#install bioconductor

source("http://bioconductor.org/biocLite.R")

biocLite()

#Read the data

Data\_TC<-read.csv("combined\_RNA.csv",header=T)

head(Data\_TC)

colnames(Data\_TC)

gene\_ID<-Data\_TC[,1]

Data\_TC2<-Data\_TC[,-1]

head(Data\_TC2)

colnames(Data\_TC2)

#Data Labels

colnames(Data\_TC2)<-rep(c("PTC\_LNM","PTC\_withoutLNM","Normal"),c(205,213,59))

n1=205;

n2=213;

n3=59;

nc=477;

DataLabel<-c(rep(1,n1),rep(2,n2),rep(3,n3))

dim(Data\_TC2)

colnames(Data\_TC2)

rownames(Data\_TC2)

rownames(Data\_TC2)<-gene\_ID

--------------------------------------------------------------------------

##Group data##

#PTC\_LNM vs PTC\_withoutLNM

colnames(Data\_TC2)

PTC\_LNMvsPTC\_withoutLNM<-Data\_TC2[,c(1:205,206:418)]

colnames(PTC\_LNMvsPTC\_withoutLNM)

#change format to matrix format

PTC\_LNMvsPTC\_withoutLNM\_mat<-as.matrix(PTC\_LNMvsPTC\_withoutLNM)

head(PTC\_LNMvsPTC\_withoutLNM\_mat)

#data with gene\_id as a row

PTC\_LNMvsPTC\_withoutLNM\_log<-log2(PTC\_LNMvsPTC\_withoutLNM\_mat+1)

head(PTC\_LNMvsPTC\_withoutLNM\_log)

dim(PTC\_LNMvsPTC\_withoutLNM\_log)

colnames(PTC\_LNMvsPTC\_withoutLNM\_log)

colnames(PTC\_LNMvsPTC\_withoutLNM\_log)<-rep(c("PTC\_LNM","PTC\_withoutLNM"),c(205,213))

n1=205;

n2=213;

nc=418;

DataLabelPTC\_LNMvsPTC\_withoutLNM<-c(rep(1,n1),rep(2,n2))

## Statistical Analysis

pValue\_t<-NULL

for (i2 in 1:dim(PTC\_LNMvsPTC\_withoutLNM\_log)[1])

{

DataYY<-data.frame(YY=PTC\_LNMvsPTC\_withoutLNM\_log[i2,],FactorLevels=factor(DataLabelPTC\_LNMvsPTC\_withoutLNM))

pValue\_t[i2]<-t.test(YY~FactorLevels,paired=F,data=DataYY,var.equal=FALSE)[[3]]}

pValuePTC\_LNMvsPTC\_withoutLNM.adj<-p.adjust(pValue\_t,method="BH")

sig\_ttestPTC\_LNMvsPTC\_withoutLNM<-which(pValuePTC\_LNMvsPTC\_withoutLNM.adj<0.05)

length(sig\_ttestPTC\_LNMvsPTC\_withoutLNM)

markersPTC\_LNMvsPTC\_withoutLNM<-PTC\_LNMvsPTC\_withoutLNM\_log[sig\_ttestPTC\_LNMvsPTC\_withoutLNM,]

write.csv(markersPTC\_LNMvsPTC\_withoutLNM,"8611\_PTC\_LNMvsPTC\_withoutLNM.csv",row.names=T)

## Statistical Analysis and exporting p-value

pValue\_t<-NULL

for (i2 in 1:dim(PTC\_LNMvsPTC\_withoutLNM\_log)[1])

