#GO、KEGG enrichment analysis of DEGs

library(ggplot2)

library(ggpubr)

library(clusterProfiler)

library(org.Hs.eg.db)

library(stats)

library(data.table)

library(dplyr)

DEG\_data1 <- read.csv("DEGS.csv",sep=",",header = T)

gene.df <- bitr(DEG\_data1$gene, fromType = "SYMBOL",

toType = c("ENTREZID", "SYMBOL"),

OrgDb = org.Hs.eg.db)

write.table(gene.df,file="gene.df.csv",sep=",",quote=F,col.names=T)

gene <- gene.df$ENTREZID

GO\_all <-ego\_ALL <- enrichGO(gene = gene,

OrgDb=org.Hs.eg.db,

keyType = "ENTREZID",

ont = "ALL",

pAdjustMethod = "BH",

minGSSize = 1,

pvalueCutoff = 0.05,

qvalueCutoff = 0.05,

readable = TRUE)

GO\_all

GO\_result <- data.frame(GO\_all)

go\_enrichment\_pathway <- GO\_result %>% group\_by(ONTOLOGY) %>% top\_n(n = 5, wt = -p.adjust)

write.table(go\_enrichment\_pathway,file="go\_enrichment\_pathway.csv",sep=",",quote=F,row.names = F)

PLOTYELLOW<-ggplot(go\_enrichment\_pathway, aes(x=reorder(Description, Count), y=Count)) +

geom\_point(aes(size=Count,color=-log10(p.adjust))) +

scale\_size\_continuous(range=c(2, 6)) +

facet\_grid(ONTOLOGY~., scale = 'free\_y', space = 'free\_y')+

coord\_flip() +

theme\_minimal() +

scale\_color\_gradient(low = "pink",high ="red")+

labs(color=expression(-log10(p.adjust),size="Count"),

x="Gene Ratio",y="Gene\_Number",title="GO Enrichment")+

theme\_bw()

ggsave(filename="GO.png",

width=10,

height=5,

units="in",

dpi=300)

ggsave("GO of DEGS.pdf",width=10,height=4)

data <- read.csv("GO\_enrichment\_pathway.csv",header=TRUE)

GO\_term\_order=factor(as.integer(rownames(data)),labels=data$Description)

ggplot(data=data, aes(x=GO\_term\_order,y=Count, fill=ONTOLOGY)) + geom\_bar(stat="identity", width=0.8) + coord\_flip() +theme\_bw()

ek <- enrichKEGG(gene = gene.df$ENTREZID,

organism = "hsa",

pvalueCutoff =0.05,

qvalueCutoff = 0.05)

write.table(ek,file="eRICHk.txt",sep="\t",quote=F,row.names = F)

ek2 = setReadable(ek,

OrgDb = "org.Hs.eg.db",

keyType = "ENTREZID")

head(ek2@result$geneID)

write.table(ek2,file="ek2.txt",sep="\t",quote=F,row.names = F)

library("enrichplot")

pdf(file="ek\_barplot-YELLOW.pdf",width = 7,height = 3)

barplot(ek, x = "GeneRatio", color = "p.adjust",

showCategory =5)

dev.off()

library(tidyverse)

ek.rt = read.table("eRICHk.txt",header=TRUE,sep="\t",quote = "")

ek.rt <- separate(data=ek.rt, col=GeneRatio, into = c("GR1", "GR2"), sep = "/")

ek.rt <- separate(data=ek.rt, col=BgRatio, into = c("BR1", "BR2"), sep = "/")

ek.rt <- mutate(ek.rt, enrichment\_factor = (as.numeric(GR1)/as.numeric(GR2))/(as.numeric(BR1)/as.numeric(BR2)))

ek.rt10 <- ek.rt %>% filter(row\_number() >= 1,row\_number() <= 10)

p <- ggplot(ek.rt10,aes(enrichment\_factor, fct\_reorder(factor(Description), enrichment\_factor))) +

geom\_point(aes(size=Count,color=-1\*log10(pvalue))) +

scale\_color\_gradient(low="pink",high ="red") +

labs(color=expression(-log[10](p\_value)),size="Count",

x="Enrichment Factor",y="KEGG term",title="KEGG enrichment") +

theme\_bw()

p

ggsave("er.rt10\_KEGG.pdf",width=8,height=4)# 保存为pdf格式

ggsave("er.rt10\_KEGG.png",width=8,height=4,dpi=300)# 保存为png格式

#gsva

library(ggplot2)

library(ComplexHeatmap)

library(clusterProfiler)

library(GSVA)

library(GSEABase)

library(dplyr)

library(data.table)

library(tidyverse)

