

# 09.rnf43.bp.z.Rmd

HBG

11/1/2018

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

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 <- round((obs - mean)/std, 3)

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

write.table(z, file="human.rnf43.P.bp.z.csv", sep=",", row.names=F, col.names=F, quote=F)

library('gplots')

##
## Attaching package: 'gplots'
## The following object is masked from 'package:stats':
##
##      lowess

z <- t(z)

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

colnames(z) <- bp.go.cat

colors = c(seq(min(z), -10.1, length=100), seq(-9.9, 9.9, length=100), seq(10.1, max(z), length=100))
#colors = c(seq(min(z), max(z), length=300))
my_palette <- colorRampPalette(c("blue2", "white", "red2"))(n = 299)

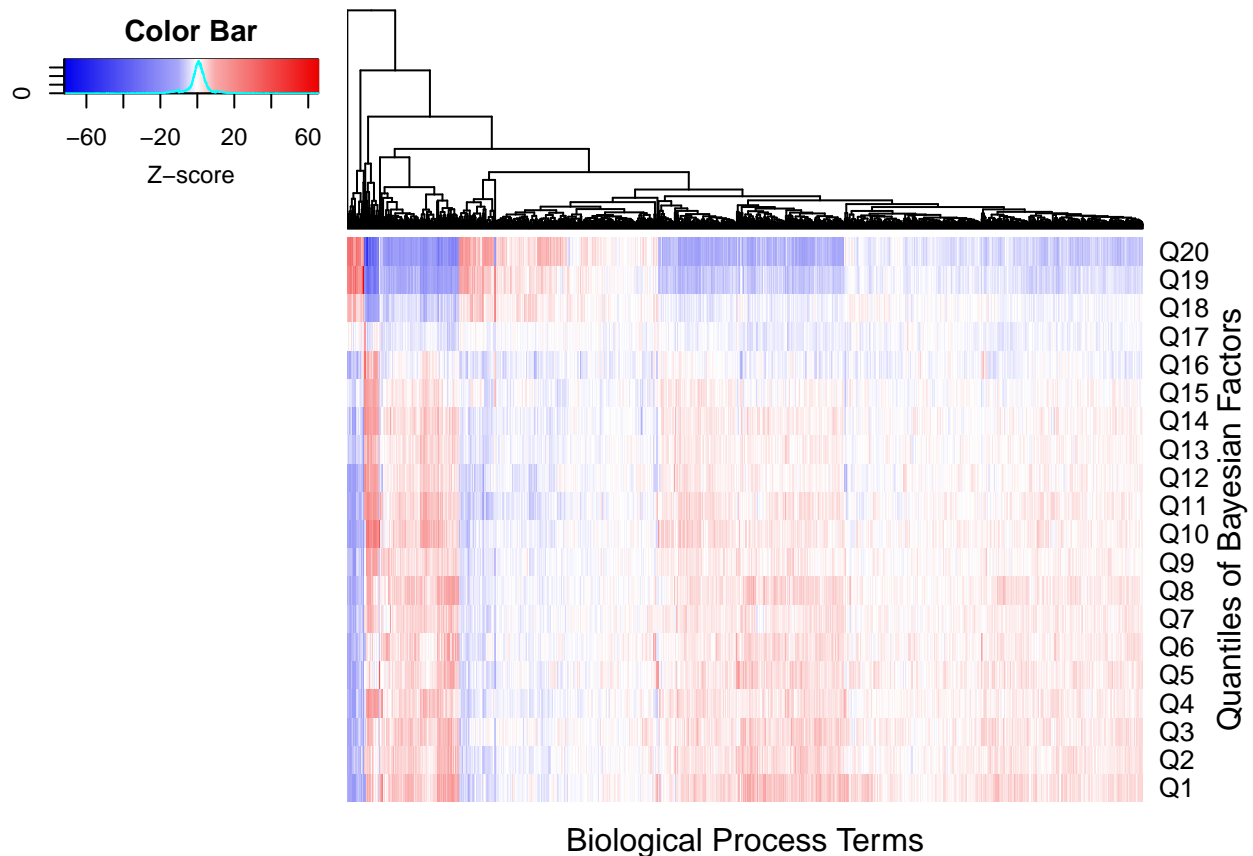
png(filename = "human.rnf43.P.bp.heatmap.png", width=6, height=5.5, res=1200, unit="in")
```

