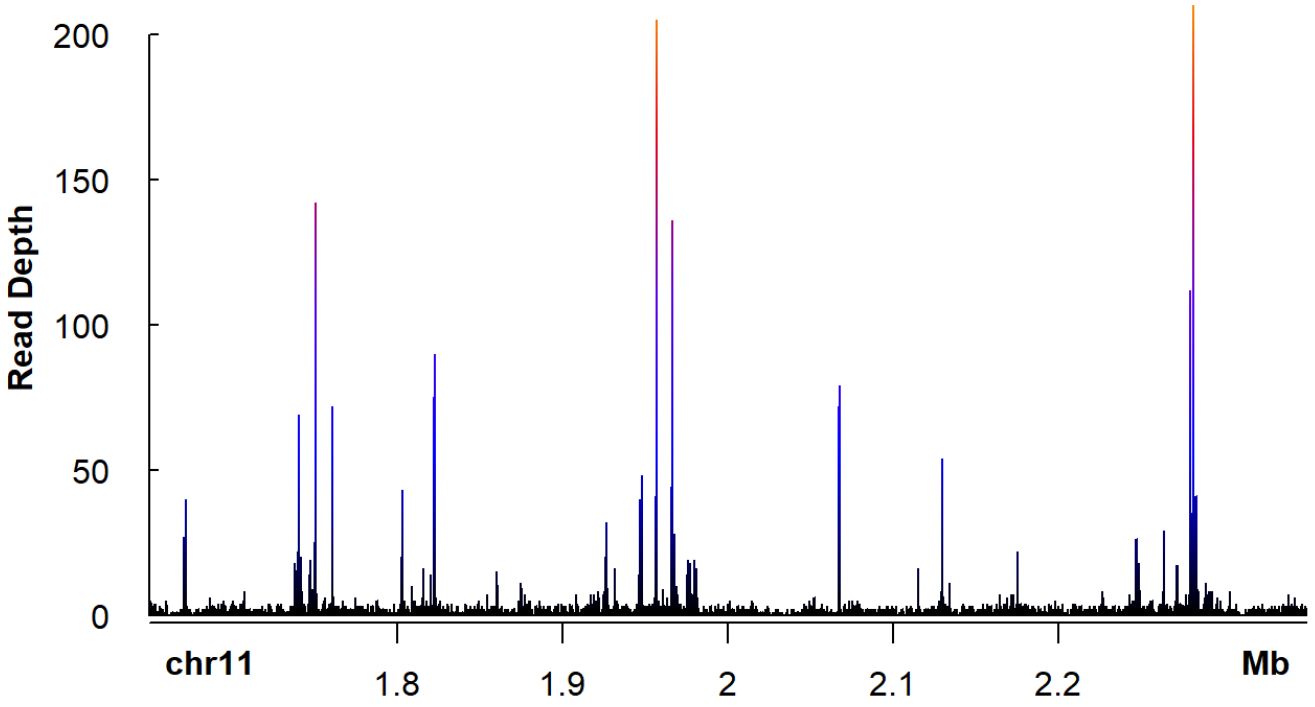
```
#-----------------------------------FUNCTIONS-------------------------------
-----------
#plotBedgraph() plots signal tracks using bedgraph files
head(Sushi_DNaseI.bedgraph) #exploring the dataset
```

```
##   chrom   start     end value
## 1 chr11 1640504 1640664     1
## 2 chr11 1640904 1641004     1
## 3 chr11 1641004 1641064     2
## 4 chr11 1641064 1641164     1
## 5 chr11 1645224 1645384     1
## 6 chr11 1645504 1645664     1
```

```
tail(Sushi_DNaseI.bedgraph)
```

```
##      chrom   start     end value
## 5709 chr11 2357884 2358044     1
## 5710 chr11 2358504 2358664     1
## 5711 chr11 2358824 2358964     1
## 5712 chr11 2359424 2359564     1
## 5713 chr11 2359704 2359864     1
## 5714 chr11 2359884 2359964     2
```

```
# plotBedgraph requirments: bedgraph data, chromosome ID, chromosome star
t, chromosome end, and a value that represents coverage depth


chrom = "chr11" #chromosome name
chromstart = 1650000 #chromosome position
chromend = 2350000

plotBedgraph(Sushi_DNaseI.bedgraph,chrom,chromstart,chromend,colorbycol=
 SushiColors(5))
#colorbycol() places a heatmap of 5 colors along the lines

labelgenome(chrom,chromstart,chromend,n=5,scale="Mb") #labelgenome() anno
tates the plot; n= no. of tick marks, scale= megabases

mtext("Read Depth",side=2,line=2.75,cex=1,font=2) #adding text to the y-a
xis
axis(side=2,las=2,tcl=.2) #tcl: tick mark length & direction; las:label o
rientation
```




```
#?mtext
#?axis
```

```
#------------------------------ZOOM AND MULTIPANELING------------------
----------------
#FUNCTIONS: multipaneling with layout(base) and zoom(sushi)
#create a layout for multiple plots
p<-layout(matrix(c(1,1,1,1,
                   1,1,1,1,
                   2,2,2,2,
                   2,2,2,2,
                   3,3,3,3,
                   3,3,3,3
                   ),
                 6,4,byrow=TRUE))
layout.show(p)
```

1
2
3

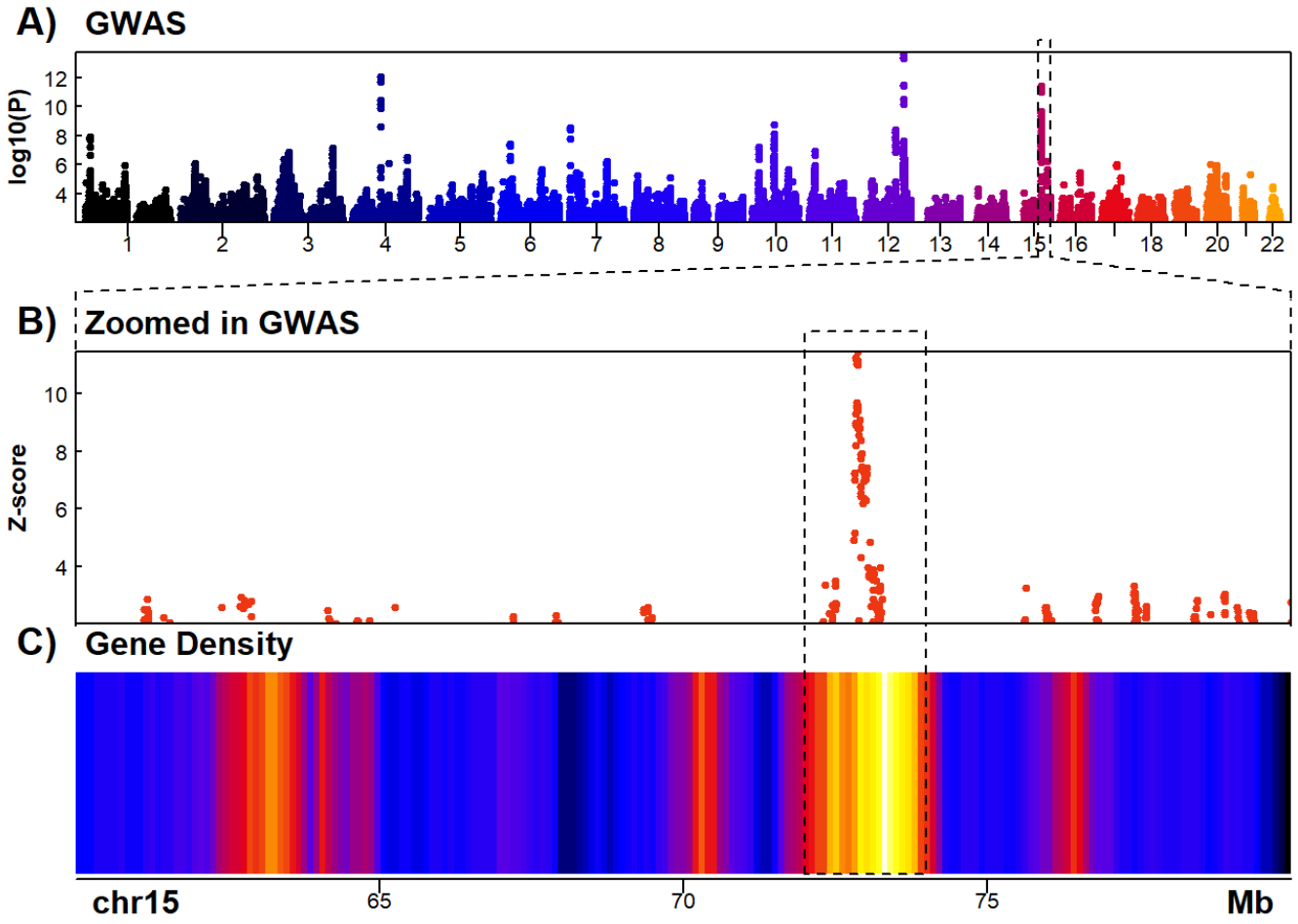
```
par(mgp=c ( 3,.3,0 )) #setting the margin line thickness for axis labels,
 title, and axis line

##FUNCTIONS: Plot Manhattan Plot using plotManhattan()
