Package 'GPA'

February 10, 2020

Version 1.1-0
Date 2017-06-30
Title GPA (Genetic analysis incorporating Pleiotropy and Annotation)
Author Dongjun Chung, Emma Kortemeier
Maintainer Dongjun Chung <chungd@musc.edu></chungd@musc.edu>
Depends R (>= 3.0.0), methods, graphics, Rcpp, parallel, ggplot2, ggrepel, plyr, vegan, DT, shiny, shinyBS
Suggests gpaExample
LinkingTo Rcpp
Description This package provides functions for fitting GPA, a statistical framework to prioritize GWAS results by integrating pleiotropy information and annotation data. In addition, it also includes ShinyGPA, an interactive visualization toolkit to investigate pleiotropic architecture. License GPL (>= 2)
<pre>URL http://dongjunchung.github.io/GPA/</pre>
NeedsCompilation yes
R topics documented:
GPA-package 2 assoc 2 aTest 6 fitAll 5 GPA 6 GPA-class 11 pTest 12 shinyGPA 14
Index 10

2 GPA-package

GPA-package

GPA (Genetic analysis incorporating Pleiotropy and Annotation)

Description

This package provides functions for fitting GPA, a statistical approach to prioritizing GWAS results by integrating pleiotropy information and annotation data, along with ShinyGPA, a visualization toolkit to investigate the pleiotropic architecture using GWAS results.

Details

Package: GPA
Type: Package
Version: 1.1.0
Date: 2017-06-30
License: GPL (>= 2)

LazyLoad: yes

This package contains a main class, GPA, which represents GPA model fit.

This package contains four main methods for the GPA framework (Chung et al., 2014), GPA, assoc, pTest, and aTest. GPA method fits the GPA model and assoc method implements association mapping. pTest and aTest methods implement hypothesis testing for pleiotropy and annotation enrichment, respectively.

This package contains two main methods for the ShinyGPA visualization toolkit (Kortemeier et al., 2017), fitAll and shinyGPA. fitAll function generates all the intermediate results needed to run shinyGPA opens the ShinyGPA interface, which takes the results generated from fitAll as input.

Author(s)

Dongjun Chung, Emma Kortemeier

Maintainer: Dongjun Chung <chungd@musc.edu>

References

Chung D*, Yang C*, Li C, Gelernter J, and Zhao H (2014), "GPA: A statistical approach to prioritizing GWAS results by integrating pleiotropy information and annotation data," PLoS Genetics, 10: e1004787. (* joint first authors)

Kortemeier E, Ramos PS, Hunt KJ, Kim HJ, Hardiman G, and Chung D (2017), "ShinyGPA: An interactive and dynamic visualization toolkit for genetic studies."

See Also

GPA, assoc, pTest, aTest, GPA, fitAll, shinyGPA.

GPA-package 3

```
## Not run:
library(GPA)
# simulation setting
nBin <- 20000
pi1 < -0.2
common <- 0.5
betaAlpha <- c( 0.6, 0.6 )
annP <- c(0.2, 0.4, 0.4, 0.4)
seed <- 12345
# simulation setting
nCommon <- round( pi1 * common * nBin )</pre>
nUniq <- round( pi1 * ( 1 - common ) * nBin )</pre>
nBg <- nBin - 2 * nUniq - nCommon
# M * K matrix of GWAS p-value
set.seed( seed )
pvec1 <- c( rbeta( nCommon, betaAlpha[1], 1 ), rbeta( nUniq, betaAlpha[1], 1 ),</pre>
runif( nUniq ), runif( nBg ) )
pvec2 <- c( rbeta( nCommon, betaAlpha[2], 1 ), runif( nUniq ),</pre>
rbeta( nUniq, betaAlpha[2], 1 ), runif( nBg ) )
pmat <- cbind( pvec1, pvec2 )</pre>
# M * D matrix of annotation
ann <- c(
sample( c(1,0), nCommon, replace=TRUE, prob = c( annP[4], 1 - annP[4] ) ),
sample( c(1,0), nUniq, replace=TRUE, prob = c(annP[2], 1 - annP[2])),
sample( c(1,0), nUniq, replace=TRUE, prob = c( annP[3], 1 - annP[3] ) ),
sample( c(1,0), nBg, replace=TRUE, prob = c(annP[1], 1 - annP[1]) )
# GPA without annotation data
fit.GPA.noAnn <- GPA( pmat, NULL )</pre>
cov.GPA.noAnn <- cov( fit.GPA.noAnn )</pre>
# GPA with annotation data
fit.GPA.wAnn <- GPA( pmat, ann )</pre>
cov.GPA.wAnn <- cov( fit.GPA.wAnn )</pre>
# GPA under pleiotropy H0
fit.GPA.pleiotropy.H0 <- GPA( pmat, NULL, pleiotropyH0=TRUE )</pre>
# association mapping
```

