

# Influence of breeding strategy on genetic diversity of individuals

Poroshina A.A.\*, Sherbakov D.Yu.

Limnological Institute of the Siberian Branch of the Russian Academy of Sciences, Ulan-Batorskaya Str., 3, Irkutsk, 664033, Russia

**ABSTRACT.** Using a computer simulation model, we tried to investigate how the transition from sexual reproduction to asexual reproduction will affect the population of diploid organisms with a neutral character of molecular evolution. At the same time, special attention was paid to the specificity of microsatellite markers. In this paper, we develop fast and inexpensive methods for assessing the changes in populations that occur with a change in reproductive strategy.

**Keywords:** microsatellite markers, object-oriented modeling, asexual reproduction, parthenogenesis

## 1. Introduction

The most widespread method of asexual reproduction in animals is parthenogenesis that is also found among highly organized organisms (Slatkin, 1987). Even in stable natural populations of multicellular organisms, the phenomenon of the formation of offspring due to asexual reproduction is rather common when there is no recombination and unification of the genetic material of different organisms. This process is the most widespread among plants and primitive animals (annelids, coelenterates, etc.). Often, the appearance of a new generation in such populations occurs according to a mixed type, some individuals are formed as a result of sexual reproduction, and some - during asexual reproduction. There are situations when populations, depending on environmental conditions, can switch from sexual reproduction to asexual reproduction and vice versa (Delmotte et al., 2002; Combosch and Vollmer, 2013). In such cases, violations of the patterns of genetic diversity should be expected.

In populations having a sexual process with neutral molecular evolution, fixation of new alleles due to gene drift is established by the Hardy-Weinberg principle.

The Hardy-Weinberg principle is the position of population genetics, which states that in a population of an infinitely large size where natural selection does not work there are no mutation process, exchange of individuals with other populations and gene drift. All crosses are random, for which the frequencies of genotypes or a gene (if there are two alleles of this

gene in the population) will be kept constant from generation to generation and will correspond to equation (Wigginton et al., 2005).

The formation of genetic diversity in populations with sexual reproduction occurs due to the recombination and independent segregation of genetic material in several generations. This leads to the formation of new combinations of alleles. Some of the new alleles that have arisen as a result of spontaneous mutations are fixed or eliminated from the population due to gene drift and/or natural selection. In vegetatively propagating populations, the process of recombination and segregation of alleles is absent. For this reason, during vegetative reproduction, the genotype is completely transferred from one individual to another. In this case, genetic diversity is formed only due to the mutational process (Page and Holmes, 2009; Wang and Bradburd, 2014) and competition/drift of clones. The mutation appears independently in different alleys.

In populations with a sexual process with neutral molecular evolution, fixation of new alleles occurs due to gene drift, and the equilibrium of frequencies by the allele is maintained following the Hardy-Weinberg principle.

Violation of this ratio, as a rule, indicates the impact on the population of any processes that dramatically change genetic diversity. Such processes can be as follows: natural selection, fixing new alleles, expansion or decline in numbers ("bottleneck"), and migration. The bottleneck effect is common in many invertebrates because their number is greatly reduced when the temperature regime changes (autumn-winter

\*Corresponding author.

E-mail address: [a.poroshina@lin.irk.ru](mailto:a.poroshina@lin.irk.ru) (A.A. Poroshina)

**Received:** November 9, 2021; **Accepted:** November 22, 2021;

**Available online:** November 29, 2021

© Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



period). The sharp drop in numbers leads to significant difficulties in the analysis of various genetic markers. In highly organized animals, this effect can lead to the low viability of the species. Individuals also have to resort to inbreeding, and this, in turn, leads to inbred depressions within the species.

In our study, using a computer simulation model, we attempted to investigate how the transition from sexual reproduction to vegetative reproduction will affect the population of diploid organisms with a neutral character of molecular evolution. At the same time, special attention was paid to the specificity of microsatellite markers. This specificity is a consequence of the multi-allele nature of microsatellite repeats and the high rate of their mutation.

## 2. Materials and methods

Two population scenarios were considered and compared:

1. the formation of genetic diversity and the frequency of occurrence of homozygous and heterozygous organisms in a population with sexual reproduction;
2. the formation of allelic diversity and frequencies of occurrence of homozygous and heterozygous organisms in a population with asexual reproduction.

Comparing the results of the two model experiments, it is possible to understand how inconsistent are the frequencies of alleles, homo and heterozygous individuals in two contrasting population scenarios.

