

Modelling Different Management Strategies for an Extinct-In-The-Wild Species – The Sihek

(*Todiramphus cinnamominus*)

Bachelor Thesis, Faculty of Science, University of Bern

handed in by Emma Ochsner

FS2025

Supervised by Prof. Dr. Claudia Bank & PD Dr. Stefano Canessa

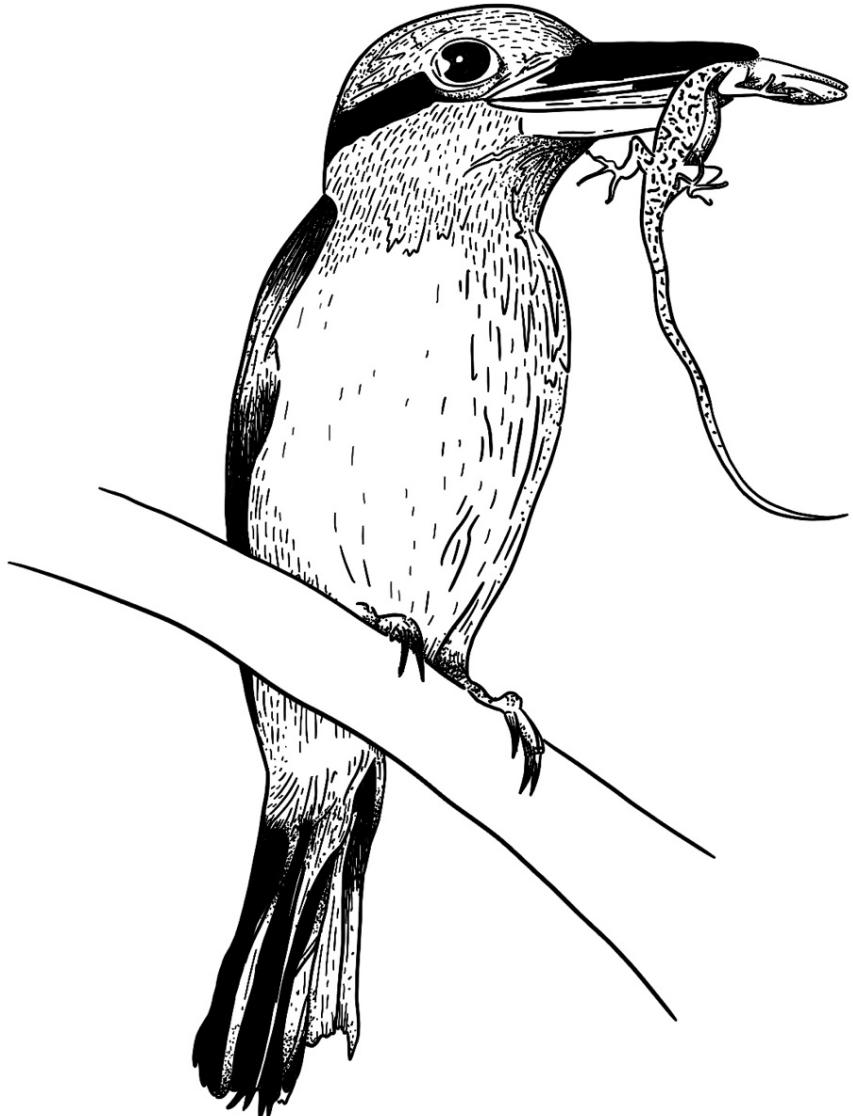


Table of Contents

I.	Abstract.....	3
II.	Introduction	3
III.	Biological question.....	5
IV.	Methods.....	6
	Simulation framework.....	6
	Mate Choices	7
	Equilibrium and early bottleneck	7
	Model 1	8
	Matings.....	8
	Model 2	9
	Model 2.1.....	9
V.	Results	10
	Model 1	10
	Model 2	14
VI.	Discussion.....	16
	Limitations.....	17
VII.	References	18
VIII.	Author Contributions.....	19
IX.	Supplementary Material.....	20

I. Abstract

Extinct-in-the-wild species often go back to a small number of founding individuals and thus face inherently higher risks of extinction, partly due to reduced genetic diversity. Therefore, when managing such species, it is vital to continuously and carefully assess the viability of the ex-situ population and its future health, depending on different management strategies. Here I used computer simulations in SLiM to predict future perspectives of a high-profile extinct-in-the-wild species - the sihek (Guam kingfisher, *Todiramphus cinnamominus*). I simulated future dynamics of the population and its levels of heterozygosity, depending on different mate choice scenarios, namely random mating, mating selecting for highest relatedness, and mating selecting for lowest relatedness. The results indicated that random mating is likely to yield higher levels of heterozygosity compared to mating strategies based on lowest relatedness. This underlines that preserving genetic diversity may be more important than strictly minimizing inbreeding in a small population. Furthermore, the results suggest that when considering reintroductions, one should aim to maximize the heterozygosity between selected individuals and not only minimize their pairwise relatedness.

II. Introduction

Small populations face inherently higher risks of extinction¹. When a population declines, its risk of inbreeding rises. Inbreeding leads to genetically more similar individuals and a loss of heterozygosity¹. With the decline of a population, genetic drift – allele frequencies which randomly fluctuate over time – will become more dominant as well². This can lead to the random loss of adaptive alleles, as well as the fixation of deleterious ones. An increase in homozygosity and the expression of deleterious alleles often leads to inbreeding depression, characterized by a decrease of individual fitness². Further consequences are lower resistance to diseases and higher susceptibility to environmental and demographic stochasticity¹. This results in small populations having a reduced capacity to adapt to and evolve in novel environments³. For example, a study by Ortego et al.⁴ showed that small and isolated populations of lesser kestrels (*Falco naumanni*) are more prone to extinction than larger, better-connected ones, mostly due to stochastic processes. Small population sizes in this bird led to reduced genetic diversity⁴. Ultimately this can end in the extinction of the species^{1,3}, due to reduced short- and long-term viability of the population². Thus, maintaining genetic diversity in a population is crucial for its stability and resilience⁵.

The consequences of small population sizes are of particular concern for extinct-in-the-wild species. They often go back to a small number of founding individuals and are maintained at population sizes far below a necessary threshold that ensures demographic security⁶. This is the case for the sihek, also known as Guam Kingfisher (*Todiramphus cinnamominus*), a bird species currently listed as Extinct in the Wild in the IUCN Red List^{7, 8}.

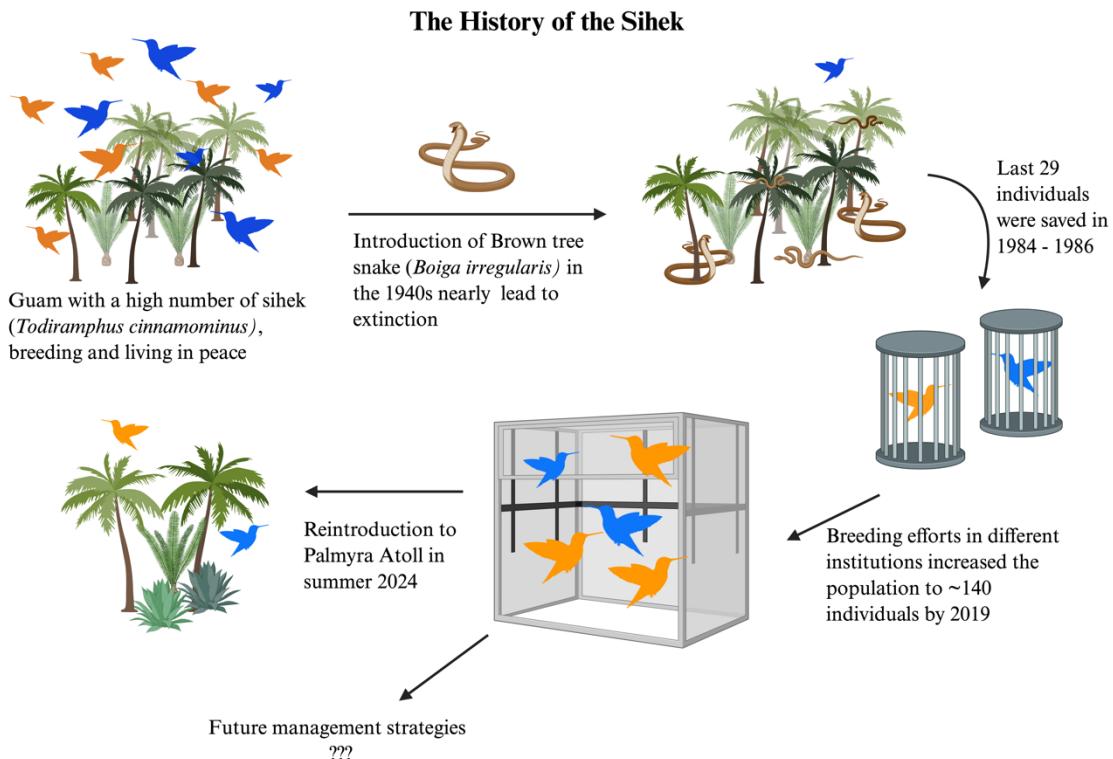


Figure 1: A graphical representation of the history of the sihek. Done in BioRender by Emma Ochsner, March 2025.

The sihek is endemic to the island of Guam, the southernmost island in the Marianas group in the Micronesian sub-region of the Western Pacific. The bird population underwent a strong decline in the 1940s, due to the accidental introduction of brown tree snakes (*Boiga irregularis*) to this island. Between 1984 and 1986, researchers were able to capture 29 wild sihek, just in time before the whole population on Guam went extinct two years later⁸. These 29 individuals are genetically based on 16 individuals and founded the ex-situ population, which is currently managed by a consortium of breeding facilities in the USA and Guam⁸. By 2019, the population had increased to around 140 individuals across 25 institutions but faced space and resource constraints to further growth. Trask et al. were able to detect impacts of inbreeding depression in adult survival and reproductive success in this species, which directly translates into an effect on population viability⁸. They suggested that the most effective way to mitigate inbreeding effects is to increase breeding efforts, which could stabilize and even grow the population. Only then, harvests for releases into the wild would be possible⁸. To increase the viability of the population⁸, the sihek recovery group started planning conservation introductions, including Palmyra Atoll in the Line Islands, aiming to eventually reintroduce the species on Guam (Figure 1). Within this management project, there is an urgent need for assessing the viability of (sub)populations, including their ability to support harvests for releases⁸.

A mathematical model to understand and predict such dynamics can greatly help in decision making. Tracking the heterozygosity in the sihek population can serve as an indicator for its future health. To

address this, computer-based simulations were developed to assess genome-wide levels of heterozygosity across the entire population, providing a basis for further management steps. The model could also be used for other endangered species, with species-specific adaptations of different variables.

III. Biological question

Assessing levels of heterozygosity in the sihek population regarding different management strategies with the help of computer simulations could be a big help in decision-making.

Many ambitious projects are currently underway to conserve biodiversity. For managers of those projects, for example when considering reintroductions into the wild, it is important to know which decisions will lead to which outcomes. To this end, conservationists are increasingly integrating mathematics and computer simulations in their planning.

For an extinct-in-the-wild species, the sihek, we aim to determine whether and how

- different management strategies affect levels of heterozygosity in the whole ex-situ managed population.
- there is an optimal strategy to increase the genetic diversity of the small population and thus make it more resilient to stochastic events.

IV. Methods

Computer simulations were used to assess genome-wide expected heterozygosity in an extinct-in-the-wild population in regard to different management strategies.

For the models, two different coding languages were used, SLiM (Version 4.3) to simulate the population dynamics and R Studio (Version 4.4.1) to analyze the data and visualize results.

