script.r

tcarroll

Tue Feb 3 08:41:31 2015

```r
filesToRead <- dir("../ExpressionResults/",pattern = "*\\.txt",full.names=T)
fileRead <- vector("list",length=length(filesToRead))
for(i in 1:length(filesToRead)){
  fileRead[[i]] <- read.delim(filesToRead[i],header=F,sep="\t")
  colnames(fileRead[[i]]) <- c("GeneNames",basename(filesToRead[i]))
}
mergedTable <- NULL
for(i in fileRead){
  if(is.null(mergedTable)){
    mergedTable <- i
  }else{
    mergedTable <- merge(mergedTable,i,by=1,all=T)
  }

  print(nrow(mergedTable))
}
```

```
## [1] 2000
## [1] 2000
## [1] 2000
## [1] 2000
## [1] 2000
## [1] 2000
## [1] 2000
## [1] 2000
## [1] 2000
## [1] 2000
```

```r
mergedTable[1:3,]
```

```
##   GeneNames ExpressionResults_Sample1.txt ExpressionResults_Sample10.txt
## 1    Gene_1                      3.448466                       7.665488
## 2   Gene_10                      5.314180                       7.813501
## 3  Gene_100                      5.591612                       5.186500
##   ExpressionResults_Sample2.txt ExpressionResults_Sample3.txt
## 1                      5.250063                      5.968927
## 2                      5.361170                      5.305980
## 3                      6.840497                      5.197710
##   ExpressionResults_Sample4.txt ExpressionResults_Sample5.txt
## 1                      6.868251                      5.367100
## 2                      6.742855                      5.957786
## 3                      5.922931                      6.813154
##   ExpressionResults_Sample6.txt ExpressionResults_Sample7.txt
## 1                      5.189686                      3.882930
## 2                      6.293098                      7.361497
## 3                      6.228178                      5.831575
```

```
##   ExpressionResults_Sample8.txt ExpressionResults_Sample9.txt
## 1                      5.329258                      6.167451
## 2                      6.649428                      6.213910
## 3                      6.653152                      3.992555
```

```r
Annotation <- read.table("../ExpressionResults/Annotation.ann",sep="\t",h=T)
annotatedExpression <- merge(Annotation,mergedTable,by=1,all.x=F,all.y=T)
annotatedExpression[1:2,]
```

```
##   GeneName   Ensembl    Pathway ExpressionResults_Sample1.txt
## 1   Gene_1  Ens_1001 DNA_Binding                      3.448466
## 2  Gene_10 Ens_10010 DNA_Binding                      5.314180
##   ExpressionResults_Sample10.txt ExpressionResults_Sample2.txt
## 1                       7.665488                      5.250063
## 2                       7.813501                      5.361170
##   ExpressionResults_Sample3.txt ExpressionResults_Sample4.txt
## 1                      5.968927                      6.868251
## 2                      5.305980                      6.742855
##   ExpressionResults_Sample5.txt ExpressionResults_Sample6.txt
## 1                      5.367100                      5.189686
## 2                      5.957786                      6.293098
##   ExpressionResults_Sample7.txt ExpressionResults_Sample8.txt
## 1                      3.882930                      5.329258
## 2                      7.361497                      6.649428
##   ExpressionResults_Sample9.txt
## 1                      6.167451
## 2                      6.213910
```

```r
table(Annotation$Pathway)
```

```
## 
##  DNA_Binding    Glycolysis          TGFb WntSignaling
##        1000           500           300          200
```

```r
summary(Annotation$Pathway)
```

```
##  DNA_Binding    Glycolysis          TGFb WntSignaling          NA's
##        1000           500           300          200          3000
```

```r
cor(annotatedExpression[,grep("ExpressionResults",colnames(annotatedExpression))])
```

```
##                                ExpressionResults_Sample1.txt
## ExpressionResults_Sample1.txt                      1.0000000
## ExpressionResults_Sample10.txt                     0.8620963
## ExpressionResults_Sample2.txt                      0.8662458
## ExpressionResults_Sample3.txt                      0.8681681
## ExpressionResults_Sample4.txt                      0.8616757
## ExpressionResults_Sample5.txt                      0.8643865
## ExpressionResults_Sample6.txt                      0.8712424
## ExpressionResults_Sample7.txt                      0.8625237
## ExpressionResults_Sample8.txt                      0.8618539
```