{

DataYY<-data.frame(YY=PTC\_LNMvsPTC\_withoutLNM\_log[i2,],FactorLevels=factor(DataLabelPTC\_LNMvsPTC\_withoutLNM))

pValue\_t[i2]<-t.test(YY~FactorLevels,paired=F,data=DataYY,var.equal=FALSE)[[3]]}

pValuePTC\_LNMvsPTC\_withoutLNM.adj<-p.adjust(pValue\_t,method="BH")

sig\_ttestPTC\_LNMvsPTC\_withoutLNM<-which(pValuePTC\_LNMvsPTC\_withoutLNM.adj<0.05)

pvaluedata<-pValuePTC\_LNMvsPTC\_withoutLNM.adj[sig\_ttestPTC\_LNMvsPTC\_withoutLNM]

length(sig\_ttestPTC\_LNMvsPTC\_withoutLNM)

rownames(sig\_ttestPTC\_LNMvsPTC\_withoutLNM)

row\_sig\_pvalue<-rownames(PTC\_LNMvsPTC\_withoutLNM\_log[sig\_ttestPTC\_LNMvsPTC\_withoutLNM,])

sig\_markers<-cbind(row\_sig\_pvalue,pvaluedata)

head(sig\_markers)

write.csv(sig\_markers,"pvalue\_8611\_PTC\_LNMvsPTC\_withoutLNM.csv",row.names=T)

-----------------------------------------------------------------------------------------------------------------------------------------

##Group data##

#PTC\_LNM vs normal thyroid

colnames(Data\_TC2)

PTC\_LNMvsNormal<-Data\_TC2[,c(1:205,419:477)]

colnames(PTC\_LNMvsNormal)

#change format to matrix format

PTC\_LNMvsNormal\_mat<-as.matrix(PTC\_LNMvsNormal)

head(PTC\_LNMvsNormal\_mat)

#data with gene\_id as a row

PTC\_LNMvsNormal\_log<-log2(PTC\_LNMvsNormal\_mat+1)

head(PTC\_LNMvsNormal\_log)

dim(PTC\_LNMvsNormal\_log)

colnames(PTC\_LNMvsNormal\_log)

colnames(PTC\_LNMvsNormal\_log)<-rep(c("PTC\_LNM","Normal"),c(205,59))

n1=205;

n2=59;

nc=264;

DataLabelPTC\_LNMvsNormal<-c(rep(1,n1),rep(2,n2))

## Statistical Analysis

pValue\_t<-NULL

for (i2 in 1:dim(PTC\_LNMvsNormal\_log)[1])

{

DataYY<-data.frame(YY=PTC\_LNMvsNormal\_log[i2,],FactorLevels=factor(DataLabelPTC\_LNMvsNormal))

pValue\_t[i2]<-t.test(YY~FactorLevels,paired=F,data=DataYY,var.equal=FALSE)[[3]]}

pValuePTC\_LNMvsNormal.adj<-p.adjust(pValue\_t,method="BH")

sig\_ttestPTC\_LNMvsNormal<-which(pValuePTC\_LNMvsNormal.adj<0.05)

length(sig\_ttestPTC\_LNMvsNormal)

markersPTC\_LNMvsNormal<-PTC\_LNMvsNormal\_log[sig\_ttestPTC\_LNMvsNormal,]

write.csv(markersPTC\_LNMvsNormal,"14192\_PTC\_LNMvsNormal.csv",row.names=T)

## Statistical Analysis and exporting p-value

pValue\_t<-NULL

for (i2 in 1:dim(PTC\_LNMvsNormal\_log)[1])

{

DataYY<-data.frame(YY=PTC\_LNMvsNormal\_log[i2,],FactorLevels=factor(DataLabelPTC\_LNMvsNormal))

pValue\_t[i2]<-t.test(YY~FactorLevels,paired=F,data=DataYY,var.equal=FALSE)[[3]]}

pValuePTC\_LNMvsNormal.adj<-p.adjust(pValue\_t,method="BH")

sig\_ttestPTC\_LNMvsNormal<-which(pValuePTC\_LNMvsNormal.adj<0.05)

pvaluedata<-pValuePTC\_LNMvsNormal.adj[sig\_ttestPTC\_LNMvsNormal]

length(sig\_ttestPTC\_LNMvsNormal)

rownames(sig\_ttestPTC\_LNMvsNormal)

row\_sig\_pvalue<-rownames(PTC\_LNMvsNormal\_log[sig\_ttestPTC\_LNMvsNormal,])

sig\_markers<-cbind(row\_sig\_pvalue,pvaluedata)

head(sig\_markers)

write.csv(sig\_markers,"pvalue\_14192\_PTC\_LNMvsNormal.csv",row.names=T)

