

# P vs NP, differential measurement

HBG

11/5/2018

This script use the differential measure (z-score) between tumor/non-tumor cells

```
### 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.PvNP.bp.txt", header=F, sep=","))), nrow=bp.d
obs <- c(hspin)

perm <- c()
for (i in 1:100) {
  name <- paste("ms02.human", "/", "rnf43.PvNP", "/", "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.pvnp <- matrix(zscore, nrow=bp.dim, ncol=conn.dim)

#write.table(z, file="human.rnf43.PvNP.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.pvnp <- t(z.pvnp)

rownames(z.pvnp) <- c("5%", "10%", "15%", "20%", "25%", "30%", "35%", "40%", "45%", "50%",
  "55%", "60%", "65%", "70%", "75%", "80%", "85%", "90%", "95%", "100%")
```

```

colnames(z.pvnp) <- bp.go.cat

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

png(filename = "17.PvsNP.z.hm.png",width=6, height=5.5, res=1200, unit="in")
heatmap.2(z.pvnp, col=my_palette, breaks=colors, dendrogram='none', Rowv=F, Colv=F,
  trace='none', offsetRow = 0, offsetCol = 0,
  xlab="Biological Process Terms", ylab="PDAC vs Normal Cells, Quantiles",
  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", symbreaks=T, symm=F, symkey = F)
dev.off()

## pdf
## 2

#correlations to quantile 20 (95%-100%) genes
y.val <- c()
for (i in 1:20) {
  y.val2 <- c()
  for (j in 1:20) {
    cor.val <- cor.test(z.pvnp[j,], z.pvnp[i,])
    y.val2 <- rbind(y.val2, cor.val$estimate)
  }
  y.val <- cbind(y.val, y.val2)
}

rownames(y.val) <- c("5%", "10%", "15%", "20%", "25%", "30%", "35%", "40%", "45%", "50%",
  "55%", "60%", "65%", "70%", "75%", "80%", "85%", "90%", "95%", "100%")
colnames(y.val) <- c("5%", "10%", "15%", "20%", "25%", "30%", "35%", "40%", "45%", "50%",
  "55%", "60%", "65%", "70%", "75%", "80%", "85%", "90%", "95%", "100%")
my_palette2 <- colorRampPalette(c("blue2", "white", "red2"))(9)
png(filename = "17.PvsNP.z.hm.cor.matrix.png",width=6, height=5.5, res=1200, unit="in")
heatmap.2(y.val, col=my_palette2, dendrogram='none', Rowv=F, Colv=F,
  trace='none', offsetRow = 0, offsetCol = 0,
  xlab="Quantiles of Bayes Factors", ylab="Quantiles of Bayes Factors",
  margins = c(3.5,3.5), key.title = "Color Bar", key.xlab="Pearson Correlation", 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", symbreaks=T, symm=F, symkey = F)
dev.off()

## pdf
## 2

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
## 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
##

```

```

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

hspin <- matrix(as.numeric(unlist(read.table("human.rnf43.NP.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.NP", "/", "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.np <- matrix(zscore, nrow=bp.dim, ncol=conn.dim)
z.np <- t(z.np)

enrich <- names(z.pvnp[20, which(z.pvnp[20,] > 5)])
suppress <- names(z.pvnp[20, which(z.pvnp[20,] < -5)])
enriched.terms <- c("GO.ID", "GO.Term", "Z.Tumor", "Z.Normal", "Z.Differential")
for (i in 1:length(enrich)) {
  id <- as.character(enrich[i])
  term <- Term(GO.ID(id))
  seri <- which(bp.go.cat %in% id)
  z.tumor <- z.p[20,seri]
  z.normal <- z.np[20,seri]
  z.diff <- z.pvnp[20,seri]
  enriched.terms <- rbind(enriched.terms, c(id, term, z.tumor, z.normal, z.diff))
}
enriched.terms

##
## enriched.terms "GO.ID"
## "GO:0001892"
## "GO:0006099"