4 assoc

```
assoc.GPA.wAnn <- assoc( fit.GPA.wAnn, FDR=0.05, fdrControl="global" )</pre>
# hypothesis testing for pleiotropy
test.pleiotropy <- pTest( fit.GPA.noAnn, fit.GPA.pleiotropy.H0 )</pre>
# hypothesis testing for annotation enrichment
test.annotation <- aTest( fit.GPA.noAnn, fit.GPA.wAnn )</pre>
# simulator function
simulator <- function( risk.ind, nsnp=20000, alpha=0.6 ) {</pre>
  m <- length(risk.ind)</pre>
  p.sig <- rbeta( m, alpha, 1 )</pre>
  pvec <- runif(nsnp)</pre>
  pvec[ risk.ind ] <- p.sig</pre>
  return(pvec)
}
# run simulation
set.seed(12345)
nsnp <- 10000
alpha <- 0.4
pmat <- matrix( NA, nsnp, 5 )</pre>
pmat[,1] <- simulator( c(1:2000), nsnp=nsnp, alpha=alpha )</pre>
pmat[,2] <- simulator( c(501:2500), nsnp=nsnp, alpha=alpha )</pre>
pmat[,3] <- simulator( c(4001:6000), nsnp=nsnp, alpha=alpha )</pre>
pmat[,4] <- simulator( c(4501:7500), nsnp=nsnp, alpha=alpha )</pre>
pmat[,5] \leftarrow simulator(c(8001:10000), nsnp=nsnp, alpha=alpha)
# Fit GPA for all possible pairs of GWAS datasets
out <- fitAll( pmat )</pre>
# Run the ShinyGPA app using the ouput from fitAll()
shinyGPA(out)
## End(Not run)
```

assoc 5

Description

Association mapping.

Usage

```
assoc( object, ... )
## S4 method for signature 'GPA'
assoc( object, FDR=0.05, fdrControl="global", pattern=NULL )
```

Arguments

object GPA model fit. FDR FDR level.

fdrControl Method to control FDR. Possible values are "global" (global FDR control) and

"local" (local FDR control). Default is "global".

pattern Pattern for association mapping. By default (i.e., pattern=NULL), assoc returns

a binary matrix indicating association of SNPs for each phenotypes. If a pattern is specified, a corresponding binary vector is provided. See the details about

how users can specify the pattern.

... Other parameters to be passed through to generic assoc.

Details

assoc uses the direct posterior probability approach of Newton et al. (2004) to control global FDR in association mapping.

Users can specify the pattern using 1 and * in pattern argument, where 1 and * indicate phenotypes of interest and phenotypes that are not of interest, respectively. For example, when there are three phenotypes, pattern="111" means a SNP associated with all of three phenotypes, while pattern="11*" means a SNP associated with the first two phenotypes (i.e., association with the third phenotype is ignored (averaged out)).

Value

If pattern=NULL, returns a binary matrix indicating association of SNPs for each phenotype, where its rows and columns match those of input p-value matrix for function GPA. Otherwise, returns a binary vector indicating association of SNPs for the phenotype combination of interest.

Author(s)

Dongjun Chung

References

Chung D*, Yang C*, Li C, Gelernter J, and Zhao H (2014), "GPA: A statistical approach to prioritizing GWAS results by integrating pleiotropy information and annotation data," PLoS Genetics, 10: e1004787. (* joint first authors)

Newton MA, Noueiry A, Sarkar D, and Ahlquist P (2004), "Detecting differential gene expression with a semiparametric hierarchical mixture method," *Biostatistics*, Vol. 5, pp. 155-176.