-----------------------------------------------------------------------------------------------------------------------------------------

##Group data##

#PTC\_withoutLNM vs normal thyroid

colnames(Data\_TC2)

PTC\_withoutLNMvsNormal<-Data\_TC2[,c(206:418,419:477)]

colnames(PTC\_withoutLNMvsNormal)

#change format to matrix format

PTC\_withoutLNMvsNormal\_mat<-as.matrix(PTC\_withoutLNMvsNormal)

head(PTC\_withoutLNMvsNormal\_mat)

#data with gene\_id as a row

PTC\_withoutLNMvsNormal\_log<-log2(PTC\_withoutLNMvsNormal\_mat+1)

head(PTC\_withoutLNMvsNormal\_log)

dim(PTC\_withoutLNMvsNormal\_log)

colnames(PTC\_withoutLNMvsNormal\_log)

colnames(PTC\_withoutLNMvsNormal\_log)<-rep(c("PTC\_withoutLNM","Normal"),c(213,59))

n1=213;

n2=59;

nc=272;

DataLabelPTC\_withoutLNMvsNormal<-c(rep(1,n1),rep(2,n2))

## Statistical Analysis

pValue\_t<-NULL

for (i2 in 1:dim(PTC\_withoutLNMvsNormal\_log)[1])

{

DataYY<-data.frame(YY=PTC\_withoutLNMvsNormal\_log[i2,],FactorLevels=factor(DataLabelPTC\_withoutLNMvsNormal))

pValue\_t[i2]<-t.test(YY~FactorLevels,paired=F,data=DataYY,var.equal=FALSE)[[3]]}

pValuePTC\_withoutLNMvsNormal.adj<-p.adjust(pValue\_t,method="BH")

sig\_ttestPTC\_withoutLNMvsNormal<-which(pValuePTC\_withoutLNMvsNormal.adj<0.05)

length(sig\_ttestPTC\_withoutLNMvsNormal)

markersPTC\_withoutLNMvsNormal<-PTC\_withoutLNMvsNormal\_log[sig\_ttestPTC\_withoutLNMvsNormal,]

write.csv(markersPTC\_withoutLNMvsNormal,"13392\_PTC\_withoutLNMvsNormal.csv",row.names=T)

## Statistical Analysis and exporting p-value

pValue\_t<-NULL

for (i2 in 1:dim(PTC\_withoutLNMvsNormal\_log)[1])

{

DataYY<-data.frame(YY=PTC\_withoutLNMvsNormal\_log[i2,],FactorLevels=factor(DataLabelPTC\_withoutLNMvsNormal))

pValue\_t[i2]<-t.test(YY~FactorLevels,paired=F,data=DataYY,var.equal=FALSE)[[3]]}

pValuePTC\_withoutLNMvsNormal.adj<-p.adjust(pValue\_t,method="BH")

sig\_ttestPTC\_withoutLNMvsNormal<-which(pValuePTC\_withoutLNMvsNormal.adj<0.05)

pvaluedata<-pValuePTC\_withoutLNMvsNormal.adj[sig\_ttestPTC\_withoutLNMvsNormal]

length(sig\_ttestPTC\_withoutLNMvsNormal)

rownames(sig\_ttestPTC\_withoutLNMvsNormal)

row\_sig\_pvalue<-rownames(PTC\_withoutLNMvsNormal\_log[sig\_ttestPTC\_withoutLNMvsNormal,])

sig\_markers<-cbind(row\_sig\_pvalue,pvaluedata)

head(sig\_markers)

write.csv(sig\_markers,"pvalue\_13392\_PTC\_withoutLNMvsNormal.csv",row.names=T)