Package ‘etma’

April 10, 2016

Type Package
Title Epistasis Test in Meta-Analysis
Version 1.1-1
Date 2016-04-10
Author Chin Lin
Maintainer Chin Lin <xup6fup@gmail.com>
Description Traditional meta-regression based method has been developed for using meta-
analysis data, but it faced the challenge of inconsistent estimates. This package pur-
pose a new statistical method to detect epistasis using incomplete information sum-
mary, and have proven it not only successfully let consistency of evidence, but also in-
crease the power compared with traditional method (Detailed tutorial is shown in website).
License GPL (>= 3)
LazyData TRUE
Depends R (>= 2.10), graphics, stats, utils
Suggests knitr
VignetteBuilder knitr
NeedsCompilation no
Repository CRAN
Date/Publication 2016-04-10 16:59:50

R topics documented:
etma-package ......................................................... 2
data.GST .......................................................... 3
data.PAH .......................................................... 4
data.RAS .......................................................... 5
ETMA ............................................................... 6
print.ggint .......................................................... 9
summary.ggint ...................................................... 10
Index 12
Description

Traditional meta-regression based method has been developed for using meta-analysis data, but it faced the challenge of inconsistent estimates. This packagepurpose a new statistical method to detect epistasis using incomplete information summary, and have proven it not only successfully let consistency of evidence, but also increase the power compared with traditional method (Detailed tutorial is shown in website).

Author(s)

Chin Lin
Maintainer: Chin Lin <xup6fup@gmail.com>

References


See Also

ETMA, data.GST, data.PAH, data.RAS

Examples

#Detailed tutorial is shown in website <http://rpubs.com/chinlin/ETMA>
#The simple toy example (just test this algorithm)
#Note: the computing time in this example is about 3-5 secs

data(data.RAS)
ggint.toy=ETMA(case.ACE.0,case.ACE.1,ctrl.ACE.0,ctrl.ACE.1,
        case.AGT.0,case.AGT.1,ctrl.AGT.0,ctrl.AGT.1,
        data=data.RAS,iterations.step1=100,iterations.step2=300,
        start.seed=1,show.detailed.plot=FALSE,show.final.plot=FALSE)
print(ggint.toy)
summary(ggint.toy)

#The fastest complete example (Note: the computing time in this example is about 15 mins)
#Note: the complete example need about 20,000/200,000 learning time in step 1/2, respectively.
#data(data.PAH)
#ggint2=ETMA(case.CYP1A1.0,case.CYP1A1.1,ctrl.CYP1A1.0,ctrl.CYP1A1.1,
#        case.GSTM1.0,case.GSTM1.1,ctrl.GSTM1.0,ctrl.GSTM1.1,
#        data=data.PAH,start.seed=1,show.detailed.plot=TRUE,show.p.matrix=TRUE)
#
# print(ggint2)
The data of GST's family interaction on cancer.

Description

This data used the data from a meta-analysis included about 500 studies investigating the associations between GSTM1/GSTT1 and cancer, and saved the studies both reported the genotype of GSTM1 and GSTT1. Finally, there are 360 studies (375 populations) included in this data.

Usage

data("data.GST")

Format

A data frame with 375 observations on the following 12 variables.

Study  the first author and published year in included studies.
Ethnicity  the ethnicity of each population.
Country  the location of study.
Cancer  the subtype of cancer in each study.
case.GSTM1.0  the number of GSTM1 functional type in cases.
case.GSTM1.1  the number of GSTM1 functional type in controls.

ctrl.GSTM1.0  the number of GSTM1 null type in cases (risk type).
case.GSTM1.1  the number of GSTM1 null type in controls (risk type).

case.GSTT1.0  the number of GSTT1 functional type in cases.
case.GSTT1.1  the number of GSTT1 functional type in controls.

ctrl.GSTT1.0  the number of GSTT1 null type in cases (risk type).
case.GSTT1.1  the number of GSTT1 null type in controls (risk type).

References

S-Transferase M1, T1 Polymorphisms with Cancer: Evidence from a Meta-Analysis. PLoS ONE
8(11): e78707. doi:10.1371/journal.pone.0078707

See Also

ETMA

Examples

data(data.GST)
head(data.GST)

data.PAH  The data of metabolic pathway of PAH interaction on oral cancer.

