Package 'exnta'

December 13, 2022

Title Conduct the analysis of the features from the high-resolution mass spectrometry.

Version 1.0.0

Description To conduct the analysis of the features from the high-resolution mass spectrometry. It mainly aims to screen and annotate the significant features associated with the health outcomes.

License GPL (>= 3)

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.2

Imports httr,vroom,ggplot2,readxl,gridExtra

NeedsCompilation no

Author bin Wang [aut, cph, cre]

Maintainer bin Wang <binwang@pku.edu.cn>

R topics documented:

FindCovaNta *Find covariates*

Description

Find covariates

Usage

```
FindCovaNta(PID, OutPath = "default", VarsY, VarsC_Prior = "default",
    VarsC_Fixed = NULL, Method = "single.factor", Thr = 0.1)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoNTA
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsC_Prior	chr. Potential covariates needing further statistical test. The default value is all covariate variables listed in the data file.
VarsC_Fixed	chr. Covariate variables fixed in the model by users.
Method	chr. Methods for screening the covariates, including two options, i.e. "single.factor" and "two.stage".
Thr	num. Threshold of the P-value for screening the covariates. It is ranging 0.05-0.25. The defaults value is 0.1.

Value

A list containing the selected covariates.

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
   res1 = LoadNTA(PID = res$PID, UseExample = "example#1")
   res2 = FindCovaNta(PID=res$PID, VarsY = "Y1", VarsC_Prior = "default",
   VarsC_Fixed = "C2", Method = "single.factor", Thr = 0.1)
```

FuncExit *End the module analysis*

Description

End the module analysis

Usage

```
FuncExit(PID)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by any initial functions.

Value

Exit status

Author(s)

Bin Wang (corresponding author)

Examples

```
res = InitNTA()
   res = LoadNTA(PID = res$PID,UseExample = "example#1")
   FuncExit(PID = res$PID)
```

InitNTA *Initialize ExpoNTA module*

Description

Initialize ExpoNTA module analysis. It can generate an R6 class object.

Usage

```
InitNTA()
```

Details

ExpoNontarget module is designed to conduct the analysis of the features from the high-resolution mass spectrometry. It mainly aims to screen and annotate the significant features associated with the health outcomes.

Value

An R6 class object.

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
```

`LoadNTA` *Load data file for ExpoNTA module*

Description

Load data file for ExpoNTA module

Usage

```
LoadNTA(PID, UseExample = "default", DataPath=NULL, VocaPath=NULL)
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by ExpoNTA
`UseExample`	chr. Method of uploading data. If "default",user should upload their own data files, or use "example#1" provided by this module.
`DataPath`	chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".
`VocaPath`	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
   res = LoadNTA(PID = res$PID, UseExample = "example#1")
```

NtaAnno *Annotate the non-targeted features*

Description

Annotate the non-targeted features

Usage

```
NtaAnno(PID, OutPath = "default", VarsY, VarsX  = "default",
   VarsN  = "single.factor", FdrCorrect = F, AdductPos = "all", AdductNeg = "all",Accuracy)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoNTA
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
FdrCorrect	lgl. T (or TRUE) and F (or FALSE). Whether to correct the multiple hypotheses by false positive rate (FDR) method.
AdductPos	chr. Adducts formed in the positive mode using LC-MS. The default is "all", i.e., all the adducts in positive mode will be chosen, including "M+3ACN+2H,M+ACN+H,M+NH4,M+2ACN H,M+ACN+Na,M+H+Na, M+2K-H,M+H+NH4,2M+K,M+K,2M+NH4,2M+Na,M+2Na,M+DMSO+H, M+2H+Na,M+ACN+2H,M+H,2M+H,M+CH3OH+H,M+H+2Na,M+Na,2M+ACN+H,2M+ACN+Na,M
AdductNeg	chr. Adducts formed in the negative mode using LC-MS. The default is "all", i.e., all the adducts in positive mode will be chosen, including "M+FA-H,M+Hac-H,M+Br,3M-H,2M+Hac-H,M+K-2H,2M+FA-H,M-H,M-H2O-H,M+Na-2H, M-2H,M+TFA-H,M+Cl,M-3H,2M-H"
Accuracy	num. Accuracy threshold to match the target compounds. The default is 1 ppm.

Value

A list containing non-target analysis results

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
   res1 = LoadNTA(PID = res$PID, UseExample = "example#1")
   res2 = NtaCros(PID=res$PID, VarsY = "Y1", VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = TRUE, SelMethod = "all",
   StepwizeThr = 0.1, RF_ImpThr = 0.9, IncCova = "F", Family = "gaussian",
   RepMsr = "F", Corstr = "ar1")
   res3 = NtaAnno(PID=res$PID,VarsY = "Y1",VarsX = "default",VarsN = "single.factor",
   FdrCorrect = "F",AdductPos = "M+H",AdductNeg = "M-H",Accuracy = 1)
```