COAD <- read.csv("bindgeo\_exp.csv", check.names = FALSE, header = TRUE, row.names = 1)

#COAD <- log2(COAD+1)

gene\_set <- getGmt("c2.cp.kegg.v7.4.symbols.gmt")

group\_list <- read.csv("group.csv", check.names = FALSE, header = TRUE, row.names = 1)

table(group\_list)

annotation <- data.frame(group\_list)

rownames(annotation) <- colnames(COAD)

head(annotation)

gsva\_result<- gsva(as.matrix(COAD), gene\_set, method = "gsva",min.sz=1,

max.sz=Inf,kcdf="Gaussian")

library(dendextend)

library(circlize)

library(RColorBrewer)

colors = colorRampPalette(rev(brewer.pal(n = 7, name ="RdYlBu")))(100)

values <- seq(-0.8, 0.8, length.out = 101)[-101]

col\_fun = colorRamp2(values, colors)

top\_annotation<-HeatmapAnnotation(df=annotation,col=list(group=c("Normal"="blue","OA"="red")))

Heatmap(gsva\_result, name = "GSVA", col = col\_fun,cluster\_rows = T,cluster\_columns = F,show\_row\_names = T,

show\_column\_names = F,column\_split = annotation$group,)

annotation <- annotation[order(annotation$group == "RA", decreasing = TRUE), , drop = FALSE]

exp="RA"

ctr="NC"

design <- model.matrix(~0+factor(group\_list$group))

colnames(design) <- levels(factor(group\_list$group))

rownames(design) <- colnames(gsva\_result)

contrast.matrix <- makeContrasts(contrasts=paste0(exp,'-',ctr), #"exp/ctrl"

levels = design)

fit1 <- lmFit(gsva\_result,design)

fit2 <- contrasts.fit(fit1, contrast.matrix)

efit <- eBayes(fit2)

summary(decideTests(efit,lfc=1, p.value=0.05))

tempOutput <- topTable(efit, coef=paste0(exp,'-',ctr), n=Inf)

degs <- na.omit(tempOutput)

keep <- rownames(degs[degs$adj.P.Val<0.25& degs$P.Value<0.05, ])

length(keep)

dat <- gsva\_result[keep[1:20],]

pheatmap(dat,cluster\_rows = T,cluster\_cols = F,

color=colorRampPalette(c("navy","white","firebrick3"))(100),

show\_colnames = F,border\_color = NA,scale = "row",show\_rownames =T,

annotation\_col = group\_list,fontsize=5.0,name = "GSVA")

#gsea analysis

library(clusterProfiler)

library(org.Hs.eg.db)

KEGG\_database <- 'hsa'

library(tidyverse)

GSEoa= read.table("DEG.csv",sep = ",",header = T)

library(dplyr)

GSEoa<- GSEoa %>% filter(abs(logFC) > 0) %>% filter(adj.P.Val < 0.05)

names(GSEoa)[1] <- "SYMBOL"

GSEoa1 = GSEoa[,c(1,2)]

genename <- GSEoa1$SYMBOL

En\_id <- mget(genename,

org.Hs.egSYMBOL2EG,

ifnotfound=NA)

En\_id <- as.character(En\_id)

a=data.frame(ENTREZID=En\_id,logFC=GSEoa1$logFC)

a$logFC<-sort(a$logFC,decreasing = T)

geneList = a[,2]

names(geneList) = as.character(a[,1])

geneList

library(stats)

GSEA\_KEGG <- gseKEGG(geneList, organism = 'hsa', pvalueCutoff = 0.05)

ridgeplot(GSEA\_KEGG)

gseaplot2(GSEA\_KEGG,1)

gseaplot2(GSEA\_KEGG, title = GSEA\_KEGG$Description[1], geneSetID = 1)

gseaplot2(GSEA\_KEGG,1:5)

gseaplot2(GSEA\_KEGG,geneSetID = 1,

title = "",

color = "green",

base\_size = 12,

rel\_heights = c(1.5, 0.5, 1),

subplots = 1:3,

pvalue\_table = T,

ES\_geom = "line"

)