```
heatmap.2(z, 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 <- heatmap.2(z, 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.col <- as.hclust(hm$colDendrogram)
pdf("human.rnf43.P.bp.tree.pdf", width=150, height=4, paper='special')
plot(hc.col, xlab="BP Terms", main="Z-scores, Hierarchical Clustering", cex=.8)
dev.off()
```

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

## GO analysis

```
library(GO.db)
```

```
## Loading required package: AnnotationDbi
## Loading required package: stats4
## Loading required package: BiocGenerics
## Loading required package: parallel
##
## Attaching package: 'BiocGenerics'
## The following objects are masked from 'package:parallel':
##
##   clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##   clusterExport, clusterMap, parApply, parCapply, parLapply,
##   parLapplyLB, parRapply, parSapply, parSapplyLB
## The following objects are masked from 'package:stats':
##
##   IQR, mad, sd, var, xtabs
## The following objects are masked from 'package:base':
##
##   anyDuplicated, append, as.data.frame, basename, cbind,
##   colMeans, colnames, colSums, dirname, do.call, duplicated,
##   eval, evalq, Filter, Find, get, grep, grepl, intersect,
##   is.unsorted, lapply, lengths, Map, mapply, match, mget, order,
##   paste, pmax, pmax.int, pmin, pmin.int, Position, rank, rbind,
##   Reduce, rowMeans, rownames, rowSums, sapply, setdiff, sort,
##   table, tapply, union, unique, unsplit, which, which.max,
##   which.min
## Loading required package: Biobase
## Welcome to Bioconductor
##
##   Vignettes contain introductory material; view with
##   'browseVignettes()'. To cite Bioconductor, see
##   'citation("Biobase")', and for packages 'citation("pkgname)".
##
## Attaching package: 'Biobase'
## The following objects are masked from 'package:matrixStats':
##
##   anyMissing, rowMedians
## Loading required package: IRanges
## Warning: package 'IRanges' was built under R version 3.5.1
## Loading required package: S4Vectors
```

```

##
## Attaching package: 'S4Vectors'

## The following object is masked from 'package:gplots':
##
##     space

## The following object is masked from 'package:base':
##
##     expand.grid
##

bp.go.term.tree <- bp.go.cat[hc.col$order]
write.table(bp.go.term.tree, file="human.rnf43.P.bp.tree.go.csv", sep=",", col.names=F, row.names = F, c

##the terms 2:28 are the branch enriched in Q17-Q20 (especially Q19 and Q20), but suppressed in Q1-Q16
pair <- c("GO.ID", "GO.Term")
for (i in 2:28) {
  go.term <- as.character(bp.go.term.tree[i])
  detail <- Term(GOID(go.term))
  pair <- rbind(pair, c(go.term, detail))
}

pair

##
## pair "GO.ID"
##      "GO:0006405"
##      "GO:0016070"
##      "GO:0008380"
##      "GO:0006406"
##      "GO:0031124"
##      "GO:0006368"
##      "GO:0006370"
##      "GO:0042795"
##      "GO:0006364"
##      "GO:0000184"
##      "GO:0006614"
##      "GO:0006413"
##      "GO:0006412"
##      "GO:0006397"
##      "GO:0045292"
##      "GO:0000381"
##      "GO:0006369"
##      "GO:0000375"
##      "GO:0006363"
##      "GO:0006361"
##      "GO:0060964"
##      "GO:0006383"
##      "GO:0002181"
##      "GO:0032543"
##      "GO:0070126"
##      "GO:0070125"
##      "GO:0000398"
##      GO:0006405

```