In this paper, we present models suitable for describing speciation processes in ancient lakes. To obtain the data that mimic the results of real studies, all the described individual models include a neutral mutating microsatellite sequence (vector  $n1 \dots nm$ ,  $n \in \{10, 11, 12\}$  (if the allele has three states)) that can be transmitted to descendants in the form of a copy with an asexual method of reproduction or be presented in the form of two copies that independently segregate in several generations.

The question arises of how to correctly process and extract the necessary information about genetic diversity if we work with a sample characterizing the distribution of alleles in a population with a mixed type of reproduction. To do this, it is reasonable to move from analyzing the frequencies of occurrence of alleles to analyzing genotypes. The genotype in this case will be any combination of alleles that characterize the individual. During reproduction, in this case, if the mutation has not occurred, then the offspring inherits the genotype entirely from the parent; with mutation, a new genotype will appear in the offspring.

The main characteristics of genetic diversity in a population, in this case, will be the number of genotypes and the frequency of their occurrence in the population. The transition from loci to genotypes can be performed for both diploid and polyploid organisms.

Thus, the population is considered rather a population of diploids or polyploids than a population of virtual haploid organisms. For the samples transformed in this way, the entire set of methods is applicable, which is used for the analysis of haploid populations or markers of mitochondrial DNA or Y chromosome transmitted through the maternal or paternal lines. For such virtual haploids, it is possible to calculate the indicator of the expected heterozygosity ( $He$ ) according to the formula (1). Then, in this formula,  $n$  is the number of genotypes, and  $a_i$  is the frequency of occurrence of the  $i$ -genotype number in the sample.

$$He = 1 - \sum_{i=1}^n a_i^2 \quad (1)$$

It is no longer possible to interpret  $He$  as the expected heterozygosity; in this case, there will be a probability that two randomly taken individuals from the population will have different genotypes. The more  $He$  is shown for the vegetative population, the more diverse it is. It cannot be interpreted as an indicator of the genetic diversity of a population.

The observed heterozygosity ( $Ho$ ) was estimated directly as the proportion of heterozygous individuals in the sample. Statistical comparison of the  $He$  and  $Ho$  samples calculated in the framework of experiments with different reproductive strategies (comparison of mean values) was carried out using the nonparametric  $\chi^2$  test.

The disadvantage of this criterion is that it does not take into account the effective population size. The effective population size ( $Ne$ ) is the number of parents responsible for the genetic makeup of the next generation. The effective population size ( $Ne$ ) that reflects only the number of breeding individuals and can be inferred from genetic data through the theta parameter can be used to estimate the effective population size based on data on population genetic polymorphism.

$$\theta = 4Ne\mu \quad (2)$$

where  $\mu$  is the mutation rate expressed as changes in the length of the microsatellite locus per generation.

One of the ways to study the processes of formation of genetic diversity is mathematical and computer modeling.

We described the general pattern of violation of the Hardy-Weinberg principle using the methods of object-oriented simulation modeling. Such a model is a program that creates a set of virtual "organisms" with several properties. The parameters that will affect the simulation result are selected for the study. The program gives us data similar to real ones.

The model algorithm was implemented in the Python programming language.

The genotype of the organism was a diallelic locus (diploid organism). Each of the two loci could be in its state. Their length varied in the following ranges:

1. 10 to 12 (3 alleles)
2. 10 to 14 (5 alleles)
3. 10 to 18 (9 alleles)

At the starting moment of the calculation (zero generation), the state of organisms was set randomly. The states were selected from a range of possible values. Organisms could be both homozygous and heterozygous.

Scenarios with different values of the probability of sexual reproduction ( $h$ ) were also implemented: 0.01, 0.25, 0.5, 0.75, and 0.99. At the minimum value, the population is reproduced exclusively asexually. At 0.25, sexual reproduction is optional. Sexual reproduction is inherent in populations with  $h = 0.99$ .

Mutations occurred with a given probability,  $m$  (0.005, 0.0005 and 0.00005). Upon mutation in the offspring, the old position of the locus was changed to a new one chosen at random from a set of possible positions. This value works only for microsatellite sequences because they mutate much more often than the rest of the genome. Mutations occur mainly due to the sliding of DNA polymerase along the homopolymer sequence by several nucleotides, which is multiple of the repeat length.

The population size during the entire simulation was limited to a fixed value,  $N$  (10 tons). Since the Hardy-Weinberg principle assumes neutral molecular evolution, it was necessary to set the  $N$  value, with which the results would not be distorted and, at the same time, did not greatly increase the running time of the software pipeline. The value of this parameter was selected experimentally.

The number of reproductive cycles is equal to the  $T$  value (50 000).

