

Package ‘glads’

January 1, 2024

Type Package

Title Genomic Landscape of Divergence Simulation

Version 0.1.1.2

Author Claudio S. Quilodran, Kristen Ruegg, Ashley T. Sendell-Price, Sonya Clegg, Tim Coulson, Eric Anderson.

Maintainer Claudio S. Quilodran <claudio.quilodran@unige.ch>

Description The glads package is an individual-based framework for forward demographic, genetic and genomic simulations. The main aim of the R package is therefore to simulate the divergence of populations further in time, elucidating genomic patterns that may be generated by a range of demographic and genetic processes.

License GPL-2

LazyData TRUE

Imports dplyr,
magrittr,
pegas,
progress,
Rcpp (>= 0.12.7),

LinkingTo Rcpp

RoxygenNote 7.2.3

Encoding UTF-8

R topics documented:

glads-package	2
evolve	2
fitness	8
initial.struct	10
phenotype	11
struct2pegas	12
Index	13

glads-package

Genomic Landscape of Divergence Simulation

Description

The package glads is an individual-based framework for forward demographic, genetic and genomic simulations. The main aim of the R package is to simulate the divergence of populations further in time, elucidating genomic patterns that may be generated by a range of demographic and genetic processes.

Details

The list of functions can be displayed with `library(help = glads)`.

The main function of the package glads is "evolve", which allows the forward simulation of the genetic structure of populations under various scenarios of drift and selection. The supported data are allelic frequencies, SNPs or sequences of DNA. User may either simulate a theoretical dataset or directly work with their own empirical observations. In any case, the observation should be arranged in a two-dimensional array that represents a pair of homologous chromosomes. Each element of the array is an integer defining the copy of a given allele at a given locus.

evolve

Evolution of genetic and genomic landscapes

Description

This function simulates the evolution of individual-based populations forward in time.

Usage

```
evolve(
  x,
  time,
  type = c("constant", "dynamic", "additive", "custom"),
  recombination = c("map", "average"),
  recom.rate,
  loci.pos = NULL,
  chromo_mb = NULL,
  init.sex = NULL,
  migration.rate = NULL,
  mutation.rate = NULL,
  param.z = NULL,
  param.w = NULL,
  fun = c(phenotype = NULL, fitness = NULL)
)
```

Arguments

<code>x</code>	List of objects of class "struct" with the initial genetic structure of populations.
<code>time</code>	Number of simulated generations.
<code>type</code>	Type of simulated evolution. Current options are 'constant', 'dynamic', 'additive', and 'custom' (see Details).
<code>recombination</code>	Type of recombination between homologous chromosomes. Current options are 'map' and 'average' (see Details).
<code>recom.rate</code>	A numerical value for the recombination type 'average', or a vector of $nl - 1$ elements, with the recombination rate between neighbouring loci for type 'map' (nl : number of loci) (see Details).
<code>loci.pos</code>	A vector with the position of loci. The default value is NULL, but it is required for the recombination type 'average'.
<code>chromo_mb</code>	A numerical value with the size of the simulated chromosome (in megabase). The default value is NULL, but it is required for the recombination type 'average'.
<code>init.sex</code>	A list of vectors defining the sex of the initial individuals in the populations (1: females and 2: males). The default value is NULL and assigns the sex of the first generation randomly.
<code>migration.rate</code>	A single value setting the migration rate between all populations, or a squared matrix of order equal to the number of populations, with the migration rate between them (migration['from', 'to']). It is ignored for single population simulations. The default value is NULL with no migration between populations.
<code>mutation.rate</code>	A numerical value setting the mutation rate per site. It is currently restricted to biallelic SNPs (genetic structures with values 1 or 2). The default value is NULL.
<code>param.z</code>	A list with the parameter values for the function computing phenotypes. The list of parameters for each population should be included as a list of list (see Example)
<code>param.w</code>	A list with the parameter values for the fitness function. The list of parameters for each population should be included as a list of list (see Example)
<code>fun</code>	A character vector with the names of the custom phenotype and fitness functions. The default value is NULL, but it is required for the 'custom' type of evolution

Details

This function returns a list of populations composed of two-dimensional arrays that represents a pair of homologous chromosomes. Rows represent individuals and columns the different loci. Each element of the array is an integer defining the copy of a given allele at a given locus.

