

# Causal Inference Methods

## Causal inference - BRCA

hiddenICP with BRCA with  $1/3 + 1/3 + 1/3$

```
# Reset the environment
rm(list = ls ())

library(InvariantCausalPrediction)
library(AER)
library(CancerSubtypes)
library(miRBaseConverter)
library(miRLAB)

##=====
# Set environment variables here
directoryPath <- "C:/Users/phavy022/MyDoc/02CausalInferenceMethods/Experiment-Final/InOut/"
# directoryPath <- "R/CausalCompare/"
topk_miR <- 30
topk_mR <- 1500
topk <- 15000
k <- 500
##=====

source(paste(directoryPath, "CausalCompare.R", sep=""))

# 1. Get data set
load(paste(directoryPath, "01BRCA/BRCA_matchedData2.RData", sep=""))
# Get significant miRs and mRs from matchedData
l <- getDataByMAD(matchedData, topk_miR, topk_mR)

# 2. Estimate causal effects
nmiR <- length(l$miRs)
nmR <- length(l$mRs)
standardizedData <- scale(l$d)
temp = matrix(nrow = nmR, ncol = nmiR)

## define experimental settings
ExpInd <- c(rep(0,nrow(standardizedData)))
noOfsample <- nrow(l$d)
ExpInd[1:round(noOfsample/3)] <- 1
ExpInd[(round(noOfsample/3)+1):round((2*noOfsample)/3)] <- 2
ExpInd[(round((2*noOfsample)/3)+1):noOfsample] <- 3

temp_maximinCoefficients = matrix(nrow = nmR, ncol = nmiR)
temp_pvalues = matrix(nrow = nmR, ncol = nmiR)
# Estimate causal effects
for (i in 1:nmR) {
  Y <- standardizedData[,nmiR+i]
  X <- standardizedData[,c(1:nmiR)]
```

```

icp <- hiddenICP(X, Y, ExpInd, alpha = 0.1, mode = "asymptotic", intercept=FALSE)
for (k in 1:nmiR) {
  temp[i,k] <- icp$betahat[k]
  temp_maximinCoefficients[i,k] <- icp$maximinCoefficients[k]
  temp_pvalues[i,k] <- icp$pvalues[k]
}
}

colnames(temp) <- l$miRs
row.names(temp) <- l$mRs
## Write the result
# write.csv(temp, file = paste(directoryPath, "ICP_result.csv", sep = ""))
write.csv(temp, file = paste(directoryPath, "hiddenICP_BRCA13_result.csv", sep = ""))

# 3. Validate
#colnames(temp) <- convertmiRs(colnames(temp))

# Get the result
temp <- read.csv(file = paste(directoryPath, "hiddenICP_BRCA13_result.csv", sep = ""), row.names = 1)
temp <- data.matrix(temp)
nmiR <- ncol(temp)

fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(topk in c(500, 1000, 1500, 2000)) {
  top = Extopk(temp, topk)
  write.csv(top, file = paste(directoryPath, "hiddenICP_BRCA13_top_", topk, ".csv", sep = ""))
  topConfirmed = Validation(top, fileName)
  write.csv(topConfirmed[[1]], file = paste(directoryPath, "hiddenICP_BRCA13_topConfirmed_",
                                          topk, ".csv", sep = ""))
}

##=====
# Validate for each miRNA
# colnames(temp) <- convertmiRs(colnames(temp))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(k in c(50, 100, 150, 200)) {
  topConfirmed = NULL
  top = NULL
  for(i in 1:nmiR) {
    miRiTopk = bRank(temp, i, k, FALSE)
    top <- rbind(top, miRiTopk)
    topConfirmedi = Validation(miRiTopk, fileName)
    if(nrow(topConfirmedi[[1]]) != 0) {
      topConfirmed <- rbind(topConfirmed, topConfirmedi[[1]])
    }
  }
  write.csv(top, file = paste(directoryPath, "hiddenICP_BRCA13_top_eachmiR_", k, ".csv",
                              sep = ""), row.names = FALSE)
  write.csv(topConfirmed, file = paste(directoryPath, "hiddenICP_BRCA13_topConfirmed_eachmiR_",
                                       k, ".csv",
                                       sep = ""), row.names = FALSE)
}

