#ICAclust: a tool for clustering of temporal RNA-Seq data

rm(list = ls())

#Required package

#install.packages("fastICA")

library(fastICA)

#Specifying data Path

setwd("your path")

#The first and second columns should be the genes and Ids)

data <- read.table("your\_data\_set.txt", header = T)

#Real Data (The first and second colum are genes and Ids)

data <- read.table("http://www.seraolab.com/resources/ICAclust/RNA\_seq\_Pig.txt",

header = T)

#Function to clustering genes

ICAclust<-function(data=data, cluster\_method="ward.D", mojena=3.75, times = c(20,40,70,90))

{

X=data

Y=X[,3:ncol(X)]

X.ICA<-fastICA(Y,n.comp=ncol(Y), alg.typ = "parallel", fun = "exp", alpha = 1.0,

method = "C", row.norm = FALSE, maxit = 5000, tol = 1e-03, verbose = TRUE)

hc.ICA<-hclust(dist(X.ICA$S), method = cluster\_method, members=NULL)

dendro<-as.dendrogram(hc.ICA)

plot(hc.ICA,labels = NULL, hang = 0.0,

axes = TRUE, frame.plot = FALSE, ann = TRUE,

main = "Cluster Dendrogram",

sub = NULL, xlab = "", ylab = "Height")

hc.ICA$order

mojema=mean(hc.ICA$height)+mojena\*sd(hc.ICA$height)

g=length(hc.ICA$height[hc.ICA$height>mojema]) + 1

rect.hclust(hc.ICA, k =g)

group\_number<-g

group<-cutree(hc.ICA, k=g)

number\_gene<-table(group)

grupos1<-as.data.frame(X$geneID)

comp<-cbind(group,grupos1)

dadosnovos <-cbind(group,X[,-c(1,2)])

dadosnovoss <-cbind(group,X[,2],X[,-c(1,2)])

names(dadosnovoss)<-c("group",names(X)[-1])

dadosnovos1 <-as.data.frame(dadosnovos)

a1=by(dadosnovos1, dadosnovos1[,"group"],function(x) colMeans(x))

x1<-do.call("rbind",as.list(a1))

write.csv(dadosnovoss, file = "groups.csv")

write.table(dadosnovoss, file = "groups.txt")

write.csv(x1, file = "averagevalues.csv")

write.table(x1, file = "averagevalues.txt")

result<-list( "nGroups" = group\_number, "gbGroup"= number\_gene)

return(result)

}

#Using ICAclust function

result <- ICAclust(data=data, cluster\_method="ward.D", mojena=3.75, times = c(20,40,70,90))

result