NtaCros *Association analysis*

Description

Association analysis for non-targeted data

Usage

```
NtaCros(PID,OutPath = "default", VarsY, VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = T, SelMethod = "all", StepwizeThr = 0.1,
    RF_ImpThr = 0.9, IncCova = F, Family, RepMsr = F, Corstr= "ar1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoNTA
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
FdrCorrect	lgl. T (or TRUE) and F (or FALSE). Whether to correct the multiple hypotheses by false positive rate (FDR) method.
SelMethod	chr. Methods to select the significant features. Options include "stepwise" (multiple linear regress using stepwise algorithm), "lasso" (multiple linear regress using LASSO regularization algorithm), "random.forest" (random forest), and "all" (combination of the above three method)..
StepwizeThr	num. Threshold of the P value for stepwise regression to screen important variables. It ranges 0.05-0.25 with the default value of 0.1.
RF_ImpThr	num. Threshold of the total importance for the variables to a random forest model. It ranges 0.5-1.0 with the default value of 0.9.

IncCova	lgl. T (or TRUE) and F (or FALSE). Whether to include the covariate selected in the step of "FindCovaNta".
Family	chr. The link function for the regression model according the data type of outcomes, including "gaussian" for continuous variable, "binomial" for binary variable, and "poisson" for counting variable
RepMsr	lgl. T (or TRUE) and F (or FALSE). Whether existing repeated measurement of the subjects.
Corstr	chr. If "RepMsr" = T, the generalized estimating equations (GEE) will be used. For GEE, three correlation structure options are "exchangeable" "ar1" "unstructured".

Value

A list containing non-target analysis association results

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
   res1 = LoadNTA(PID = res$PID, UseExample = "example#1")
   res2 = NtaCros(PID=res$PID, VarsY = "Y1", VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = T, SelMethod = "all",
   StepwizeThr = 0.1, RF_ImpThr = 0.9, IncCova = "F", Family = "gaussian",
   RepMsr = "F", Corstr = "ar1")
```

VizNtaAnno *Visualize annotation results*

Description

Visualize annotation results of non-targeted data

Usage

```
VizNtaAnno(PID, OutPath = "default", VarsY, VarsN ,Accuracy = 1,
    Brightness = "light",Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoNTA
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.

VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
Accuracy	num. Upper limit of accuracy to match the target molecular. The default is 1.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing plots of non-target annotation analysis results

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
   res1 = LoadNTA(PID = res$PID, UseExample = "example#1")
   res2 = NtaCros(PID=res$PID, VarsY = "Y1", VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = T, SelMethod = "all",
   StepwizeThr = 0.1, RF_ImpThr = 0.9, IncCova = "F", Family = "gaussian",
   RepMsr = "F", Corstr = "ar1")
   res3 = NtaAnno(PID=res$PID,VarsY = "Y1",VarsX = "default",VarsN = "single.factor",
   FdrCorrect = "F",AdductPos = "M+H",AdductNeg = "M-H",Accuracy = 1)
   res4 = VizNtaAnno(PID=res$PID, VarsY = "Y1",VarsN = "single.factor",
   Accuracy = 1,Brightness = "light",Palette = "default1")
```

VizNtaCros *Visualize the results of Association analysis*

Description

Visualize the results of Association analysis for non-targeted data

Usage

```
VizNtaCros(PID, OutPath = "default", VarsY,VarsN, Layout = "volcano",
    EffectThr = 0.5, Brightness = "light", Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoNTA
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.

VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
Layout	chr. Visualization layout. Available options include "forest" and "volcano".
EffectThr	num. Threshold of the total importance for the variables to a random forest model. It ranges 0.5-1.0 with the default value of 0.9.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing plots of non-target analysis association results

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitNTA()
   res1 = LoadNTA(PID = res$PID, UseExample = "example#1")
   res2 = NtaCros(PID=res$PID, VarsY = "Y1", VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = T, SelMethod = "all",
   StepwizeThr = 0.1, RF_ImpThr = 0.9, IncCova = "F", Family = "gaussian",
   RepMsr = "F", Corstr = "ar1")
   res3 = VizNtaCros(PID=res$PID, VarsY = "Y1",VarsN = "single.factor",
   Layout = "volcano", EffectThr = 0.5, Brightness = "light",Palette = "default1")
```

Index