Different types of evolution are available for simulations:

- 'constant' a constant population size over time. There is no selection, equal sex ratio and each breeding pair generates two offspring. This case represent neutral evolution in which variation are due to recombination, mutation and migration. Variations in the population sizes are also possible due to the effect of migration.
- 'dynamic' a dynamic population size over time. This type of evolution introduces to type 'constant' an unequal sex ratio, a variable number of offspring and a density-dependent demographic effect to avoid exponential growth. A list with the parameters for the phenotype function (`param.z=list(sex.ratio)`) and for the fitness function (`param.w=list(mean.fitness, d.d)`) are required:

- sex.ratio: a numerical value defining the sex ratio of populations. This value should be included within a list in 'param.z'.
- mean.fitness: an integer with the mean number of offsprings generated by breeding pair. This value should be included within a list in 'param.w'.
- d.d: numerical value introducing the density-dependent demographic effect to avoid exponential growth. This value should be included within a list in 'param.w'.
The sex of individuals (1: females and 2: males) are generated by a binomial distribution with probability of being a male equal to the 'sex.ratio'. The fitness of individuals (in number of offspring) are obtained from: $Poisson(\lambda) - N * d.d$, in which λ is equal to 'mean.fitness' and N represents the population size.
- 'additive' additive phenotype evolution. This function introduce quantitative phenotypes and convergent or divergent selection as implemented in Quilodr  n et al. (2019). A list with the parameters for the phenotype function (param.z=list(sex.ratio, fitness.pos, bvs, add.loci, e.v)) and for the fitness function (param.w=list(b0,b1,b2,b3, d.v, add.loci)) are required.

The following parameters are needed for the computation of phenotypes (z):

- sex.ratio: A numerical value defining the sex ratio of populations. This value should be included within a list in 'param.z'.
- fitness.pos: A vector with the position of the additive loci participating in the computation of phenotypes. This value should be included within a list in 'param.z'.
- bvs: A matrix with the breeding values of alleles on each loci. The number of rows is equal to the number of additive loci, while the number of columns is equal to the maximum number of alleles in a locus. This object should be included within a list in 'param.z'.
- add.loci: An integer with the total number of additive loci participating in the computation of phenotypes. This object should be included within a list in 'param.z'.
- e.v: A numerical value defining a stochastic environmental variant in the computation of phenotypes. This value should be included within a list in 'param.z'.
The default phenotype function (z) focus on an additive genetic genotype-phenotype map. Therefore, the sum of values of alleles at each locus gives a breeding value (b_v) for each individual at a given locus. The sum of breeding values bvs across loci gives a breeding value for the phenotypes (z), which is computed as follows:

$$z = \sum_{v=1}^{n_a} b_v + \varepsilon_e(0, \sigma_e)$$

Where n_a is equal to 'add.loci' and σ_e is equal to 'e.v'. The environmental contribution ε_e is assumed to be stochastic and normally distributed, with a mean of 0 and standard variation 'e.v'. This function returns a data.frame with rows equal to the number of individuals and two columns ('sex' and 'z')

The default fitness function (ω) computes a Gaussian relationship between z and ω . The following parameters are needed:

- b0: A numerical value defining the maximum number of offspring generated by breeding pair. This value should be included within a list in 'param.w'.
- b1: A numerical value defining the phenotypic optima. In the gaussian relationship between z and ω , the value of 'b1' represent the z value expected to produce the maximum number of offspring. This value should be included within a list in 'param.w'. The difference in phenotypic optima between the populations drives the strength of 'divergent selection'. Populations exposed to equal phenotypic optima are considered to be under 'concordant selection'.
- b2: A numerical value defining the variance of the Gaussian curve. This value should be included within a list in 'param.z'.

