

Package ‘junctions’

August 13, 2018

Type Package

Title The Breakdown of Genomic Ancestry Blocks in Hybrid Lineages

Version 1.1

Date 2018-08-11

URL <https://github.com/thijsjanzen/junctions>

Description Individual based simulations of hybridizing populations, where the accumulation of junctions is tracked. Furthermore, mathematical equations are provided to verify simulation outcomes. Both simulations and mathematical equations are based on Janzen (2018) <doi:10.1101/058107>.

License GPL (>= 2)

Imports Rcpp (>= 0.12.13)

LinkingTo Rcpp

SystemRequirements C++11

Suggests testit, testthat, knitr

VignetteBuilder knitr

NeedsCompilation yes

Author Thijs Janzen [aut, cre]

Maintainer Thijs Janzen <thijsjanzen@gmail.com>

Repository CRAN

Date/Publication 2018-08-13 13:40:02 UTC

R topics documented:

junctions-package	2
calculate_mat	2
calc_k	3
estimate_time	4
number_of_junctions	5
sim_fin_chrom	6
sim_inf_chrom	7
time_error	8

junctions-package *Extending The Theory of Junctions*

Description

The theory of junctions is extended by this package by including the effect of a finite number of recombination sites along the chromosome. The package provides functions to calculate the estimated number of junctions, depending on the time since the onset of hybridization, population size, number of recombination sites, initial heterozygosity and the number of crossovers per meiosis.

Details

This package provides individual based simulations in order to simulate the accumulation of junctions over time, both for chromosomes with a finite and an infinite number of recombination sites. Furthermore, the package provides mathematical tools to verify the outcomes of the individual based simulations.

Update version 1.1: updated underlying random number generator for picking recombination sites. The previous generator had limited precision, which could generate duplicate recombination sites. This update fixes that

Author(s)

Maintainer: Thijs Janzen <thijsjanzen@gmail.com>

References

Janzen, T. , Nolte, A. W. and Traulsen, A. (2018), The breakdown of genomic ancestry blocks in hybrid lineages given a finite number of recombination sites. *Evolution*, 72: 735-750. doi:10.1111/evo.13436

calculate_mat *Function to calculate the maximum accurate time*

Description

Function that calculates the maximum time after hybridization after which the number of junctions can still be reliably used to estimate the onset of hybridization. This is following equation 15 in Janzen et al. 2018.

Usage

```
calculate_mat(N = Inf, R = Inf, H_0 = 0.5, C = 1)
```

Arguments

N	Population size
R	Number of genetic markers
H_0	Initial frequency of heterozygosity (e.g. in Hardy-Weinberg this equates to 2pq)
C	Mean number of crossovers per meiosis (e.g. size in Morgan of the chromosome)

Value

The maximum accurate time

Author(s)

Thijs Janzen

Examples

```
calculate_mat(N = Inf, R = 1000, H_0 = 0.5, C = 1)
```

calc_k

Calculate the limit of the number of junctions

Description

Calculate the average number of junctions after an infinite number of generations, provided information on the initial heterozygosity, population size and the number of generations.

Usage

```
calc_k(N = Inf, R = Inf, H_0 = 0.5, C = 1)
```

Arguments

N	Population size
R	Number of genetic markers
H_0	Initial frequency of heterozygosity (e.g. in Hardy-Weinberg this equates to 2pq)
C	Mean number of crossovers per meiosis (e.g. size in Morgan of the chromosome)

Value

The number of junctions for at time = infinity

Author(s)

Thijs Janzen

Examples

```
k <- calc_k(N = 100, R = 1000, H_0 = 0.5, C = 1)
```

estimate_time	<i>Estimate the time since the onset of hybridization, using the number of junctions</i>
---------------	--

Description

Estimate the time since the onset of hybridization, following equation 14 in Janzen et al. 2018

Usage

```
estimate_time(J = NA, N = Inf, R = Inf, H_0 = 0.5, C = 1)
```

Arguments

J	The number of junctions
N	Population size
R	Number of genetic markers
H_0	Initial frequency of heterozygosity (e.g. in Hardy-Weinberg this equates to 2pq)
C	Mean number of crossovers per meiosis (e.g. size in Morgan of the chromosome)

Value

The number of generations passed since the onset of hybridization

Author(s)

Thijs Janzen

Examples

```
J <- number_of_junctions(N = 100, R = 1000, H_0 = 0.5, C = 1, t = 200)
estimate_time(J = J, N = 100, R = 1000, H_0 = 0.5, C = 1) #should be 200 again
```

number_of_junctions *Calculate the average number of junctions*

Description

Calculate the average number of junctions after t generations, provided information on the initial heterozygosity, population size and the number of generations.

Usage

```
number_of_junctions(N = Inf, R = Inf, H_0 = 0.5, C = 1, t = 100)
```

Arguments

N	Population size
R	Number of genetic markers
H_0	Initial frequency of heterozygosity (e.g. in Hardy-Weinberg this equates to $2pq$)
C	Mean number of crossovers per meiosis (e.g. size in Morgan of the chromosome)
t	A single time point or a vector of time points for which the number of junctions is estimated.