6 aTest

See Also

```
GPA, GPA.
```

Examples

```
## Not run:
fit.GPA.wAnn <- GPA( pmat, ann )
cov.GPA.wAnn <- cov( fit.GPA.wAnn )
assoc.GPA.wAnn <- assoc( fit.GPA.wAnn, FDR=0.05, fdrControl="global" )
## End(Not run)</pre>
```

aTest

Hypothesis testing for annotation enrichment

Description

Hypothesis testing for annotation enrichment.

Usage

```
aTest(fitWithoutAnn, fitWithAnn, vDigit=1000)
```

Arguments

fitWithoutAnn GPA model fit without using annotation data. fitWithAnn GPA model fit with using annotation data.

vDigit Number of digits for reporting parameter estimates and standard errors. For

example, setting it to 1000 means printing out values up to three digits below

zero.

Details

aTest implements the hypothesis testing for annotation enrichment. It requires two GPA model fits, one fitted with using annotation data and one fitted without using annotation data, and evaluates annotation enrichment for risk-associated SNPs using the likelihood ratio test.

Value

Returns a list with components:

q q estimates.

statistics Statistics of the test for annotation enrichment.

p-value of the test for annotation enrichment.

Author(s)

Dongjun Chung

fitAll 7

References

Chung D*, Yang C*, Li C, Gelernter J, and Zhao H (2014), "GPA: A statistical approach to prioritizing GWAS results by integrating pleiotropy information and annotation data," PLoS Genetics, 10: e1004787. (* joint first authors)

See Also

```
pTest, GPA, GPA.
```

Examples

```
## Not run:
# GPA without annotation data
fit.GPA.noAnn <- GPA( pmat, NULL )
# GPA with annotation data
fit.GPA.wAnn <- GPA( pmat, ann )
# hypothesis testing for annotation enrichment
test.annotation <- aTest( fit.GPA.noAnn, fit.GPA.wAnn )
## End(Not run)</pre>
```

fitAll

Fit GPA model for all possible pairs of GWAS datasets

Description

Fit GPA model and the GPA model under H0 for all possible pairs of GWAS datasets.

Usage

```
fitAll( pmat,
  maxIter=2000, stopping="relative", epsStopLL=1e-10,
  parallel=FALSE, nCore=8 )
```

Arguments

p-value matrix from GWAS data, where row and column correspond to SNP and

phenotype, respectively.

maxIter Maximum number of EM iteration. Default is 2000.

stopping Stopping rule for EM iteration. Possible values are "absolute" (based on ab-

solute difference in log likelihood), "relative" (based on relative difference in log likelihood), or "aitken" (based on Aitken acceleration-based stopping

rule). Default is "relative".

8 fitAll

epsStopLL Threshold to stop the EM iteration. Default is 1e-100.

parallel Utilize multiple CPUs for parallel computing using "parallel" package? Pos-

sible values are TRUE (utilize multiple CPUs) or FALSE (do not utilize multiple

CPUs). Default is FALSE (do not utilize multiple CPUs).

nCore Number of CPUs when parallel computing is utilized.

Details

fitAll function fits the GPA model and the GPA model under H0 for all possible pairs of GWAS datasets. Its output can be used as an input for the shinyGPA function.

Value

A list with 6 elements, including pmat (original GWAS p-value matrix), combs (a matrix of GWAS pair indices), combList (a matrix of GWAS pair indices), pTestPval (a matrix of pleiotropy test p-values), fitGPA (a list of the GPA fit for each pair), and fitH0 (a list of the GPA fit under H0 for each pair).

Author(s)

Dongjun Chung, Emma Kortemeier

References

Kortemeier E, Ramos PS, Hunt KJ, Kim HJ, Hardiman G, and Chung D (2017), "ShinyGPA: An interactive and dynamic visualization toolkit for genetic studies."