At the initial moment, the program creates 10  $t$  individuals with random states of alleles. An organism can have both a homozygous state of an allele and a heterozygous one. Next, the program selects organisms that will begin the reproduction process in this generation. Depending on the indicator of parameter  $h$ , the individual chooses the breeding method and sex (only if the breeding method is sexual). Depending on parameter  $m$ , the state of the locus in the offspring may change to another value selected from the range of possible ones. Individuals that have reached the maximum possible age are eliminated from the population. The number of offsprings that appeared in a given generation can exceed the established value of 10 tons, so their number is reduced randomly to the value of parameter  $N$ . This mechanism reduces the probability of fixing new variations of allelic states in the population. After that, the program enters a cycle and repeats the algorithm for 50  $t$ . cycles (Fig. 1).

Every  $t = 100$  generations, and information on the genotypes of  $n = 1000$  individuals was saved. The information was saved in the form of a text table. According to the tables of the stored intermediate results for the populations, the indicators of the expected heterozygosity ( $He$ ) were calculated. The observed heterozygosity ( $Ho$ ) was estimated directly as the proportion of heterozygous individuals in the sample. The population parameter  $\theta$  and the average  $He / Ho$  values were also estimated.

The script for calculating the  $Ho / He$  parameters

and the population parameter,  $\theta$ , were written in the R programming language. Data visualization was carried out using the ggplot2 and violplot data packages.

The number of individuals who have passed to the next reproductive cycle amounts to 10,000.

### 3. Results and discussion

The simplest case was a population of strictly constant size, in which polymorphism of a microsatellite marker was simulated, being limited to only three alleles. The physical size of the population accounted for 10,000 organisms, and the duration of the computer experiment was 50,000 generations. The life cycle of an organism was one step of the simulation, respectively; each of the organisms could participate in one reproductive cycle.

Within the framework of this study, models were analyzed, in which the fate of only one bite was monitored; therefore, the estimated population parameters belonged to the one-locus group. The same applies to all other cases considered in this study.

Figure 2 shows the result of one of the experiments, in which the polymorphic locus had three possible states. Such cases are not typical of microsatellites. They usually have more allelic states. The red color indicates expected heterozygosity, and the blue color indicates observed heterozygosity. The observed heterozygosity varies from  $\sim 0.5$  to 0.9, following the scenario with asexual reproduction. In the second case (sexual reproduction prevails in the population), the range of values is greater, from 0.4 to 1. The analysis revealed no significant differences between the samples. The same situation is observed when measuring the  $\theta$ -parameter.

The results of the scenario, in which the polymorphic locus had nine allelic states are shown in Figure 3. This scenario is more typical of microsatellites because they usually have a large number of allelic conditions. In the case of sexual reproduction, the values of the observed heterozygosity are distributed over a wider range. No significant differences are also found in the  $Ho / He$  parameters, but when measuring the  $\theta$ -parameter, the blue line in two cases goes beyond the graph. This indicates very high values of the observed heterozygosity. In the scenario with asexual