Description

This data used the data from a meta-analysis included about 50 studies investigating the associations
between CYP1A1/GSTM1 and oral cancer, and saved the studies both reported the genotype of
GSTM1 and CYP1A1 rs4646903. Finally, there are 13 studies included in this data.

Usage

data("data.PAH")
**data.RAS**

**Format**
A data frame with 13 observations on the following 11 variables.

- **Athour** the first author in included studies.
- **Year** the published year in included studies.
- **Country** the location of study.
- **case.CYP1A1.0** the number of AA genotype (rs4646903) in cases.
- **case.CYP1A1.1** the number of AC/CC genotype (rs4646903) in cases (risk type).
- **ctrl1.CYP1A1.0** the number of AA genotype (rs4646903) in controls.
- **ctrl1.CYP1A1.1** the number of AC/CC genotype (rs4646903) in controls (risk type).
- **case.GSTM1.0** the number of GSTM1 functional type in cases.
- **case.GSTM1.1** the number of GSTM1 null type in cases (risk type).
- **ctrl1.GSTM1.0** the number of GSTM1 functional type in controls.
- **ctrl1.GSTM1.1** the number of GSTM1 null type in controls (risk type).

**References**

**See Also**
- ETMA

**Examples**

```r
data(data.PAH)
head(data.PAH)
```

---

**data.RAS**

The data of RAS interaction on chronic kidney disease.

**Description**
This used the data from a meta-analysis included about 100 studies investigating the associations between ACE insertion/deletion (I/D) and chronic kidney disease, and re-collected the studies included AGT M235T information. Finally, there are 34 studies included in this data.

**Usage**

```r
data("data.RAS")```
**Format**

A data frame with 34 observations on the following 12 variables.

- **Author** the first author in included studies.
- **Year** the published year in included studies.
- **Race** the race of each population.
- **Type** the subtype of chronic kidney disease in each study.
- **case.ACE.0** the number of I allele (rs4340) in cases.
- **case.ACE.1** the number of D allele (rs4340) in cases (risk allele).
- **ctrl1.ACE.0** the number of I allele (rs4340) in controls.
- **ctrl1.ACE.1** the number of D allele (rs4340) in controls (risk allele).
- **case.AGT.0** the number of M allele (rs699) in cases.
- **case.AGT.1** the number of T allele (rs699) in cases (risk allele).
- **ctrl1.AGT.0** the number of M allele (rs699) in controls.
- **ctrl1.AGT.1** the number of T allele (rs699) in controls (risk allele).

**References**


**See Also**

ETMA

**Examples**