```
## pair "GO.Term"
##      "RNA export from nucleus"
##      "RNA metabolic process"
##      "RNA splicing"
##      "mRNA export from nucleus"
##      "mRNA 3'-end processing"
##      "transcription elongation from RNA polymerase II promoter"
##      "7-methylguanosine mRNA capping"
##      "snRNA transcription by RNA polymerase II"
##      "rRNA processing"
##      "nuclear-transcribed mRNA catabolic process, nonsense-mediated decay"
##      "SRP-dependent cotranslational protein targeting to membrane"
##      "translational initiation"
##      "translation"
##      "mRNA processing"
##      "mRNA cis splicing, via spliceosome"
##      "regulation of alternative mRNA splicing, via spliceosome"
##      "termination of RNA polymerase II transcription"
##      "RNA splicing, via transesterification reactions"
##      "termination of RNA polymerase I transcription"
##      "transcription initiation from RNA polymerase I promoter"
##      "regulation of gene silencing by miRNA"
##      "transcription by RNA polymerase III"
##      "cytoplasmic translation"
##      "mitochondrial translation"
##      "mitochondrial translational termination"
##      "mitochondrial translational elongation"
##      "mRNA splicing, via spliceosome"
```

```
write.table(pair, file="human.rnf43.P.bp.branch2.csv", col.names=F, row.names=F, sep="\t", quote=F)
```