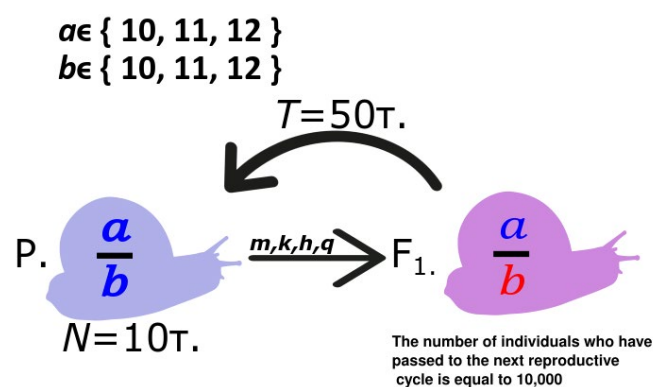
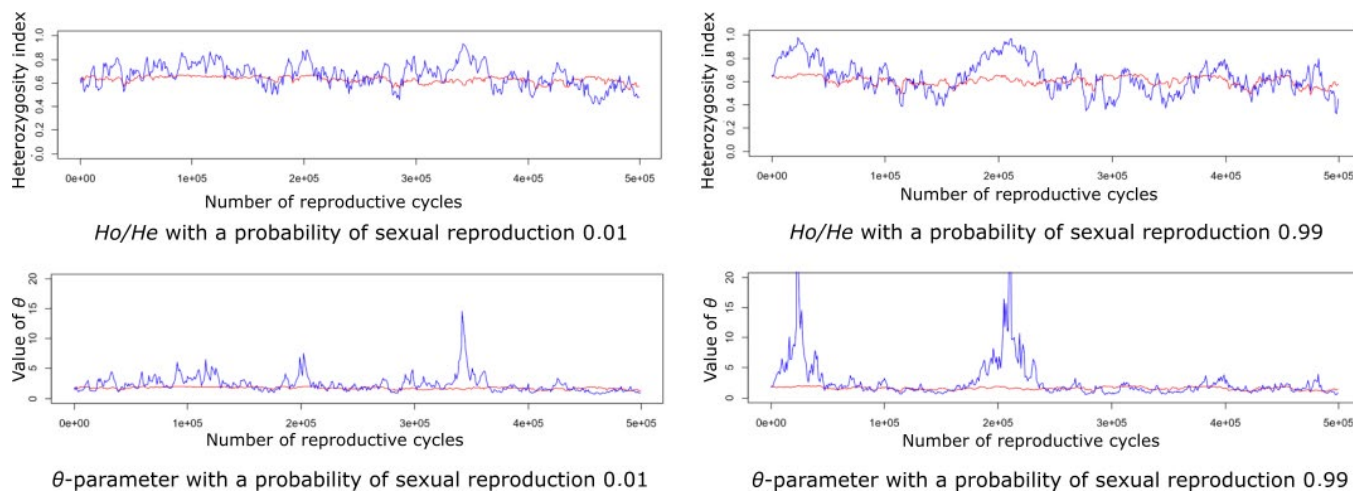


Fig.1. Reproduction of the evolutionary scenario with three allelic states.



**Fig.2.** Results of analysis of  $H_o / H_e$  and  $\theta$ -parameters for three allelic states ( $m = 0.0005$ ). The upper figure is the dynamics of the distribution of observed and expected heterozygosity in the population with asexual and sexual reproduction. The lower figure is the density of distribution of the  $\theta$  value of the observed and expected heterozygosity in the population with asexual and sexual reproduction.

reproduction, there is a strong excess of the observed heterozygosity over the expected. Also, when measuring the  $\theta$ -parameter, more peaks are going beyond the graph than in the case of a predominance of sexual reproduction. Therefore, the lower the likelihood of sexual reproduction, the higher the level of observed heterozygosity. Comparing the results for the three and nine states, the spread in the expected heterozygosity is higher for the nine states. This is due to a large number of variations in heterozygous states.

An analysis was also carried out with different values of the probability of mutations. With a high value of probability of mutation ( $m = 0.005$ ), no significant differences were found between sexual and asexual reproduction. There is also no difference between scenarios with three and nine states of alleles.

At a low value of the mutation probability ( $m = 0.00005$ ), in the scenario with sexual reproduction, the observed heterozygosity is 1, for  $\theta H_o = 499$  (maximum possible value) from 4000 to 50,000 generations (Fig. 4.).

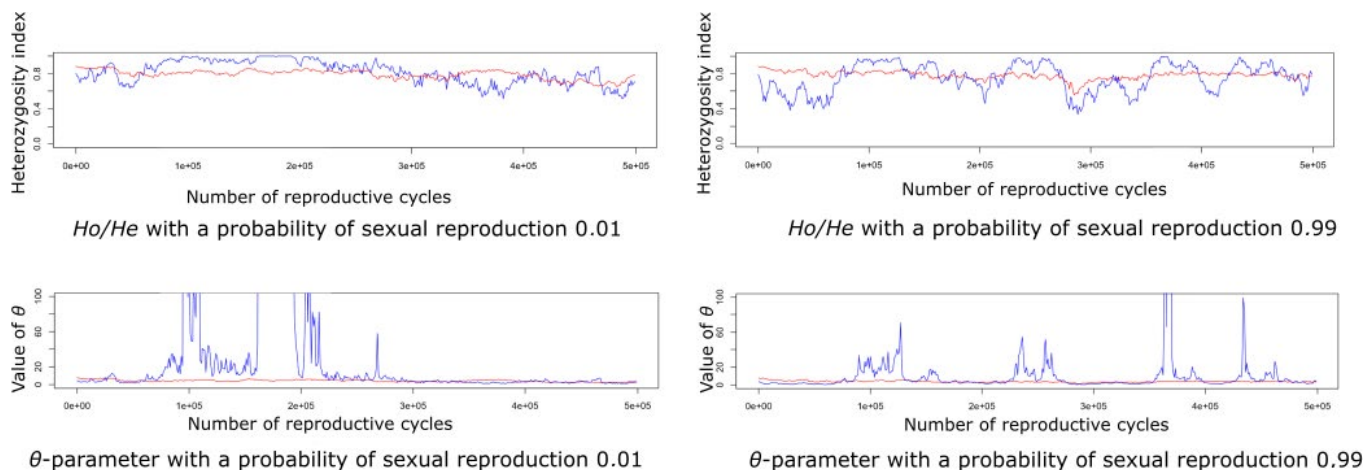
This is because all individuals in the population have become heterozygotes, but due to the low  $m$

value, new variations of states do not have time to appear. All organisms are heterozygous for the first and third allelic states from 4000 until the very end of the program.