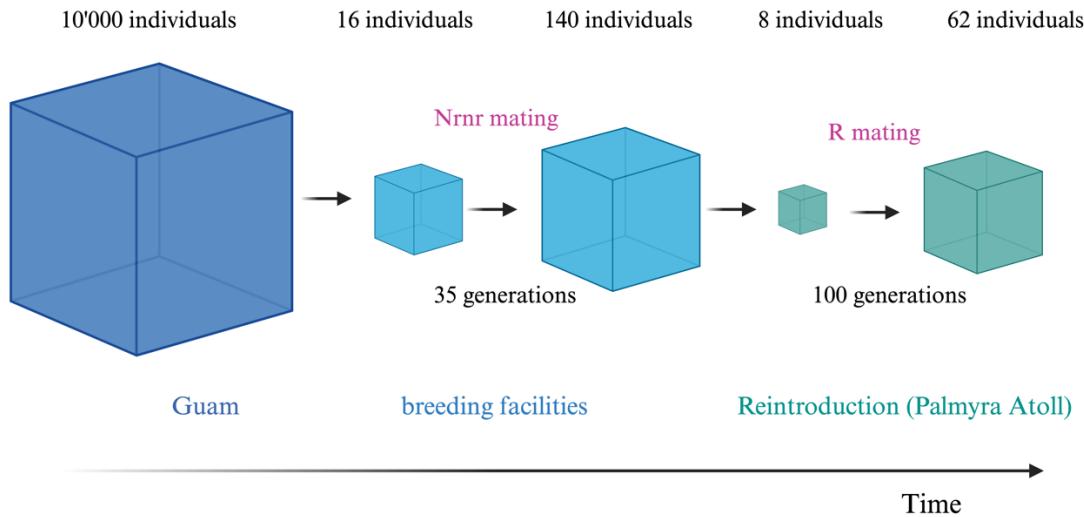


Figure 2: Graphical visualization of the simulation framework. The model “Equilibrium and early bottleneck” simulates the first step of this figure. Model 1 investigates levels of heterozygosity within the breeding facilities, depending on different mating scenarios. Additionally, Model 2 and Model 2.1 focus on different reintroduction scenarios, while integrating the whole population history. The values indicated above the cubes correspond to the values used for the simulation of Model 2 and Model 2.1. Done in BioRender by Emma Ochsner, May 2025.

Simulation framework

For all models, a base model provided by the SLiM manual (14.1 Relatedness, inbreeding, and heterozygosity)⁹ was used, namely a Wright-Fisher (WF) model. The assumptions of a WF model (SLiM specific) are (1) discrete generations, (2) diploid organisms, (3) offspring are generated by recombination of parental chromosomes with addition of new mutations and (4) the probability of an individual being chosen as a parent is proportional to their fitness⁹. Per default, all individuals have the same fitness⁹, thus the assumptions match the ones of the original Wright-Fisher model¹⁰. Some modifications of the base model were made, namely the recombination rate, the mutation rate and the generation time. The mode was changed to a sexual model and a command (`initializeSLiMOptions(keepPedigrees = T)`) was implemented, to be able to track an individuals’ ancestry over the generations. These settings were defined at the beginning of each simulation, using the `initialize()` command in SLiM. To assess levels of heterozygosity in the population, the command `calcHeterozygosity()` was used in all the simulations, which calculates the expected mean heterozygosity across the population⁹.

Table of parameters:

- Mutation Rate (=1e-7), based on mutation rates of the common kingfisher ¹¹
- Recombination Rate (=1e-7) for most simulations, some models were additionally run with recombination rates of 1e-6, 1e-8 and 1e-9 as comparison (see *supplementary material*).
- Mutation type: m1 (m1 has a dominance coefficient of 0.5 and is neutral (fitness 0.0))
- Genomic element type: g1 (initialized with mutation type m1, base mutation rate and a length of 10'000 bases)

Mate Choices

The last assumption of the WF model was changed, using different *mateChoice()* commands in SLiM, which allowed the modification of how individuals are selected as parents. In SLiM, the *mateChoice()* command always chooses the first parent randomly with a probability proportional to the individual's fitness. It then selects a second parent based on the criteria defined in the command ⁹. In a sexual model, the first parent is always the female. Since only neutral mutations were implemented, the fitness of each female is the same, thus they all have the same chance of being picked. The second parent is the male and its probability of being chosen depends on the previously selected female. Three different mate choice patterns were implemented:

- Nrnr (Non-random mating, selecting for non-relatedness): In this model, the male is selected based on lowest relatedness to the previously chosen female. If multiple males show the same degree of relatedness to the specific female, one of those gets picked at random. This was done with the *sample()* command in SLiM, which reshuffles all males with the lowest relatedness value into a new list. The first male of this list is then chosen as “best mate”.
- Nrr (Non-random mating, selecting for relatedness): In this callback, the same procedure as in the Nrnr model was used, only that it selects the males based on highest relatedness values.
- R (Random mating): To match the conditions of the previous models, a *mateChoice()* callback was implemented as well, even if the reproduction would be random per default. The female gets picked randomly and then a list is created for all males, using the *sample()* command in SLiM which reshuffles the individuals with replacement. Then the first male of this list is chosen as “best mate”, for the previously selected female.

Equilibrium and early bottleneck

Before starting with the first model, an additional one was created, to simulate the bottleneck the species underwent while declining to extinction on Guam. The simulation started with a large population of 10'000 individuals, which mated for 100'000 generations – enough to ensure that mutation-drift equilibrium was reached, which is guaranteed after ten times the population size ($10N$ generations) ⁹. This was done to create a population of individuals with distinct genomes, since per default newly created individuals all have the same genome ⁹. From this large population, 16 individuals were

randomly chosen to create the founding population (simulation of the bottleneck). These were then saved in a *.tree* file and in all next simulations reloaded into the models as the starting population.

In an additional analysis, the relatedness values between the males and females of this starting population of 16 individuals were analyzed.

As a comparison, the bottleneck was rerun and this time 1000 individuals were randomly chosen as founders. These individuals were saved for later simulations, to compare different founding population sizes.

Model 1

Model 1 investigates the influence of the three different *mateChoice()* callbacks on genome-wide heterozygosity in the population. To start the simulation, the 16 individuals from before were loaded into the model as starting population. They reproduced for 1000 generations with a carrying capacity of 1000 individuals and an exponential growth rate of 1.08, approximating the growth in the ex-situ population during the period 1986-2024 ⁸. Every fifth generation, the expected genome-wide heterozygosity over the whole population was calculated (*calcHeterozygosity()* command in SLiM ⁹) and printed out in a *.csv* file. 1000 repetitions for every *mateChoice()* callback were run on a server and the results visualized in a plot, using R studio (*Figure 3* and *Figure S-4 in supplementary material* for visualization of the median and the 25/75 % quartiles).

Two simulations where the population size did not change over 100 generations were carried out in addition, to better analyze the genetics of the population. For this simulation, the mean relatedness of the whole population was calculated in every generation as well, to compare it with levels of heterozygosity (*Figure 4*). This was done by looping over all pairs of individuals in the population, adding their pairwise relatedness values and then dividing by the number of pairs. The *relatedness()* command is a built-in function of SLiM, which relies on pedigree data ⁹.

Matings

Another code piece was created, to study the number of individuals reproducing each generation in the NrnR and R model. The *initialize()*, as well as the *mateChoice()* commands are the same as I used in the other models. The population sizes did not change, once with a starting population of 16 individuals and as a comparison once with 1000, implemented from the *.tree* files. The output of this model is the individual IDs of every mating in each generation. The generated files of the 1000 repetitions were loaded into R and the number of distinct IDs were collected. By subtracting this number from the total number of females or males, the number of individuals which did not mate in each generation was calculated and the mean over all repetitions visualized in a plot (*Figure 5*).

Model 2

The second model investigates the outcome of different reintroduction scenarios, especially the influence of relatedness and heterozygosity among founding individuals on population heterozygosity. A simplified “history” of the sihek was implemented into the model, to better reflect the true situation (*Figure 2*). At the beginning of this model, the *.tree* file was loaded into the simulation, containing the 16 individuals with different genomes. The *mateChoice()* command, selecting for lowest relatedness between the mates, was implemented for the first 35 generations. It is the same command used for all the other simulations (Nrnr) and should reflect the breeding in captivity. An exponential growth rate of 1.08 was used during this time to reach a population size of 140 individuals. After these 35 generations, the population size was randomly set to 8 individuals, to simulate the reintroduction scenario. Since a WF model was used for the simulations, it was not possible to choose specific individuals for reintroductions. These 8 individuals then mated for another 100 generations randomly, since in the wild there is no option for pairing specific birds. The overall relatedness and the mean heterozygosity of these 8 individuals was calculated and saved in a file. At the end of the simulation, the mean relatedness and the mean heterozygosity was printed out again, to compare starting and end values. An exponential growth rate of 1.25 was used in this second part of the model, since otherwise there would not have been a change in population size. The carrying capacity was set to 62 individuals, to match the expected carrying capacity of Palmyra Atoll ¹². Once the data of 1000 repetitions were simulated, a correlation analysis was performed with R Studio, using the command *cor.test*, and the method *pearson*. The results were then plotted in a graph (*Figure 6*).

Model 2.1

Model 2.1 is an extension of Model 2, but the output of the model was changed to genome-wide heterozygosity of the whole population in every generation. The change in heterozygosity levels was then plotted in R Studio (*Figure 7*).

V. Results

Model 1

Figure 3 shows the simulations of the first model. The population size increases from 16 individuals to a carrying capacity of 1000, with an exponential growth rate of 1.08, meaning that after 53 generations, the population size of 1000 is reached. The graph shows, that during the start of the simulation, the levels of heterozygosity drop drastically. The steepest decline is observed in the model with non-random mating selecting for highest relatedness. The highest heterozygosity is reached when having random mating, with only a small difference to the mating where the lowest relatedness is selected. This effect is due to a faster decline at the beginning of the simulation, where the blue line drops quicker than the orange one (*Supplementary material, Figure S-3*). Figure 3 shows the model output for a recombination rate of 1e-7, while results for other rates (1e-6, 1e-8 and 1e-9) are presented in the supplementary material (*Figure S-5 – S-10*).

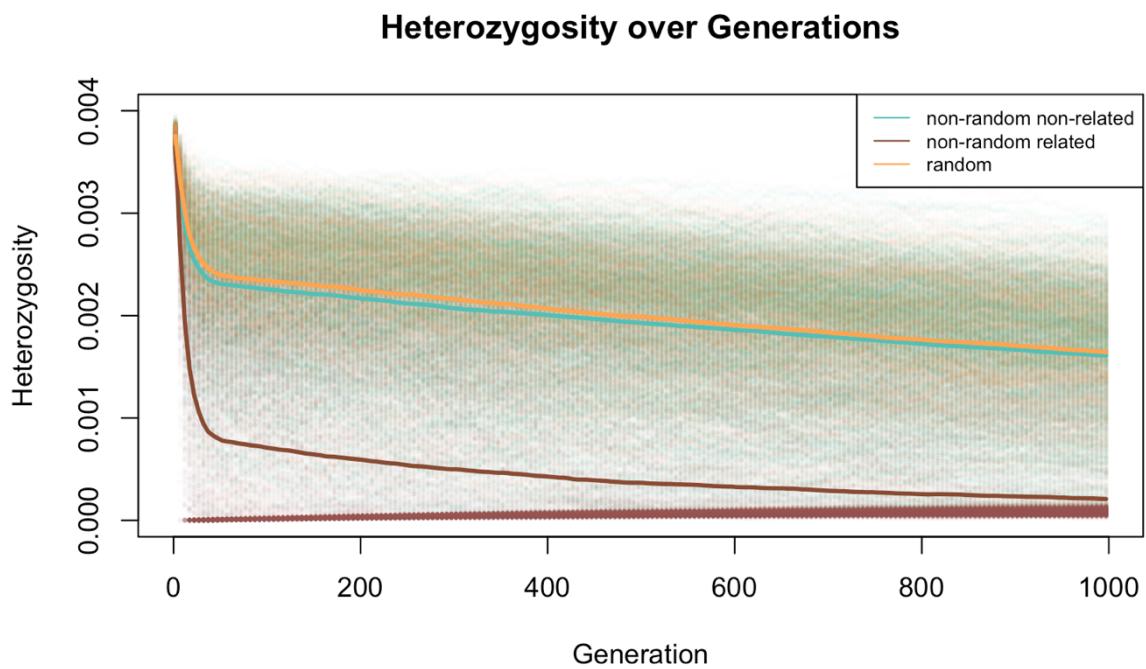


Figure 3: Random mating yields on average higher heterozygosity than when mate choice favors the least related individuals in Model 1. Different levels of heterozygosity across generations are visible, depending on different mate choice scenarios. The x-axis shows the generations and the y-axis the genome-wide heterozygosity over the whole population. The population size goes from 16 to 1000 individuals, with an exponential growth rate of 1.08 and a recombination rate of 1e-7. The heterozygosity was calculated for every fifth generation and indicated by small dots. The different colors represent the different mateChoice() callbacks and the lines show the mean trend over all 1000 replicates. Indicated in orange is random mating, in blue mating selecting for the lowest relatedness and in brown the mating where highest relatedness is selected.

