Package 'GenWin'

November 25, 2022

Title Spline Based Window Boundaries for Genomic Analyses

Version 1.0

Date 2022-11-25

Author Timothy M. Beissinger <timbeissinger@gmail.com>

Maintainer Timothy M. Beissinger <timbeissinger@gmail.com>

Description Defines window or bin boundaries for the analysis of genomic data. Boundaries are based on the inflection points of a cubic smoothing spline fitted to the raw data. Along with defining boundaries, a technique to evaluate results obtained from unequally-sized windows is provided. Applications are particularly pertinent for, though not limited to, genome scans for selection based on variability between populations (e.g. using Wright's fixations index, Fst, which measures variability in subpopulations relative to the total population).

Depends R (>= 3.1.1)

Imports pspline, stats, graphics

Encoding UTF-8

License MIT + file LICENSE

LazyData true

RoxygenNote 7.2.2

NeedsCompilation no

R topics documented:

---

chr6   *Maize Fst Data*

Description

Data from 85,388 SNPs on chromosome 6 of a maize, in a selection experiment involving 30 generations of selection for increased number of ears per plant The variables are:

- Chromosome. Chromosome 6, always.
- Position. SNP position on maize reference version 2.
- C0Freq. Frequency of minor allele in cycle 0.
- C30Freq. Frequency of same allele in C30 (not necessarily minor)
- Fst. Fst value based on C0 and C30 frequencies.

Usage

```
data(chr6)
```

Format

A data frame with 85,388 rows and 5 columns.

References

Beissinger et al. (2014) Genetics 196(3):829-840 (http://www.ncbi.nlm.nih.gov/pubmed/24381334)

---

splineAnalyze	*Spline-based window analysis*

Description

Defines window boundaries for analyzing genetic data and evaluates the specified windows

Usage

```
splineAnalyze(
  Y,
  map,
  smoothness = 100,
  s2 = NA,
  mean = NA,
  plotRaw = FALSE,
  plotWindows = FALSE,
  method = 3
)
```

Arguments

Y	A vector of estimates of some parameter, such as Fst, computed at individual markers. One chromosome should be analyzed at a time.
map	A vector of positions for each marker with a corresponding parameter estimate.
smoothness	The level of resolution (in base pairs) for computing the spline and its derivatives

s2	The variance of parameter estimates, to be used for computing Wstats. Default is to compute this automatically, but it may be manually specified, e.g. so that the value across chromosomes may be utilized.
mean	The mean of parameter estimates, to be used for computing Wstats. Default is to comput this automatically, but it may be manually specified, e.g. so that the value across chromosomes may be utilized.
plotRaw	Whether or not to produce a plot of raw data, with the fitted spline
plotWindows	Whether or not to include a plot of Wstat values over the computed windows
method	The method for controlling amount of smoothing: 1, 2, 3, or 4. See documentation of smooth.Pspline for description. Usual choices are either 3 for generalized cross validation or 4 for ordinary cross validation.

Value

rawSpline The fitted spline object

breaks The spline-suggested window breaks

windowData A table of mean parameter estimates and Wstats computed over spline-suggested windows

Examples

```
data(chr6)
sub6 <- chr6[55000:63000,]
chr6Spline <- splineAnalyze(Y=sub6$Fst,map=sub6$Position,smoothness=100,
plotRaw=TRUE,plotWindows=TRUE,method=4)

## Not run:
chr6Spline <- splineAnalyze(Y=chr6$Fst,map=chr6$Position,smoothness=100,
plotRaw=TRUE,plotWindows=TRUE,method=4)

## End(Not run)
```

Index