```

```
##=====
```

## hiddenICP with BRCA - PAM50

```
# Reset the environment
rm(list = ls ())

library(InvariantCausalPrediction)
library(AER)
library(CancerSubtypes)
library(miRBaseConverter)
library(miRLAB)
library(genefu)

##=====
# Set environment variables here
directoryPath <- "C:/Users/phavy022/MyDoc/02CausalInferenceMethods/Experiment-Final/InOut/"
# directoryPath <- "R/CausalCompare/"
dataDir = "bioclassifier_data"
inputFileName = "inputFile.txt"
topk_miR <- 30
topk_mR <- 1500
##=====

source(paste(directoryPath, "CausalCompare.R", sep=""))

# 1. Get data set
load(paste(directoryPath, "01BRCA/BRCA_matchedData2.RData", sep=""))
# Get significant miRs and mRs from matchedData
l <- getDatabyMAD(matchedData, topk_miR, topk_mR)

# 2. Estimate causal effects
####-----
# Classify samples based on cancer subtypes using Pam50
# Get 50 mRNAs from Pam50
data(pam50)
str(pam50)
fiftymRNAs <- pam50$centroids.map$probe.centroids
# In matchedData$mRNAs, there are not CDCA1, KNTC2, ORC6L which are in Pam50
# However, CDCA1 ~ NUF2, KNTC2 ~ NDC80, ORC6L ~ ORC6
# In matchedData$mRNAs, replace NUF2 by CDCA1, NDC80 by KNTC2, ORC6 by ORC6L
colnames(matchedData$mRNAs)[which(colnames(matchedData$mRNAs) == "NUF2")] <- "CDCA1"
colnames(matchedData$mRNAs)[which(colnames(matchedData$mRNAs) == "NDC80")] <- "KNTC2"
colnames(matchedData$mRNAs)[which(colnames(matchedData$mRNAs) == "ORC6")] <- "ORC6L"
# Get data of 50 mRNAs of Pam50 from matchedData
fiftymRNAsData <- matchedData$mRNAs[, fiftymRNAs]
prepareData(fiftymRNAsData, directoryPath, dataDir, inputFileName)
####
# input variables for the subtype prediction script
####
library(ctc)
library(heatmap.plus)
```

```

paramDir<- paste(directoryPath, "bioclassifier_R", sep = "") # the location of unchanging files
# such as the function library and main program
inputDir<- paste(directoryPath, dataDir, sep = "") # the location of the data matrix,
# and where output will be located

inputFile<- inputFileNames # the input data matrix as a tab delimited text file
short<-"outputFile" # short name that will be used for output files

calibrationParameters<- NA #the column of the "mediansPerDataset.txt" file to use for calibration;
#NA will force centering within the test set & -1 will not do any
#adjustment (when adjustment performed by used)

hasClinical<-FALSE #may include tumor size as second row, with 'T' as the gene name,
#and encoded as binary (0 for size <= 2cm or 1 for size > 2cm)
#set this variable to FALSE if tumor size is not available

collapseMethod<-"mean" # can be mean or iqr (probe with max iqr is selected)
# typically, mean is preferred for long oligo and
# iqr is preferred for short oligo platforms

####
# run the assignment algorithm
####
source(paste(paramDir,"subtypePrediction_functions.R",sep="/"))
source(paste(paramDir,"subtypePrediction_distributed.R",sep="/"))

result <- read.table(paste(directoryPath, "/", dataDir, "/", "outputFile_pam50scores.txt", sep = ""))
result <- result[-1, -1]
###-----
nmiR <- length(l$miRs)
nmR <- length(l$mRs)
standardizedData <- scale(l$d)
temp = matrix(nrow = nmR, ncol = nmiR)
## define experimental settings
ExpInd <- c(rep(0,nrow(standardizedData)))
noOfsample <- nrow(l$d)
ExpInd[which(result[, 6] == "Basal")] <- 1
ExpInd[which(result[, 6] == "Her2")] <- 2
ExpInd[which(result[, 6] == "LumA")] <- 3
ExpInd[which(result[, 6] == "LumB")] <- 4
ExpInd[which(result[, 6] == "Normal")] <- 5

temp_maximinCoefficients = matrix(nrow = nmR, ncol = nmiR)
temp_pvalues = matrix(nrow = nmR, ncol = nmiR)
# Estimate causal effects
for (i in 1:nmR) {
  Y <- standardizedData[,nmiR+i]
  X <- standardizedData[,c(1:nmiR)]

  icp <- hiddenICP(X, Y, ExpInd, alpha = 0.1, mode = "asymptotic", intercept=FALSE)
  for (k in 1:nmiR) {
    temp[i,k] <- icp$betahat[k]
  }
}