```
library(ggplot2)
```

```
branch2 <- data.frame(z[,hc.col$order[2:28]])
```

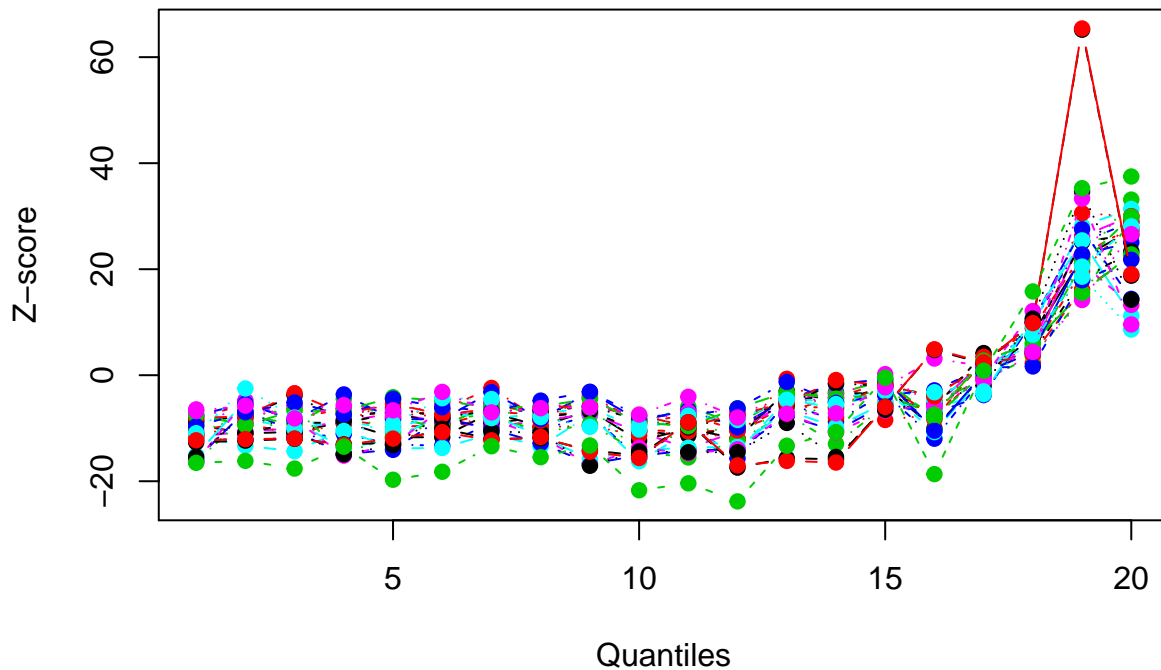
```
branch2
```

	GO.0006405	GO.0016070	GO.0008380	GO.0006406	GO.0031124	GO.0006368
## Q1	-9.349	-10.496	-11.000	-12.094	-11.520	-11.585
## Q2	-10.974	-8.909	-7.623	-12.149	-13.333	-10.600
## Q3	-10.689	-11.393	-11.035	-11.858	-14.318	-10.028
## Q4	-8.830	-8.365	-7.019	-11.053	-10.329	-8.768
## Q5	-11.713	-9.760	-10.705	-13.066	-13.809	-7.156
## Q6	-11.686	-9.408	-9.533	-13.300	-13.726	-8.845
## Q7	-6.989	-6.974	-5.522	-8.960	-11.040	-9.654
## Q8	-9.686	-7.357	-9.926	-10.400	-13.429	-8.662
## Q9	-6.607	-7.140	-5.798	-7.487	-7.470	-7.228
## Q10	-12.354	-10.294	-14.759	-15.617	-16.219	-10.684
## Q11	-10.101	-10.684	-10.019	-9.196	-10.730	-13.838
## Q12	-11.994	-10.326	-10.534	-15.714	-14.460	-10.519
## Q13	-2.957	-6.842	-6.836	-5.132	-4.626	-5.041
## Q14	-4.924	-4.136	-6.488	-8.698	-6.161	-3.104
## Q15	-2.202	-1.379	-1.673	-3.458	-2.677	-5.045
## Q16	-10.048	-8.343	-10.435	-11.973	-10.738	-5.770
## Q17	0.035	2.796	0.311	-0.416	0.040	2.442
## Q18	7.198	8.835	5.384	9.888	9.263	4.772
## Q19	22.188	22.746	21.631	25.139	28.166	26.187

## Q20	27.319	27.901	33.127	28.886	31.361	30.011
##	GD.0006370	GD.0042795	GD.0006364	GD.0000184	GD.0006614	GD.0006413
## Q1	-11.115	-11.447	-16.284	-15.927	-15.098	-16.372
## Q2	-10.926	-9.271	-7.859	-6.711	-4.735	-4.756
## Q3	-8.914	-11.071	-6.240	-10.603	-8.885	-8.318
## Q4	-9.044	-9.132	-14.830	-14.860	-13.640	-15.162
## Q5	-7.106	-4.496	-11.420	-14.076	-12.228	-13.089
## Q6	-9.186	-9.293	-8.344	-7.991	-7.249	-6.418
## Q7	-8.944	-8.255	-7.339	-10.164	-7.445	-8.428
## Q8	-8.052	-9.779	-10.802	-12.821	-11.513	-11.559
## Q9	-6.766	-9.552	-9.260	-16.684	-15.844	-16.954
## Q10	-10.589	-11.357	-14.475	-15.167	-14.135	-13.713
## Q11	-11.355	-10.629	-15.513	-14.213	-13.673	-14.711
## Q12	-9.987	-11.979	-12.285	-13.986	-13.497	-13.973
## Q13	-5.212	-0.682	-7.370	-7.643	-6.649	-7.215
## Q14	-3.404	-7.991	-12.948	-10.517	-9.792	-10.798
## Q15	-3.409	-8.443	-5.789	-5.519	-5.179	-6.657
## Q16	-6.227	-7.426	-5.027	-7.168	-6.188	-5.504
## Q17	4.179	3.476	-0.798	-3.745	-3.563	-0.478
## Q18	7.137	2.277	3.547	10.207	10.782	12.100
## Q19	25.228	30.565	14.771	27.535	25.431	21.062
## Q20	26.450	29.993	29.684	14.405	11.217	13.262
##	GD.0006412	GD.0006397	GD.0045292	GD.0000381	GD.0006369	GD.0000375
## Q1	-15.400	-7.833	-8.555	-7.575	-7.569	-6.440
## Q2	-7.211	-5.685	-7.775	-5.350	-8.117	-8.096
## Q3	-10.742	-8.215	-9.105	-8.474	-10.155	-5.757
## Q4	-14.927	-6.275	-4.485	-3.626	-8.568	-6.277
## Q5	-13.187	-7.941	-8.210	-8.279	-10.293	-7.691
## Q6	-7.412	-6.556	-9.756	-8.332	-7.005	-5.240
## Q7	-10.618	-2.400	