Value

The number of junctions for at time = t

Author(s)

Thijs Janzen

Examples

```
jt <- number_of_junctions(N = 100, R = 1000, H_0 = 0.5, C = 1, t = 1000)
jt2 <- number_of_junctions(N = 100, R = 1000, H_0 = 0.5, C = 1, t = 0:1000)
```

`sim_fin_chrom`*Individual Based Simulation of the accumulation of junctions*

Description

Individual based simulation of the accumulation of junctions for a chromosome with an finite number of recombination sites.

Usage

```
sim_fin_chrom(pop_size, initial_heterozygosity, total_runtime, morgan, seed, R)
```

Arguments

<code>pop_size</code>	Population Size
<code>initial_heterozygosity</code>	Initial frequency of heterozygosity (e.g. in Hardy-Weinberg this equates to $2pq$)
<code>total_runtime</code>	Maximum time after which the simulation is to be stopped
<code>morgan</code>	Size of the chromosome in Morgan
<code>seed</code>	Seed of the pseudo-random number generator
<code>R</code>	Number of genetic markers

Value

`avgJunctions` vector of the average number of junctions at time = $[0, total_runtime]$

Author(s)

Thijs Janzen

Examples

```
N = 100 #population size
R = 100
H_0 = 0.5 #initial heterozygosity
maxT = 1000 #run time
C = 1 #number of recombinations per meiosis
seed = 42

v <- sim_fin_chrom(N, H_0, maxT, C, seed, R)
```

sim_inf_chrom *Individual Based Simulation of the accumulation of junctions*

Description

Individual based simulation of the accumulation of junctions for a chromosome with an infinite number of recombination sites.

Usage

```
sim_inf_chrom(pop_size, initial_heterozygosity, total_runtime, morgan, markers, seed)
```

Arguments

pop_size	Population Size
initial_heterozygosity	Initial frequency of heterozygosity (e.g. in Hardy-Weinberg this equates to $2pq$)
total_runtime	Maximum time after which the simulation is to be stopped
morgan	Size of the chromosome in Morgan
markers	The number of genetic markers superimposed on the chromosome. If markers is set to -1, no markers are superimposed (faster simulation)
seed	Seed of the pseudo-random number generator

Value

avgJunctions	vector of the average number of junctions at time = $[0, total_runtime]$
detectedJunctions	vector of the detected number of junctions at time = $[0, total_runtime]$

Author(s)

Thijs Janzen

Examples

```
N = 100 #population size
H_0 = 0.5 #initial heterozygosity
maxT = 1000 #run time
C = 1 #number of recombinations per meiosis
numMarkers = 100
seed = 42

v <- sim_inf_chrom(N, H_0, maxT, C, numMarkers, seed)

plot(v$avgJunctions, type = "l", xlab = "Generations",
     ylab = "Number of Junctions", main = "Example Infinite Chromosome")
lines(v$detectedJunctions, col = "blue")
```

```
legend("bottomright", c("Real number", "Number detected"),
      lty = 1, col = c("black", "blue"))
```

time_error

Estimate the error in the time estimate

Description

Calculate the error in the estimate of the onset of hybridisation, following Equations 3 & 4 in the Supplementary information of Janzen et al. 2018.

Usage

```
time_error(J = NA, N = Inf, R = Inf, H_0 = 0.5, C = 1, t = 1, relative = TRUE)
```

Arguments

J	The number of junctions at time t
N	Population Size
R	Number of genetic markers
H_0	Frequency of heterozygosity at t = 0
C	Mean number of crossovers per meiosis (e.g. size in Morgan of the chromosome)
t	Inferred time
relative	Boolean flag, if TRUE: return the relative error, if FALSE: return error in generations

Value

the error in the time estimate

Author(s)

Thijs Janzen

Examples

```
time_error(J = 100, N = Inf, R = 1000, H_0 = 0.5, C = 1)
```


Index

- *Topic **analytic**
 - calc_k, 3
 - calculate_mat, 2
 - estimate_time, 4
 - number_of_junctions, 5
 - time_error, 8
- *Topic **estimate**
 - estimate_time, 4
- *Topic **finite chromosome**
 - sim_fin_chrom, 6
- *Topic **infinite chromosome**
 - sim_inf_chrom, 7
- *Topic **junctions**
 - calc_k, 3
 - junctions-package, 2
 - number_of_junctions, 5
- *Topic **maximum accurate time**
 - calculate_mat, 2
- *Topic **simulation**
 - sim_fin_chrom, 6
 - sim_inf_chrom, 7
- *Topic **time error**
 - time_error, 8
- *Topic **time**
 - estimate_time, 4
 - time_error, 8

calc_k, 3

calculate_mat, 2

estimate_time, 4

junctions (junctions-package), 2

junctions-package, 2

number_of_junctions, 5

sim_fin_chrom, 6

sim_inf_chrom, 7

time_error, 8