```r
data(data.RAS)
head(data.RAS)
```

---

**ETMA**  
*Epistasis Test in Meta-Analysis (ETMA)*

**Description**

This function is a Markov chain Monte Carlo (MCMC) based method, called "Epistasis Test in Meta-Analysis (ETMA)"), using the genotype summary data for estimating a consistent estimate of epistasis in meta-analysis.
Usage

ETMA(case.x1.0, case.x1.1, ctrl.x1.0, ctrl.x1.1,  
    case.x2.0, case.x2.1, ctrl.x2.0, ctrl.x2.1,  
    data = NULL, sig.level = 0.05, max.step.EM = 20,  
    iterations.step1 = 20000, iterations.step2 = 200000, start.seed = NULL,  
    show.detailed.plot = TRUE, show.final.plot = TRUE,  
    show.p.matrix = FALSE, progress.bar = TRUE)

Arguments

case.x1.0  the number of wild type of SNP1 in case group.
case.x1.1  the number of mutation type of SNP1 in case group.
ctrl.x1.0  the number of wild type of SNP1 in control group.
ctrl.x1.1  the number of mutation type of SNP1 in control group.
case.x2.0  the number of wild type of SNP2 in case group.
case.x2.1  the number of mutation type of SNP2 in case group.
ctrl.x2.0  the number of wild type of SNP2 in control group.
ctrl.x2.1  the number of mutation type of SNP2 in control group.
data  an optional data frame, list or environment (or object coercible by as.data.frame to a data frame) containing the variables in the model. If not found in data, the variables are taken from environment(formula), typically the environment from which lm is called.
sig.level  the significance level used to calculate confidence intervals.
max.step.EM  the maximum number of iterations if convergence is too slow.
iterations.step1  the length of chain to obtain the study-level parameters \([p(disease|base), p(SNP1=1), p(SNP2=1)]\) in step 1.
iterations.step2  the length of chain to obtain the global-level parameters \([OR(SNP1), OR(SNP2), OR(interaction)]\) in step 2.
start.seed  the start seed of this algorithm (if you want your results can be reproduced). A NULL value means a random seed in this algorithm.
show.detailed.plot  a logical indicating whether showing the MCMC plot in each step.
show.final.plot  a logical indicating whether showing the MCMC plot in the last step.
show.p.matrix  a logical indicating whether a p.matrix should be printed.
progress.bar  a logical indicating whether a progress bar should be presented.

Value

b  the beta values of each SNP and interaction term (Sequence is SNP1, SNP2, and interaction).
vcov  
the variance covariance matrix of beta value.

LKK  
the log of likelihood value in the last step.

se  
the standard errors of each SNP and interaction term (Sequence is SNP1, SNP2, and interaction).

df  
the degree of freedom in this analysis.

OR  
the odds ratios of each SNP and interaction term (Sequence is SNP1, SNP2, and interaction).

ci.l  
the lower bounds of confidence interval based on a specific significance level (please see the sig.level).

ci.u  
the upper bounds of confidence interval based on a specific significance level (please see the sig.level).

t  
the t value of each each SNP and interaction term (Sequence is SNP1, SNP2, and interaction).

pval  
the p value of each each SNP and interaction term (Sequence is SNP1, SNP2, and interaction).

sig.level  
the significance level for calculating the confidence interval.

p.matrix  
the p matrix in iterations process.

Author(s)

Chin Lin <xup6fup@gmail.com>

References


See Also

data.GST, data.PAH, data.RAS

Examples

#Detailed tutorial is shown in website <http://rpubs.com/chinlin/ETMA>
#The simple toy example (just test this algorithm)
#Note: the computing time in this example is about 3-5 secs

data(data.RAS)
ggint.toy=ETMA(case.ACE.0,case.ACE.1,ctrl.ACE.0,ctrl.ACE.1,  
case.AGT.0,case.AGT.1,ctrl.AGT.0,ctrl.AGT.1,  
data=data.RAS,iterations.step1=100,iterations.step2=300,  
start.seed=1,show.detailed.plot=FALSE,show.final.plot=FALSE)

print(ggint.toy)
summary(ggint.toy)

#Following examples are complete examples.
#They need 20,000/200,000 learning time in step 1/step 2, respectively (default).
Please note they need more than 15 mins, and one of example need about 3 hrs.
The complete learning time is necessary in real data analysis.
Please use default setting as following to analysis your data.

Example 1 (Note: the computing time in this example is about 3 hrs)

```
data(data.GST)
ggint1=ETMA(case.GSTM1.0,case.GSTM1.1,ctrl.GSTM1.0,ctrl.GSTM1.1,
            case.GSTM1.0,case.GSTM1.1,ctrl.GSTM1.0,ctrl.GSTM1.1,
            data=data.GST,start.seed=1,show.detailed.plot=TRUE,show.p.matrix=TRUE)
# print(ggint1)
# summary(ggint1)
```

Example 2 (Note: the computing time in this example is about 15 mins)

```
data(data.PAH)
ggint2=ETMA(case.CYP1A1.0,case.CYP1A1.1,ctrl.CYP1A1.0,ctrl.CYP1A1.1,
            case.GSTM1.0,case.GSTM1.1,ctrl.GSTM1.0,ctrl.GSTM1.1,
            data=data.PAH,start.seed=1,show.detailed.plot=TRUE,show.p.matrix=TRUE)
# print(ggint2)
# summary(ggint2)
```

Example 3 (Note: the computing time in this example is about 15 mins)

```
data(data.RAS)
ggint3=ETMA(case.ACE.0,case.ACE.1,ctrl.ACE.0,ctrl.ACE.1,
            case.AGT.0,case.AGT.1,ctrl.AGT.0,ctrl.AGT.1,
            data=data.RAS,start.seed=1,show.detailed.plot=TRUE,show.p.matrix=TRUE)
# print(ggint3)
# summary(ggint3)
```

---

**print.ggint**  
*Print Method for 'ggint' Objects*

**Description**

Print method for objects of class "ggint".

**Usage**