```

##	"G0:0006629"
##	"G0:0006886"
##	"G0:0006888"
##	"G0:0006890"
##	"G0:0006891"
##	"G0:0006897"
##	"G0:0006986"
##	"G0:0007005"
##	"G0:0007030"
##	"G0:0007411"
##	"G0:0007568"
##	"G0:0007623"
##	"G0:0008202"
##	"G0:0008286"
##	"G0:0009408"
##	"G0:0009755"
##	"G0:0010628"
##	"G0:0010718"
##	"G0:0010906"
##	"G0:0015031"
##	"G0:0015721"
##	"G0:0015991"
##	"G0:0016192"
##	"G0:0016241"
##	"G0:0019216"
##	"G0:0019886"
##	"G0:0030182"
##	"G0:0030216"
##	"G0:0030521"
##	"G0:0030522"
##	"G0:0032007"
##	"G0:0032008"
##	"G0:0032526"
##	"G0:0033572"
##	"G0:0034220"
##	"G0:0034614"
##	"G0:0034620"
##	"G0:0035116"
##	"G0:0035556"
##	"G0:0035567"
##	"G0:0039702"
##	"G0:0042026"
##	"G0:0042147"
##	"G0:0042407"
##	"G0:0042632"
##	"G0:0043066"
##	"G0:0043154"
##	"G0:0043312"
##	"G0:0043401"
##	"G0:0043524"
##	"G0:0043547"
##	"G0:0045165"
##	"G0:0045600"
##	"G0:0045669"

```

## "GO:0045892"
## "GO:0045893"
## "GO:0045944"
## "GO:0046034"
## "GO:0048167"
## "GO:0048208"
## "GO:0048589"
## "GO:0050767"
## "GO:0050790"
## "GO:0050821"
## "GO:0051085"
## "GO:0060021"
## "GO:0060070"
## "GO:0061024"
## "GO:0071773"
## "GO:0090090"
## "GO:0090263"
## "GO:0120162"
## "GO:1901998"
## "GO:1904837"
## "GO:1904886"
## GO:0001892
## enriched.terms "GO.Term"
## "embryonic placenta development"
## "tricarboxylic acid cycle"
## "lipid metabolic process"
## "intracellular protein transport"
## "ER to Golgi vesicle-mediated transport"
## "retrograde vesicle-mediated transport, Golgi to ER"
## "intra-Golgi vesicle-mediated transport"
## "endocytosis"
## "response to unfolded protein"
## "mitochondrion organization"
## "Golgi organization"
## "axon guidance"
## "aging"
## "circadian rhythm"
## "steroid metabolic process"
## "insulin receptor signaling pathway"
## "response to heat"
## "hormone-mediated signaling pathway"
## "positive regulation of gene expression"
## "positive regulation of epithelial to mesenchymal transition"
## "regulation of glucose metabolic process"
## "protein transport"
## "bile acid and bile salt transport"
## "ATP hydrolysis coupled proton transport"
## "vesicle-mediated transport"
## "regulation of macroautophagy"
## "regulation of lipid metabolic process"
## "antigen processing and presentation of exogenous peptide antigen via MHC class II"
## "neuron differentiation"
## "keratinocyte differentiation"
## "androgen receptor signaling pathway"