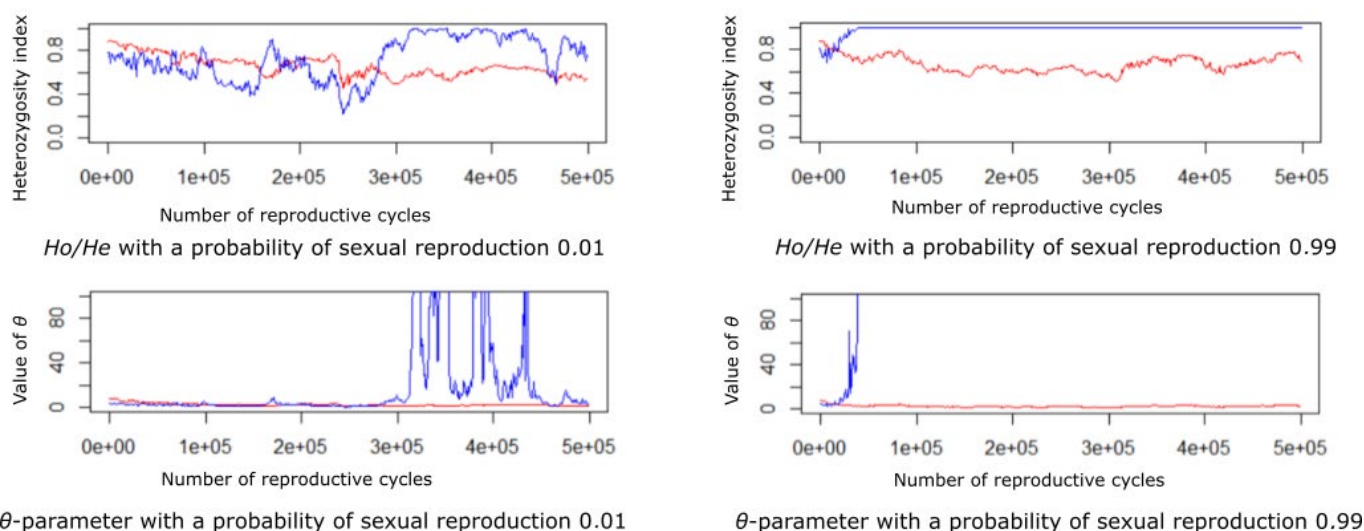
In the case of asexual reproduction, the values of the observed heterozygosity are distributed over a wider range. There are also several peaks in the  $\theta$ -parameter. Throughout the simulation, genetic diversity was maintained in the population and asexual reproduction, contributing to this.

These results appeared due to the exposed parameters of the model. Since organisms reproduced sexually are eliminated from the population after the reproduction act, a low value of parameter  $m$  does not play any role in the occurrence of genetic diversity. With asexual reproduction, the individual survives. If it acquires new variations in allelic states, they can become fixed in the population. In summary, we can conclude that asexual reproduction helps preserve genetic diversity in the population.

Animals that can reproduce themselves in two ways generally resort to asexual reproduction when the population is in constant and favorable conditions.



**Fig.3.** Results of analysis of  $H_o / H_e$  and  $\theta$ -parameters for nine allelic states ( $m = 0.0005$ ). The upper figure is the dynamics of the distribution of observed and expected heterozygosity in the population with asexual and sexual reproduction. The lower figure is the density of distribution of the  $\theta$  value of the observed and expected heterozygosity in the population with asexual and sexual reproduction.



**Fig.4.** Results of analysis of  $H_o / H_e$  and  $\theta$ -parameters for nine allelic states ( $m = 0.00005$ ). The upper figure is the dynamics of the distribution of observed and expected heterozygosity in the population with asexual and sexual reproduction. The lower figure is the density of distribution of the  $\theta$  value of the observed and expected heterozygosity in the population with asexual and sexual reproduction.