Figure 4 shows mean relatedness (A & C) and mean heterozygosity (B & D) of 1000 replicates over 100 generations for different mate choice scenarios and different population sizes. The Nrr mating always yields the highest relatedness in the population, as well as the lowest heterozygosity over the simulations. The relatedness in the Nrnr model is similar to the Nrr model when the population size is 16 (A), but not when having 1000 individuals (C), where it shows the same levels of relatedness as the R model. Heterozygosity wise, the Nrnr model shows the same results as the random model when having a population size of 1000 individuals (D), but leads to lower levels of heterozygosity when having 16 individuals (B). The single data points, as well as the median and the 25/75 % quartiles of the simulations are visible in additional graphs in the supplementary material (*Figure S-11 – S-12*).

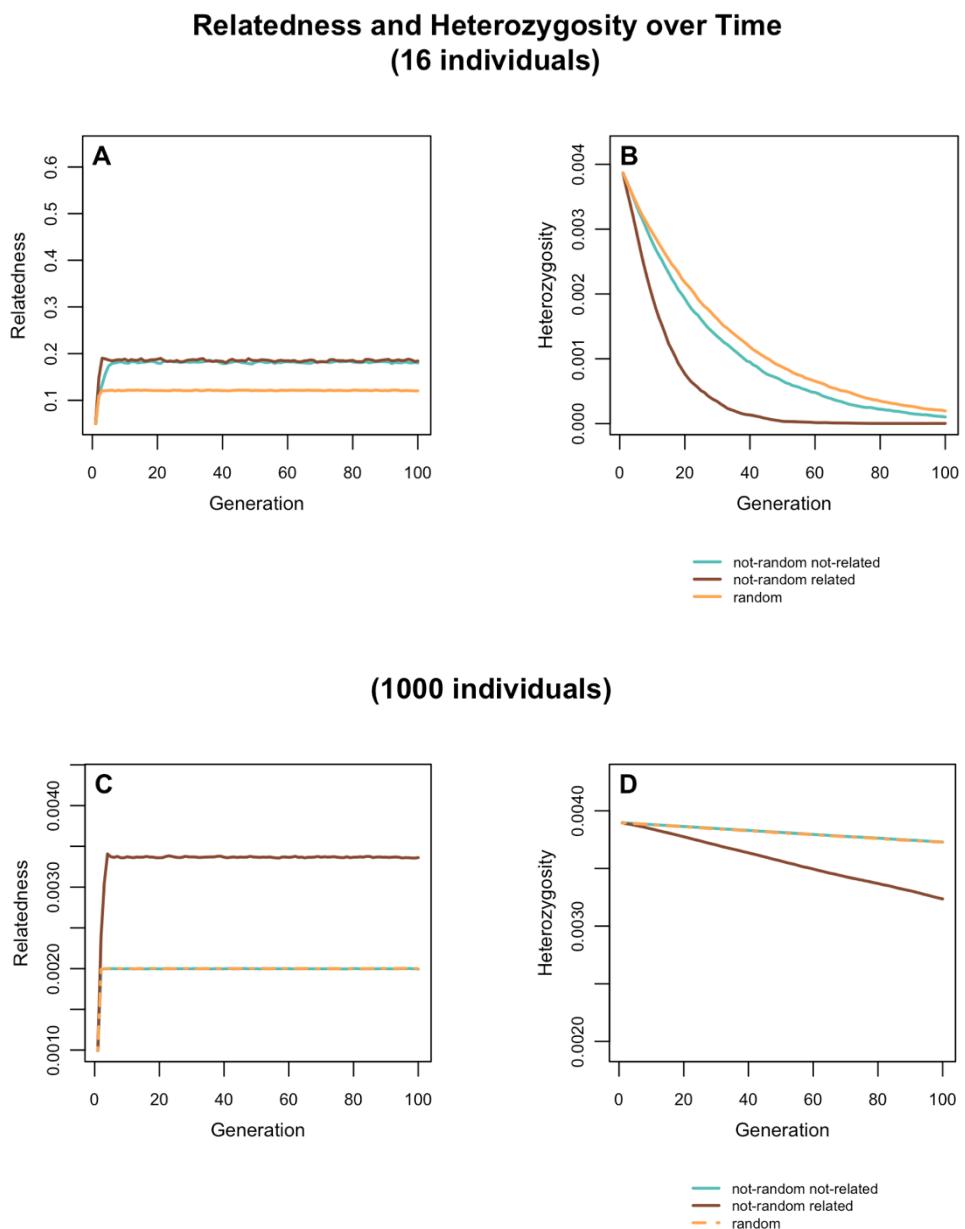


Figure 4: If the population size is set to 16 individuals (A & B), random mating yields the highest heterozygosity as well as the lowest relatedness values, whereas the same simulation with a starting population of 1000 individuals shows no difference in heterozygosity and relatedness between random

mating and mating selecting for lowest relatedness. Relatedness (A & C) and heterozygosity (B & D) in the population, based on different mate choice scenarios are plotted over 100 generations – Once with a population size of 16 individuals (A & B) and once for 1000 individuals (C & D) with no change over time. The data was generated using Model 1 in SLiM. The lines indicate the mean over 1000 replicates. The different colors represent the data simulated with the different mateChoice() callbacks. Orange shows random mating, blue the mating scenario selecting for least related individuals and brown the one selecting mates with highest relatedness. The recombination rate is 1e-7.

Figure 5 shows the number of individuals that do not mate per generation. The data was simulated with Model 1 in SLiM. Two mating callbacks were compared here, namely the model with random mating and the one where mating selects least related individuals. The simulations were done for population sizes which did not change over time (A & B with a population size of 16 and C & D with 1000 individuals). When the population size is 16 and the mate choice happens randomly (B), there is on average one individual per generation that does not mate. In the model where the mate choice favors lowest relatedness (A), there is one female and three males which do not mate per generation. Increasing the population size to 1000, the difference between the two modes of mating vanishes, with an average number of 68 females and males that do not mate per generation (C & D).

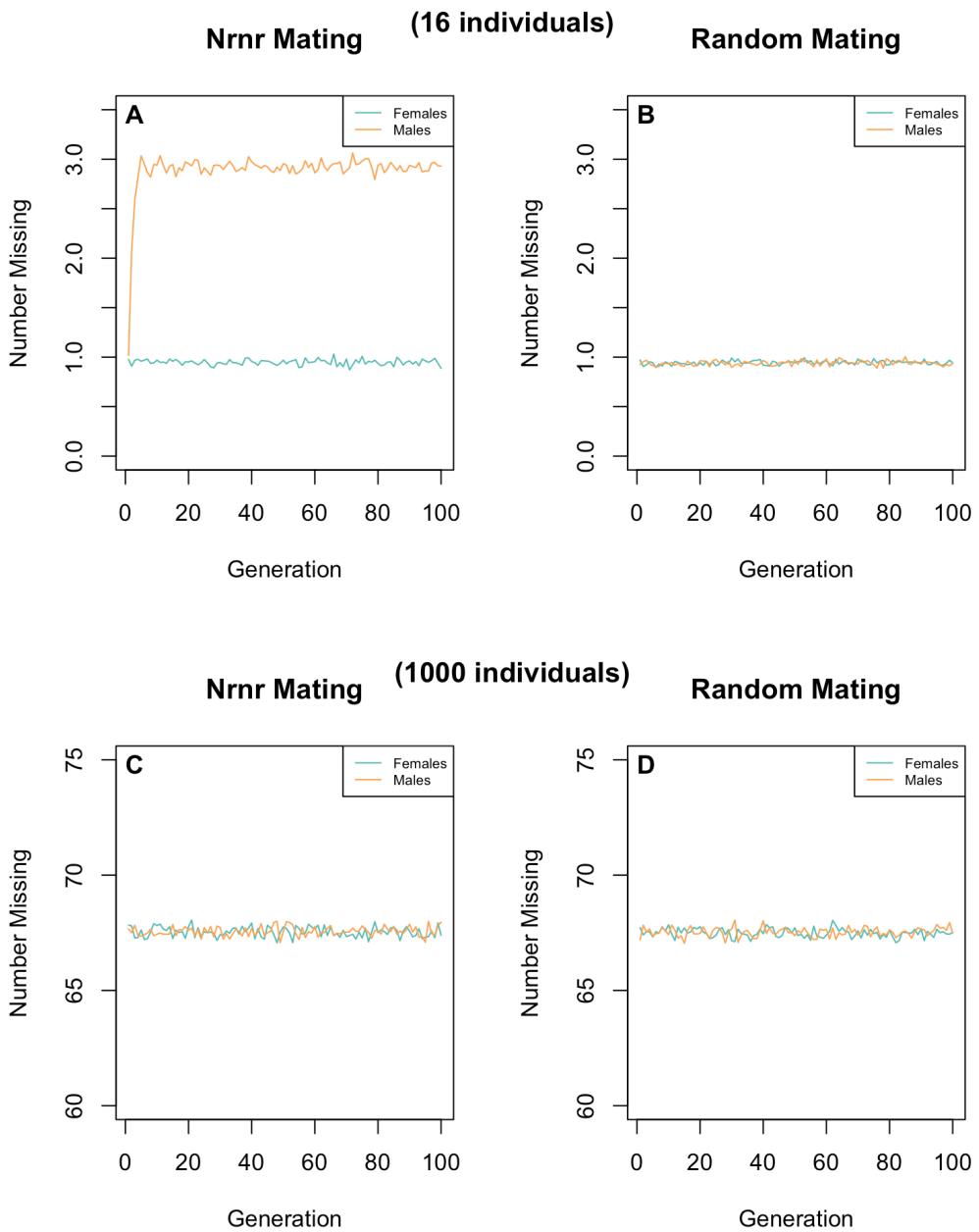


Figure 5: If the population size is set to 16 individuals, a difference in number of females and males which mate per generation in the NrnR model is observable, which vanishes with a population size of 1000 individuals. The mean number of females and males that do not mate per generation in Model 1 with NrnR (A & C) and R (B & D) mate choice implemented is visualized. The x-axis shows the generations and the y-axis the number of individuals which do not mate per generation. Blue represents the females and orange the males. The population size is 16 (A & B) or 1000 (C & D) with an even sex ratio and did not change over time. The individuals mated for 100 generations, in the graphs on the left (A & C) with a mate choice selecting for least related individuals and on the right side (B & D) with random mating implemented. The visualized data is the mean over 1000 replications.

Model 2

Figure 6 shows the output of the second model. The population increases from 16 to 140 individuals in the first 35 generations, drops down to 8 and increases again to a number of 62 over the next 100 generations. The plots below show the correlation between starting relatedness and starting heterozygosity (8 randomly selected individuals after the bottleneck) to end heterozygosity after 100 generations of random mating. There is no statistically significant correlation between starting relatedness and end heterozygosity (Pearson correlation coefficient $r=-0.0157$, $p=0.6202$). Starting heterozygosity shows a positive correlation to end heterozygosity which is statistically significant (Pearson correlation coefficient $r=0.2843$, $p=2.2e-16$). The correlation is lower than 1, meaning that heterozygosity decreases over the simulation, thus is lower after 100 generations than in the starting population of 8 individuals. This analysis was also performed for other recombination rates (1e-6, 1e-8 and 1e-9), showing robust results (*supplementary material, Figure S-14 – S-19*).