```

```

}

colnames(temp) <- l$miRs
row.names(temp) <- l$mRs
## Write the result
# write.csv(temp, file = paste(directoryPath, "ICP_result.csv", sep = ""))
write.csv(temp, file = paste(directoryPath, "hiddenICP_BRCA_pam50_result.csv", sep = ""))

# 3. Validate
#colnames(temp) <- convertmiRs(colnames(temp))

# Get the result
temp <- read.csv(file = paste(directoryPath, "hiddenICP_BRCA_pam50_result.csv", sep = ""),
                 row.names = 1)
temp <- data.matrix(temp)
nmiR <- ncol(temp)

fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(topk in c(500, 1000, 1500, 2000)) {
  top = Extopk(temp, topk)
  write.csv(top, file = paste(directoryPath, "hiddenICP_BRCA_pam50_top_", topk, ".csv", sep = ""))
  topConfirmed = Validation(top, fileName)
  write.csv(topConfirmed[[1]], file = paste(directoryPath, "hiddenICP_BRCA_pam50_topConfirmed_",
                                           topk, ".csv", sep = ""))
}

##=====
# Validate for each miRNA
# colnames(temp) <- convertmiRs(colnames(temp))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(k in c(50, 100, 150, 200)) {
  topConfirmed = NULL
  top = NULL
  for(i in 1:nmiR) {
    miRiTopk = bRank(temp, i, k, FALSE)
    top <- rbind(top, miRiTopk)
    topConfirmedi = Validation(miRiTopk, fileName)
    if(nrow(topConfirmedi[[1]]) != 0) {
      topConfirmed <- rbind(topConfirmed, topConfirmedi[[1]])
    }
  }
  write.csv(top, file = paste(directoryPath, "hiddenICP_BRCA_pam50_top_eachmiR_", k, ".csv",
                              sep = ""), row.names = FALSE)
  write.csv(topConfirmed, file = paste(directoryPath, "hiddenICP_BRCA_pam50_topConfirmed_eachmiR_",
                                       k, ".csv",
                                       sep = ""), row.names = FALSE)
}

##=====

```

pcalg - ida (idaFast) (& pc) - BRCA

```
# Reset the environment
rm(list = ls())

# Load libraries
library(robustbase)
library(pcalg)
library(Rgraphviz)
library(CancerSubtypes)
library(miRBaseConverter)
library(miRLAB)

##=====
# Set environment variables here
directoryPath <- "C:/Users/phavy022/MyDoc/02CausalInferenceMethods/Experiment-Final/InOut/"
# directoryPath <- "R/CausalCompare/"
# directoryPath <- "D:/UniSA/CausalCompare/"
# topk_miR <- 30
# topk_mR <- 500
# topk <- 7500
# k <- 250
topk_miR <- 30
topk_mR <- 1500
topk <- 15000
k <- 500
method <- "min" # value can be "min", "max" or "median"
pcmethod <- "original"
alpha <- 0.05
##=====

# Load functions
source(paste(directoryPath, "CausalCompare.R", sep=""))

# 1. Get data set
#load(paste(directoryPath, "01BRCA/BRCA_matchedData.RData", sep=""))
#matchedData$miRs <- processmiRName(matchedData$miRs)
#matchedData$mRNAs <- processmRNAName(matchedData$mRNAs)
#save(matchedData, file=paste(directoryPath, "01BRCA/BRCA_matchedData2.RData", sep=""))
load(paste(directoryPath, "01BRCA/BRCA_matchedData2.RData", sep=""))
# Get significant miRs and mRs from matchedData
l <- getDataByMAD(matchedData, topk_miR, topk_mR)

# 2. Estimate causal effects
nmiR <- length(l$miRs)
nmR <- length(l$mRs)
standardizedData <- scale(l$d)
suffStat <- list(C = cor(standardizedData), n = nrow(standardizedData))
pc.fit <- pc(suffStat, indepTest=gaussCIttest, p=ncol(standardizedData), alpha = alpha,
             skel.method=pcmethod)

# Run idaFast
temp = matrix(nrow = nmR, ncol = nmiR)
for(i in 1:nmiR) {
```

```

eff.estF <- idaFast(i, c((nmiR+1):(nmiR + nmR)), cov(standardizedData), pc.fit@graph)
for(j in 1:nmR) {
  absValue <- abs(eff.estF[j,])
  if(method == "min" || method == "max") {
    if(method == "min") {
      index <- which(absValue==min(absValue, na.rm = TRUE))
    } else { # "max"
      index <- which(absValue==max(absValue, na.rm = TRUE))
    }
    temp[j,i] <- eff.estF[j,index[1]]
  } else { # "median"
    temp[j,i] <- median(eff.estF[j,])
  }
}
}
colnames(temp) <- l$miRs
row.names(temp) <- l$mRs
## Write the result
write.csv(temp, file = paste(directoryPath, "pc_idaFast_BRCA_result.csv", sep = ""))

# 3. Validate
# Validate the results of the topk targets of miRNAs predicted by the method
# colnames(temp) <- convertmiRs(colnames(temp))

# Get the result
temp <- read.csv(file = paste(directoryPath, "pc_idaFast_BRCA_result.csv", sep = ""), row.names = 1)
temp <- data.matrix(temp)
nmiR <- ncol(temp)

fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(topk in c(500, 1000, 1500, 2000)) {
  top = Extopk(temp, topk)
  write.csv(top, file = paste(directoryPath, "pc_idaFast_BRCA_top_", topk, ".csv", sep = ""))
  topConfirmed = Validation(top, fileName)
  write.csv(topConfirmed[[1]], file = paste(directoryPath, "pc_idaFast_BRCA_topConfirmed_",
                                           topk, ".csv", sep = ""))
}

##=====
# Validate for each miRNA
# colnames(temp) <- convertmiRs(colnames(temp))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(k in c(50, 100, 150, 200)) {
  topConfirmed = NULL
  top = NULL
  for(i in 1:nmiR) {
    miRiTopk = bRank(temp, i, k, FALSE)
    top <- rbind(top, miRiTopk)
    topConfirmedi = Validation(miRiTopk, fileName)
    if(nrow(topConfirmedi[[1]]) != 0) {
      topConfirmed <- rbind(topConfirmed, topConfirmedi[[1]])
    }
  }
}