```
## ExpressionResults_Sample9.txt                        0.8686308
##                               ExpressionResults_Sample10.txt
## ExpressionResults_Sample1.txt                         0.8620963
## ExpressionResults_Sample10.txt                        1.0000000
## ExpressionResults_Sample2.txt                         0.8662233
## ExpressionResults_Sample3.txt                         0.8681699
## ExpressionResults_Sample4.txt                         0.8701285
## ExpressionResults_Sample5.txt                         0.8703278
## ExpressionResults_Sample6.txt                         0.8683173
## ExpressionResults_Sample7.txt                         0.8698833
## ExpressionResults_Sample8.txt                         0.8660971
## ExpressionResults_Sample9.txt                         0.8685996
##                               ExpressionResults_Sample2.txt
## ExpressionResults_Sample1.txt                         0.8662458
## ExpressionResults_Sample10.txt                        0.8662233
## ExpressionResults_Sample2.txt                         1.0000000
## ExpressionResults_Sample3.txt                         0.8724322
## ExpressionResults_Sample4.txt                         0.8707259
## ExpressionResults_Sample5.txt                         0.8715251
## ExpressionResults_Sample6.txt                         0.8685840
## ExpressionResults_Sample7.txt                         0.8697635
## ExpressionResults_Sample8.txt                         0.8697318
## ExpressionResults_Sample9.txt                         0.8757107
##                               ExpressionResults_Sample3.txt
## ExpressionResults_Sample1.txt                         0.8681681
## ExpressionResults_Sample10.txt                        0.8681699
## ExpressionResults_Sample2.txt                         0.8724322
## ExpressionResults_Sample3.txt                         1.0000000
## ExpressionResults_Sample4.txt                         0.8679524
## ExpressionResults_Sample5.txt                         0.8660294
## ExpressionResults_Sample6.txt                         0.8699042
## ExpressionResults_Sample7.txt                         0.8637106
## ExpressionResults_Sample8.txt                         0.8656122
## ExpressionResults_Sample9.txt                         0.8704524
##                               ExpressionResults_Sample4.txt
## ExpressionResults_Sample1.txt                         0.8616757
## ExpressionResults_Sample10.txt                        0.8701285
## ExpressionResults_Sample2.txt                         0.8707259
## ExpressionResults_Sample3.txt                         0.8679524
## ExpressionResults_Sample4.txt                         1.0000000
## ExpressionResults_Sample5.txt                         0.8749856
## ExpressionResults_Sample6.txt                         0.8745045
## ExpressionResults_Sample7.txt                         0.8670132
## ExpressionResults_Sample8.txt                         0.8703234
## ExpressionResults_Sample9.txt                         0.8697739
##                               ExpressionResults_Sample5.txt
## ExpressionResults_Sample1.txt                         0.8643865
## ExpressionResults_Sample10.txt                        0.8703278
## ExpressionResults_Sample2.txt                         0.8715251
## ExpressionResults_Sample3.txt                         0.8660294
## ExpressionResults_Sample4.txt                         0.8749856
## ExpressionResults_Sample5.txt                         1.0000000
## ExpressionResults_Sample6.txt                         0.8645176
## ExpressionResults_Sample7.txt                         0.8707010
```

```