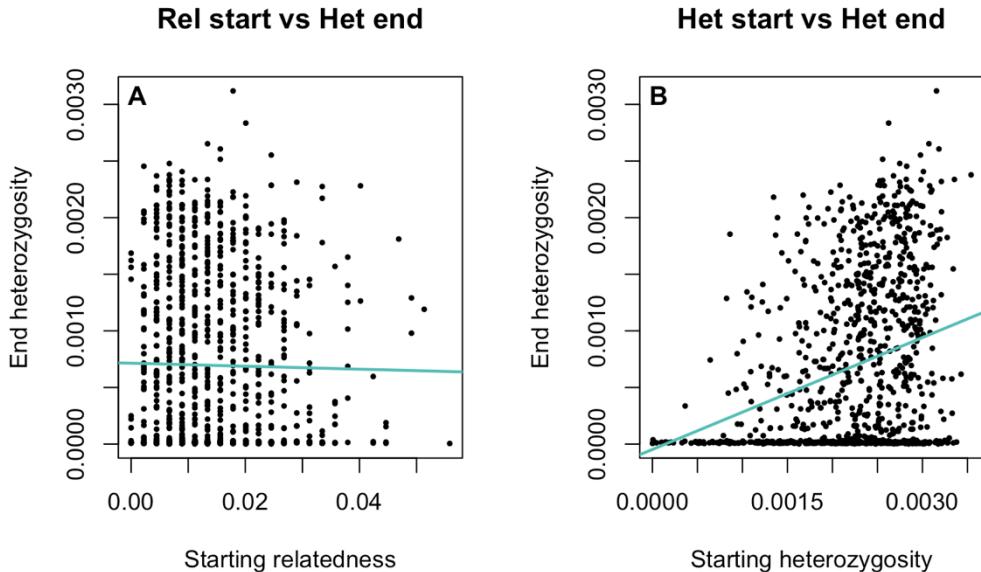


Figure 6: Starting heterozygosity shows a statistically significant correlation to end heterozygosity (B), but for starting relatedness to end heterozygosity there was no correlation found (A). The graphs show the correlation analysis between starting relatedness (x-axis of graph A) and starting heterozygosity (x-axis of graph B) to end heterozygosity (y-axis) for 1000 repetitions. The black dots represent the data points from the single repetitions and the blue line shows their correlation. Left side: Pearson correlation coefficient $r=-0.0157$, $p=0.6202$, right side: Pearson correlation coefficient $r=0.2843$, $p=2.2e-16$.

In addition to Figure 6, a correlation analysis was made for starting relatedness to starting heterozygosity, which can be seen in the supplementary material (*Figure S-13*). There is a statistically significant negative correlation between the two variables (Pearson correlation coefficient $r=-0.0692$, $p=0.0286$); however, the effect size is extremely small.

Figure 7 represents the simulations done with Model 2 with a changed output (*Model 2.1*). The levels of heterozygosity within the whole population across the generations are visible. The population size

changes from 16 individuals to 140 in the first 35 generations of the simulation, with an exponential growth rate of 1.08, is then set to 8 and increases again to 62 individuals over 100 generations, with an exponential growth rate of 1.25. After the bottleneck in generation 35 (indicated as red line), there is a steep decline in heterozygosity, which eventually flattens after some generations. The trend remains declining over the whole simulation. Many repetitions ended with a heterozygosity of 0, as indicated by the 25% quartile.

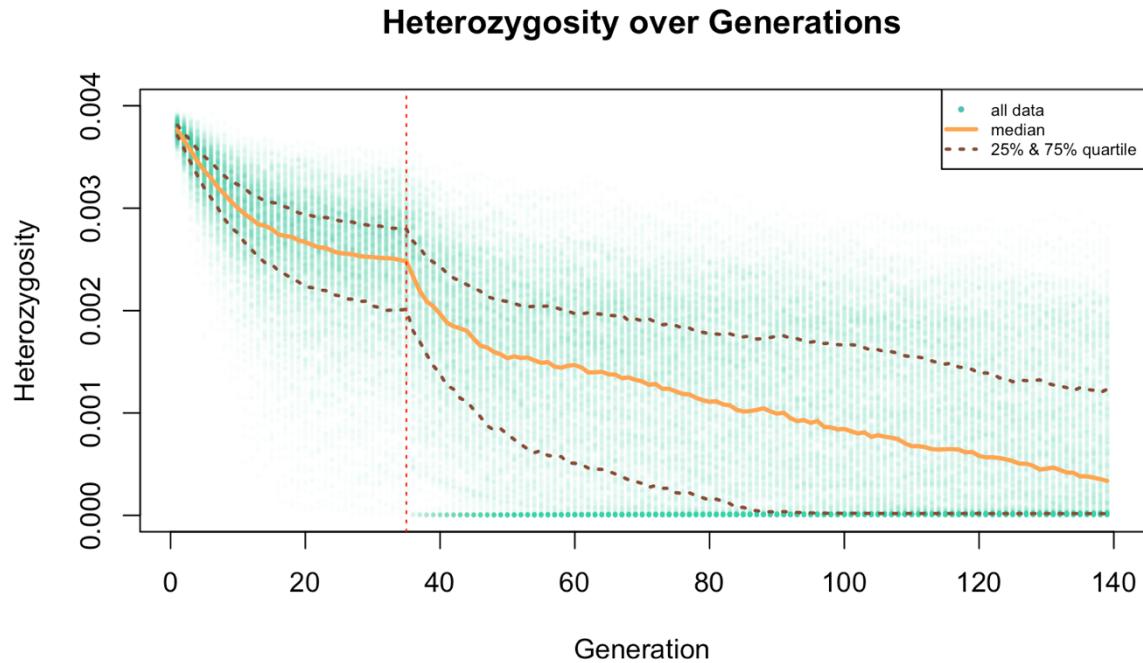


Figure 7: The overall trend of heterozygosity is declining over the generations, but shows an especially steep fall off in small population sizes (start of the simulation and after the bottleneck in generation 35). The heterozygosity across generations for Model 2.1 is plotted, where the history of the sihek is implemented. On the x-axis are the generations and on the y-axis the heterozygosity over the whole population. The single dots represent values from one of the 1000 repetitions, the orange line the median and the dashed lines show the 25% and the 75% quartiles. The population size starts with 16 individuals, increases to 140 over 35 generations, is then set to 8 individuals and grows again to a number of 62. The dashed, red line indicates the timepoint of the bottleneck (generation 35), which set the population size from 140 to 8 individuals.

VI. Discussion

Computer simulations were used to predict future levels of heterozygosity in an extinct-in-the-wild species – the sihek, depending on different mate choice strategies. Contrary to expectations, the simulations of the first model revealed that heterozygosity is higher for the random mating, than when mate choice favors the least related individuals. This pattern is especially pronounced when the population size is small and vanishes with a population size of 1000 individuals (*Figure 4*). This suggests that strictly relatedness-based mating in a small population might not be the best strategy to increase its genetic diversity.

A comparable pattern was also observed by Ivy and Lacy (2012)¹³. They found that under certain circumstances, especially in small or demographically constrained populations, strategies that focus on minimizing mean kinship could perform worse than random mating regarding genetic diversity. In Ivy and Lacy's Static Minimizing Kinship (MK) model, only the individuals with low mean kinship got to produce offspring for a given generation. The genetic diversity of the population therefore declined more quickly than under random mating, because the selected parents in this model did not represent the full range of genetic diversity as well as randomly selected parents would have¹³.

These results were also observed in our Nrnr model, where a lowered effective number of breeders is observable, compared to the R model (*Figure 5*). In a population with 16 individuals with a random mate choice pattern, on average one individual does not get to mate per generation. Conversely when the mate choice command favors the least related individuals, on average three males and one female do not mate per generation. There is a difference in male and female effective number of breeders, because when using the *mateChoice()* command in SLiM, the female is randomly chosen first, then a suitable male is chosen based on the implemented criteria. When the population size is set to 1000, there is no observable difference in the number of breeding individuals between the different models, eventually leading to the steeper decline in heterozygosity at the beginning of Figure 3 for the Nrnr model. Once a higher population size is reached, the decline in heterozygosity is similar for both models, supporting the findings that there is only a difference in effective number of breeders when the number of individuals is small. This is particularly evident when comparing the different population sizes (*Figure 4*): when running the model with 1000 individuals and no change in population size, there is no difference in heterozygosity between R and Nrnr mating; conversely a large difference is visible when running it with 16 individuals. The pattern is likewise reflected in terms of relatedness across the population.

These findings match genetic theory, which predicts that genetic diversity increases with effective population size¹⁴.

The main recommendation we can draw from the results of the first model is that during breeding in captivity – when mating can be controlled – one should focus on maintaining the genetic diversity within the population. It is important to guarantee that all individuals in the population are genetically represented in the next generation, potentially even more important than strictly minimizing inbreeding.

The results of the second model illustrate that higher heterozygosity in the group of individuals selected for reintroduction will lead to a higher heterozygosity within the reintroduced population after 100 generations. Figure 6 shows a statistically significant positive correlation between starting heterozygosity and end heterozygosity of the reintroduced population. Even if there is a rather small difference in starting heterozygosity (*generation 35 in Figure 7*), this variation increases over generations, which highlights how important the choice of individuals is for reintroductions. Despite major fluctuations in heterozygosity over generations, individuals with higher initial heterozygosity increase the chances of success in a reintroduction (*supplementary material, Figure S-20*). In our model, relatedness did not seem to play an important role (*Figure 6*); however, starting relatedness values showed very little variation, suggesting these results should be interpreted with caution. There is a statistically significant negative correlation between starting relatedness and starting heterozygosity ($r=-0.0692$, $p=0.0286$, *supplementary material, Figure S-13*). However, the strength of the correlation is very weak, suggesting that the effect may be negligible. This statistical significance is likely a result of the large sample size (1000), which increases power to detect even trivial associations¹⁵. Additionally, running the correlation analysis with other recombination rates revealed no significant correlation (*supplementary material, Figures S-15, S-17, S-19*). Thus, based on these simulations only, selecting individuals with the lowest relatedness does not automatically imply that they will have high heterozygosity. Nevertheless, selecting individuals who are not related is important for reintroductions, to minimize the risk of inbreeding¹⁶. In summary, given the results of the second model, we recommend that reintroduction planning not only considers the relatedness but also the genetic variation of candidate individuals.

Limitations

As with any model, our analysis represents a simplification of reality and should be implemented with caution. While we could parameterize the model with the mutation rate for a related species, the common kingfisher¹¹, we found no comparable recombination rate in the literature, using instead SLiM's default value of 1e-7 for all simulations⁹. However, results remained consistent when we ran the models again with different values of recombination rate (1e-6, 1e-8 and 1e-9; *supplementary material, Figures S-1, S-2 and S-14 – S-19*). Thus, the fact that the real recombination rate was not used in the simulations can be neglected and is not expected to change the outcomes of the observations. A number of future extensions to the model would be worth exploring. For example, we implemented only one genomic element and heterozygosity as the only metric for genetic diversity; future models could build on this by incorporating more complex genetic elements to better reflect the genetic structure of the sihek. Further there was no uneven sex ratio, as observed in the real sihek population⁸. Although generations are discrete and non-overlapping in our model, in reality sihek would be able to mate over more than just one generation. Another limitation is that we only incorporated neutral mutations, meaning all individuals show the same fitness.

An interesting extension of the first model would be to look at the different levels of heterozygosity with a changed number of breeders for the random mating model. Forcing the R model to operate with the same, lowered number of breeding individuals would enable us to better compare the outcomes of the two different mating scenarios.

Finally, it would be valuable to compare the results of the NrnR model with the actual pedigree data of the ex-situ sihek populations. This would enable the simulations of model 2 to be more accurate and closer to reality.