#set the margins
par(mar=c(3,4,3,2)) # here we are setting the number of lines of margins
 on four sides of the plot (bottom, left, top, right)

#set the genomic regions

chrom2= "chr15"
chromstart2= 73000000
chromend2= 89500000

#make manhattan plot
plotManhattan(bedfile = Sushi_GWAS.bed, pvalues = Sushi_GWAS.bed[,5], gen
ome = Sushi_hg18_genome, cex=0.75)

#zoom into a regions using zoomregion()

zoomsregion(region=c(chromstart2, chromend2), chrom = chrom2, genome=Sush
i_hg18_genome,extend = c(0.07,0.2), wideextend = 0.2, offsets = c(0, 0))
#?zoomsregion
#add labels
labelgenome(genome=Sushi_hg18_genome, n=4, scale = "Mb", edgeblankfractio
n = 0.20)
#?labelgenome
#add y-axis
axis(side=2, las=2, tcl=.2)
mtext("log10(P)", side=2, line=1.75, cex= .75, font=2)

#addplot label

labelplot("A)"," GWAS")


##Zoomed in manhattan plot

par(mar=c(0.1, 4,2,2))

#set genomic regions
```

```
chrom= "chr15"
chromstart= 60000000
chromend= 80000000
chromstart2= 72000000
chromend2= 74000000

#make manhattan plot
plotManhattan(bedfile = Sushi_GWAS.bed, chrom=chrom2,
              chromstart = chromstart,
              chromend = chromend,
              pvalues =Sushi_GWAS.bed$pval.GC.DBP,
              col= SushiColors(6) (nrow(Sushi_hg18_genome))[15],
              cex=0.75)


#add another zoom in, creates an outline around the region you want to zo
om into
zoomsregion(region=c(chromstart2, chromend2),
chrom=chrom2,
genome=NULL,
extend = c(0.075,1),
wideextend = 0.2,
offsets = c(0.0,0))

#add a zoom box

zoombox(passthrough=TRUE, topextend = 2)
#?zoombox

#add y-axis
axis(side=2, las=2, tcl=.2)
mtext("Z-score", side=2, line = 1.75, cex = .75, font = 2)


#add plot labels
labelplot("B)"," Zoomed in GWAS")


##FUNCTIONS: plotBed() for gene density heat maps
#plotBed() can also plot gene density heat maps along a chromosome but fi
rst we need to load in dataset of genes from ensembl
par(mar=c(3,4,1.8,2)) #set the margins
```

```
#set genomic regions
chrom = "chr15"
chromstart = 60000000
chromend = 80000000
chrom_biomart = gsub("chr","",chrom) #the biomart package is built into s