- b3: A numerical value defining the intensity of the density-dependence on the fitness of individuals in a population of size N . This value should be included within a list in 'param.w'.
- d.v: A numerical value defining a stochastic demographic variant in the fitness of individuals. This value should be included within a list in 'param.w'.
- add.loci: An integer with the total number of additive loci participating in the computation of phenotypes. This object should be included within a list in 'param.w'.

The default fitness function (ω) has the form:

$$\omega = b_0 \exp^{-\frac{1}{2} \left(\frac{z - b_1 n_a}{b_2 n_a} \right)^2} - b_3 N + \varepsilon_d(0, \sigma_d)$$

Where n_a is equal to 'add.loci', N is the population size and σ_d is equal to 'd.v'. The demographic variant ε_d is assumed to be stochastic and normally distributed, with a mean of 0 and standard variation 'd.v'. This function returns a vector with the fitness value (ω) of individuals.

- 'custom' is a custom computation of phenotypes and fitness functions. These functions will define the type of selection fitting particular case studies. The name of each user defined function should be introduced in the 'fun' parameter as a character vector with two elements e.g. c('phenotype', 'fitness'). Assuming a per-generation time step, the potential number of offspring produced by each individual depends on its phenotype $\omega = f(z)$, which in turn depends on the individual genotype and on the environment $z = g(G, E)$. G is a numeric value determined by an individual's genotype, representing the genetic value of the genotype. In the case of an additive genetic map, the genetic value of a genotype will be a breeding value. E represents the effect of the environment on phenotypic expression, and this enables the effects of plasticity on phenotypic expression to be captured. The list of parameters 'param.z' and 'param.w' include all needed parameters for the custom 'phenotype' and 'fitness' functions. These lists should not include variables. For the phenotype function, the variable genetic structure of individuals is already included as the object 'struct', which should also be the first argument of the custom phenotype function, followed by "...". This means that all parameters used in the custom phenotype function are only included in 'param.z'. The custom phenotype function should return a matrix with two columns, named "z" and "sex". The first column is the resulting phenotype value and the second column represents the assignation of sex to each individual. For the fitness function, the variable individual phenotype value, sex of individuals and the population size are already included in the environment. The function should start with these three elements ("z", "sex", "n"), followed by "...". The user does not need to use all three of these variables. All parameters needed for the custom fitness function should be included in 'param.w'. This function should return a vector 'w' with fitness values for the individuals (see Examples).

The recombination between homologous chromosomes are either of type 'map' or 'average'. The first case needs a vector with the recombination rate (ρ) between neighbour loci of length equal to $nl - 1$ (nl : number of loci). The probability of having a crossover (1) or not (0) is uniformly distributed at a rate defined by the value of ρ between loci (i.e. positions with a probability smaller than ρ recombine). The uniform distribution allows each position with the same values of ρ to have an equal chance of crossover across all iterations. There is no recombination between homologous chromosomes when $\rho = 0$, both loci are completely linked (e.g. within an inversion or situated close to centromeres), while with a value of $\rho = 0.5$, the recombination rate is completely random (i.e. both loci are very distant on the same chromosome or are located on different chromosomes). A value of $\rho < 0.5$ means the loci are physically linked. The second case, when recombination is of type 'average', a numerical value with the average recombination rate per base pair should be supplied, the loci position and the size of the chromosome in megabase are also required. The crossover points are exponentially distributed as a Poisson process (see Example).

Value

A list of objects of class "struct" or array.

References

Quilodrán, C. S., Ruegg, K., Sendell-Price, A. T., Anderson, E., Coulson, T. and Clegg, S. (2020). The many population genetic and demographic routes to islands of genomic divergence. *Methods in Ecology and Evolution* 11(1):6-21.. doi:[10.1111/2041210X.13324](https://doi.org/10.1111/2041210X.13324).