```

```

## "intracellular receptor signaling pathway"
## "negative regulation of TOR signaling"
## "positive regulation of TOR signaling"
## "response to retinoic acid"
## "transferrin transport"
## "ion transmembrane transport"
## "cellular response to reactive oxygen species"
## "cellular response to unfolded protein"
## "embryonic hindlimb morphogenesis"
## "intracellular signal transduction"
## "non-canonical Wnt signaling pathway"
## "viral budding via host ESCRT complex"
## "protein refolding"
## "retrograde transport, endosome to Golgi"
## "cristae formation"
## "cholesterol homeostasis"
## "negative regulation of apoptotic process"
## "negative regulation of cysteine-type endopeptidase activity involved in apoptotic pr
## "neutrophil degranulation"
## "steroid hormone mediated signaling pathway"
## "negative regulation of neuron apoptotic process"
## "positive regulation of GTPase activity"
## "cell fate commitment"
## "positive regulation of fat cell differentiation"
## "positive regulation of osteoblast differentiation"
## "negative regulation of transcription, DNA-templated"
## "positive regulation of transcription, DNA-templated"
## "positive regulation of transcription by RNA polymerase II"
## "ATP metabolic process"
## "regulation of synaptic plasticity"
## "COPII vesicle coating"
## "developmental growth"
## "regulation of neurogenesis"
## "regulation of catalytic activity"
## "protein stabilization"
## "chaperone cofactor-dependent protein refolding"
## "roof of mouth development"
## "canonical Wnt signaling pathway"
## "membrane organization"
## "cellular response to BMP stimulus"
## "negative regulation of canonical Wnt signaling pathway"
## "positive regulation of canonical Wnt signaling pathway"
## "positive regulation of cold-induced thermogenesis"
## "toxin transport"
## "beta-catenin-TCF complex assembly"
## "beta-catenin destruction complex disassembly"
##
## enriched.terms "Z.Tumor" "Z.Normal" "Z.Differential"
## "-3.737" "-4.385" "5.34"
## "13.279" "10.205" "6.085"
## "-8.63" "-11.448" "5.16"
## "-5.797" "-12.979" "9.586"
## "0.447" "-5.762" "10.926"
## "5.891" "1.758" "8.685"

