

```

setwd("C:/Users")
library(readxl)

dat<-read_excel("C4-2_DMSO_123_vs_MDV-R_123.xlsx")
dat$log2(fold_change)`<-as.numeric(dat$log2(fold_change))
note1<-read.delim("hg38-cancer-circrna.txt",header = FALSE)
note2<-read.csv("human_hg38_All_circRNA.csv",header = FALSE,quote = ",")
note1<-note1[note1$V18%in%c("lncRNA","mRNA,lncRNA"),]
note1<-unique(note1$V15)
note1<-intersect(note1,note2$V3)
data1<-dat[dat$gene%in%note1,]
data1<-data1[data1$p_value<0.001,]
dat1<-dat[dat$log2(fold_change)`>5,]
dat2<-dat[dat$log2(fold_change)`<(-5),]

gene<-union(dat1$gene,dat2$gene)

data<-data1
data$log2(fold_change)`<-as.numeric(data$log2(fold_change))
data<-data[abs(data$log2(fold_change))<20,]
data$log2(fold_change)`<-(-data$log2(fold_change))
data$change <- factor(ifelse(data$p_value < 0.001 & abs(data$log2(fold_change)) >= 5,
ifelse(data$log2(fold_change)` >= 5 , 'Up', 'Down'), 'NoSignifi'),levels=c('Up', 'Down', 'NoSignifi'))
data$gene[data$change=="NoSignifi"]<-NA
data$p_value[data$p_value==0]<-2.3e-18

ggplot(data = data, aes(x = `log2(fold_change)`, y = -log10(p_value), color = change)) +
  geom_point(size = 2.5) +
  geom_hline(yintercept = 3,lty=3,col="#AEB6BF",lwd=0.8) +
  geom_vline(xintercept=c(-5,5),lty=3,col="#AEB6BF",lwd=0.8) +
  geom_text_repel(aes(label = gene),
    size = 4.5,
    color = "black",
    segment.color = "black", show.legend = FALSE )+
  theme_classic(base_size = 15) +
  theme(panel.grid.minor = element_blank(),axis.text = element_text(colour = "black"),
    panel.grid.major = element_blank(),legend.position = "top") +
  labs(x=expression(Log[2](FoldChange)),y=expression(-Log[10](padj)))+
  scale_color_manual(name = "", values=c( "#FC4E2A","#4393C3","#00000033"), limits = c("Up",
"Down", "NoSignifi"))

library(GEOquery)

```

```

library(Biobase)
library(limma)
library(tidyverse)
library(affy)
library(limma)
library(AnnoProbe)
library(stringr)

#read GEO data
Sys.setenv("VROOM_CONNECTION_SIZE"=131072*600)
gset<-getGEO(GEO='GSE70769', destdir=getwd(),getGPL = F)
e2<-gset[[1]]
expma<-as.data.frame(exprs(e2))

expr<-as.data.frame(expma)
expr$ID<-rownames(expr)

GPL="GPL10558"
gpl <- getGEO(GPL,destdir = getwd()) %>% Table()
colnames(gpl)

probe<-gpl %>% select("ID","Symbol")

#Convert the probe
exp<-merge(expr,probe,by= "ID")
exp<-exp[,-1]
exp<-exp[exp$Symbol!="",]

#group
phen<-pData(e2)
phen$Name<-as.data.frame(str_split_fixed(phen$title, "_", 3))[,3]

phen1<-data.frame(os1=phen$`biochemical relapse (bcr):ch1`,
                  os.time1=as.numeric(phen$`time to bcr (months):ch1`)*30,
                  os.time2=as.numeric(phen$`total follow up (months):ch1`)*30,
                  TNM=phen$`pathology stage:ch1`,
                  T1=phen$`clinical stage:ch1`)
phen1$os<-0
phen1$os[phen1$os1=="Y"]<-1
phen1$os.time<-phen1$os.time2
phen1$os.time[phen1$os==1]<-phen1[phen1$os==1,"os.time1"]
phen1$TNM_T<-NA

```

```
phen1$TNM_T1<-NA
phen1[grep("T1",phen1$TNM),"TNM_T"]<-"T1"
phen1[grep("T2",phen1$TNM),"TNM_T"]<-"T2"
phen1[grep("T3",phen1$TNM),"TNM_T"]<-"T3"