See Also

[initial.struct](#)

Examples

```
## Not run:
## We start with a population of 20 individuals and 10 biallelic SNPs
initial.population.size <- 20
n.loci <- 10
n.alleles.per.locus <- 2

start1 <- initial.struct(initial.population.size,n.loci,n.alleles.per.locus)

#####
# Type of evolution = "constant" #
#####

## We set a recombination map and the number of generations to be simulated
recom.map <- rep(0.5, n.loci-1) #all loci are independent
n.gens <- 10
pop <- evolve(list(start1), n.gens, "constant", "map", recom.map)

## A similar simulation but with recombination type "average".
## We need to specify the position of loci and the size of the chromosome (MB)
loci.pos<- sample(80000: 12000000, 10) #random loci positions
chromo_mb<-12000000 #chromosome size
crossover <- 1/100000000.0 #average recombination rate (1 cM/MB)
pop <- evolve(list(start1), n.gens, "constant", "average", crossover, loci.pos, chromo_mb)

# We include a mutation rate for the simulation of biallelic loci
pop <- evolve(list(start1), n.gens, "constant", "average", crossover,
              loci.pos, chromo_mb, mutation.rate = 0.0001)

# A second population is included to incorporate the effect of migration
start2 <- initial.struct(initial.population.size,n.loci,n.alleles.per.locus)
pop <- evolve(list(start1, start2), n.gens, "constant", "average", crossover,
              loci.pos, chromo_mb, migration.rate = 0.01)

# The sex of individuals may be set for the starting generations.
sex.start1 <- sample(1:2, initial.population.size, replace=T)
sex.start2 <- sample(1:2, initial.population.size, replace=T)
init.sex <- list(sex.start1, sex.start2)
pop <- evolve(list(start1, start2), n.gens, "constant", "average", crossover,
              loci.pos, chromo_mb, init.sex = init.sex)
```

```
#####
# Type of evolution = "dynamic" #
#####

## We set the parameters for the computation of phenotypes and for the
## fitness function of the type 'dynamic'
sex.ratio <- 0.5
mean.fitness <- 3 #mean number of offsprings per breeding pair
d.d <- 0.01 #density-dependent demographic effect

set.seed(1) #setting the seed for reproducible random numbers
pop <- evolve(list(start1), n.gens, "dynamic", "map", recom.map, param.z = list(list(sex.ratio)),
              param.w = list(list(mean.fitness, d.d)))

#####
# Type of evolution = "additive" #
#####

## We set the parameters for the computation of phenotypes and for
## the fitness function of the type 'additive'
sex.ratio <- 0.5
fitness.pos <- 1:n.loci #all simulated loci are additive
add.loci <- n.loci

# next line is breeding value of additive loci
bvs <- t(array( seq(0,1, length = n.alleles.per.locus) ,c(n.alleles.per.locus, add.loci)))

e.v <- 0.01 # stochastic environmental variant
param.z <- list(sex.ratio, fitness.pos, bvs, add.loci, e.v)

b0 <- 6 # maximum number of offspring
b1 <- 1 # phenotypic optima
b2 <- 0.75 # variance of the Gaussian curve
b3 <- 0.01 # intensity of the density-dependence
d.v = 1 # stochastic demographic variant
param.w <- list(b0,b1,b2,b3, d.v, add.loci)

set.seed(1) #setting the seed for reproducible random numbers
pop<-evolve(list(start1), n.gens, "additive", "map", recom.map,
            param.z = list(param.z), param.w = list(param.w))

#####
# Type of evolution = "custom" #
#####

## We set a custom 'phenotype' and 'fitness' function.
## In this example, the custom functions are very similar to the default 'additive' ones.

phenotype2 <- function(struct, ...){
  pop.struct <- struct[ , fitness.pos, ]
  temp <- dim(pop.struct)
  mat <- matrix(1:temp[2],temp[1],temp[2],byrow=TRUE)

  # the next line gives an array that gives the locus index at each position in pop.struct
  loci.n <- array(mat,c(temp[1],temp[2],2))

  new.n <- (pop.struct-1)*add.loci+loci.n
```

```

bvv <- as.vector(bvs) # turn to bvs
outp <- array(bvv[new.n],c(temp[1],temp[2],2))
z <- apply(outp,1,sum)-add.loci+rnorm(temp[1],0,e.v)
sex <- rbinom(nrow(struct), 1, sex.ratio)+1
return(cbind(z = z, sex = sex))
}

fitness2 <- function(z, sex, n, ...){
  ng <- n.loci
  a <- b0
  b <- b1
  c <- b2
  w <- round( a*exp(-((z-b*ng)^2)/(2*(c*ng)^2) ) - b3*n + rnorm(n,0,d.v) , 0)
  w <- ifelse(w<0,0,w)
  return(w)
}

set.seed(1) #setting the seed for reproducible random numbers
pop<-evolve(list(start1), n.gens, "custom", "map", recom.map,
            param.z =list(sex.ratio = sex.ratio, fitness.pos = fitness.pos,
                          bvs = bvs, add.loci = add.loci, e.v = e.v),
            param.w = list(b0 = b0, b1 = b1, b2 = b2, b3 = b3, d.v = d.v, n.loci = add.loci),
            fun=c("phenotype2", "fitness2"))

## End(Not run)