## ExpressionResults_Sample8.txt                            0.8716815
## ExpressionResults_Sample9.txt                            0.8764503
##                              ExpressionResults_Sample6.txt
## ExpressionResults_Sample1.txt                            0.8712424
## ExpressionResults_Sample10.txt                           0.8683173
## ExpressionResults_Sample2.txt                            0.8685840
## ExpressionResults_Sample3.txt                            0.8699042
## ExpressionResults_Sample4.txt                            0.8745045
## ExpressionResults_Sample5.txt                            0.8645176
## ExpressionResults_Sample6.txt                            1.0000000
## ExpressionResults_Sample7.txt                            0.8721725
## ExpressionResults_Sample8.txt                            0.8725429
## ExpressionResults_Sample9.txt                            0.8722234
##                              ExpressionResults_Sample7.txt
## ExpressionResults_Sample1.txt                            0.8625237
## ExpressionResults_Sample10.txt                           0.8698833
## ExpressionResults_Sample2.txt                            0.8697635
## ExpressionResults_Sample3.txt                            0.8637106
## ExpressionResults_Sample4.txt                            0.8670132
## ExpressionResults_Sample5.txt                            0.8707010
## ExpressionResults_Sample6.txt                            0.8721725
## ExpressionResults_Sample7.txt                            1.0000000
## ExpressionResults_Sample8.txt                            0.8584981
## ExpressionResults_Sample9.txt                            0.8675662
##                              ExpressionResults_Sample8.txt
## ExpressionResults_Sample1.txt                            0.8618539
## ExpressionResults_Sample10.txt                           0.8660971
## ExpressionResults_Sample2.txt                            0.8697318
## ExpressionResults_Sample3.txt                            0.8656122
## ExpressionResults_Sample4.txt                            0.8703234
## ExpressionResults_Sample5.txt                            0.8716815
## ExpressionResults_Sample6.txt                            0.8725429
## ExpressionResults_Sample7.txt                            0.8584981
## ExpressionResults_Sample8.txt                            1.0000000
## ExpressionResults_Sample9.txt                            0.8692940
##                              ExpressionResults_Sample9.txt
## ExpressionResults_Sample1.txt                            0.8686308
## ExpressionResults_Sample10.txt                           0.8685996
## ExpressionResults_Sample2.txt                            0.8757107
## ExpressionResults_Sample3.txt                            0.8704524
## ExpressionResults_Sample4.txt                            0.8697739
## ExpressionResults_Sample5.txt                            0.8764503
## ExpressionResults_Sample6.txt                            0.8722234
## ExpressionResults_Sample7.txt                            0.8675662
## ExpressionResults_Sample8.txt                            0.8692940
## ExpressionResults_Sample9.txt                            1.0000000
```

```r
indexGroupOne <- grep("[1-5].txt",colnames(annotatedExpression))
indexGroupTwo <- grep("[6-9,0].txt",colnames(annotatedExpression))
ttestResults <- apply(annotatedExpression,1,function(x) t.test(as.numeric(x[indexGroupOne]),as.numeric(

str(ttestResults[[1]])
```

```
## List of 9
```

```
## $ statistic   : Named num -0.318
##  ..- attr(*, "names")= chr "t"
## $ parameter   : Named num 7.92
##  ..- attr(*, "names")= chr "df"
## $ p.value     : num 0.759
## $ conf.int    : atomic [1:2] -2.2 1.67
##  ..- attr(*, "conf.level")= num 0.95
## $ estimate    : Named num [1:2] 5.38 5.65
##  ..- attr(*, "names")= chr [1:2] "mean of x" "mean of y"
## $ null.value  : Named num 0
##  ..- attr(*, "names")= chr "difference in means"
## $ alternative: chr "two.sided"
## $ method      : chr "Welch Two Sample t-test"
## $ data.name   : chr "as.numeric(x[indexGroupOne]) and as.numeric(x[indexGroupTwo])"
## - attr(*, "class")= chr "htest"
```

```r
testResult <- sapply(ttestResults,function(x) c(log2(x$estimate[2]) - log2(x$estimate[1]), x$statistic,
testResult <- t(testResult)
colnames(testResult) <- c("logFC","tStatistic","pValue")
annotatedResult <- cbind(annotatedExpression[,1:3],testResult)
annotatedResult <- annotatedResult[order(annotatedResult$tStatistic),]
annotatedResult[1:2,]
```

```
##      GeneName    Ensembl     Pathway     logFC tStatistic      pValue
## 1491  Gene_54  Ens_10054 DNA_Binding 0.2631082    -5.3950 0.0007161659
## 1716 Gene_742 Ens_100742  Glycolysis 0.2311226    -4.5044 0.0029789595
```

```r
#write.table(annotatedResult,file="annotatedResults.csv",sep=",",row.names=F,col.names=F)
```