See Also

GPA, pTest, and shinyGPA.

```
## Not run:
# simulator function

simulator <- function( risk.ind, nsnp=20000, alpha=0.6 ) {
    m <- length(risk.ind)
    p.sig <- rbeta( m, alpha, 1 )
    pvec <- runif(nsnp)
    pvec[ risk.ind ] <- p.sig

    return(pvec)
}

# run simulation

set.seed(12345)
nsnp <- 10000</pre>
```

GPA 9

```
alpha <- 0.4
pmat <- matrix( NA, nsnp, 5 )

pmat[,1] <- simulator( c(1:2000), nsnp=nsnp, alpha=alpha )
pmat[,2] <- simulator( c(501:2500), nsnp=nsnp, alpha=alpha )
pmat[,3] <- simulator( c(4001:6000), nsnp=nsnp, alpha=alpha )
pmat[,4] <- simulator( c(4501:7500), nsnp=nsnp, alpha=alpha )
pmat[,5] <- simulator( c(8001:10000), nsnp=nsnp, alpha=alpha )

# Fit GPA for all possible pairs of GWAS datasets

out <- fitAll( pmat )

# Run the ShinyGPA app using the ouput from fitAll()
shinyGPA(out)

## End(Not run)</pre>
```

GPA

Fit GPA model

Description

Fit GPA model.

Usage

```
GPA( gwasPval, annMat=NULL, pleiotropyH0=FALSE, empiricalNull=FALSE, maxIter=2000, stopping="relative", epsStopLL=1e-10, initBetaAlpha=0.1, initPi=0.1, initQ=0.75, lbPi=NA, lbBetaAlpha=0.001, lbQ=0.001, lbPval=1e-30, vDigit=1000, verbose=1)
```

Arguments

gwasPval	p-value matrix from GWAS data, where row and column correspond to SNP and phenotype, respectively.
annMat	Binary matrix from annotation data, where row and column correspond to SNP and annotation, respectively.
pleiotropyH0	Fit GPA under the null hypothesis of pleiotropy test? Possible values are TRUE (under the null hypothesis of pleiotropy test) or FALSE (usual assumption for GPA). Default is FALSE.
empiricalNull	Empirically estimate null distribution for GPA? Possible values are TRUE (empirical null distribution) or FALSE (theoretical null distribution). Default is FALSE.
maxIter	Maximum number of EM iteration. Default is 2000.

10 GPA

stopping Stopping rule for EM iteration. Possible values are "absolute" (based on ab-

solute difference in log likelihood), "relative" (based on relative difference in log likelihood), or "aitken" (based on Aitken acceleration-based stopping

rule). Default is "relative".

epsStopLL Threshold to stop the EM iteration. Default is 1e-100.

initBetaAlpha Initial value for alpha estimate. Default is 0.1.
initPi Initial value for pi estimate. Default is 0.1.
initQ Initial value for q estimate. Default is 0.75.

lbPi Lower bound for pi estimate. If lbPi=NA, lower bound is set to 1 / [number of

SNPs]. Default is NA.

lbBetaAlpha Lower bound for alpha estimate. Default is 0.001.

Lower bound for q estimate. Default is 0.001.

lbPval Lower bound for GWAS p-value. Any GWAS p-values smaller than lbPval are

set to 1bPval. Default is 1e-30.

vDigit Number of digits for reporting parameter estimates. For example, setting it to

1000 means printing out values up to three digits below zero.

verbose Amount of progress report during the fitting procedure. Possible values are 0

(minimal output), 1, 2, or 3 (maximal output). Default is 1.

Details

GPA fits the GPA model. It requires to provide GWAS p-value to gwasPval, while users can also provide annotation data to annMat. It is assumed that number of rows of matrix provided to gwasPval equals to that provided to annMat.

pTest implements the hypothesis testing for pleiotropy. It requires two GPA model fits, one of interest and one under the null hypothesis, and they can be obtained by setting pleiotropyH0=FALSE and pleiotropyH0=TRUE, respectively.

aTest implements the hypothesis testing for annotation enrichment. It requires two GPA model fits, one fitted with using annotation data and one fitted without using annotation data, and they can be obtained by providing annotation data to annMat and not, respectively.

Value

Construct GPA class object.