```

```

write.csv(top, file = paste(directoryPath, "pc_idaFast_BRCA_top_eachmiR_", k, ".csv",
                             sep = ""), row.names = FALSE)
write.csv(topConfirmed, file = paste(directoryPath, "pc_idaFast_BRCA_topConfirmed_eachmiR_",
                                     k, ".csv",
                                     sep = ""), row.names = FALSE)
}
##=====

```

## JointIDA\_direct (beta) - BRCA

```

# Reset the environment
rm(list = ls())

# Load libraries
library(robustbase)
library(pcalg)
library(Rgraphviz)
library(CancerSubtypes)
library(miRBaseConverter)
library(miRLAB)
library(ParallelPC)
library(doParallel)

##=====
# Set environment variables here
# directoryPath <- "C:/Users/phavy022/MyDoc/02CausalInferenceMethods/Experiment-Final/InOut/"
# directoryPath <- "R/CausalCompare/"
directoryPath <- "D:/UniSA/CausalCompare/"
# topk_miR <- 30
# topk_mR <- 500
# topk <- 7500
# k <- 250
topk_miR <- 30
topk_mR <- 1500
topk <- 15000
k <- 500
method <- "min" # value can be "min", "max" or "median"
##=====

# Load functions
source(paste(directoryPath, "CausalCompare.R", sep=""))
source(paste(directoryPath, "jointIDA_direct.R", sep=""))

# 1. Get data set
load(paste(directoryPath, "01BRCA/BRCA_matchedData2.RData", sep=""))
# Get significant miRs and mRs from matchedData
l <- getDatabyMAD(matchedData, topk_miR, topk_mR)

# 2. Estimate causal effects
nmiR <- length(l$miRs)
nmR <- length(l$mRs)
standardizedData <- scale(l$d)

```



```

result_jointIDA <- jointIDA_direct(standardizedData, cause = 1:nmiR,
                                  effect = (nmiR+1):(nmiR+nmR),
                                  method=method, pcmethod = "parallel",
                                  num.cores = 8, mem.efficient = FALSE,
                                  alpha = 0.01, technique = "RRC")

## Write the result
write.csv(result_jointIDA, file = paste(directoryPath, "jointIDA_direct_BRCA_result.csv", sep = ""))

# 3. Validate
# Validate the results of the topk targets of miRNAs predicted
# by the method

# Get the result
result_jointIDA <- read.csv(file = paste(directoryPath, "jointIDA_direct_BRCA_result.csv", sep = ""),
                           row.names = 1)
result_jointIDA <- data.matrix(result_jointIDA)
nmiR <- ncol(result_jointIDA)

fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
# colnames(result_jointIDA) <- convertmiRs(colnames(result_jointIDA))
for(topk in c(500, 1000, 1500, 2000)) {
  top = Extopk(result_jointIDA, topk)
  write.csv(top, file = paste(directoryPath, "jointIDA_direct_BRCA_top_", topk, ".csv", sep = ""))
  topConfirmed = Validation(top, fileName)
  write.csv(topConfirmed[[1]], file = paste(directoryPath, "jointIDA_direct_BRCA_topConfirmed_",
                                           topk, ".csv", sep = ""))
}

##=====
# Validate for each miRNA
# colnames(result_jointIDA) <- convertmiRs(colnames(result_jointIDA))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(k in c(50, 100, 150, 200)) {
  topConfirmed = NULL
  top = NULL
  for(i in 1:nmiR) {
    miRiTopk = bRank(result_jointIDA, i, k, TRUE)
    top <- rbind(top, miRiTopk)
    topConfirmedi = Validation(miRiTopk, fileName)
    if(nrow(topConfirmedi[[1]]) != 0) {
      topConfirmed <- rbind(topConfirmed, topConfirmedi[[1]])
    }
  }
  write.csv(top, file = paste(directoryPath, "jointIDA_direct_BRCA_top_eachmiR_", k, ".csv",
                              sep = ""), row.names = FALSE)
  write.csv(topConfirmed, file = paste(directoryPath, "jointIDA_direct_BRCA_topConfirmed_eachmiR_",
                                       k, ".csv",
                                       sep = ""), row.names = FALSE)
}
##=====

