MetaPopGen tutorial 2: conservation genetics

Marco Andrello

August 7th, 2016

Contents

Genetic stochasticity can reduce the fitness of individuals by leading to inbreeding depression, loss of favourable alleles and accumulation of deleterious mutations. When the genetic effective size of a population is small, alleles can be lost by genetic drift. In this tutorial, you will create scenarios to study the loss of alleles by genetic drift.

1 Loss of alleles

1.1 Input data

1. Monoecious species with on age-class and one deme; twenty alleles; time of simulation: 50 years.

```
rm(list=ls())
library(MetaPopGen)

n <- 1
z <- 1
l <- 20
m <- l*(l+1) / 2
print(m)
T_max <- 50
```

Note the high number of possible genotypes when the number of alleles is `l = 20`.
2. The survival probability is 1 for all genotypes and constant in time. Remember the order of events in the life-cycle (see Tutorial 1): survival, reproduction, dispersal, recruitment. Setting the survival probability to 1 means that the number of individuals taking part in reproduction is equal to the number of individual at the beginning of the time-step. Since there is only one age-class, all these individuals die at the end of the time step, after they breed.
Given the high number of genotypes (210), we avoid naming the genotype dimension with character strings and use consecutive numbers fom 1 to 210.

```
sigma <- array(NA,dim=c(m,n,z,T_max))
dimnames(sigma) <- list(genotype=c(1:m),deme=c(1:n),
                        age=c(1:z),time=c(1:T_max))
sigma[,,,] <- 1
```

3. Female fecundity is equal to 20 embryo sacs (i.e. female gametes) per capita; male fecundity is equal to 100 pollen grains (i.e. male gametes)

```
phi_F <- array(NA,dim=c(m,n,z,T_max))
dimnames(phi_F) <- list(genotype=c(1:m),deme=c(1:n),
                        age=c(1:z),time=c(1:T_max))
phi_F[,,,] <- 20

phi_M <- array(NA,dim=c(m,n,z,T_max))
dimnames(phi_M) <- list(genotype=c(1:m),deme=c(1:n),
                        age=c(1:z),time=c(1:T_max))
phi_M[,,,] <- 100
```

4. The carrying capacity is 25 individuals; settler survival probability depends on the number of adults and is constant in time.

```
recr.dd <- "adults"
kappa0 <- array(25,dim=c(n,T_max))
```

5. Mutation. Each allele has a 10^{-6} probability of mutating to one of the other 19 alleles, with equal probability among alleles; and a probability of 1 - 10^{-6} of not mutating.

```
mu <- array(1e-6 / 19,dim=c(l,l))
diag(mu) <- 1 - 1e-6
```

6. Since there is only one deme, there is no dispersal to other demes. However, we still have to define a dispersal matrix, of size 1 x 1, to let all newborns settle in the deme.

```
delta <- matrix(1,nrow=1,ncol=1)
```

7. The initial number of individuals is 10 for each genotype. This means that all twenty alleles are present in the population at the beginning of the simulation.

```
N1 <- array(NA,dim=c(m,n,z))
dimnames(N1) <- list(genotype=c(1:m),deme=c(1:n),age=c(1:z))
N1[,1,1] <- rep(10,m)
```

8. Finally, define the control parameters and save the data

```
input.type = "array"
verbose <- F
save.res <- T
save.res.T <- seq(1,T_max)

save(list=ls(),file="Data.Tut2.1.RData")
```

1.2 Simulation

Create a folder for the simulation

```
rm(list=ls())
load("Data.Tut2.1.RData")
name.dir <- paste0(getwd(),"/Simulation.diversity.loss")
dir.create(name.dir)
setwd(name.dir)
```

Peform the simulation with the function `sim.metapopgen.monoecious`:

```
sim.metapopgen.monoecious(input.type=input.type, N1=N1, sigma=sigma,
                          phi_F=phi_F, phi_M=phi_M,
                          mu=mu, delta=delta,
                          recr.dd=recr.dd, kappa0=kappa0, T_max=T_max,
                          save.res=save.res, save.res.T=save.res.T)
```

1.3 Analyze the results

The simulation results are saved in a folder named with the date and time of the simulation. For me, it was the 8th of August, 2016 at 11h 59min 37sec (on a French computer! "août" = "August"). Replace the actual name of your folder in the code below:

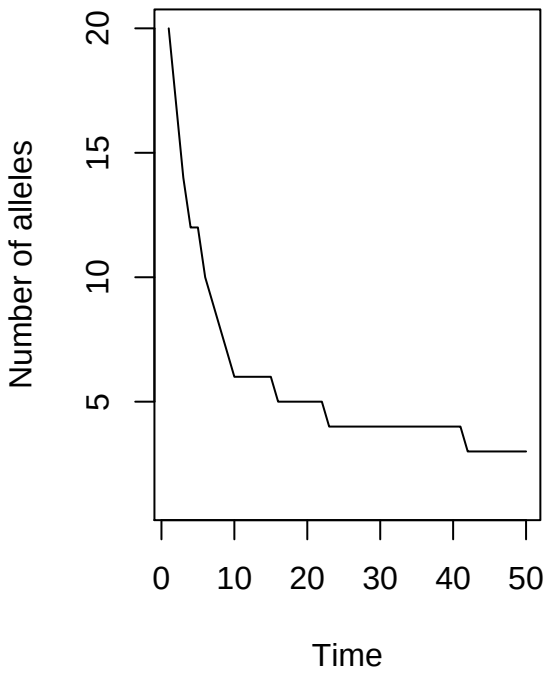
```
name.dir <- "2016-août-08-11.59.37"
```