Author(s)

Dongjun Chung

References

Chung D*, Yang C*, Li C, Gelernter J, and Zhao H (2014), "GPA: A statistical approach to prioritizing GWAS results by integrating pleiotropy information and annotation data," PLoS Genetics, 10: e1004787. (* joint first authors)

GPA-class 11

See Also

```
assoc, pTest, aTest, GPA.
```

Examples

```
## Not run:
# GPA without annotation data

fit.GPA.noAnn <- GPA( pmat, NULL )
cov.GPA.noAnn <- cov( fit.GPA.noAnn )

# GPA with annotation data

fit.GPA.wAnn <- GPA( pmat, ann )
cov.GPA.wAnn <- cov( fit.GPA.wAnn )

# GPA under the null hypothesis of pleiotropy test

fit.GPA.pleiotropy.H0 <- GPA( pmat, NULL, pleiotropyH0=TRUE )

## End(Not run)</pre>
```

GPA-class

Class "GPA"

Description

This class represents GPA model fit.

Details

When users use fdr method, users can specify the pattern using 1 and * in pattern argument, where 1 and * indicate phenotypes of interest and phenotypes that are not of interest, respectively. For example, when there are three phenotypes, pattern="111" means a SNP associated with all of three phenotypes, while pattern="11*" means a SNP associated with the first two phenotypes (i.e., association with the third phenotype is ignored (averaged out)).

Objects from the Class

Objects can be created by calls of the form new("GPA",...).

Slots

```
fit: Object of class "list", representing the fitted GPA model.
setting: Object of class "list", representing the setting for GPA model fitting.
gwasPval: Object of class "matrix", representing the p-value matrix from GWAS data.
annMat: Object of class "matrix", representing the annotation matrix.
```

12 GPA-class

Methods

show signature(object = "GPA"): provide brief summary of the object.

print signature(x = "GPA"): provide the matrix of posterior probability that a SNP belongs to each combination of association status.

fdr signature(object = "GPA", pattern=NULL): provide local FDR. By default (i.e., pattern=NULL), it returns a matrix of local FDR that a SNP is not associated with each phenotype (i.e., marginal FDR), where the order of columns is same as that in input GWAS data. If a pattern is specified, a vector of corresponding local FDR is provided. See the details about how users can specify the pattern.

cov signature(object = "GPA", silent=FALSE, vDigitEst=1000, vDigitSE=1000): provide the covariance matrix for parameter estimates of GPA model. If silent=TRUE, it suppresses the summary output. vDigitEst and vDigitSE control number of digits for reporting parameter estimates and standard errors. For example, setting it to 1000 means printing out values up to three digits below zero.

estimates signature(object = "GPA"): extract parameter estimates from GPA model fit.

se signature(object = "GPA"): extract standard errors for parameter estimates from GPA model fit.

Author(s)

Dongjun Chung

References

Chung D*, Yang C*, Li C, Gelernter J, and Zhao H (2014), "GPA: A statistical approach to prioritizing GWAS results by integrating pleiotropy information and annotation data," PLoS Genetics, 10: e1004787. (* joint first authors)

See Also

```
GPA, pTest, aTest.
```

```
showClass("GPA")
## Not run:
fit.GPA.wAnn <- GPA( pmat, ann )
fit.GPA.wAnn
pp.GPA.wAnn <- print( fit.GPA.wAnn )
fdr.GPA.wAnn <- fdr( fit.GPA.wAnn )
fdr11.GPA.wAnn <- fdr( fit.GPA.wAnn, pattern="11" )
fdr1..GPA.wAnn <- fdr( fit.GPA.wAnn, pattern="1*" )
cov.GPA.wAnn <- cov( fit.GPA.wAnn )
est.GPA.wAnn <- estimates( fit.GPA.wAnn )
se.GPA.wAnn <- se( fit.GPA.wAnn )
## End(Not run)</pre>
```

pTest 13

pTest Hypothesis testing for pleiotropy	
---	--

Description

Hypothesis testing for pleiotropy.

Usage

```
pTest( fit, fitH0, vDigit=1000 )
```

Arguments

fit Fit of the GPA model of interest.

fitH0 GPA model fit under the null hypothesis of pleiotropy test.

vDigit Number of digits for reporting parameter estimates and standard errors. For

example, setting it to 1000 means printing out values up to three digits below

zero.

Details

pTest implements the hypothesis testing for pleiotropy. It requires two GPA model fits, one of interest and one under the null hypothesis (obtained by setting pleiotropyH0=TRUE when running GPA function), and evaluates genetical correlation among multiple phenotypes using the likelihood ratio test.