```

## Lasso in miRLAB with BRCA

```
# Reset the environment
rm(list = ls ())

library(miRBaseConverter)
library(miRLAB)

##=====
# Set environment variables here
directoryPath <- "C:/Users/phavy022/MyDoc/02CausalInferenceMethods/Experiment-Final/InOut/"
# directoryPath <- "R/CausalCompare/"
topk_miR <- 30
topk_mR <- 1500
topk <- 15000
k <- 500
##=====

source(paste(directoryPath, "CausalCompare.R", sep=""))

# 1. Get data set
load(paste(directoryPath, "01BRCA/BRCA_matchedData2.RData", sep=""))
# Get significant miRs and mRs from matchedData
l <- getDataByMAD(matchedData, topk_miR, topk_mR)

#l$miRs <- convertmiRs(l$miRs)
#colnames(l$d) <- c(l$miRs, l$mRs)

# 2. Estimate causal effects
nmiR <- length(l$miRs)
nmR <- length(l$mRs)
standardizedData <- scale(l$d)

write.table(standardizedData, file = paste(directoryPath, "Lasso_BRCA_input.csv", sep = ""),
            sep = ",", row.names = FALSE)

las=Lasso(paste(directoryPath, "Lasso_BRCA_input.csv", sep = ""), cause = 1:nmiR,
          effect = (nmiR+1):(nmiR+nmR))

## Write the result
write.csv(las, file = paste(directoryPath, "Lasso_BRCA_result.csv", sep = ""))

# 3. Validate
# colnames(las) <- convertmiRs(colnames(las))

# Get the result
las <- read.csv(file = paste(directoryPath, "Lasso_BRCA_result.csv", sep = ""), row.names = 1)
las <- data.matrix(las)
row.names(las) <- l$mRs
nmiR <- ncol(las)

fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(topk in c(500, 1000, 1500, 2000)) {
```

```

top = Extopk(las, topk)
write.csv(top, file = paste(directoryPath, "Lasso_BRCA_top_", topk, ".csv", sep = ""))
topConfirmed = Validation(top, fileName)
write.csv(topConfirmed[[1]], file = paste(directoryPath, "Lasso_BRCA_topConfirmed_", topk,
                                          ".csv", sep = ""))
}

##=====
# Validate for each miRNA
# colnames(las) <- convertmiRs(colnames(las))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(k in c(50, 100, 150, 200)) {
  topConfirmed = NULL
  top = NULL
  for(i in 1:nmiR) {
    miRiTopk = bRank(las, i, k, FALSE)
    top <- rbind(top, miRiTopk)
    topConfirmedi = Validation(miRiTopk, fileName)
    if(nrow(topConfirmedi[[1]]) != 0) {
      topConfirmed <- rbind(topConfirmed, topConfirmedi[[1]])
    }
  }
  write.csv(top, file = paste(directoryPath, "Lasso_BRCA_top_eachmiR_", k, ".csv",
                              sep = ""), row.names = FALSE)
  write.csv(topConfirmed, file = paste(directoryPath, "Lasso_BRCA_topConfirmed_eachmiR_", k, ".csv",
                                       sep = ""), row.names = FALSE)
}

##=====

```

## Pearson in miRLAB with BRCA

```

# Reset the environment
rm(list = ls ())

library(miRBaseConverter)
library(miRLAB)

##=====
# Set environment variables here
directoryPath <- "C:/Users/phavy022/MyDoc/02CausalInferenceMethods/Experiment-Final/InOut/"
# directoryPath <- "R/CausalCompare/"
topk_miR <- 30
topk_mR <- 1500
topk <- 15000
k <- 500
##=====

source(paste(directoryPath, "CausalCompare.R", sep = ""))

# 1. Get data set
load(paste(directoryPath, "01BRCA/BRCA_matchedData2.RData", sep = ""))

