

Package ‘womblsoft’

April 19, 2009

Type Package

Title Wombling Computation

Version 2.0

Date 2007-01-04

Author Ameline Crida

Maintainer Ameline Crida <ameline.crida@ens-lyon.org>

Description Analyses individually geo-referenced multilocus genotypes for the inferences of genetic boundaries between populations. It is based on the wombling method that estimates the systemic function by looking for the local variation of the allele frequencies. The systemic function estimation is based on the local polynomial regression, and a binomial test assess the significance of boundaries. The method applies to codominant or dominant markers and allows for missing data.

License GPL (>= 2)

Repository CRAN

Date/Publication 2007-08-02 16:12:13

R topics documented:

Binomial Tests	2
CandidateBoundaries	2
CandidateBoundariesRk	3
Data	4
Mirror	5
Ranks	6
Wombling	6
WomblingRk	7

Index	9
--------------	----------

Description

These functions take the value of the CandidateBoundaries and CandidateBoundariesRk functions as input, and test to see if the number of candidate boundary elements at each point of the grid is significant with regards to the binomial distribution. The pvalue is chosen by the user. The resulting matrices of the boundaries and their directions are written in text files. And a map is also produced, showing the boundaries in grey.

Usage

```
BinomialTestDominant(data, cbe, pvalue = 0.05, output_bounds = "bounds.txt", output_dir = "direction.txt")
BinomialTestCodominant(data, cbe, pvalue = 0.05, output_bounds = "bounds.txt", output_dir = "direction.txt")
BinomialTestContingency(data, cbe, pvalue = 0.05, output_bounds = "bounds.txt", output_dir = "direction.txt")
```

Arguments

data	the list of the six items provided by the Data and the Mirror functions.
cbe	the candidate boundary elements, results of the functions CandidateBoundaries.
pvalue	pvalue of the binomial test. Default value is 0.05.
output_bounds, output_dir	file names for writing the two result matrices. Default values are "bounds.txt" and "direction.txt" in the current directory.

Value

none

See Also

[Data](#) and [CandidateBoundaries](#) and [CandidateBoundariesRk](#) for the format of the data and the cbe variables.

Description

These functions count for each point of the grid, the number of candidate boundary elements at the given percentile pB at each point of the grid, and computes the overall direction.

Usage

```
CandidateBoundariesCodominant(data, h, pB = 0.3)
CandidateBoundariesDominant(data, h, pB = 0.3)
CandidateBoundariesContingency(data, h, pB = 0.3)
```

Arguments

`data` the list of the six items provided by the Data and the Mirror functions.

`h` bandwidth.

`pB` percentile of points to be in the candidate boundary for each allele.

Value

a list of three components : the first is the matrix of the number of potential bounds at each point, the second component is the sum of the corresponding directions, and the third is the threshold.

See Also

[Data](#) for the format of the data variable.

CandidateBoundariesRk

Number of alleles that contains a potential boundary at any point of the grid.

Description

These functions count for each point of the grid, with a moving bandwidth, the number of candidate boundary elements at the given percentile `pB` at each point of the grid, and computes the overall direction.

Usage

```
CandidateBoundariesCodominantRk(data, ranks, k, pB = 0.3)
CandidateBoundariesDominantRk(data, ranks, k, pB = 0.3)
```

Arguments

`data` the list of the six items provided by the Data and the Mirror functions.

`ranks` the vector of the ranks of distances of individuals to the grid points, result of the function Ranks.

`k` number of individuals in the window.

`pB` percentile of points to be in the candidate boundary for each allele.

Value

a list of three components : the first is the matrix of the number of potential bounds at each point, the second component is the sum of the corresponding directions, and the third is the threshold.

See Also

[Data](#) and [Ranks](#) for the format of the data and the ranks variables.

Data

Data preparation

Description

Read a text file with coordinates and markers in columns and individuals in rows.

Usage

```
DataDominant(input_file, conversion, nb_x, nb_y, output_coords="coord_km.txt")
DataCodominant(input_file, conversion, nb_x, nb_y, output_coords="coord_km.txt")
DataContingency(input_file, conversion, nb_x, nb_y, output_coords="coord_km.txt")
```

Arguments

<code>input_file</code>	Path of the input text file. For dominant or codominant data, each row contains the name of the individual, the two coordinates (either abscissa and ordinates, or longitude and latitude), and the genetic data in succession. For contingency table, each row corresponds to a sampled point, with the name of the point, its coordinates, and the number of individuals for each modality of each variable.
<code>conversion</code>	0 if the coordinates are cartesian, 1 if they are in degree and therefore need to be converted to cartesian.
<code>nb_x, nb_y</code>	number of pixels in width and length of the grid.
<code>output_coords</code>	the name of the file where the kilometer coordinates will be saved in. Default value is "coord_indiv.txt".

Value

a list of six items :

<code>spatial</code>	coordinates of individuals a matrix with one line per individual, and two columns containing abscissa and ordinates of individuals, (x,y).
<code>genetic_encoded</code>	the genetic data, containing one column per locus. If data are dominant, it's the same table as the input file.
<code>grid</code>	a list of the vector of x, and the vector of y.