phen1[grep("T4",phen1$TNM),"TNM_T"]<-"T4"
```

```
phen1[grep("T1",phen1$T1),"TNM_T1"]<-"T1"
phen1[grep("T2",phen1$T1),"TNM_T1"]<-"T2"
phen1[grep("T3",phen1$T1),"TNM_T1"]<-"T3"
phen1[grep("T4",phen1$T1),"TNM_T1"]<-"T4"
```

```
rownames(phen1)<-rownames(phen)
phen1<-phen1[,6:9]
phen1<-phen1[!is.na(phen1$os.time),]
```

```
#Average the same gene names
exp<-exp[,c(phen$geo_accession,"Symbol")]
names(exp)<-c(rownames(phen),"Symbol")
exp1<-aggregate(exp[,1:94],by=list(exp[,95]),mean)
exp1<-exp1[,-1]
rownames(exp1)<-exp1$Group.1
exp1<-exp1[,-1]
exp1<-as.data.frame(log2(exp1+1))
```

```
exp1<-exp1[,rownames(phen1)]
```

```
#save
GSE70769_Expr<-exp1
GSE70769_Traits<-phen1
save(GSE70769_Expr,GSE70769_Traits,file = 'GSE70769.Rdata')
```

```
load("GSE70769.Rdata")
GSE70769_Expr<-as.data.frame(t(GSE70769_Expr["BMP4",]))
GSE70769_Traits$BMP4<-GSE70769_Expr$BMP4
res.cut1 <-surv_cutpoint(GSE70769_Traits, time = "os.time", event = "os",variables ="BMP4")
res.cat1 <- surv_categorize(res.cut1)
GSE70769_Traits$G<-res.cat1[,3]
fit <- survfit(Surv(os.time,os) ~G, data = GSE70769_Traits)
```

```
p<-ggsurvplot(fit,
  data = GSE70769_Traits,
```

```

conf.int=FALSE,
pval.size=7,
#risk.table=TRUE,
legend.title = "GSE70769",
legend.labs = c("High", "Low"),
xlab = "Follow up time(d)",
palette = c("#FC4E2A", "#4393C3"),
ggtheme=theme_classic(base_size = 16),
pval = T)

p<-p$plot+theme(axis.text = element_text(colour = "black"))

dat1<-GSE70769_Traits[!is.na(GSE70769_Traits$TNM_T),]
ggplot(dat1, aes(x=TNM_T, y=BMP4, fill=TNM_T)) +
  geom_violin()+
  geom_boxplot(width=0.2,outlier.size=0)+
  theme_classic(base_size = 15)+
  theme(legend.position = "none",
        legend.title=element_blank(),
        axis.text = element_text(colour = "black"))+
  scale_fill_manual(values=c("#FC4E2A", "#4393C3"))+
  scale_color_manual(values=c("#FC4E2A", "#4393C3"))+
  labs(x="",y="BMP4")

gset<-getGEO(GEO='GSE74367', destdir=getwd(),getGPL = F)
e2<-gset[[1]]

raw_dir <- "G:/GSE74367_RAW"
raw_datas <- paste0(raw_dir,"/",list.files(raw_dir,pattern = "GSM\\d*"))

data.raw <- read.maimages(raw_datas,
  source = "agilent",
  green.only = F,
  other.columns = "glsWellAboveBG")

data.bg <- backgroundCorrect(data.raw,method = "normexp")

data.norm <- normalizeWithinArrays(data.bg, method = "loess")

data.norm <- normalizeBetweenArrays(data.norm, method = "Aquantile")

```

```

head(data.norm$M)
head(data.norm$A)

data.exp <- as.data.frame(data.norm$A)
data.exp$ProbeName <- data.norm$genes$GeneName
data.exp <- aggregate(data.exp[,1:76],by = list(data.exp[,77]),FUN = mean)
rownames(data.exp)<-data.exp$Group.1
data.exp <- data.exp[,-1]

phen<-pData(e2)
phen1<-data.frame(GEO=phen$geo_accession,Group=as.data.frame(str_split_fixed(phen$title,"
",3)[,3]))
names(phen1)<-c("GEO","Group")
GSE74367_Traits<-phen1
GSE74367_Expr<-data.exp
names(GSE74367_Expr)<-phen1$GEO

save(GSE74367_Expr,GSE74367_Traits,file = "GSE74367.Rdata")

data1<-data.frame("BMP4"=t(GSE74367_Expr["BMP4",]),"group"=GSE74367_Traits$Group)