```

```

# Get significant miRs and mRs from matchedData
l <- getDatabyMAD(matchedData, topk_miR, topk_mR)

#l$miRs <- convertmiRs(l$miRs)
#colnames(l$d) <- c(l$miRs, l$mRs)

# 2. Estimate causal effects
nmiR <- length(l$miRs)
nmR <- length(l$mRs)
standardizedData <- scale(l$d)

write.table(standardizedData, file = paste(directoryPath, "Pearson_BRCA_input.csv", sep = ""),
            sep = ",", row.names = FALSE)

results=Pearson(paste(directoryPath, "Pearson_BRCA_input.csv", sep = ""), cause = 1:nmiR,
                effect = (nmiR+1):(nmiR+nmR))

## Write the result
write.csv(results, file = paste(directoryPath, "Pearson_BRCA_result.csv", sep = ""))

# 3. Validate
# colnames(results) <- convertmiRs(colnames(results))

# Get the result
results <- read.csv(file = paste(directoryPath, "Pearson_BRCA_result.csv", sep = ""), row.names = 1)
results <- data.matrix(results)
row.names(results) <- l$mRs
nmiR <- ncol(results)

fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(topk in c(500, 1000, 1500, 2000)) {
  top = Extopk(results, topk)
  write.csv(top, file = paste(directoryPath, "Pearson_BRCA_top_", topk, ".csv", sep = ""))
  topConfirmed = Validation(top, fileName)
  write.csv(topConfirmed[[1]], file = paste(directoryPath, "Pearson_BRCA_topConfirmed_",
                                           topk, ".csv", sep = ""))
}

##=====
# Validate for each miRNA
# colnames(results) <- convertmiRs(colnames(results))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(k in c(50, 100, 150, 200)) {
  topConfirmed = NULL
  top = NULL
  for(i in 1:nmiR) {
    miRiTopk = bRank(results, i, k, FALSE)
    top <- rbind(top, miRiTopk)
    topConfirmedi = Validation(miRiTopk, fileName)
    if(nrow(topConfirmedi[[1]]) != 0) {
      topConfirmed <- rbind(topConfirmed, topConfirmedi[[1]])
    }
  }
}

```

```

write.csv(top, file = paste(directoryPath, "Pearson_BRCA_top_eachmiR_", k, ".csv",
                             sep = ""), row.names = FALSE)
write.csv(topConfirmed, file = paste(directoryPath, "Pearson_BRCA_topConfirmed_eachmiR_", k, ".csv",
                                     sep = ""), row.names = FALSE)
}

##=====

```

## Borda for hiddenICP 1/3, Pearson and Lasso with BRCA

```

# Reset the environment
rm(list = ls ())

library(miRBaseConverter)
library(miRLAB)
library(robustbase)
library(pcalg)
library(Rgraphviz)
library(CancerSubtypes)

##=====
# Set environment variables here
directoryPath <- "C:/Users/phavy022/MyDoc/02CausalInferenceMethods/Experiment-Final/InOut/"
# directoryPath <- "R/CausalCompare/"
##=====

source(paste(directoryPath, "CausalCompare.R", sep=""))

# 1. Get the results hiddenICP, Pearson and Lasso
hid <- read.csv(paste(directoryPath, "hiddenICP_BRCA13_result.csv", sep = ""), row.names = 1)
pea <- read.csv(paste(directoryPath, "Pearson_BRCA_result.csv", sep = ""), row.names = 1)
las <- read.csv(paste(directoryPath, "Lasso_BRCA_result.csv", sep = ""), row.names = 1)

hid <- data.matrix(hid)
pea <- data.matrix(pea)
las <- data.matrix(las)

row.names(pea) <- row.names(hid)
row.names(las) <- row.names(hid)

# 2. Borda
borda=Borda(list(hid, pea, las))

## Write the result
write.csv(borda, file = paste(directoryPath, "Borda_hiddenICP_13_BRCA_result.csv", sep = ""))

# 3. Validate
# colnames(borda) <- convertmiRs(colnames(borda))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(topk in c(500, 1000, 1500, 2000)) {
  top = Extopk(borda, topk)
  write.csv(top, file = paste(directoryPath, "Borda_hiddenICP_13_BRCA_top_", topk, ".csv", sep = ""))
}

```

```

topConfirmed = Validation(top, fileName)
write.csv(topConfirmed[[1]], file = paste(directoryPath, "Borda_hiddenICP_13_BRCA_topConfirmed_",
                                          topk, ".csv", sep = ""))
}

##=====
# Validate for each miRNA
# colnames(borda) <- convertmiRs(colnames(borda))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
nmiR <- ncol(hid)
for(k in c(50, 100, 150, 200)) {
  topConfirmed = NULL
  top = NULL
  for(i in 1:nmiR) {
    miRiTopk = bRank(borda, i, k, FALSE)
    top <- rbind(top, miRiTopk)
    topConfirmedi = Validation(miRiTopk, fileName)
    if(nrow(topConfirmedi[[1]]) != 0) {
      topConfirmed <- rbind(topConfirmed, topConfirmedi[[1]])
    }
  }
  write.csv(top, file = paste(directoryPath, "Borda_hiddenICP_13_BRCA_top_eachmiR_", k, ".csv",
                              sep = "", row.names = FALSE))
  write.csv(topConfirmed, file = paste(directoryPath, "Borda_hiddenICP_13_BRCA_topConfirmed_eachmiR_",
                                       k, ".csv",
                                       sep = ""), row.names = FALSE))
}

##=====