Switching to another mode of reproduction occurs under unfavorable or stressful situations as well as with rapid changes in environmental parameters that are important for the survival of the species. For example, a representative of hybridization of the same species in river ecosystems is reproduced only by parthenogenesis, while in coastal waters of the seas they switch to sexual reproduction. In other words, the latter strategy prevails in nature when the need for increased genetic diversity arises.

There were no significant differences in the values of both observed and expected heterozygosity, even taking into account the effective population size,  $m = 0.005$ , between populations where one or the other reproductive strategy predominates. Consequently, at a low mutation rate, regardless of the method of reproduction, under stressful conditions, various combinations and variations of genotypes should be fixed at the same rate.

During asexual reproduction, populations are usually dormant. The population does not need to adapt to new environmental conditions; therefore, there is no need for increased genetic diversity. When loci have a fast and slow rate of mutation, genetic diversity will be preserved during asexual reproduction.

An analysis was also carried out using “violin” diagrams. Comparing scenarios with low and high

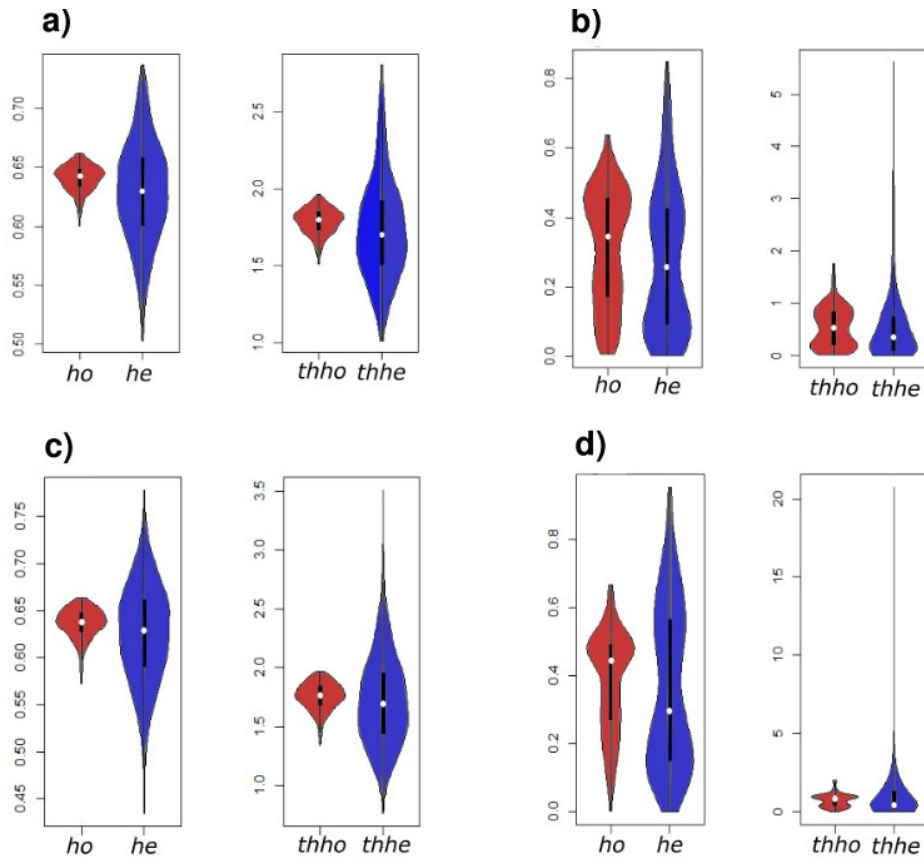
mutation levels, it can be concluded that the values of the parameter of the observed heterozygosity are distributed over a wider range in the scenario with three allelic states, with a value of  $m = 0.00005$ . Also, the median values of expected and observed heterozygosity are higher in the case of a high  $m$  value (Fig. 5, Fig. 6.).

In the scenario with nine allelic states, the main differences are observed in the  $\theta$ -parameter. There are discrepancies in the median values between sexual and asexual reproduction with  $m = 0.00005$ . The parameter values are distributed over a wider range in the case of asexual reproduction. The expected heterozygosity is higher with  $m = 0.005$ . In the scenario with a high probability of mutation during asexual reproduction, there are discrepancies in the median values for the expected and observed heterozygosity. In the case of sexual reproduction, no significant differences are found.