VII. References

1. Nguyen, T. N. *et al.* Dynamics of reduced genetic diversity in increasingly fragmented populations of Florida SCRUB JAYS , *Aphelocoma coerulescens*. *Evol. Appl.* **15**, 1018–1027 (2022).
2. Ouborg, N. J., Pertoldi, C., Loeschke, V., Bijlsma, R. (Kuke) & Hedrick, P. W. Conservation genetics in transition to conservation genomics. *Trends Genet.* **26**, 177–187 (2010).
3. Willi, Y., Van Buskirk, J. & Hoffmann, A. A. Limits to the Adaptive Potential of Small Populations. *Annu. Rev. Ecol. Evol. Syst.* **37**, 433–458 (2006).
4. Ortego, J., Aparicio, J. M., Cordero, P. J. & Calabuig, G. Individual genetic diversity correlates with the size and spatial isolation of natal colonies in a bird metapopulation. *Proc. R. Soc. B Biol. Sci.* **275**, 2039–2047 (2008).
5. Pannell, J. R. & Charlesworth, B. Effects of metapopulation processes on measures of genetic diversity. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **355**, 1851–1864 (2000).
6. Smith, D. *et al.* Extinct in the wild: The precarious state of Earth’s most threatened group of species. *Science* **379**, eadd2889 (2023).
7. BirdLife International. 2024. Todiramphus cinnamominus. The IUCN Red List of Threatened Species 2024: e.T22725862A261269865. <https://dx.doi.org/10.2305/IUCN.UK.2024-2.RLTS.T22725862A261269865.en>.
8. Trask, A.E., Ferrie, G.M., Wang, J. *et al.* Multiple life-stage inbreeding depression impacts demography and extinction risk in an extinct-in-the-wild species. *Sci Rep* **11**, 682 (2021).
9. Haller, B. C. & Messer, P. W. SLiM: An Evolutionary Simulation Framework.
10. Ishida, Y. & Rosales, A. The origins of the stochastic theory of population genetics: The Wright-Fisher model. *Stud. Hist. Philos. Sci. Part C Stud. Hist. Philos. Biol. Biomed. Sci.* **79**, 101226 (2020).
11. Eliason, C. M. *et al.* Genomic signatures of convergent shifts to plunge-diving behavior in birds. *Commun. Biol.* **6**, 1011 (2023).
12. Ager, T. Sihek Translocation Scoping Trip to Palmyra Atoll: Summary Report.
13. Ivy, J. A. & Lacy, R. C. A Comparison of Strategies for Selecting Breeding Pairs to Maximize Genetic Diversity Retention in Managed Populations. *J. Hered.* **103**, 186–196 (2012).

14. Frankham, R. Relationship of Genetic Variation to Population Size in Wildlife. *Conserv. Biol.* **10**, 1500–1508 (1996).
15. Sullivan, G. M. & Feinn, R. Using Effect Size—or Why the *P* Value Is Not Enough. *J. Grad. Med. Educ.* **4**, 279–282 (2012).
16. Hogg, C. J. *et al.* Founder relationships and conservation management: empirical kinships reveal the effect on breeding programmes when founders are assumed to be unrelated. *Anim. Conserv.* **22**, 348–361 (2019).

VIII. Author Contributions

Conceived and designed the analysis: Prof. Dr. Claudia Bank, PD Dr. Stefano Canessa, Russel Jasper, Emma Ochsner

Collected the data: Emma Ochsner

Contributed data, analysis tools and ideas: Prof. Dr. Claudia Bank, PD Dr. Stefano Canessa, Russell Jasper, Emma Ochsner (with the help of ChatGPT (spring 2025), regarding some specific commands in R).

Performed the analysis: Emma Ochsner

Wrote the paper: Emma Ochsner

IX. Supplementary Material		
Table of Contents		
I.	Starting Relatedness and Heterozygosity	21
	Starting relatedness.....	21
	Starting heterozygosity.....	22
II.	Model 1.....	23
	Recombination rate 1e-7	25
	Recombination rate 1e-6	26
	Recombination rate 1e-8	27
	Recombination rate 1e-9	27
	Relatedness and heterozygosity.....	29
III.	Model 2.....	31
	Starting heterozygosity vs end heterozygosity	31
	Recombination rate of 1e-6	32
	Recombination rate of 1e-8	33
	Recombination rate 1e-9	34
IV.	Model 2.1.....	35

I. Starting Relatedness and Heterozygosity

Starting relatedness

Female id	Male id	Relatedness (1e-7)		Relatedness (1e-6)	Relatedness (1e-8)	Relatedness (1e-9)
0	8	0.0		0.0	0.25	0.25
0	9	0.0		0.25	0.0	0.0
0	10	0.0		0.0	0.25	0.25
0	11	0.0		0.0	0.25	0.0
0	12	0.25		0.0	0.0	0.25
0	13	0.0		0.25	0.0	0.0
0	14	0.0		0.0	0.0	0.0
0	15	0.0		0.0	0.0	0.25
1	8	0.0		0.0	0.0	0.25
1	9	0.0		0.0	0.0	0.0
1	10	0.0		0.0	0.25	0.25
1	11	0.0		0.25	0.5	0.0
1	12	0.0		0.0	0.0	0.25
1	13	0.5		0.25	0.0	0.0
1	14	0.0		0.0	0.0	0.0
1	15	0.0		0.0	0.0	0.25
2	8	0.0		0.25	0.0	0.0
2	9	0.0		0.0	0.0	0.0
2	10	0.0		0.0	0.0	0.0
2	11	0.0		0.0	0.0	0.0
2	12	0.25		0.25	0.25	0.0
2	13	0.0		0.0	0.0	0.25
2	14	0.0		0.0	0.0	0.0
2	15	0.0		0.0	0.0	0.0
3	8	0.0		0.25	0.0	0.0
3	9	0.0		0.25	0.0	0.25
3	10	0.0		0.0	0.0	0.0
3	11	0.0		0.0	0.0	0.0
3	12	0.0		0.25	0.25	0.0
3	13	0.0		0.25	0.0	0.0
3	14	0.0		0.0	0.0	0.0
3	15	0.0		0.0	0.0	0.0
4	8	0.0		0.0	0.0	0.25
4	9	0.0		0.0	0.0	0.0
4	10	0.0		0.5	0.25	0.0
4	11	0.0		0.0	0.0	0.25

4	12	0.0		0.25	0.0	0.25
4	13	0.0		0.0	0.0	0.0
4	14	0.0		0.0	0.0	0.25
4	15	0.0		0.0	0.0	0.0
5	8	0.0		0.0	0.0	0.0
5	9	0.0		0.25	0.0	0.0
5	10	0.0		0.0	0.25	0.0
5	11	0.0		0.0	0.5	0.0
5	12	0.0		0.0	0.0	0.0
5	13	0.0		0.25	0.0	0.25
5	14	0.25		0.25	0.0	0.0
5	15	0.0		0.0	0.0	0.0
6	8	0.0		0.25	0.0	0.0
6	9	0.0		0.0	0.0	0.0
6	10	0.0		0.0	0.25	0.0
6	11	0.0		0.25	0.25	0.0
6	12	0.25		0.0	0.0	0.0
6	13	0.0		0.25	0.0	0.0
6	14	0.25		0.0	0.0	0.0
6	15	0.0		0.0	0.0	0.0
7	8	0.0		0.0	0.0	0.25
7	9	0.0		0.25	0.0	0.0
7	10	0.0		0.0	0.0	0.0
7	11	0.0		0.25	0.0	0.25
7	12	0.0		0.0	0.0	0.25
7	13	0.25		0.0	0.25	0.0
7	14	0.0		0.25	0.0	0.5
7	15	0.0		0.5	0.0	0.0

Table S-1: Pairwise relatedness between males and females within the different starting populations (varying recombination rates). The starting population consists of 16 individuals with equal sex ratio. The table shows four different starting populations for the different recombination rates used during the early bottleneck. 1e-7 is the default recombination rate and used for most of my simulations.

Starting heterozygosity

Recomb. rate	1e-7	1e-6	1e-8	1e-9
Mean het	0.00386514	0.00380518	0.00282129	0.00451131

Table S-2: Genome-wide heterozygosity across all individuals in the starting populations of 16 individuals with different recombination rates.

II. Model 1

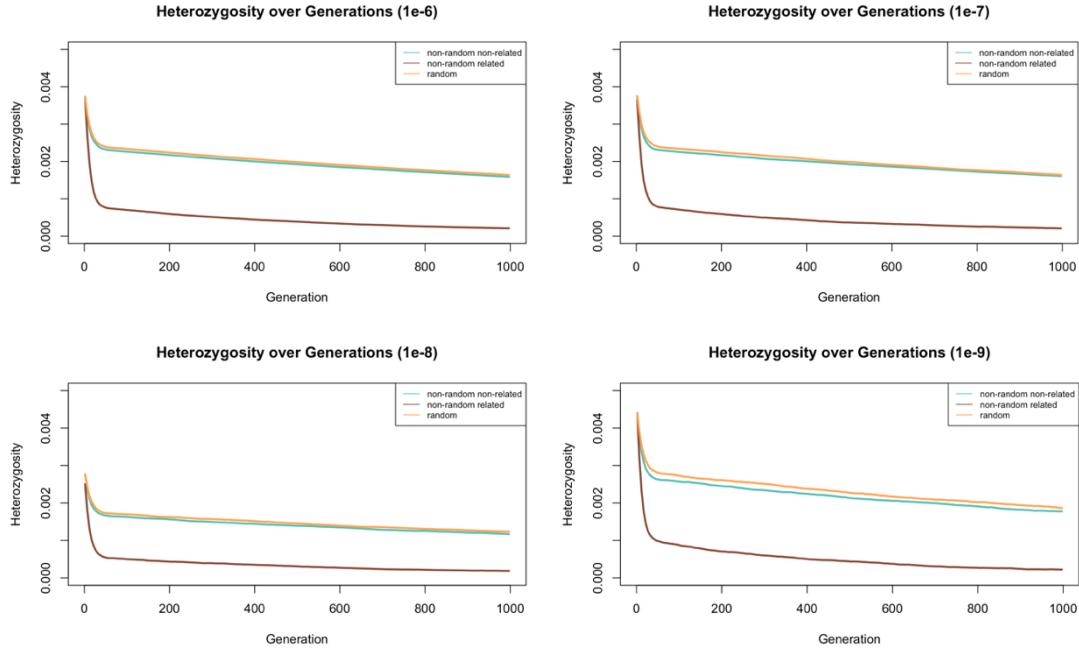


Figure S-1: The trend of random mating yielding highest heterozygosity remains consistent when using different recombination rates. The decline in heterozygosity, depending on different mate choice scenarios is plotted – The data was simulated with Model 1 and different recombination rates (topleft: $1e-6$, topright: $1e-7$, bottomleft: $1e-8$, bottomright: $1e-9$). The distinct starting points of the different graphs, go back to the different starting values in heterozygosity and relatedness (Table S-1 and S-2). On the x-axis are the generations and on the y-axis the genome wide heterozygosity over the whole population. The population size increases from 16 to 1000 individuals, with an exponential growth rate of 1.08. The mean of the 1000 replicates is plotted and the different colors indicate the different mate choice callbacks. Orange represents the values, generated by the model with random mating, blue shows mating selecting for least related individuals and brown mating selecting for the most related individuals.

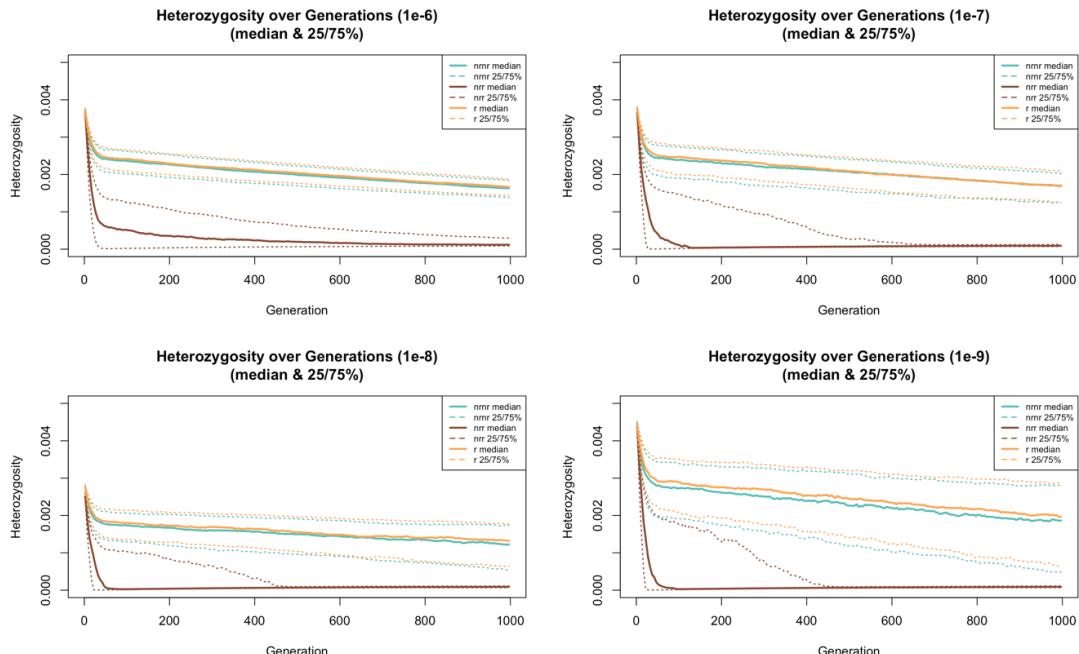


Figure S-2: The trend of random mating yielding highest heterozygosity remains consistent when using different recombination rates. The decline in heterozygosity is plotted, depending on different mate

choice scenarios –The data was simulated with Model 1 and different recombination rates (topleft: 1e-6, topright: 1e-7, bottomleft: 1e-8, bottomright: 1e-9). The distinct starting points of the different graphs, go back to the different starting values in heterozygosity and relatedness (Table S-1 and S-2). On the x-axis are the generations, on the y-axis the genome wide heterozygosity over the whole population. The population size increases from 16 to 1000 individuals, with an exponential growth rate of 1.08. The different colors indicate the different mate choice callbacks. Orange represents the values, generated by the model with random mating, blue shows mating selecting for least related individuals and brown mating selecting for the most related individuals. In this graph the median, as well as the 25% and 75% quartile are shown.