```

##	"2.799"	"0.137"	"6.315"
##	"-22.059"	"-24.493"	"6.103"
##	"12.104"	"10.536"	"7.843"
##	"1.064"	"-1.279"	"6.726"
##	"3.117"	"-0.006"	"5.662"
##	"-23.115"	"-24.623"	"5.243"
##	"-4.834"	"-6.943"	"5.078"
##	"2.215"	"1.863"	"6.318"
##	"-5.911"	"-6.765"	"5.298"
##	"-4"	"-9.397"	"11.969"
##	"6.423"	"4.803"	"5.007"
##	"-12.518"	"-13.146"	"6.035"
##	"-6.513"	"-7.699"	"7.425"
##	"-9.127"	"-10.561"	"6.344"
##	"-1.066"	"-2.308"	"6.071"
##	"-4.142"	"-8.97"	"10.599"
##	"-0.11"	"-0.828"	"8.931"
##	"6.554"	"2.113"	"8.124"
##	"4.109"	"0.105"	"6.026"
##	"3.269"	"-1.876"	"8.35"
##	"-8.752"	"-11.681"	"8.712"
##	"-3.916"	"-7.812"	"7.511"
##	"-2.41"	"-4.362"	"5.489"
##	"-8.678"	"-10.3"	"5.817"
##	"-1.831"	"-1.435"	"6.25"
##	"-4.464"	"-6.244"	"6.099"
##	"-5.061"	"-7.337"	"5.109"
##	"3.171"	"-1.486"	"6.338"
##	"-6.807"	"-9.567"	"5.534"
##	"5.408"	"0.851"	"7.969"
##	"-4.452"	"-9.392"	"9.001"
##	"-4.406"	"-7.034"	"5.005"
##	"13.712"	"11.881"	"7.435"
##	"-5.599"	"-6.037"	"5.197"
##	"-21.99"	"-23.086"	"6.969"
##	"-6.001"	"-8.781"	"6.302"
##	"0.172"	"-0.008"	"5.363"
##	"13.97"	"11.975"	"7.06"
##	"-2.8"	"-5.906"	"5.734"
##	"-4.115"	"-5.801"	"5.082"
##	"-8.123"	"-11.269"	"5.024"
##	"-5.518"	"-6.714"	"5.883"
##	"-7.962"	"-9.372"	"5.164"
##	"3.505"	"0.389"	"6.718"
##	"-5.947"	"-7.538"	"8.016"
##	"-6.514"	"-8.333"	"5.717"
##	"-14.937"	"-20.518"	"6.022"
##	"-7.727"	"-9.86"	"9.338"
##	"-5.972"	"-8.827"	"5.009"
##	"-8.566"	"-10.341"	"5.196"
##	"-12.017"	"-10.634"	"5.571"
##	"-11.599"	"-11.275"	"9.389"
##	"-5.972"	"-4.889"	"7.805"
##	"11.101"	"10.325"	"5.521"



```
##          "-9.24"    "-10.642"   "5.525"
##          "2.219"    "-3.349"    "6.983"
##          "-7.837"   "-8.903"    "5.017"
##          "-6.482"   "-7.004"    "6.041"
##          "-0.171"   "-1.489"    "5.342"
##          "4.114"    "2.46"      "5.504"
##          "14.031"   "12.581"   "7.574"
##          "-5.453"   "-7.336"   "5.482"
##          "-8.112"   "-11.894"  "11.8"
##          "-6.643"   "-9.363"   "8.481"
##          "-8.76"    "-9.21"    "5.011"
##          "-7.665"   "-9.836"   "7.306"
##          "-3.1"     "-4.759"   "6.628"
##          "-13.806"  "-15.767"  "5.867"
##          "9.632"    "5.991"    "5.577"
##          "-0.481"   "-1.023"   "8.535"
##          "-3.366"   "-6.601"   "7.25"
```

```
write.table(enriched.terms, file="17.PvNP.bp.enriched.csv", sep="\t", col.names = F, row.names = F, quo
```

```
suppressed.terms <- c("GO.ID", "GO.Term", "Z.Tumor", "Z.Normal", "Z.Differential")
for (i in 1:length(suppress)) {
  id <- as.character(suppress[i])
  term <- Term(GO.ID(id))
  seri <- which(bp.go.cat %in% id)
  z.tumor <- z.p[20,seri]
  z.normal <- z.np[20,seri]
  z.diff <- z.pvnp[20,seri]
  suppressed.terms <- rbind(suppressed.terms, c(id, term, z.tumor, z.normal, z.diff))
}
suppressed.terms
```

```
##
## suppressed.terms "GO.ID"
##          "GO:0000027"
##          "GO:0000184"
##          "GO:0000209"
##          "GO:0000245"
##          "GO:0000375"
##          "GO:0000381"
##          "GO:0000387"
##          "GO:0000398"
##          "GO:0001580"
##          "GO:0002181"
##          "GO:0006120"
##          "GO:0006260"
##          "GO:0006283"
##          "GO:0006289"
##          "GO:0006334"
##          "GO:0006351"
##          "GO:0006357"
##          "GO:0006361"
##          "GO:0006363"
##          "GO:0006364"
##          "GO:0006368"
```

```

## "GO:0006369"
## "GO:0006370"
## "GO:0006378"
## "GO:0006396"
## "GO:0006397"
## "GO:0006405"
## "GO:0006406"
## "GO:0006412"
## "GO:0006413"
## "GO:0006614"
## "GO:0006935"
## "GO:0006955"
## "GO:0006956"
## "GO:0007186"
## "GO:0007193"
## "GO:0007204"
## "GO:0007218"
## "GO:0008380"
## "GO:0008543"
## "GO:0016070"
## "GO:0031124"
## "GO:0032981"
## "GO:0042795"
## "GO:0045292"
## "GO:0048025"
## "GO:0050911"
## "GO:0060326"
## "GO:0060964"
## "GO:0070098"
## GO:0000027
## suppressed.terms "GO.Term"
## "ribosomal large subunit assembly"
## "nuclear-transcribed mRNA catabolic process, nonsense-mediated decay"
## "protein polyubiquitination"
## "spliceosomal complex assembly"
## "RNA splicing, via transesterification reactions"
## "regulation of alternative mRNA splicing, via spliceosome"
## "spliceosomal snRNP assembly"
## "mRNA splicing, via spliceosome"
## "detection of chemical stimulus involved in sensory perception of bitter taste"
## "cytoplasmic translation"
## "mitochondrial electron transport, NADH to ubiquinone"
## "DNA replication"
## "transcription-coupled nucleotide-excision repair"
## "nucleotide-excision repair"
## "nucleosome assembly"
## "transcription, DNA-templated"
## "regulation of transcription by RNA polymerase II"
## "transcription initiation from RNA polymerase I promoter"
## "termination of RNA polymerase I transcription"
## "rRNA processing"
## "transcription elongation from RNA polymerase II promoter"
## "termination of RNA polymerase II transcription"
## "7-methylguanosine mRNA capping"

