

P.v.NP zscore differences

HBG

10/30/2018

```
### this block perform z-score calculations
library("microbenchmark")
library("matrixStats")

## the GO terms for biological processes (BP)
GOcategory.file <- read.csv("../Data/human.bp.term.csv",header=TRUE, stringsAsFactors=F)
bp.go.cat <- GOcategory.file$GO.term
bp.dim <- length(bp.go.cat)
conn.dim <- 20
hspin <- matrix(as.numeric(unlist(read.table("human.rnf43.all.bp.txt", header=F, sep=","))), nrow=bp.dim, ncol=conn.dim)
obs <- c(hspin)

perm <- c()
for (i in 1:100) {
  name <- paste("ms02.human", "/", "rnf43.heatmap", "/", "ms02.", i, ".bp.matrix.csv", sep="")
  mat <- matrix(as.numeric(unlist(read.table(name, header=F, sep=","))), nrow=bp.dim, ncol=conn.dim)
  perm <- rbind(perm, c(mat))
}

mean <- colMeans(perm)
std <- colSds(perm)

zscore.panc <- round((obs - mean)/std, 3)

z.panc <- matrix(zscore.panc, nrow=bp.dim, ncol=conn.dim)

### non-panc
hspin.np <- matrix(as.numeric(unlist(read.table("human.rnf43.NP.bp.txt", header=F, sep=","))), nrow=bp.dim, ncol=conn.dim)
obs.np <- c(hspin.np)

perm.np <- c()
for (i in 1:100) {
  name.np <- paste("ms02.human", "/", "rnf43.NP", "/", "ms02.", i, ".bp.matrix.csv", sep="")
  mat.np <- matrix(as.numeric(unlist(read.table(name.np, header=F, sep=","))), nrow=bp.dim, ncol=conn.dim)
  perm.np <- rbind(perm.np, c(mat.np))
}

mean.np <- colMeans(perm.np)
std.np <- colSds(perm.np)

zscore.np <- round((obs.np - mean.np)/std.np, 3)

z.np <- matrix(zscore.np, nrow=bp.dim, ncol=conn.dim)

deltz <- z.panc - z.np
```

```

library('gplots')

##
## Attaching package: 'gplots'
## The following object is masked from 'package:stats':
##
##      lowess
deltz <- t(deltz)

rownames(deltz) <- c("Q1","Q2","Q3","Q4","Q5","Q6","Q7","Q8","Q9","Q10",