All these graphs are shown again below as single graphs, for better visualization. The single data points of the 1000 repetitions are visible as dots.

Recombination rate 1e-7

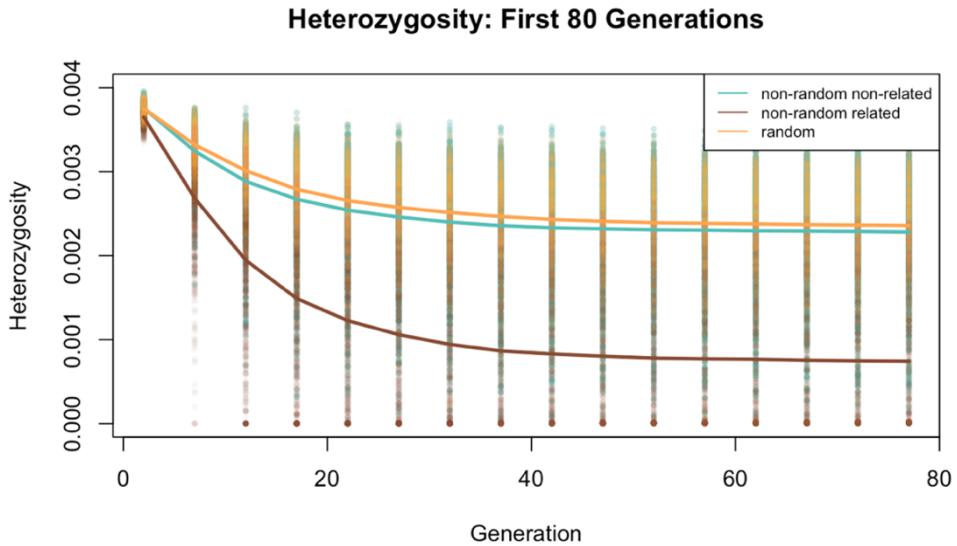


Figure S-3: Output of the first model, with focus on the initial decline in heterozygosity, when the population size is small. On the x-axis are the generations, on the y-axis the genome wide heterozygosity over the whole population. The population size increases from 16 to 1000 individuals, with an exponential growth rate of 1.08 and a recombination rate of 1e-7. The different colors indicate the different mate choice callbacks. Orange represents the values, generated by the model with random mating, blue shows mating selecting for the least related individuals and brown mating selecting for the most related individuals. This graph only shows the first 80 generations of the simulation.

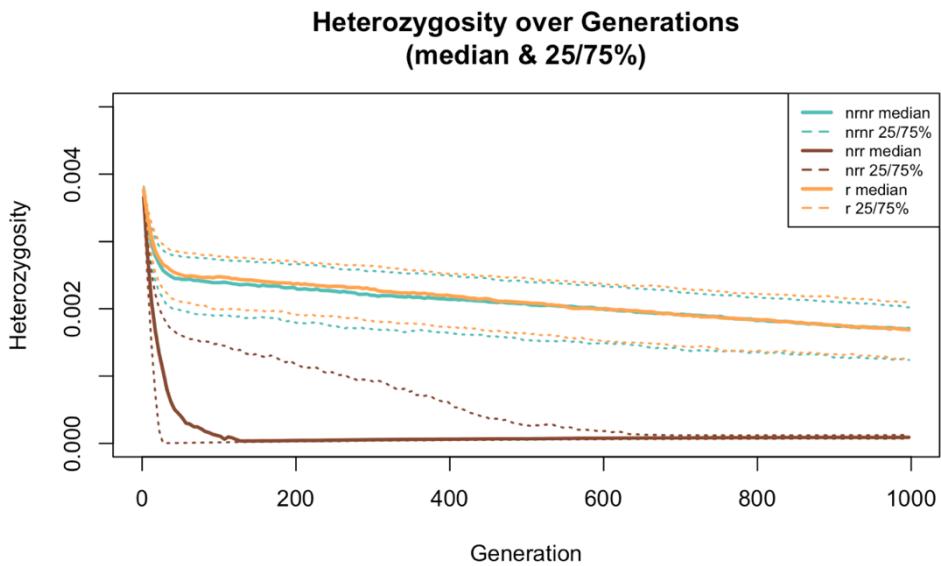


Figure S-4: Random mating yields on average higher heterozygosity than when mate choice favors the least related individuals. The Output of Model 1 is visualized, showing different levels of heterozygosity across generations, depending on different mate choice scenarios. On the x-axis are the generations, on the y-axis the genome wide heterozygosity over the whole population. The population size goes from 16 to 1000 individuals, with an exponential growth rate of 1.08 and a recombination rate of 1e-7. The different colors indicate the different mate choice callbacks. Orange represents the values, generated by the model with random mating, blue shows mating selecting for least related individuals and brown mating selecting for the most related individuals. In this graph the median, as well as the 25% and 75% quartiles are shown.

Recombination rate 1e-6

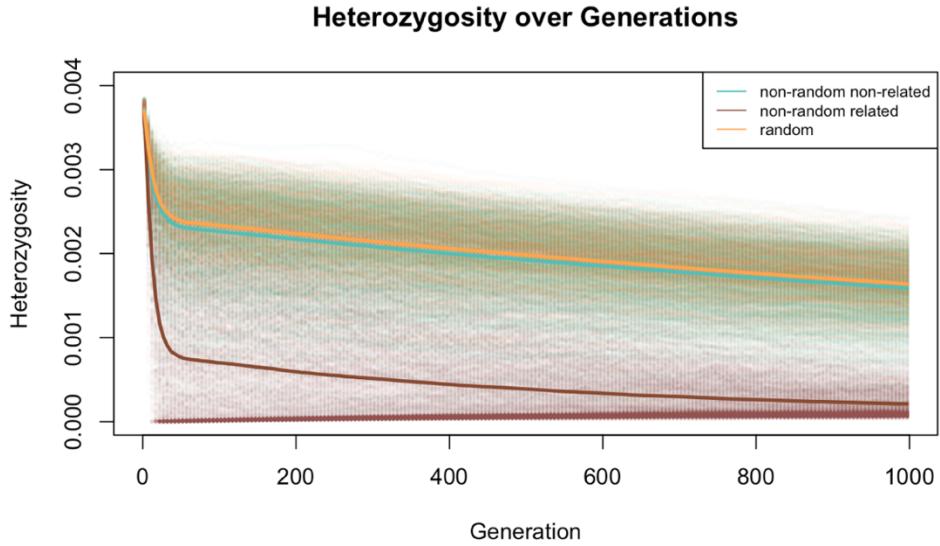


Figure S-5: Random mating yields on average higher heterozygosity than when mate choice favors the least related individuals. The Output of Model 1 is visualized, showing different levels of heterozygosity across generations, depending on different mate choice scenarios. The x-axis shows the generations and the y-axis the genome-wide heterozygosity over the whole population. The population size goes from 16 to 1000 individuals, with an exponential growth rate of 1.08 and a recombination rate of 1e-6. The different colors indicate the different mateChoice() callbacks. Orange represents the values, generated by the model with random mating, blue shows the mating selecting for least related individuals and brown mating selecting for most related individuals.

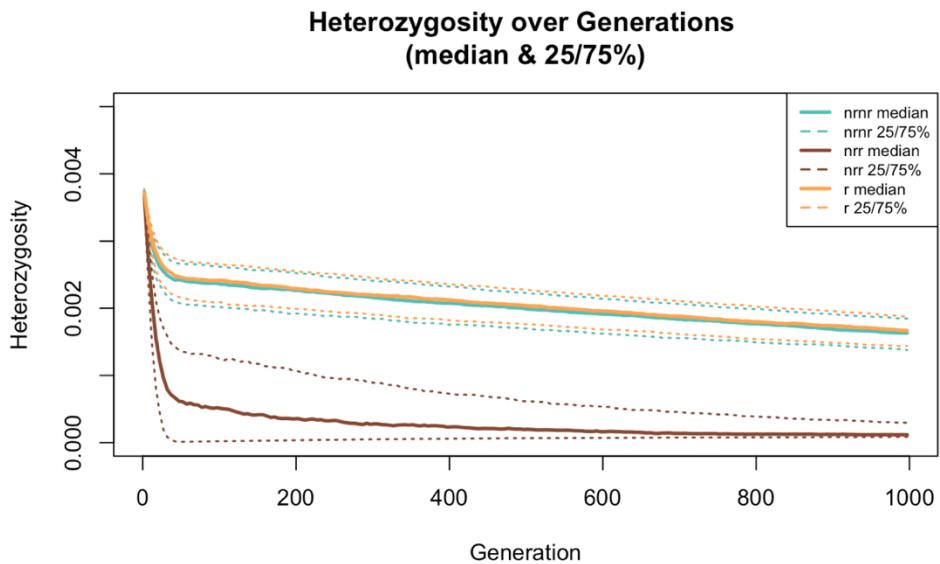


Figure S-6: Random mating yields on average higher heterozygosity than when mate choice favors the least related individuals. The Output of Model 1 is visualized, showing different levels of heterozygosity across generations, depending on different mate choice scenarios. On the x-axis are the generations, on the y-axis the genome wide heterozygosity over the whole population. The population size goes from 16 to 1000 individuals, with an exponential growth rate of 1.08 and a recombination rate of 1e-6. The different colors indicate the different mate choice callbacks. Orange represents the values, generated by the model with random mating, blue shows mating selecting for least related individuals and brown mating selecting for the most related individuals. In this graph the median, as well as the 25% and 75% quartiles are shown.

Recombination rate 1e-8

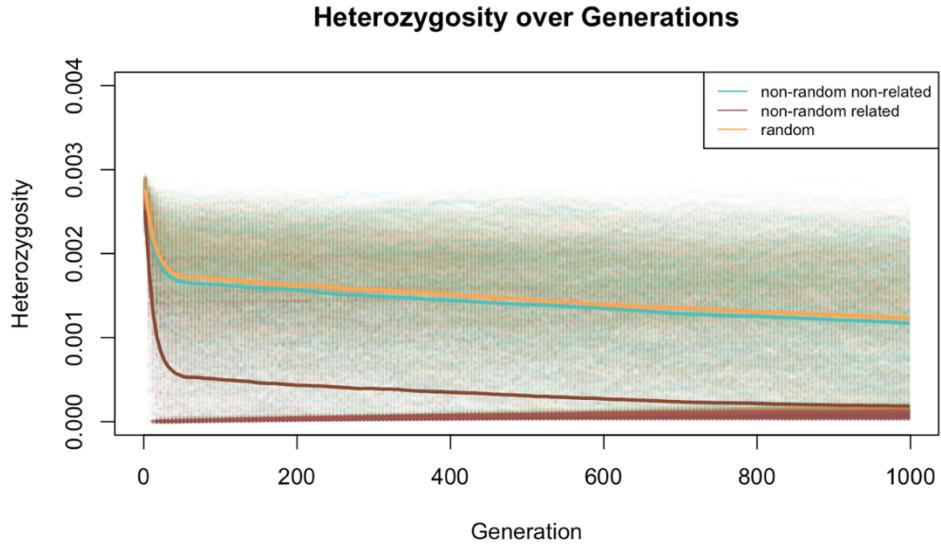


Figure S-7: Random mating yields on average higher heterozygosity than when mate choice favors the least related individuals. The Output of Model 1 is visualized, showing different levels of heterozygosity across generations, depending on different mate choice scenarios. The x-axis shows the generations and the y-axis the genome-wide heterozygosity over the whole population. The population size goes from 16 to 1000 individuals, with an exponential growth rate of 1.08 and a recombination rate of 1e-8. The different colors indicate the different mateChoice() callbacks. Orange represents the values, generated by the model with random mating, blue shows the mating selecting for least related individuals and brown mating selecting for the most related individuals.

