Introduction

```
library(ECIbootstrap)
```

ECI bootstrap test

In the following we will give an easy example on how to use the ECI bootstrap test to perform equivalence tests on molecular data such as differential gene expression data. The package expects raw molecular data (e.g. gene expression values) as input as well as information about the subjects (two groups) and a function to analyse the raw data such as differential gene expression.

Here, we present an expample for differential gene expression. We simulated two diverse gene expression studies of 1000 genes and each of the studies has 20 subjects in the control group as well as the case group. To load the data just use below code:

```
study1 <- readRDS(system.file("extdata", "study1.rds", package = "ECIbootstrap"))
study2 <- readRDS(system.file("extdata", "study2.rds", package = "ECIbootstrap"))
```

Investigating the data shows that each study has two matricies: one matrix for the information about the sampless of the study in form of ID and Type or the sample, and another matrix with gene expression values where the rows represent the genes and the columns represent the samples. It is important that the two matricies follow the same order of samples.

```
# sample information
head(study1$samples)
#>   ID   Type
#> 1  1 normal
#> 2  2 normal
#> 3  3 normal
#> 4  4 normal
#> 5  5 normal
#> 6  6 normal
# gene expression values
study1$geneExpr[1:6,1:6]
#>          1          2        3         4         5          6
#> 1 4.969364  3.0467652 4.056499  5.973114  5.231125  2.2793616
#> 2 1.564175  0.9669992 3.764061  1.988641  3.430212  1.7487344
#> 3 9.657263  4.4865308 4.558670  4.535304 11.096905 10.9294638
#> 4 9.874502  8.2142366 9.243403 12.726420  9.914938 11.0585898
#> 5 5.394700 11.1557491 9.948781  9.464547  6.168237  8.0743529
#> 6 2.483195  2.6080564 2.161167  3.717456  3.469493  0.4189962
```

To perform the boostrap test we provide the sample information, the raw gene expression values and the function to analyze the data to the ECIboostrapTest function. In the presented example the gene expression values are derived from micro array expression data and are modified in a way that they are compatible with the limma R package for differential gene expression. The provided differential gene expression function uses limma to perform differential gene expression and returns the log2 fold change as variable “ES” and the p-value as variable “pval”.

```
#ECIbootstrapTest
results <- ECIbootstrapTest(data1 = study1$geneExpr, data2 = study2$geneExpr,
                            targets1 = study1$samples, targets2 = study2$samples,
                            analysisFunc = diffExpr)
results$EffectSize1[1:6,]
#>           ES         pval        sd
#> 1  0.1622046 6.398255e-01 0.3441517
#> 2  2.9818240 9.758131e-11 0.3500582
#> 3 -3.0494723 1.026247e-02 1.1351322
#> 4  0.7606090 1.015438e-01 0.4544572
#> 5 -0.1205273 7.606522e-01 0.3931208
#> 6 -0.3650246 2.959671e-01 0.3449658
results$EffectSize2[1:6,]
#>           ES         pval        sd
#> 1 -1.6833575 1.056390e-05 0.3370787
#> 2  5.4052484 1.336022e-14 0.4712629
#> 3  0.7863174 4.674081e-01 1.0723254
#> 4  2.3154457 1.866270e-07 0.3726730
#> 5 -1.8148429 1.298495e-04 0.4311536
#> 6 -0.7468948 1.946506e-02 0.3075531
results$ECIboostrap[1:6,]
#>           ECI       CI.LL     CI.UL   p_value   q_value
#> 1 -0.03470561 -0.68052156 0.1533605 0.3336218 0.4397157
#> 2  0.55165346  0.41891678 0.7502263 0.0000000 0.0000000
#> 3 -0.13733072 -0.92287743 0.1749296 0.2524320 0.3901020
#> 4  0.29513709 -0.02477378 0.7803113 0.1119885 0.2717741
#> 5  0.01589555 -0.17475131 0.6629621 0.3704741 0.4622318
#> 6  0.34407700 -0.19190557 0.9158933 0.2603477 0.3944663
```

The function returns a list with the returned values for the raw data analysis (e.g. differential gene expression analysis) for both studies and a matrix with the results for the ECI bootstrap test with ECI values, the Bca confidence interval, p-pvalue, and q-value of each ECI value.