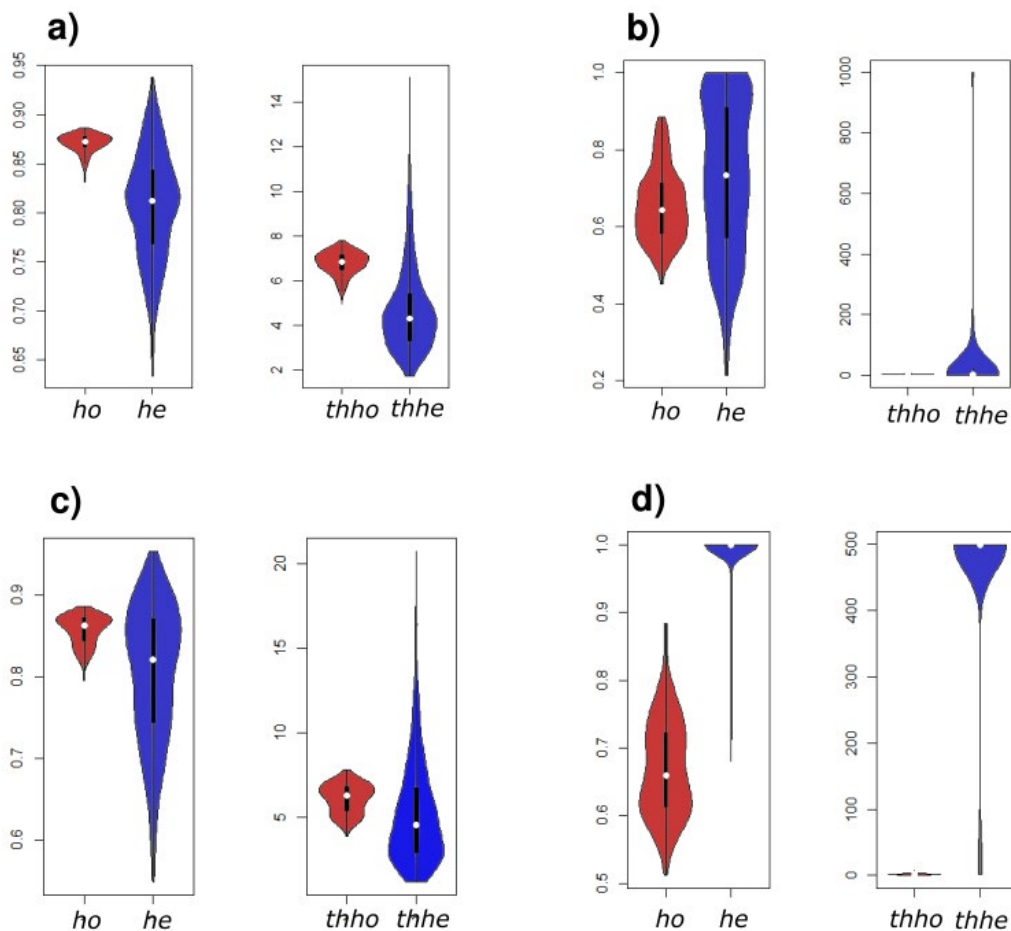
The mean values were also calculated for three, five and nine states of the locus ( $m = 0.005$ ) (Table 1). Differences between observed and expected heterozygosity are traced only for the  $\theta$ -parameter and in the case of nine allelic states. The lower the value of the probability of sexual reproduction, the greater the value of the observed heterozygosity, taking into account the effective population size. Differences are also observed for five allelic states. There are also

**Table 1.** Mean values for three, five and nine states of alleles ( $m = 0.0005$ ).

$m = 0.0005$												
	$h_e$			$h_o$			$th_{he}$			$th_{ho}$		
$h$	3	5	9	3	5	9	3	5	9	3	5	9
0.01	0.629	0.717	0.792	0.665	0.764	0.818	1.714	2.592	4.107	2.353	5.669	72.08
0.25	0.603	0.719	0.779	0.616	0.735	0.914	1.547	2.682	3.759	2.019	5.491	78.47
0.5	0.599	0.737	0.773	0.584	0.843	0.893	1.530	2.896	3.623	1.756	8.145	48.91
0.75	0.607	0.691	0.772	0.634	0.787	0.806	1.575	2.425	3.658	2.473	11.75	112.49
0.99	0.6060558	0.713	0.783	0.636	0.806	0.754	1.561	2.553	3.805	2.690	15.03	113.142



**Fig.5.** Results of analysis of parameters  $H_o / H_e$  and  $\theta$  for three allelic states. a) and b) show the results of the program with the probability of sexual reproduction equal to 0.01. In cases c) and d)  $h = 0.99$ . In diagrams a) and c)  $m = 0.005$ , b) and d)  $m = 0.00005$ .