Value

Returns a list with components:

pi pi estimates.

piSE Standard errors for pi estimates. statistics Statistics of the pleiotropy test. pvalue p-value of the pleiotropy test.

Author(s)

Dongjun Chung

References

Chung D*, Yang C*, Li C, Gelernter J, and Zhao H (2014), "GPA: A statistical approach to prioritizing GWAS results by integrating pleiotropy information and annotation data," PLoS Genetics, 10: e1004787. (* joint first authors)

14 shinyGPA

See Also

```
aTest, GPA, GPA.
```

Examples

```
## Not run:
# GPA without annotation data
fit.GPA.noAnn <- GPA( pmat, NULL )
# GPA under the null hypothesis of pleiotropy test
fit.GPA.pleiotropy.H0 <- GPA( pmat, NULL, pleiotropyH0=TRUE )
# hypothesis testing for pleiotropy
test.pleiotropy <- pTest( fit.GPA.noAnn, fit.GPA.pleiotropy.H0 )
## End(Not run)</pre>
```

shinyGPA

Run ShinyGPA app

Description

Run ShinyGPA app.

Usage

```
shinyGPA( out=NULL )
```

Arguments

out

output of fitAll function.

Details

shinyGPA runs the ShinyGPA app. It takes the output of the fitAll function, which fits the GPA model for all possible pairs of GWAS datasets, as input.

Author(s)

Dongjun Chung, Emma Kortemeier

References

Kortemeier E, Ramos PS, Hunt KJ, Kim HJ, Hardiman G, and Chung D (2017), "ShinyGPA: An interactive and dynamic visualization toolkit for genetic studies."

shinyGPA 15

See Also

fitAll.

```
## Not run:
# simulator function
simulator <- function( risk.ind, nsnp=20000, alpha=0.6 ) {</pre>
  m <- length(risk.ind)</pre>
  p.sig <- rbeta( m, alpha, 1 )</pre>
  pvec <- runif(nsnp)</pre>
  pvec[ risk.ind ] <- p.sig</pre>
  return(pvec)
}
# run simulation
set.seed(12345)
nsnp <- 10000
alpha <- 0.4
pmat <- matrix( NA, nsnp, 5 )</pre>
pmat[,1] <- simulator( c(1:2000), nsnp=nsnp, alpha=alpha )</pre>
pmat[,2] <- simulator( c(501:2500), nsnp=nsnp, alpha=alpha )</pre>
pmat[,3] <- simulator( c(4001:6000), nsnp=nsnp, alpha=alpha )</pre>
pmat[,4] \leftarrow simulator(c(4501:7500), nsnp=nsnp, alpha=alpha)
pmat[,5] \leftarrow simulator(c(8001:10000), nsnp=nsnp, alpha=alpha)
# Fit GPA for all possible pairs of GWAS datasets
out <- fitAll( pmat )</pre>
# Run the ShinyGPA app using the ouput from fitAll()
shinyGPA(out)
## End(Not run)
```

Index

```
* classes
                                                     se, GPA-method (GPA-class), 11
    GPA-class, 11
                                                     shinyGPA, 2, 8, 14
* methods
                                                     show, GPA-method (GPA-class), 11
    assoc, 4
    aTest, 6
    fitAll, 7
    GPA, 9
    pTest, 13
    shinyGPA, 14
* models
    assoc, 4
    aTest, 6
    fitAll, 7
    GPA, 9
    pTest, 13
     shinyGPA, 14
* package
    GPA-package, 2
assoc, 2, 4, 11
assoc, GPA-method (assoc), 4
aTest, 2, 6, 11, 12, 14
cov (GPA-class), 11
cov, GPA-method (GPA-class), 11
estimates (GPA-class), 11
estimates, GPA-method (GPA-class), 11
fdr (GPA-class), 11
fdr, GPA-method (GPA-class), 11
fitAll, 2, 7, 15
GPA, 2, 6-8, 9, 11, 12, 14
GPA-class, 11
{\sf GPA-package}, \textcolor{red}{2}
print, GPA-method (GPA-class), 11
pTest, 2, 7, 8, 11, 12, 13
se (GPA-class), 11
```