```
# S3 method for class 'ggint'
print(x, ...)
```

**Arguments**

- `x`: an object of class "ggint".
- `...`: other arguments.
See Also
etma

Examples

#Following examples are complete examples.
#They need 20,000/200,000 learning time in step 1/step 2, respectively (default).
#Please note they need more than 15 mins, and one of example need about 3 hrs.
#The complete learning time is necessary in real data analysis.
#Please use default setting as following to analysis your data.
#
#Example 1 (Note: the computing time in this example is about 3 hrs)
#
#data(data.GST)
#ggint1=ETMA(case.GSTM1.0,case.GSTM1.1,ctrl.GSTM1.0,ctrl.GSTM1.1,
#             case.GSTM1.0,case.GSTM1.1,ctrl.GSTM1.0,ctrl.GSTM1.1,
#             data=data.GST,start.seed=1,show.detailed.plot=TRUE,show.p.matrix=TRUE)
#print(ggint1)
#
#Example 2 (Note: the computing time in this example is about 15 mins)
#
#data(data.PAH)
#ggint2=ETMA(case.CYP1A1.0,case.CYP1A1.1,ctrl.CYP1A1.0,ctrl.CYP1A1.1,
#             case.GSTM1.0,case.GSTM1.1,ctrl.GSTM1.0,ctrl.GSTM1.1,
#             data=data.PAH,start.seed=1,show.detailed.plot=TRUE,show.p.matrix=TRUE)
#print(ggint2)
#
#Example 3 (Note: the computing time in this example is about 15 mins)
#
#data(data.RAS)
#ggint3=ETMA(case.ACE.0,case.ACE.1,ctrl.ACE.0,ctrl.ACE.1,
#             case.AGT.0,case.AGT.1,ctrl.AGT.0,ctrl.AGT.1,
#             data=data.RAS,start.seed=1,show.detailed.plot=TRUE,show.p.matrix=TRUE)
#print(ggint3)
summary.ggint

Arguments

object an object of class "ggint".

other arguments.

See Also

etma

Examples

#Following examples are complete examples.
#They need 20,000/200,000 learning time in step 1/step 2, respectively (default).
#Please note they need more than 15 mins, and one of example need about 3 hrs.
#The complete learning time is necessary in real data analysis.
#Please use default setting as following to analysis your data.
#
#Example 1 (Note: the computing time in this example is about 3 hrs)
#
#data(data.GST)
#ggint1=etma(case.GSTM1.0,case.GSTM1.1,ctrl.GSTM1.0,ctrl.GSTM1.1,
 #  case.GSTM1.0,case.GSTM1.1,ctrl.GSTM1.0,ctrl.GSTM1.1,
 #  data=data.GST,start.seed=1,show.detailed.plot=TRUE,show.p.matrix=TRUE)
#print(ggint1)
#summary(ggint1)
#
#Example 2 (Note: the computing time in this example is about 15 mins)
#
#data(data.PAH)
#ggint2=etma(case.CYP1A1.0,case.CYP1A1.1,ctrl.CYP1A1.0,ctrl.CYP1A1.1,
 #  case.GSTM1.0,case.GSTM1.1,ctrl.GSTM1.0,ctrl.GSTM1.1,
 #  data=data.PAH,start.seed=1,show.detailed.plot=TRUE,show.p.matrix=TRUE)
#echo(ggint2)
#summary(ggint2)
#
#Example 3 (Note: the computing time in this example is about 15 mins)
#
#data(data.RAS)
#ggint3=etma(case.ACE.0,case.ACE.1,ctrl.ACE.0,ctrl.ACE.1,
 #  case.AGT.0,case.AGT.1,ctrl.AGT.0,ctrl.AGT.1,
 #  data=data.RAS,start.seed=1,show.detailed.plot=TRUE,show.p.matrix=TRUE)
#print(ggint3)
#summary(ggint3)
Index

*Topic **datasets**
  data.GST, 3
  data.PAH, 4
  data.RAS, 5

*Topic **package**
  etma-package, 2

data.GST, 2, 3, 8
data.PAH, 2, 4, 8
data.RAS, 2, 5, 8

ETMA, 2, 4–6, 6
etma, 10, 11
etma(etma-package), 2
etma-package, 2

print.ggint, 9
summary.ggint, 10