```

## Borda for hiddenICP PAM50, Pearson and Lasso with BRCA

```

# Reset the environment
rm(list = ls ())

library(miRBaseConverter)
library(miRLAB)
library(robustbase)
library(pcalg)
library(Rgraphviz)
library(CancerSubtypes)

##=====
# Set environment variables here
directoryPath <- "C:/Users/phavy022/MyDoc/02CausalInferenceMethods/Experiment-Final/InOut/"
# directoryPath <- "R/CausalCompare/"
##=====

source(paste(directoryPath, "CausalCompare.R", sep = ""))

# 1. Get the results hiddenICP, Pearson and Lasso
hid <- read.csv(paste(directoryPath, "hiddenICP_BRCA_pam50_result.csv", sep = ""), row.names = 1)

```

```

pea <- read.csv(paste(directoryPath, "Pearson_BRCA_result.csv", sep = ""), row.names = 1)
las <- read.csv(paste(directoryPath, "Lasso_BRCA_result.csv", sep = ""), row.names = 1)

hid <- data.matrix(hid)
pea <- data.matrix(pea)
las <- data.matrix(las)

row.names(pea) <- row.names(hid)
row.names(las) <- row.names(hid)

# 2. Borda
borda=Borda(list(hid, pea, las))

## Write the result
write.csv(borda, file = paste(directoryPath, "Borda_hiddenICP_pam50_BRCA_result.csv", sep = ""))

# 3. Validate
# colnames(borda) <- convertmiRs(colnames(borda))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(topk in c(500, 1000, 1500, 2000)) {
  top = Extopk(borda, topk)
  write.csv(top, file = paste(directoryPath, "Borda_hiddenICP_pam50_BRCA_top_", topk, ".csv",
                              sep = ""))
  topConfirmed = Validation(top, fileName)
  write.csv(topConfirmed[[1]], file = paste(directoryPath, "Borda_hiddenICP_pam50_BRCA_topConfirmed_",
                                             topk, ".csv", sep = ""))
}

##=====
# Validate for each miRNA
# colnames(borda) <- convertmiRs(colnames(borda))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
nmiR <- ncol(hid)
for(k in c(50, 100, 150, 200)) {
  topConfirmed = NULL
  top = NULL
  for(i in 1:nmiR) {
    miRiTopk = bRank(borda, i, k, FALSE)
    top <- rbind(top, miRiTopk)
    topConfirmedi = Validation(miRiTopk, fileName)
    if(nrow(topConfirmedi[[1]]) != 0) {
      topConfirmed <- rbind(topConfirmed, topConfirmedi[[1]])
    }
  }
  write.csv(top, file = paste(directoryPath, "Borda_hiddenICP_pam50_BRCA_top_eachmiR_", k, ".csv",
                              sep = ""), row.names = FALSE)
  write.csv(topConfirmed, file = paste(directoryPath,
                                       "Borda_hiddenICP_pam50_BRCA_topConfirmed_eachmiR_",
                                       k, ".csv",
                                       sep = ""), row.names = FALSE)
}

##=====

