Exercise 2. Functional annotation of ChIP-peaks

1. Using Putty (Windows) or Terminal (Mac) to connect to your assigned computer.

Create a directory /workdir/myUserID (replace myUserID with you BioHPC ID), copy the test regions, and test region sequences to the working directory.

```
cd  /workdir/myUserID

cp  /shared_data/ChIP_seq_workshop2_2017/test_results_peaks.narrowPeak ./

cp /shared_data/ChIP_seq_workshop2_2017/*.gz ./

cp /shared_data/ChIP_seq_workshop2_2017/*.bam* ./
```

2. Install Bioconductor packages and annotation

```
R

source("https://bioconductor.org/biocLite.R")

biocLite("GenomicFeatures")

biocLite("clusterProfiler")

biocLite("org.At.tair.db")

biocLite("BSgenome.Athaliana.TAIR.TAIR9")

biocLite("TxDb.Athaliana.BioMart.plantsmart28")

biocLite("rGADEM")

biocLite("ChIPseeker")

install.packages("doMC", dep=T)

install.packages("caTools", dep=T)

install.packages("utils", dep=T)

biocLite( "BSgenome" )

biocLite( "Rsamtools" )
```

```
biocLite( "ShortRead" )
library ("rGADEM")
library (TxDb.Athaliana.BioMart.plantsmart28)
library (BSgenome.Athaliana.TAIR.TAIR9)
library (org.At.tair.db)
library (clusterProfiler)
library (GenomicFeatures)
library ("ChIPseeker")
####################import peak files##############################
peak<- readPeakFile("test_results_peaks.narrowPeak", as="GRanges",header=F)
#############genomic feature distribution analysis######################
peakAnno <-
annotatePeak(peak,TxDb=TxDb.Athaliana.BioMart.plantsmart28,annoDb =
"org.At.tair.db")
inform<- as.data.frame(peakAnno)
write.table (inform,file="peak_overlap_gene_feature_out.txt",sep="\t")
pdf(file="pie_chart.pdf")
plotAnnoPie(peakAnno)
dev.off ()
pdf(file="bar_chart.pdf")
plotAnnoBar (peakAnno)
dev.off ()
pdf(file="vennpie_chart.pdf")
vennpie(peakAnno)
dev.off ()
##################Distance to TSS regions###########################
```

```
pdf(file="distance_to_TSS_out.pdf")

plotDistToTSS(peakAnno, title="Distribution of H3K4me3 loci relative to TSS")

dev.off ()

##################GO enrichment analysis##########################

keytypes(org.At.tair.db)

ego_MF <- enrichGO(gene =inform$geneId,'org.At.tair.db',keytype="TAIR",ont =
"MF",pAdjustMethod = "BH",pvalueCutoff = 0.01,qvalueCutoff = 0.05)

ego_CC <- enrichGO(gene =inform$geneId,'org.At.tair.db',keytype="TAIR",ont =
"CC",pAdjustMethod = "BH",pvalueCutoff = 0.01,qvalueCutoff = 0.05)

ego_BP <- enrichGO(gene =inform$geneId,'org.At.tair.db',keytype="TAIR",ont =
"BP",pAdjustMethod = "BH",pvalueCutoff = 0.01,qvalueCutoff = 0.05)

write.table(as.data.frame(ego_MF),file="MF_out.txt",sep="\t")

write.table(as.data.frame(ego_CC),file="CC_out.txt",sep="\t")

write.table(as.data.frame(ego_BP),file="BP_out.txt",sep="\t")

pdf (file="MF_dotplot.pdf")

dotplot(ego_MF, showCategory=30)

dev.off()

pdf (file="CC_dotplot.pdf")

dotplot(ego_CC, showCategory=30)

dev.off()

pdf (file="BP_dotplot.pdf")

dotplot(ego_BP, showCategory=30)

dev.off()

###############heatmap and average density distribution##########

promoter <- getPromoters(TxDb=TxDb.Athaliana.BioMart.plantsmart28,
upstream=3000, downstream=3000)

tagMatrix <- getTagMatrix(peak, weightCol=NULL, windows=promoter)
```

```
pdf(file="heat_average_plot.pdf")

tagHeatmap(tagMatrix, xlim=c(-3000, 3000), color="red")

plotAvgProf(tagMatrix, xlim=c(-3000, 3000), xlab="Genomic Region (5'->3')",
ylab = "Read Count Frequency")

dev.off()

q ()
```

3 ngsplot install and operation

```
tar xzvf ngsplot-2.47.tar.gz

cd ngsplot

export  PATH=/workdir/myUserID/ngsplot/bin:$PATH

export   NGSPLOT=/workdir/myUserID/ngsplot

source ~/.bash_profile

ngsplotdb.py list

ngsplotdb.py install ../ngsplotdb_Tair10_21_3.00.tar.gz

ngsplotdb.py list

ngs.plot.r -G Tair10 -R genebody -
C  ../treatment.fastq_filter_qual_bowtie2_sorted_filter_sorted.bam  -O
test.genebody  -FL 300 -L 3000

ngs.plot.r -G Tair10 -R tss -
C  ../treatment.fastq_filter_qual_bowtie2_sorted_filter_sorted.bam  -O test.TSS  -
FL 300 -L 3000
```

4 deeptools operation and file format changes

```
export PATH=/programs/deeptools-2.5.0.1/bin:$PATH
```

```
export PYTHONPATH=/programs/deeptools-2.5.0.1/lib64/python2.7/site-packages:/programs/deeptools-2.5.0.1/lib/python2.7/site-packages:$PYTHONPATH

bamCompare -b1 treatment2.fastq_filter_qual_bowtie2_sorted_filter_sorted.bam -b2 H3.fastq_trim_unique_sorted.bam --ratio log2  --smoothLength 200 --extendReads 178  -o log2ratio.bw
```