```

fitness

Fitness function

Description

This function computes a fitness value (ω) depending on the phenotype (z) of individuals. The relationship between both variables is assumed to be Gaussian.

Usage

```
fitness(z, N, b0, b1, b2, b3, d.v, add.loci)
```

Arguments

<code>z</code>	A data.frame with the phenotype value ('z') and the sex of each individual in a population.
<code>N</code>	Population size of the population.
<code>b0</code>	A numerical value defining the maximum number of offspring that may be generated by a breeding pair.
<code>b1</code>	A numerical value defining the phenotypic optima of a given population (see Details).
<code>b2</code>	A numerical value defining the variance of the Gaussian curve.
<code>b3</code>	A numerical value defining the intensity of the density-dependence on the fitness of individuals in a population of size 'N'.

d.v	A numerical value defining a stochastic demographic variant in the fitness of individuals.
add.loci	An integer with the total number of additive loci participating in the computation of phenotypes.

Details

This function is used internally in the function `evolve()` of type 'selection' in order to compute the fitness value of individuals.

The value of 'b1' represents the z value expected to produce the maximum number of offspring. The difference in phenotypic optima between the populations drives the strength of 'divergent selection'. Populations exposed to equal phenotypic optima are considered to be under 'concordant selection'.

This fitness function (ω) has the following form:

$$\omega = b_0 \exp^{-\frac{1}{2} \left(\frac{4z - b_1 n_a}{b_2 n_a} \right)^2} - b_3 N + \varepsilon_d(0, \sigma_d)$$

Where n_a is equal to 'add.loci', N is the population size and σ_d is equal to 'd.v'. The demographic variant ε_d is assumed to be stochastic and normally distributed, with a mean of 0 and standard variation 'd.v'.

Value

A vector with the fitness value (ω) for each individual.

References

Quilodrán, C. S., Ruegg, K., Sendell-Price, A. T., Anderson, E., Coulson, T. and Clegg, S. (2020). The multiple population genetic and demographic routes to islands of genomic divergence. *Methods in Ecology and Evolution*. doi:10.1111/2041210X.13324.

See Also

[evolve phenotype](#)

Examples

```
## We first create a random population with 100 individuals and 10 loci
N <- 100 # Population size
nl <- 10 # Number of additive loci
na <- 4 # Number of alleles per locus
G <- initial.struct(N,nl,na)

## Additional parameters are needed for the computation of phenotypes
bvs <- t(array( seq(0,1, length = na) ,c(na, nl)))
sex.ratio <- 0.5
e.v=0.01

## Now we compute the additive phenotype value of individuals
phen <- phenotype(G, bvs, nl, sex.ratio, e.v)

## Additional parameters are needed for the fitness function
b0 <- 6
b1 <- 0.25
b2 <- 0.5
```

```

b3 <- 0.01
d.v = 1

## The fitness values of individuals are computed as follows:
fitness(phen, N, b0, b1, b2, b3, d.v, nl)

```

initial.struct	<i>Initial Structure</i>
----------------	--------------------------

Description

This function generates an object of class "struct" with an initial genetic population structure for simulations.