```

## Borda for idaFast, Pearson and Lasso with BRCA

```
# Reset the environment
rm(list = ls ())

library(miRBaseConverter)
library(miRLAB)
library(robustbase)
library(pcalg)
library(Rgraphviz)
library(CancerSubtypes)

##=====
# Set environment variables here
directoryPath <- "C:/Users/phavy022/MyDoc/02CausalInferenceMethods/Experiment-Final/InOut/"
# directoryPath <- "R/CausalCompare/"
##=====

source(paste(directoryPath, "CausalCompare.R", sep=""))

# 1. Get the results idaFast, Pearson and Lasso
ida <- read.csv(paste(directoryPath, "pc_idaFast_BRCA_result.csv", sep = ""), row.names = 1)
pea <- read.csv(paste(directoryPath, "Pearson_BRCA_result.csv", sep = ""), row.names = 1)
las <- read.csv(paste(directoryPath, "Lasso_BRCA_result.csv", sep = ""), row.names = 1)

ida <- data.matrix(ida)
pea <- data.matrix(pea)
las <- data.matrix(las)

row.names(pea) <- row.names(ida)
row.names(las) <- row.names(ida)

# 2. Borda
borda=Borda(list(ida, pea, las))

## Write the result
write.csv(borda, file = paste(directoryPath, "Borda_idaFast_BRCA_result.csv", sep = ""))

# 3. Validate
# colnames(borda) <- convertmiRs(colnames(borda))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(topk in c(500, 1000, 1500, 2000)) {
  top = Extopk(borda, topk)
  write.csv(top, file = paste(directoryPath, "Borda_idaFast_BRCA_top_", topk, ".csv", sep = ""))
  topConfirmed = Validation(top, fileName)
  write.csv(topConfirmed[[1]], file = paste(directoryPath, "Borda_idaFast_BRCA_topConfirmed_",
                                            topk, ".csv", sep = ""))
}

##=====
# Validate for each miRNA
# colnames(borda) <- convertmiRs(colnames(borda))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
```



```

nmiR <- ncol(ida)
for(k in c(50, 100, 150, 200)) {
  topConfirmed = NULL
  top = NULL
  for(i in 1:nmiR) {
    miRiTopk = bRank(borda, i, k, FALSE)
    top <- rbind(top, miRiTopk)
    topConfirmedi = Validation(miRiTopk, fileName)
    if(nrow(topConfirmedi)[1]) != 0 {
      topConfirmed <- rbind(topConfirmed, topConfirmedi[[1]])
    }
  }
  write.csv(top, file = paste(directoryPath, "Borda_idaFast_BRCA_top_eachmiR_", k, ".csv",
                              sep = ""), row.names = FALSE)
  write.csv(topConfirmed, file = paste(directoryPath, "Borda_idaFast_BRCA_topConfirmed_eachmiR_",
                                       k, ".csv",
                                       sep = ""), row.names = FALSE)
}

##=====

```

## Borda for jointIDA, Pearson and Lasso with BRCA

```

# Reset the environment
rm(list = ls ())

library(miRBaseConverter)
library(miRLAB)
library(robustbase)
library(pcalg)
library(Rgraphviz)
library(CancerSubtypes)

##=====
# Set environment variables here
directoryPath <- "C:/Users/phavy022/MyDoc/02CausalInferenceMethods/Experiment-Final/InOut/"
# directoryPath <- "R/CausalCompare/"
##=====

source(paste(directoryPath, "CausalCompare.R", sep=""))

# 1. Get the results jointIDA, Pearson and Lasso
joi <- read.csv(paste(directoryPath, "jointIDA_direct_BRCA_result.csv", sep = ""), row.names = 1)
pea <- read.csv(paste(directoryPath, "Pearson_BRCA_result.csv", sep = ""), row.names = 1)
las <- read.csv(paste(directoryPath, "Lasso_BRCA_result.csv", sep = ""), row.names = 1)

joi <- data.matrix(joi)
pea <- data.matrix(pea)
las <- data.matrix(las)

row.names(pea) <- row.names(joi)
row.names(las) <- row.names(joi)

```

```

# 2. Borda
borda=Borda(list(joi, pea, las))

## Write the result
write.csv(borda, file = paste(directoryPath, "Borda_jointIDA_BRCA_result.csv", sep = ""))

# 3. Validate
# colnames(borda) <- convertmiRs(colnames(borda))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
for(topk in c(500, 1000, 1500, 2000)) {
  top = Extopk(borda, topk)
  write.csv(top, file = paste(directoryPath, "Borda_jointIDA_BRCA_top_", topk, ".csv", sep = ""))
  topConfirmed = Validation(top, fileName)
  write.csv(topConfirmed[[1]], file = paste(directoryPath, "Borda_jointIDA_BRCA_topConfirmed_",
                                             topk, ".csv", sep = ""))
}

##=====
# Validate for each miRNA
# colnames(borda) <- convertmiRs(colnames(borda))
fileName <- paste(directoryPath, "groundtruth.csv", sep = "")
nmiR <- ncol(joi)
for(k in c(50, 100, 150, 200)) {
  topConfirmed = NULL
  top = NULL
  for(i in 1:nmiR) {
    miRiTopk = bRank(borda, i, k, FALSE)
    top <- rbind(top, miRiTopk)
    topConfirmedi = Validation(miRiTopk, fileName)
    if(nrow(topConfirmedi[[1]]) != 0) {
      topConfirmed <- rbind(topConfirmed, topConfirmedi[[1]])
    }
  }
  write.csv(top, file = paste(directoryPath, "Borda_jointIDA_BRCA_top_eachmiR_", k, ".csv",
                              sep = ""), row.names = FALSE)
  write.csv(topConfirmed, file = paste(directoryPath, "Borda_jointIDA_BRCA_topConfirmed_eachmiR_",
                                       k, ".csv",
                                       sep = ""), row.names = FALSE)
}

##=====

```