ushi. It allows you to pull genomic information from the ensmbl website.
 The gsub() allows you to replace character strings in a dataset

#set the mart
mart<-useMart(host='may2009.archive.ensembl.org',  #host to connect to
            biomart='ENSEMBL_MART_ENSEMBL', #database to connect to
            dataset='hsapiens_gene_ensembl') #dataset to use

#get gene info
#getBM: retrieves user defined attributes from the Biomart database.
geneinfobed<-getBM(attributes =c("chromosome_name","start_position","end_
position"),
            filters= c("chromosome_name","start","end"),
            values=list(chrom_biomart,chromstart,chromend),mart=mart)


#add "chr" to the chrom column
geneinfobed[,1] = paste("chr",geneinfobed[,1],sep="")

head (geneinfobed)
```

```
##   chromosome_name start_position end_position
## 1           chr15       73372069     73372334
## 2           chr15       64580642     64580710
## 3           chr15       63375442     63375557
## 4           chr15       72570353     72570422
## 5           chr15       60903209     60903293
## 6           chr15       70130646     70130724
```

```
#plot gene density
plotBed(beddata = geneinfobed[!duplicated(geneinfobed),],
        chrom = chrom,
        chromstart = chromstart,
        chromend =chromend,
        row='supplied', #how row number should be determined
        palettes = list(SushiColors(7)), type = "density") #type can also
 be set to circles and region but here we want a density plot
#?plotBed

# add zoom in
zoomsregion(region=c(chromstart2,chromend2),
            chrom=chrom2,
            genome=NULL,
            highlight = TRUE, #TRUE indicates that you just want a box ar
ound a region of interest
            extend=c(2,0)) #vector indidcating how far zoom region extend
s above and below the plot region
#add labels
labelgenome(chrom,
            chromstart,
            chromend,
            n=3,
            scale="Mb",edgeblankfraction=0.20)


labelplot("C)"," Gene Density")
```


A) GWAS
log10(P)
12
10
8
6
4
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
18
20
22
B) Zoomed in GWAS
Z-score
10
8
6
4
C) Gene Density
chr15
65
70
75
Mb