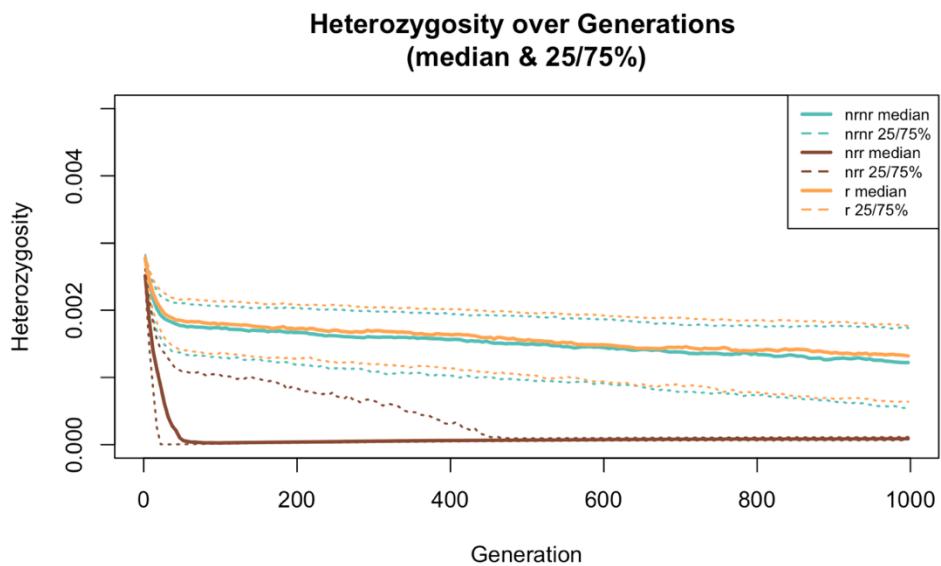


Figure S-8: Random mating yields on average higher heterozygosity than when mate choice favors the least related individuals. The Output of Model 1 is visualized, showing different levels of heterozygosity across generations, depending on different mate choice scenarios. On the x-axis are the generations, on the y-axis the genome wide heterozygosity over the whole population. The population size goes from 16 to 1000 individuals, with an exponential growth rate of 1.08 and a recombination rate of 1e-8. The different colors indicate the different mate choice callbacks. Orange represents the values, generated by the model with random mating, blue shows mating selecting for least related individuals and brown mating selecting for the most related individuals. In this graph the median, as well as the 25% and 75% quartiles are shown.

Recombination rate 1e-9

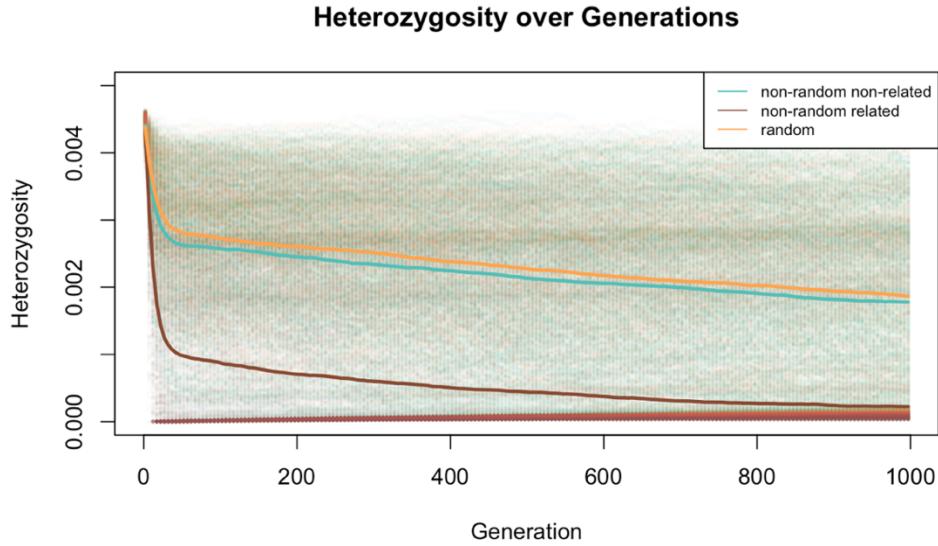


Figure S-9: Random mating yields on average higher heterozygosity than when mate choice favors the least related individuals. The Output of Model 1 is visualized, showing different levels of heterozygosity across generations, depending on different mate choice scenarios. The x-axis shows the generations and the y-axis the genome-wide heterozygosity over the whole population. The population size increases from 16 to 1000 individuals, with an exponential growth rate of 1.08 and a recombination rate of 1e-9. The different colors indicate the different mateChoice() callbacks. Orange represents the values, generated by the model with random mating, blue shows the mating selecting for least related individuals and brown mating selecting for the most related individuals.

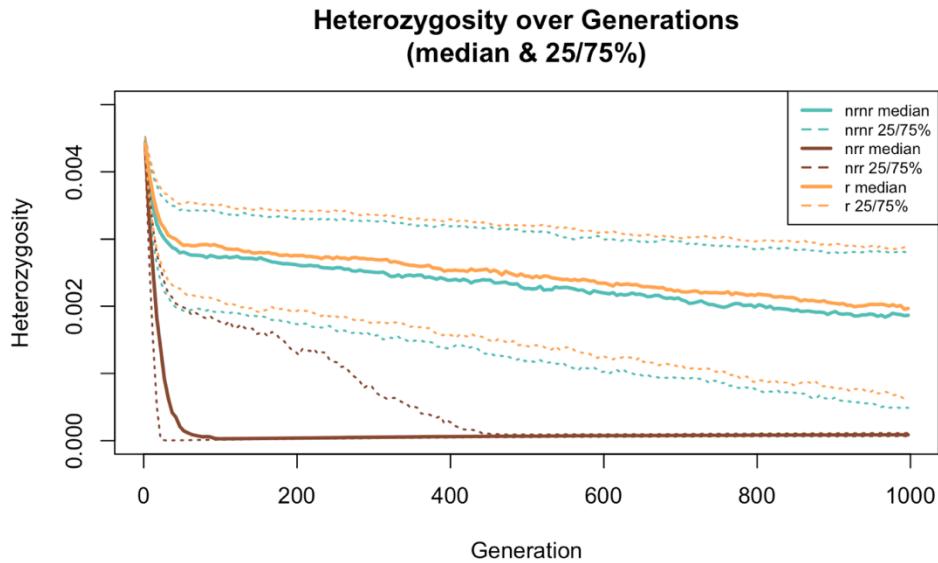


Figure S-10: Random mating yields on average higher heterozygosity than when mate choice favors the least related individuals. The Output of Model 1 is visualized, showing different levels of heterozygosity across generations, depending on different mate choice scenarios. On the x-axis are the generations, on the y-axis the genome wide heterozygosity over the whole population. The population size increases from 16 to 1000 individuals, with an exponential growth rate of 1.08 and a recombination rate of 1e-9. The different colors represent the different mate choice callbacks. Orange indicates the values, generated by the model with random mating, blue shows mating selecting for least related individuals and brown mating selecting for the most related individuals. In this graph the median, as well as the 25% and 75% quartiles are shown.

Relatedness and heterozygosity

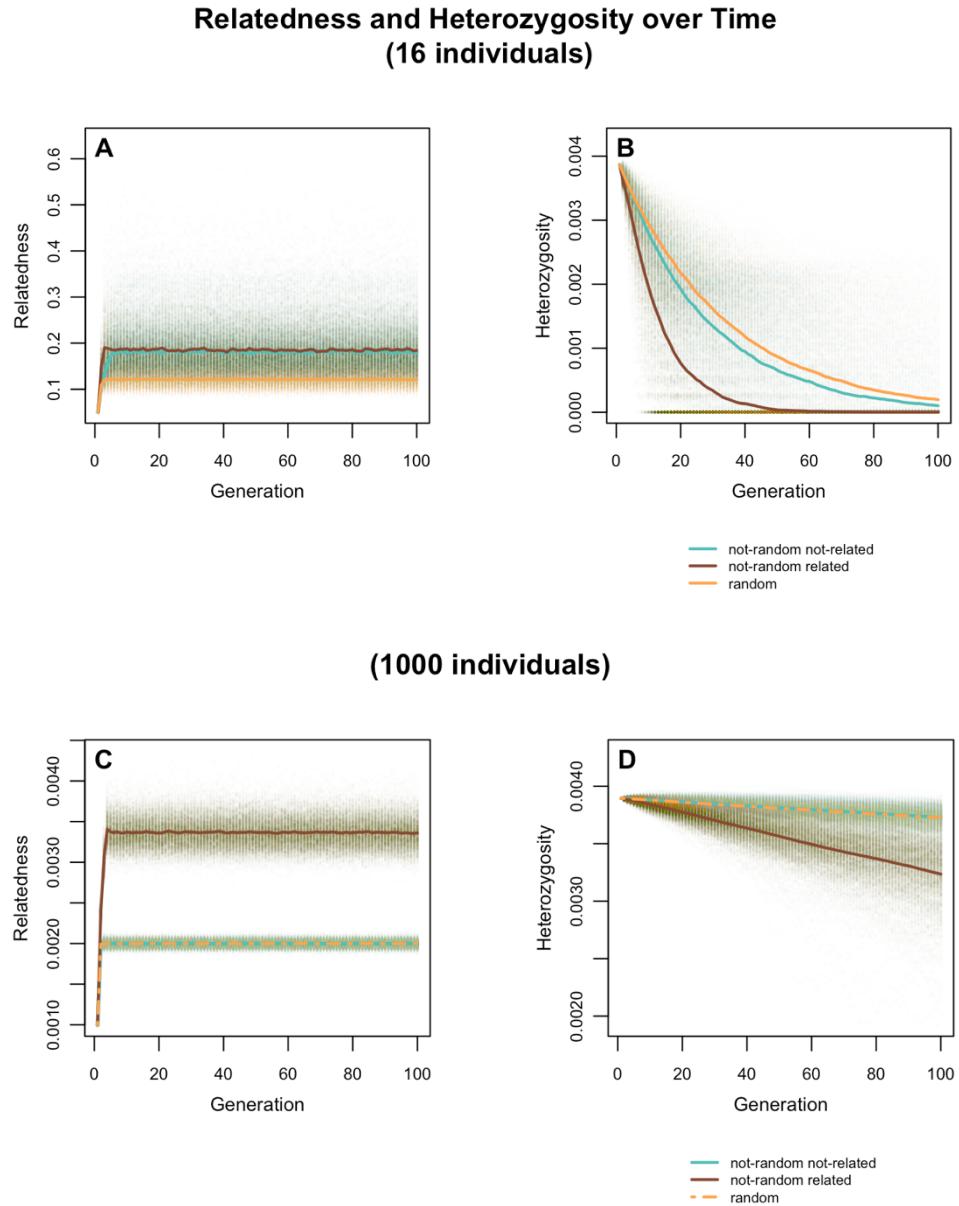
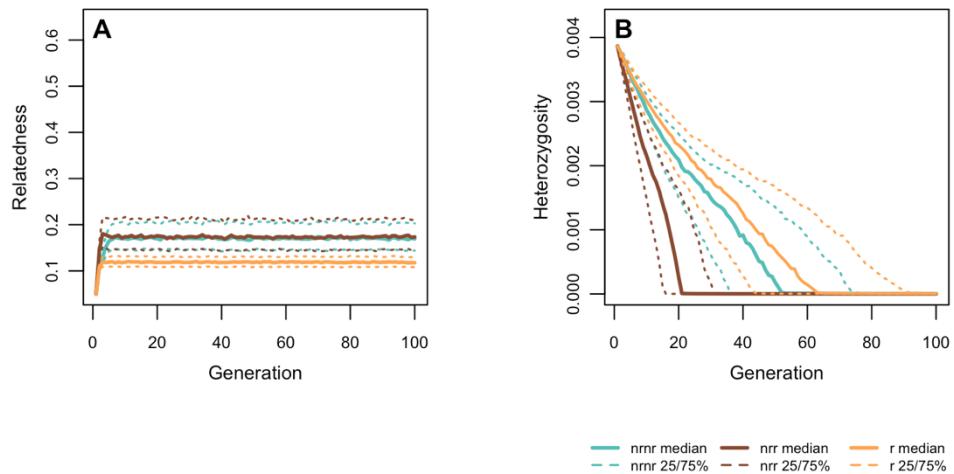


Figure S-11: If the population size is set to 16 individuals (A & B), random mating yields the highest heterozygosity as well as the lowest relatedness values, whereas the same simulation but with a starting population of 1000 individuals shows no difference in heterozygosity and relatedness between random mating and mating selecting for lowest relatedness. Relatedness (A & C) and heterozygosity (B & D) in the population, based on different mate choice scenarios is plotted over 100 generations. The data was generated using Model 1 in SLiM. The population size is set to 16 individuals (A & B) or 1000 individuals (C & D) with no change over time. The dots represent the single data points of the 1000 replications and the lines the mean. The different colors represent the data simulated with the different `mateChoice()` callbacks. Orange shows random mating, blue the mating scenario selecting for least related individuals and brown mating selecting for the most related individuals. The recombination rate is 1e-7.