**Fig.6.** Results of analysis of  $H_o / H_e$  and  $\theta$ -parameters for nine allelic states. a) and b) show the results of the program with the probability of sexual reproduction equal to 0.01. In cases c) and d)  $h = 0.99$ . In diagrams a) and c)  $m = 0.005$ , b) and d)  $m = 0.00005$ .

**Table 2.** Mean values at different mutation levels for three and nine states of alleles ( $m = 0.005, m = 0.0005$  and  $m = 0.00005$ ).

$m = 0.005$								
	<i>he</i>		<i>ho</i>		<i>thhe</i>		<i>thho</i>	
<i>h</i>	3	9	3	9	3	9	3	9
0.01	0.6410578	0.871074	0.62942	0.807004	1.788459	6.792643	1.735566	4.661963
0.99	0.6366654	0.8575254	0.626058	0.804354	1.757177	6.127904	1.726423	5.246142
$m = 0.0005$								
	<i>he</i>		<i>ho</i>		<i>thhe</i>		<i>thho</i>	
<i>h</i>	3	9	3	9	3	9	3	9
0.01	0.6297884	0.7926326	0.66569	0.818118	1.714792	4.107697	2.353728	72.08677
0.99	0.6060558	0.7834572	0.636492	0.754904	1.56149	3.805604	2.69068	113.14267
$m = 0.00005$								
	<i>he</i>		<i>ho</i>		<i>thhe</i>		<i>thho</i>	
<i>h</i>	3	9	3	9	3	9	3	9
0.01	0.3123436	0.6565852	0.277626	0.727656	0.5403204	2.23307	0.5678965	40.21082
0.99	0.385361	0.6693028	0.362432	0.988946	0.7178024	2.20507	1.033858	460.5871

minor differences in the five states of bite. This is due to the higher rate of equilibration according to the Hardy-Weinberg principle. The fewer allelic states at a locus, the faster equilibrium is established. The  $\theta Ho$  value is higher in the case of asexual reproduction. The more the locus has allelic states, the greater the indicators of expected and observed heterozygosity.

An analysis was also carried out for various values of the mutation probability (Table 2). With a high value, no significant differences were found. With a high  $m$  value, differences are observed in the  $\theta$ -parameter.

The higher the probability of mutagenesis, the greater the values of the expected and observed heterozygosity.

#### 4. Conclusions

Taking into account all of the aforesaid, we can conclude that the analysis should take into account the effective population size, the lower the value of the mutation probability and the greater the difference between the observed and expected heterozygosity. The differences are more obvious in asexual reproduction.

The average values of the  $\theta$ -parameter in apomixis populations are much higher than the average values in sexually reproducing populations. The more states at a locus, the stronger the difference between the scatter of expected and observed values. Asexual reproduction helps preserve genetic diversity in a population. This study has demonstrated that the use

of  $Ho / He$  parameters to assess the differences between breeding methods is not relevant. Research is being conducted on other population parameters, including multi-locus ones.

#### Conflict of interests

Authors declare no conflict of interests.

#### References

- Slatkin M. 1987. Gene flow and the geographic structure of natural populations. *Science* 236: 787-792. DOI: [10.1126/science.3576198](https://doi.org/10.1126/science.3576198)
- Combosch D.J., Vollmer S.V. 2013. Mixed asexual and sexual reproduction in the Indo-Pacific reef coral *Pocillopora damicornis*. *Ecology and Evolution* 3(10): 3379-3387. DOI: [10.1002/ece3.721](https://doi.org/10.1002/ece3.721)
- Delmotte F., Leterme N., Gauthier J. et al. 2002. Genetic architecture of sexual and asexual populations of the aphid *Rhopalosiphum padi* based on allozyme and microsatellite markers. *Molecular Ecology* 11(4): 711-723. DOI: [10.1046/j.1365-294X.2002.01478.x](https://doi.org/10.1046/j.1365-294X.2002.01478.x)
- Wigginton J.E., Cutler D.J., Abecasis G.R. 2005. A note on exact tests of Hardy-Weinberg equilibrium. *The American Journal of Human Genetics* 76(5): 887-893. DOI: [10.1086/429864](https://doi.org/10.1086/429864)
- Page R.D., Holmes E.C. 2009. *Molecular evolution: a phylogenetic approach*. Oxford: John Wiley & Sons.
- Wang I.J., G.S. Bradburd. 2014. Isolation by environment. *Molecular Ecology* 23(23): 5649-5662. DOI: [10.1111/mec.12938](https://doi.org/10.1111/mec.12938)