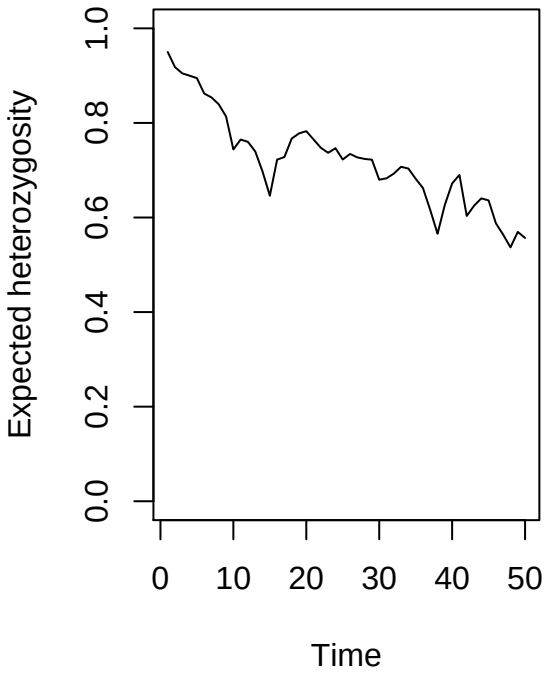
Now we calculate the allele frequency, p_i, for all alleles at each time step. We also calculate two measures of genetic diversity: the number of alleles N_A (i.e. we count all alleles with a frequency different from zero) and the expected heterozygosity $H_e = 1 - \sum p_i^2$. We write a loop on the 50 result files and use the function `freq.all` to calculate the allele frequencies

```
p <- array(NA,dim=c(l,T_max),dimnames=list(allele=c(1:l),time=c(1:T_max)))
Na <- vector()
He <- vector()
for (t in 1 : T_max){
  load(paste0(name.dir,"/N",t,".RData"))
  genotype.vector <- as.vector(N)
  p[,t] <- freq.all(genotype.vector)
  Na[t] <- length(which(p[,t]>0))
  He[t] <- 1 - sum(p[,t]^2)
}
```

We can see that the number of alleles N_A decreases over time. The heterozygosity H_e decreases too, but more slowly than N_A.

```
par(mfrow=c(1,2))
plot(Na, type="l", xlab="Time", ylab="Number of alleles", ylim=c(1,20))
plot(He, type="l", xlab="Time", ylab="Expected heterozygosity", ylim=c(0,1))
```





1.4 Exercises

1.4.1

To verify that the patterns of reduction in genetic diversity observed above are not the result of chance, replicate the simulation twenty times; plot the results of the twenty simulations over a graph; calculate the average and the variation of N_A and H_e at time $t = 50$ over the twenty replicates.

1.4.2

Study the effect of demographic parameters on the maintainance of genetic diversity. Define a population with the following characteristics:
1. Monoecious species with three age-classes and four demes; twenty alleles; time of simulation: 50 years.
2. The survival probability is 0.5 for all age-classes, genotypes and demes;
3. Female fecundity is equal to 30 embryo sacs (i.e. female gametes) per capita; male fecundity is equal to 1000 pollen grains (i.e. male gametes) per capita for all age-classes
4. The carrying capacity of local demes is 25 adults (settler survival probability depends on the total number of individuals in the deme) and is constant in time.
5. Mutation rate: 10^{-6}
6. Dispersal probabilities is 0.1 to other demes (i.e. with 4 demes: 0.1 / 3 to other demes and a 0.9 probability to remain in the same deme) .
7. The initial number of individuals is about 10 for each genotype. Perform the simulation and calculate the number of alleles N_A and the expected heterozygosity H_e over time and at the end of the simulation.

Replicate the simulation 30 times to derive a distribution of N_A and H_e at the end of the simulation ($t = 50$). Plot the curves in time for the 30 simulations.

Now explore the effects of demographic parameters on the maintainance of genetic diversity:

1. Increase the survival probability from 0.5 to 0.75, leaving all the other parameters unchanged. What do you expect? Perform 30 simulations, calculate N_A and H_e at $t = 50$ and compare with the baseline.
2. Increase the female fecundity from 30 to 45, leaving all the other parameters unchanged, as in point 1.
3. Increase the mutation rate from 10^{-6} to 10^{-2}.
4. Increase the dispersal rate to other demes, from 0.1 to 0.2 (i.e. with 4 demes: 0.2 / 3 to other demes and a 0.8 probability to remain in the same deme.

For each of the points above, plot the 30 curves of N_A and H_e, calculate the final values and plot the distributions. Try plotting the distributions of the different parameters (survival, fecundity, mutation, dispersal) one beside the other to compare them.