`cvx_vertices` the vertices of the convex hull of sampling area (same format than individuals coordinates).

`cvx_matrix` a matrix containing a 1 if the corresponding point of the grid is in the convex hull, and a 0 otherwise.

`nb_individual`
the number of individuals in the dataset.

Mirror *Provides mirror effect on the data*

Description

Duplicates individuals that are close to the border of the convex hull (at a distance smaller than the one given by the user), on the position of their symmetric about the border.

Usage

```
MirrorDominant (data, m)  
MirrorCodominant (data, m)  
MirrorContingency (data, m)
```

Arguments

`data` the list of the six items provided by the Data functions.

`m` width of the domain where individuals will be duplicated.

Value

a new data list, with the last number corresponding to the number of distinct original individuals (the new individuals provided by the mirror effect are not counted).

See Also

[Data](#) for the computing of the data list.

Ranks

Ranks computation

Description

Computes the ranks of the distances of individuals to grid points.

Usage

Ranks (data)

Arguments

data the list of the six items provided by the Data and the Mirror functions.

Value

a vector of all the ranks of the distances of individuals to grid points.

See Also

[Data](#) for the format of the data variable.

Wombling

Matrices of the systemic function and the direction of the gradient.

Description

These functions compute the wombling method (i.e. systemic function). The resulting matrix of the systemic function is written in the file systemic.txt, and the directions of the gradient in direction.txt. It also produces a coloured map showing the intensity of the gradient and its main direction.

Usage

```
WomblingDominant (data, h, output_syst="syst.txt", output_dir="dir.txt")
WomblingCodominant (data, h, output_syst="syst.txt", output_dir="dir.txt")
WomblingContingency (data, h, output_syst="syst.txt", output_dir="dir.txt")
```

Arguments

data the list of the six items provided by the Data and the Mirror functions.
h bandwidth : corresponding to the scale we perform the analysis on.
output_syst, output_dir file names of the two matrices. Default values are "syst.txt" and "dir.txt" in the current directory.

Details

at each grid point, the weight assigned to an individual is $\exp(-\frac{d_i^2}{2h^2})$ where d_i is the distance from the grid point to the individual, and h the bandwidth given by the user.

Value

none

See Also

[Data](#) for the format of the data variable.

WomblingRk

Matrices of the systemic function and the direction of the gradient.

Description

These functions compute the wombling method (i.e. systemic function) with a moving bandwidth : the windows contain always the same number of individuals : k . The resulting matrix of the systemic function is written in the file systemic.txt, and the directions of the gradient in direction.txt. It also produces a coloured map showing the intensity of the gradient and its main directions.

Usage

```
WomblingDominantRk (data, ranks, k, output_syst="syst.txt", output_dir="dir.txt")
WomblingCodominantRk (data, ranks, k, output_syst="syst.txt", output_dir="dir.txt")
```

Arguments

data	the list of the six items provided by the Data and the Mirror functions.
ranks	the vector of the ranks of distances of individuals to the grid points, result of the function Ranks.
k	number of individuals in each window.
output_syst, output_dir	file names of the two matrices. Default values are "syst.txt" and "dir.txt" in the current directory.

Details

at each grid point, the weight assigned to an individual is $\exp(-\frac{r_i}{k})$ where r_i is the rank of the individual regarding the distances of individuals to the considered grid point, and k the number of individuals given by the user.

Value

none

See Also

[Data](#) and [Ranks](#) for the format of the data and the ranks variables.

Index

*Topic **manip**

Data, 4
Mirror, 5
Ranks, 5

*Topic **regression**

CandidateBoundaries, 2
CandidateBoundariesRk, 3
Wombling, 6
WomblingRk, 7

*Topic **spatial**

Binomial Tests, 1
CandidateBoundaries, 2
CandidateBoundariesRk, 3
Mirror, 5
Wombling, 6
WomblingRk, 7

Binomial Tests, 1

BinomialTestCodominant (*Binomial Tests*), 1

BinomialTestContingency (*Binomial Tests*), 1

BinomialTestDominant (*Binomial Tests*), 1

CandidateBoundaries, 2, 2

CandidateBoundariesCodominant (*CandidateBoundaries*), 2

CandidateBoundariesCodominantRk (*CandidateBoundariesRk*), 3

CandidateBoundariesContingency (*CandidateBoundaries*), 2

CandidateBoundariesDominant (*CandidateBoundaries*), 2

CandidateBoundariesDominantRk (*CandidateBoundariesRk*), 3

CandidateBoundariesRk, 2, 3

Data, 2, 3, 4, 5–7

DataCodominant (*Data*), 4

DataContingency (*Data*), 4

DataDominant (*Data*), 4

Mirror, 5

MirrorCodominant (*Mirror*), 5

MirrorContingency (*Mirror*), 5

MirrorDominant (*Mirror*), 5

Ranks, 3, 5, 7

Wombling, 6

WomblingCodominant (*Wombling*), 6

WomblingCodominantRk (*WomblingRk*), 7

WomblingContingency (*Wombling*), 6

WomblingDominant (*Wombling*), 6

WomblingDominantRk (*WomblingRk*), 7

WomblingRk, 7