

# Package ‘REBET’

December 1, 2020

**Title** The subREgion-based BurdEn Test (REBET)

**Version** 1.8.0

**Date** 2018-08-13

**Description** There is an increasing focus to investigate the association between rare variants and diseases. The REBET package implements the subREgion-based BurdEn Test which is a powerful burden test that simultaneously identifies susceptibility loci and sub-regions.

**Imports** stats, utils

**Depends** ASSET

**Suggests** RUnit, BiocGenerics

**License** GPL-2

**biocViews** Software, VariantAnnotation, SNP

**RoxygenNote** 6.0.1

**git\_url** <https://git.bioconductor.org/packages/REBET>

**git\_branch** RELEASE\_3\_12

**git\_last\_commit** b9689ad

**git\_last\_commit\_date** 2020-10-27

**Date/Publication** 2020-11-30

**Author** Bill Wheeler [cre],  
Bin Zhu [aut],  
Lisa Mirabello [ctb],  
Nilanjan Chatterjee [ctb]

**Maintainer** Bill Wheeler <[wheelerb@imsweb.com](mailto:wheelerb@imsweb.com)>

## R topics documented:

|       |   |
|-------|---|
| data  | 2 |
| REBET | 2 |
| rebet | 3 |

|              |          |
|--------------|----------|
| <b>Index</b> | <b>5</b> |
|--------------|----------|

---

|      |                             |
|------|-----------------------------|
| data | <i>Data for the example</i> |
|------|-----------------------------|

---

### Description

Data for the example.

### Details

The data contains a binary phenotype vector response, a genotype matrix genotypes consisting of 20 rare-variant SNPs, and the sub-region annotation vector subRegions for the [rebet](#) example.

### See Also

[rebet](#)

### Examples

```
data(data, package="REBET")

# Display some of the data
table(response)
dim(genotypes)
subRegions
```

---

|       |                          |
|-------|--------------------------|
| REBET | <i>The REBET package</i> |
|-------|--------------------------|

---

### Description

An R package for the subREgion-based BurdEn Test (REBET).

### Details

In rare-variant association studies, aggregating rare and/or low frequency variants, may increase statistical power for detection of the underlying susceptibility gene or region. However, it is unclear which variants, or class of them, in a gene contribute most to the association. This subregion-based burden test (REBET) simultaneously selects susceptibility genes and identifies important underlying sub-regions. The sub-regions are predefined based on shared common biologic characteristics, such as the protein domain or possible functional impact. Based on a subset-based approach considering local correlations between combinations of test statistics of sub-regions, REBET is able to properly control the type I error rate while adjusting for multiple comparisons in a computationally efficient manner. See the reference for the details of this test. The main function in this package is [rebet](#), which performs the REBET test.

### Author(s)

Bin Zhu <bin.zhu@nih.gov>, Lisa Mirabello and Nilanjan Chatterjee

## References

Zhu, B., Mirabello, L., Chatterjee, N. (2018) A Subregion-based Burden Test for Simultaneous Identification of Susceptibility Loci and Sub-regions within Genetic Epidemiology. In press. <https://doi.org/10.1002/gepi>

---

|       |                                                |
|-------|------------------------------------------------|
| rebet | <i>The subREgion-based BurdEn Test (REBET)</i> |
|-------|------------------------------------------------|

---

## Description

A Subregion-based Burden Test for Simultaneous Identification of Susceptibility Loci and Sub-regions within

## Usage

```
rebet(response, genotypes, subRegions, responseType=NULL,
       covariates=NULL, shape1=1, shape2=1, saveMem=FALSE)
```

## Arguments

|              |                                                                                                                                                                                                |
|--------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| response     | Numerical vector of phenotypes. A binary phenotype must be coded as 0 and 1.                                                                                                                   |
| genotypes    | Matrix of genotypes with each column as a locus.                                                                                                                                               |
| subRegions   | Sub-region annotation vector with length equal to the number of columns of genotypes. In the returned object, these regions will appear as <code>paste("Region_", subRegions, sep="")</code> . |
| responseType | NULL, "continuous" or "binary". If NULL, then "continuous" or "binary" will be chosen based on Y. The default is NULL.                                                                         |
| covariates   | NULL or matrix of covariates. The default is NULL.                                                                                                                                             |
| shape1       | The shape1 parameter in the beta distribution. The default is 1.                                                                                                                               |
| shape2       | The shape2 parameter in the beta distribution. The default is 1.                                                                                                                               |
| saveMem      | TRUE or FALSE to conserve memory (see details). The default is FALSE.                                                                                                                          |

## Details

See the reference for details of this method.

Missing values in any of response, genotypes or covariates will be removed before the analysis. Setting saveMem to TRUE will allow for the analysis of a much larger number of subjects, but will take more time to compute. When saveMem is FALSE, there needs to be enough memory available to hold two or three NxN matrices, where N is the number of subjects.

This function calls the `h.traits` function in the `ASSET` package.

## Value

The object returned from `h.traits` in the `ASSET` package.

## Author(s)

Bin Zhu <bin.zhu@nih.gov>, Lisa Mirabello and Nilanjan Chatterjee

**References**

Zhu, B., Mirabello, L., Chatterjee, N. (2018) A Subregion-based Burden Test for Simultaneous Identification of Susceptibility Loci and Sub-regions within Genetic Epidemiology. In press. <https://doi.org/10.1002/gepi>

**Examples**

```
data(data, package="REBET")  
  
res <- rebet(response, genotypes, subRegions)  
h.summary(res)
```

# Index

\* **SNP, gene, rare variant**

rebet, [3](#)

\* **data**

data, [2](#)

\* **package**

REBET, [2](#)

ASSET, [3](#)

data, [2](#)

genotypes (data), [2](#)

h.traits, [3](#)

REBET, [2](#)

rebet, [2, 3](#)

response (data), [2](#)

subRegions (data), [2](#)