-5.304	-3.190	-8.092	-5.092
## Q8	-10.836	-8.536	-7.581	-8.260	-7.331	-6.577
## Q9	-17.083	-3.768	-4.282	-3.176	-6.053	-4.319
## Q10	-14.360	-8.934	-9.140	-8.522	-10.237	-8.714
## Q11	-14.533	-7.504	-5.935	-6.610	-7.919	-6.957
## Q12	-14.528	-10.399	-9.267	-9.807	-8.444	-6.805
## Q13	-8.996	-5.916	-3.743	-4.507	-4.482	-5.099
## Q14	-15.404	-6.136	-3.461	-1.661	-4.969	-4.186
## Q15	-6.512	-1.705	-1.044	-0.746	-2.257	0.203
## Q16	-4.978	-7.959	-7.985	-10.470	-4.841	-5.941
## Q17	-1.958	0.899	0.380	0.473	1.265	-1.049
## Q18	10.673	4.969	4.487	4.354	5.883	3.902
## Q19	34.644	19.657	21.077	22.782	20.488	14.185
## Q20	14.239	28.918	29.927	25.057	28.085	26.615
##	GD.0006363	GD.0006361	GD.0060964	GD.0006383	GD.0002181	GD.0032543
## Q1	-8.701	-8.102	-7.532	-9.636	-10.935	-6.732
## Q2	-7.747	-7.736	-9.379	-6.988	-2.505	-5.738
## Q3	-3.809	-3.395	-6.692	-5.150	-7.982	-8.140
## Q4	-8.606	-7.288	-6.722	-7.397	-10.514	-5.643
## Q5	-6.214	-5.367	-4.150	-4.448	-9.202	-6.622
## Q6	-7.956	-7.431	-4.808	-6.040	-4.337	-3.142
## Q7	-6.657	-6.353	-6.107	-6.751	-4.498	-7.013
## Q8	-6.696	-5.992	-5.910	-4.791	-7.903	-6.220
## Q9	-5.956	-5.853	-4.431	-3.125	-9.742	-6.050
## Q10	-8.352	-7.588	-8.938	-7.700	-9.497	-7.425

## Q11	-9.864	-9.629	-9.589	-7.569	-7.614	-4.071
## Q12	-7.634	-7.433	-6.231	-6.260	-8.263	-7.981
## Q13	-5.496	-5.219	-3.211	-1.274	-4.631	-7.256
## Q14	-1.521	-0.910	-4.426	-5.213	-5.442	-7.190
## Q15	-2.012	-1.963	-3.657	-3.725	-2.975	-2.277
## Q16	-3.859	-3.987	-7.427	-2.849	-3.206	3.145
## Q17	2.176	2.024	2.756	1.410	-3.043	1.695
## Q18	4.174	3.961	6.068	1.658	7.660	4.391
## Q19	18.091	16.189	15.564	17.913	18.580	33.312
## Q20	23.350	22.082	22.784	21.779	8.629	9.578
##	G0.0070126	G0.0070125	G0.0000398			
## Q1	-12.680	-12.273	-16.526			
## Q2	-12.345	-12.055	-16.152			
## Q3	-12.004	-11.948	-17.609			
## Q4	-13.046	-13.138	-13.484			
## Q5	-12.705	-11.955	-19.718			
## Q6	-10.103	-10.735	-18.211			
## Q7	-12.001	-12.278	-13.400			
## Q8	-11.892	-11.644	-15.462			
## Q9	-14.152	-14.416	-13.252			
## Q10	-15.152	-15.613	-21.705			
## Q11	-8.876	-8.862	-20.391			
## Q12	-17.426	-17.011	-23.800			
## Q13	-15.651	-16.171	-13.298			
## Q14	-15.861	-16.471	-10.736			
## Q15	-5.928	-5.982	-0.473			
## Q16	4.745	4.898	-18.654			
## Q17	2.241	2.435	0.845			
## Q18	9.936	9.824	15.800			
## Q19	65.173	65.405	35.278			
## Q20	18.777	19.046	37.487			

```
matplot(branch2, type="b", xlab="Quantiles", ylab="Z-score", pch=19)
```



```
pdf(file="human.rnf43.P.bp.branch2.zprofile.pdf",width=6, height=4, paper='special')
matplot(branch2, type="b", xlab="Quantiles", ylab="Z-score", pch=19)
dev.off()
```

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