data1$group[data1$group=="metastasis"]<-"mCRPC"
data1$group[data1$group=="PCa xenograft tumor"]<-"CRPC"
data1$group[data1$group=="PCa tumor"]<-"Primary PCA"
boxplot(data1$BMP4~data1$group)

data1$group<-factor(data1$group,levels = c("Primary PCA","CRPC","mCRPC"),ordered = TRUE)

dat1<-data1[data1$group!="Primary PCA",]

ggplot(dat1, aes(x=group, y=BMP4,fill=group)) +
  geom_violin( alpha = 0.7)+
  geom_boxplot(width=0.1,outlier.size=0)+
  stat_compare_means(
    comparisons = list(c("MCRPC","mCRPC")),
    method = "wilcox.test",
    label = "p.signif")+
  theme_classic(base_size = 15)+
  theme(legend.position = "none",
        plot.title = element_text(size=12,hjust=0.5),
        legend.title=element_blank(),
        axis.text = element_text(colour = "black"))+
  scale_fill_manual(values=c("#FC4E2A", "#4393C3"))+
  scale_color_manual(values=c("#FC4E2A", "#4393C3"))+
  labs(x="",y="BMP4",title = "GSE74367")

```

```
dat2<-data1[data1$group!="mCRPC",]
```

```
ggplot(dat2, aes(x=group, y=BMP4,fill=group)) +  
  geom_violin( alpha = 0.7)+  
  geom_boxplot(width=0.1,outlier.size=0)+  
  stat_compare_means(  
    comparisons = list(c("MCRPC", "Primary PCA")),  
    method = "wilcox.test",  
    label = "p.signif")+  
  theme_classic(base_size = 15)+  
  theme(legend.position = "none",  
        plot.title = element_text(size=12,hjust=0.5),  
        legend.title=element_blank(),  
        axis.text = element_text(colour = "black"))+  
  scale_fill_manual(values=c("#FC4E2A", "#F5B041"))+  
  scale_color_manual(values=c("#FC4E2A", "#F5B041"))+  
  labs(x="",y="BMP4",title = "GSE74367")
```

```
dat3<-data1[data1$group!="MCRPC",]
```

```
ggplot(dat3, aes(x=group, y=BMP4,fill=group)) +  
  geom_violin( alpha = 0.7)+  
  geom_boxplot(width=0.1,outlier.size=0)+  
  stat_compare_means(  
    comparisons = list(c("mCRPC", "Primary PCA")),  
    method = "wilcox.test",  
    label = "p.signif")+  
  theme_classic(base_size = 15)+  
  theme(legend.position = "none",  
        plot.title = element_text(size=12,hjust=0.5),  
        legend.title=element_blank(),  
        axis.text = element_text(colour = "black"))+  
  scale_fill_manual(values=c("#4393C3", "#F5B041"))+  
  scale_color_manual(values=c("#4393C3", "#F5B041"))+  
  labs(x="",y="BMP4",title = "GSE74367")
```

```
ggplot(data1, aes(x=group, y=BMP4,fill=group)) +  
  geom_violin( alpha = 0.7)+  
  geom_boxplot(width=0.1,outlier.size=0)+  
  stat_compare_means(  
    comparisons = list(c("CRPC", "Primary PCA"),c("CRPC", "mCRPC"),c("mCRPC", "Primary PCA")),  
    method = "wilcox.test",  
    label = "p.signif")+
```

```
theme_classic(base_size = 15)+  
theme(legend.position = "none",  
      plot.title = element_text(size=12,hjust=0.5),  
      legend.title=element_blank(),  
      axis.text = element_text(colour = "black"))+  
scale_fill_manual(values=c( "#4393C3", "#F5B041", "#FC4E2A"))+  
scale_color_manual(values=c("#4393C3", "#F5B041", "#FC4E2A"))+  
labs(x="",y="BMP4",title = "GSE74367")
```