                    "Q11","Q12","Q13","Q14","Q15","Q16","Q17","Q18","Q19","Q20")

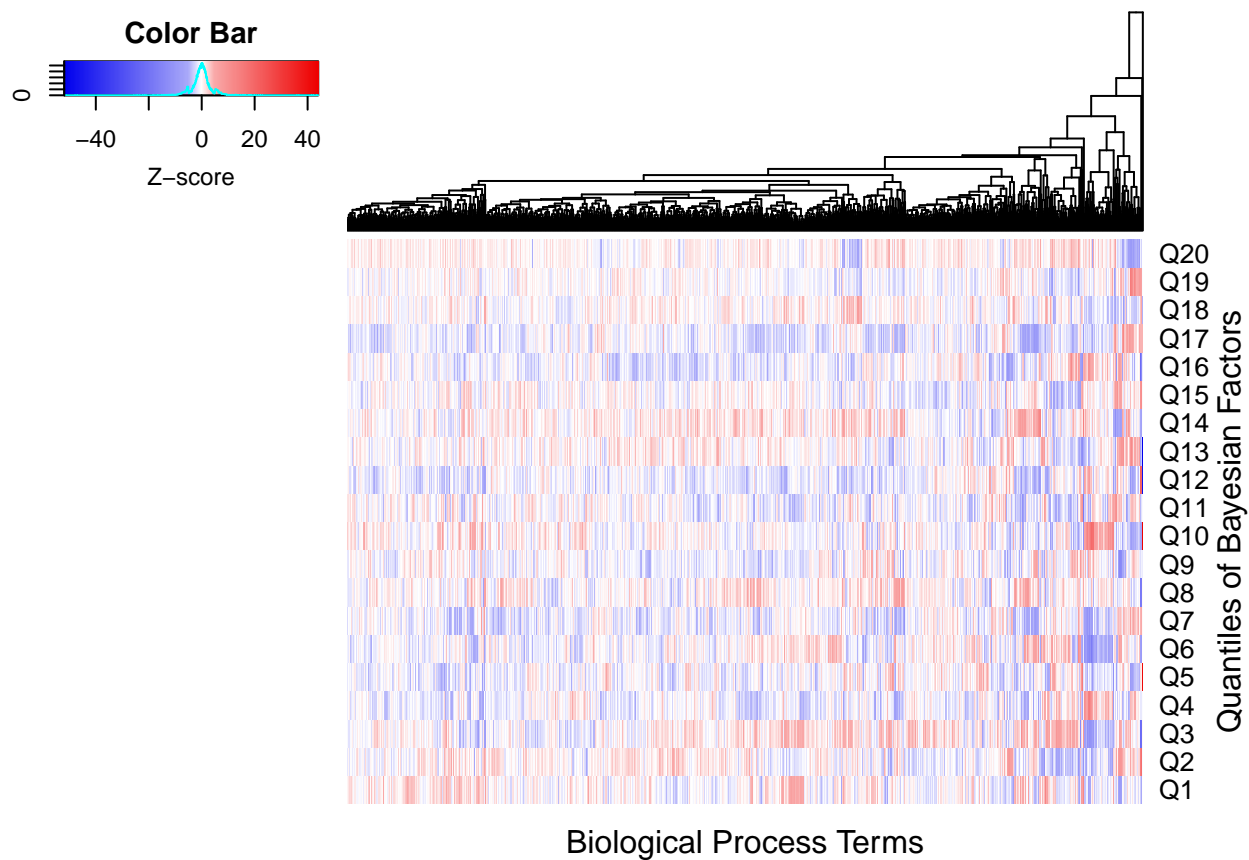
colnames(deltz) <- bp.go.cat

colors = c(seq(min(deltz),-5.1,length=100),seq(-4.9,4.9,length=100),seq(5.1,max(deltz),length=100))
my_palette <- colorRampPalette(c("blue2", "white", "red2"))(n = 299)

png(filename = "human.rnf43.deltz.bp.heatmap.png",width=6, height=5.5, res=1200, unit="in")
heatmap.2(deltz, col=my_palette, breaks=colors, Rowv=F,
          trace='none', offsetRow = 0, offsetCol = 0,
          xlab="Biological Process Terms", ylab="Quantiles of Bayesian Factors",
          margins = c(2,3.5), key.title = "Color Bar", key.xlab="Z-score", key.ylab=NA,
          revC = T,
          labCol = NA, #labRow =,
          #srtCol=45, adjCol=c(1,0),
          #lmat=rbind(c(0,3,4), c(2,1,0)), lwid=c(1.5,4,2),
          scale="none", dendrogram = "col", symbreaks=T, symm=F, symkey = F)
dev.off()

## pdf
##      2
hm.deltz <- heatmap.2(deltz, col=my_palette, breaks=colors, Rowv=F,
                      trace='none', offsetRow = 0, offsetCol = 0,
                      xlab="Biological Process Terms", ylab="Quantiles of Bayesian Factors",
                      margins = c(2,3.5), key.title = "Color Bar", key.xlab="Z-score", key.ylab=NA,
                      revC = T,
                      labCol = NA,
                      scale="none", dendrogram = "col", symbreaks=T, symm=F, symkey = F)

```



```
hc.deltz.col <- as.hclust(hm.deltz$colDendrogram)
pdf("human.rnf43.deltz.bp.tree.pdf", width=150, height=4, paper='special')
plot(hc.deltz.col, xlab="BP Terms", main="Z-scores, Hierarchical Clustering", cex=.8)
dev.off()
```

```
## pdf
## 2
```