Usage

```
initial.struct(N, nl, na)
```

Arguments

N	Number of individuals in the initial population
nl	Number of simulated loci
na	Number of alleles at each locus. This parameter set the number of alleles for all loci.

Details

This function returns a three-dimensional array. Rows represent individuals and columns the different loci. Each element of the array is an integer defining the copy of a given allele at a given locus. The third dimension of the array has two layers representing a pair of homologous chromosomes.

Value

An object of class "struct" or an array.

See Also

[evolve](#)

Examples

```

##Initial population size of 10 individuals with 5 biallelic loci
initial.population.size=10
n.loci=5
n.alleles.per.locus=2
initial.struct(initial.population.size,n.loci,n.alleles.per.locus)

```

phenotype	<i>Phenotype function</i>
-----------	---------------------------

Description

A function for the computation of additive phenotypes

Usage

```
phenotype(G, bvs, add.loci, sex.ratio, e.v)
```

Arguments

G	An object of class "struct" with the genetic structure of each individual in the population.
bvs	A matrix with the breeding values of alleles on each loci. The number of rows is equal to the number of additive loci, while the number of columns is equal to the maximum number of alleles in a locus.
add.loci	An integer with the total number of additive loci participating in the computation of phenotypes.
sex.ratio	A numerical value defining the sex ratio of populations
e.v	A numerical value defining a stochastic environmental variant in the computation of phenotypes.

Details

This function is used internally in the function `evolve()` of type 'selection' in order to compute the additive phenotype value of individuals.

This phenotype function (z) focuses on an additive genetic genotype-phenotype map. The sum of values of alleles at each locus gives a breeding value (b_v) for each individual at a given locus. The sum of breeding values bvs across loci gives a breeding value for the phenotypes (z), which is computed as follows:

$$z = \sum_{v=1}^{n_a} b_v + \varepsilon_e(0, \sigma_e)$$

Where n_a is equal to 'add.loci' and σ_e is equal to 'e.v'. The environmental contribution ε_e is assumed to be stochastic and normally distributed, with a mean of 0 and standard variation 'e.v'.

Value

A data.frame with rows equal to the number of individuals and two columns ('sex' and 'z')

References

Quilodrán, C. S., Ruegg, K., Sendell-Price, A. T., Anderson, E., Coulson, T. and Clegg, S. (2020). The multiple population genetic and demographic routes to islands of genomic divergence. *Methods in Ecology and Evolution*. doi:10.1111/2041210X.13324.

See Also

[evolve fitness](#)

Examples

```
## We first create a random population with 100 individuals and 10 loci
N <- 100 # Population size
nl <- 10 # Number of additive loci
na <- 4 # Number of alleles per locus
G <- initial.struct(N,nl,na)

## Additional parameters are needed for the computation of phenotypes
bvs <- t(array( seq(0,1, length = na) ,c(na, nl)))
sex.ratio <- 0.5
e.v=0.01

## Now we compute the additive phenotype value of individuals
phenotype(G, bvs, nl, sex.ratio, e.v)
```

struct2pegas

Conversion of class struct to class loci

Description

This function converts an object of class "struct" to an object of class "loci" that can be used by the package pegas

Usage

```
struct2pegas(x)
```

Arguments

x List of objects of class "struct" with the initial genetic structure of populations.

Examples

```
## We first create a random population with 100 individuals and 10 biallelic loci
N <- 100 # Population size
nl <- 10 # Number of additive loci
na <- 2 # Number of alleles per locus
G <- initial.struct(N,nl,na)

## We convert the object of class 'struct' into 'loci'
struct2pegas(list(G))
```

Index

evolve, [2](#), [9–11](#)

fitness, [8](#), [11](#)

glads (glads-package), [2](#)

glads-package, [2](#)

initial.struct, [6](#), [10](#)

phenotype, [9](#), [11](#)

struct2pegas, [12](#)