```

```

## "mRNA polyadenylation"
## "RNA processing"
## "mRNA processing"
## "RNA export from nucleus"
## "mRNA export from nucleus"
## "translation"
## "translational initiation"
## "SRP-dependent cotranslational protein targeting to membrane"
## "chemotaxis"
## "immune response"
## "complement activation"
## "G protein-coupled receptor signaling pathway"
## "adenylate cyclase-inhibiting G protein-coupled receptor signaling pathway"
## "positive regulation of cytosolic calcium ion concentration"
## "neuropeptide signaling pathway"
## "RNA splicing"
## "fibroblast growth factor receptor signaling pathway"
## "RNA metabolic process"
## "mRNA 3'-end processing"
## "mitochondrial respiratory chain complex I assembly"
## "snRNA transcription by RNA polymerase II"
## "mRNA cis splicing, via spliceosome"
## "negative regulation of mRNA splicing, via spliceosome"
## "detection of chemical stimulus involved in sensory perception of smell"
## "cell chemotaxis"
## "regulation of gene silencing by miRNA"
## "chemokine-mediated signaling pathway"
##
## suppressed.terms "Z.Tumor" "Z.Normal" "Z.Differential"
## "13.287" "16.114" "-5.365"
## "14.405" "16.471" "-11.518"
## "-6.78" "-7.066" "-5.401"
## "21.839" "26.653" "-9.467"
## "26.615" "33.108" "-8.647"
## "25.057" "30.565" "-8.374"
## "18.96" "23.072" "-5.126"
## "37.487" "45.921" "-25.573"
## "-27.582" "-29.217" "-7.542"
## "8.629" "10.215" "-7.447"
## "-13.098" "-10.687" "-5.106"
## "9.73" "16.622" "-5.107"
## "11.987" "16.613" "-7.57"
## "9.465" "16.451" "-5.5"
## "18.342" "21.541" "-6.729"
## "-37.573" "-35.083" "-11.756"
## "-30.787" "-29.925" "-10.889"
## "22.082" "29.988" "-6.328"
## "23.35" "31.403" "-5.902"
## "29.684" "33.13" "-9.211"
## "30.011" "36.357" "-9.318"
## "28.085" "36.305" "-9.622"
## "26.45" "33.574" "-8.672"
## "14.883" "18.587" "-5.545"
## "15.97" "20.562" "-5.283"

```

```
##      "28.918"  "37.519"  "-7.663"
##      "27.319"  "34.43"   "-11.203"
##      "28.886"  "35.153"  "-13.816"
##      "14.239"  "15.477"  "-8.634"
##      "13.262"  "14.146"  "-8.171"
##      "11.217"  "13.523"  "-10.395"
##      "-30.885" "-32.112"  "-6.414"
##      "-48.758" "-51.477"  "-6.419"
##      "-15.614" "-17.631"  "-5.348"
##      "-48.543" "-51.947"  "-9.709"
##      "-27.824" "-28.016"  "-6.158"
##      "-30.432" "-32.889"  "-5.332"
##      "-33.1"   "-32.623"  "-6.346"
##      "33.127"  "40.361"  "-11.357"
##      "15.963"  "21.579"  "-7.842"
##      "27.901"  "33.063"  "-10.41"
##      "31.361"  "40.541"  "-15.223"
##      "-16.535" "-13.475"  "-6.645"
##      "29.993"  "35.388"  "-6.399"
##      "29.927"  "34.609"  "-7.641"
##      "18.683"  "22.421"  "-5.378"
##      "-1.6"    "-1.562"  "-7.886"
##      "-23.615" "-24.658"  "-5.66"
##      "22.784"  "26.988"  "-7.021"
##      "-28.562" "-32.092"  "-7.429"
```

```
write.table(suppressed.terms, file="17.PvNP.bp.suppressed.csv", sep="\t", col.names = F, row.names = F,
```