Relatedness and Heterozygosity over Time (median & 25/75% - 16 individuals)



(median & 25/75% - 1000 individuals)

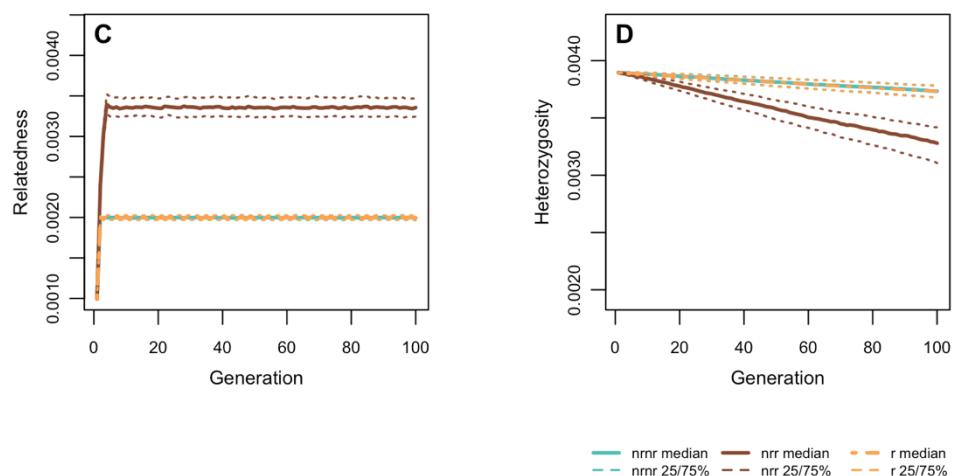


Figure S-12: If the population size is set to 16 individuals (A & B), random mating yields the highest heterozygosity as well as the lowest relatedness values, whereas the same simulation but with a starting population of 1000 individuals shows no difference in heterozygosity and relatedness between random mating and mating selecting for lowest relatedness. Relatedness (A & C) and heterozygosity (B & D) in the population, based on different mate choice scenarios is plotted over 100 generations. The data was generated using Model 1 in SLiM. The population size is set to 16 individuals (A & B) or 1000 individuals (C & D) with no change over time. The different colors represent the data simulated with the different mate choice callbacks. Orange shows random mating, blue mating selecting for least related individuals and brown mating selecting mates with highest relatedness. The median (solid line), as well as the 25% and 75% (dashed lines) of 1000 replicates are shown. The recombination rate is 1e-7.

III. Model 2

Starting heterozygosity vs end heterozygosity

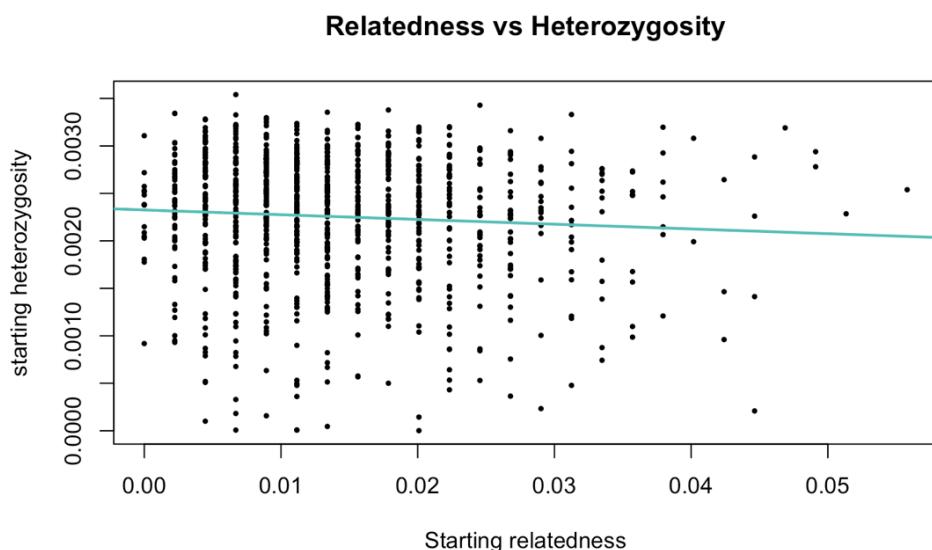


Figure S-13: There is a statistically significant negative correlation between starting heterozygosity and starting relatedness (Pearson correlation coefficient $r=-0.0692$, $p=0.0286$), although the effect size is very small. This Graph was run with Model 2 and visualizes the correlation between starting relatedness (8 introduced individuals) and starting heterozygosity. The points represent the data of 1000 repetitions and the blue line the correlation between them. The recombination rate is 1e-7.

Recombination rate of 1e-6

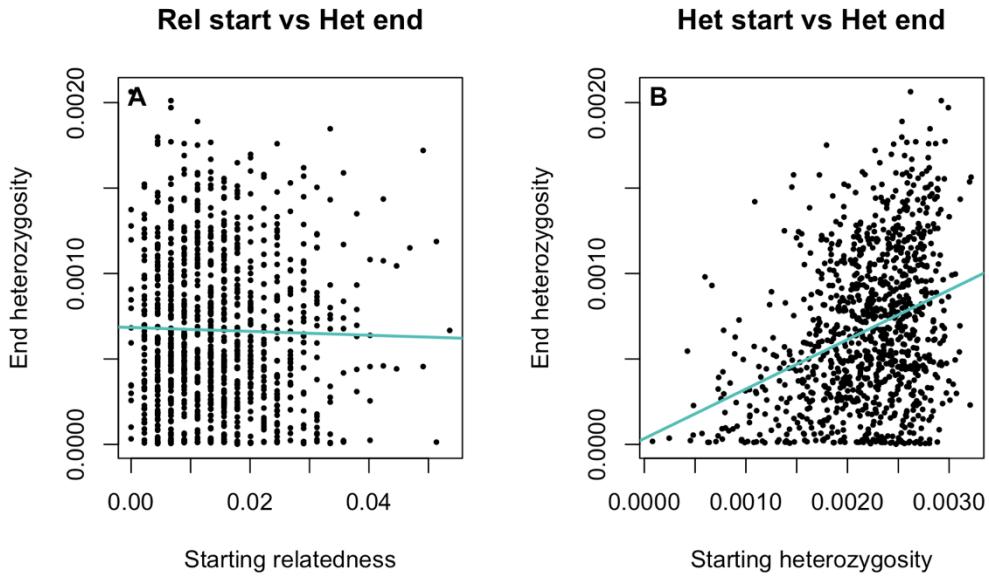


Figure S-14: Starting heterozygosity shows a statistically significant correlation to end heterozygosity (B), but for starting relatedness to end heterozygosity there was no correlation found. The graphs show the correlation analysis between starting relatedness (x-axis of graph A) and starting heterozygosity (x-axis of graph B) to end heterozygosity (y-axis) for 1000 repetitions. The black dots represent the data points from the single repetitions and the blue line shows their correlation. Left side: Pearson correlation coefficient $r=-0.0234$, $p=0.4583$, right side: Pearson correlation coefficient $r=0.3384$, $p=2.2e-16$. The recombination rate is set to 1e-6.

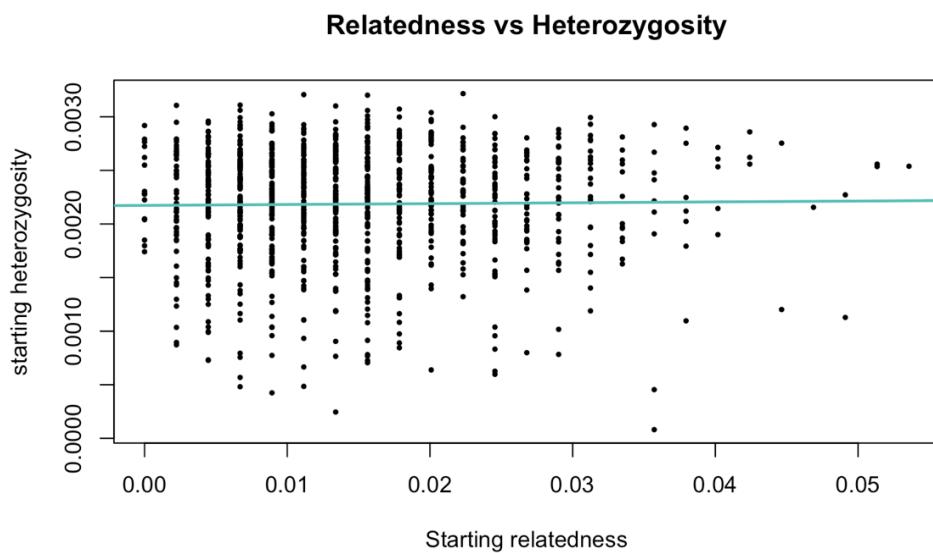


Figure S-15: There is no correlation between starting heterozygosity and starting relatedness (Pearson correlation coefficient $r=0.0147$, $p=0.6426$). This Graph was run with Model 2 and visualizes the correlation between starting relatedness (8 introduced individuals) and starting heterozygosity. The points represent the data of 1000 repetitions and the blue line the correlation between them. The recombination rate is 1e-6.

Recombination rate of 1e-8

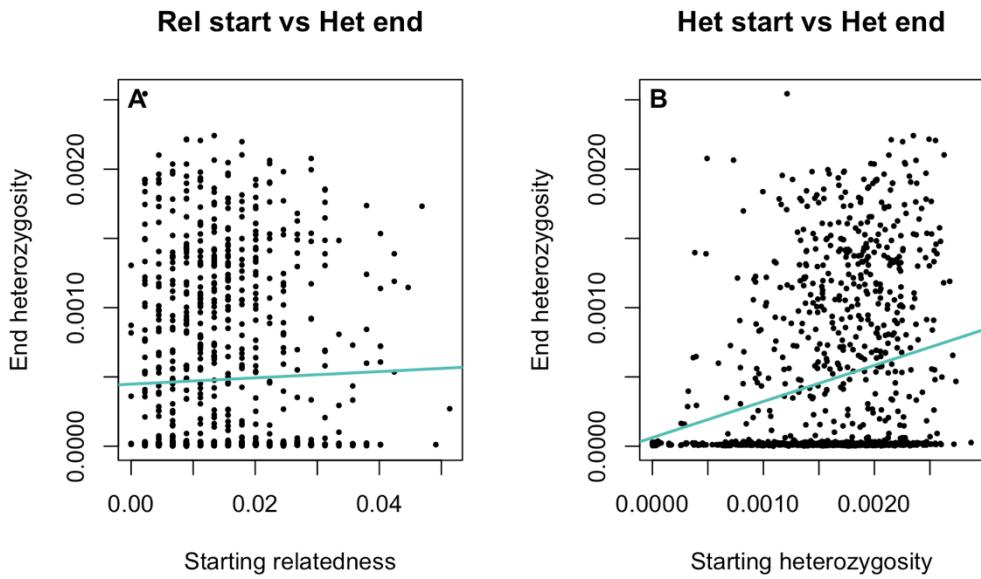


Figure S-16: Starting heterozygosity shows a statistically significant correlation to end heterozygosity (B), but for starting relatedness to end heterozygosity there was no correlation found. The graphs show the correlation analysis between starting relatedness (x-axis of graph A) and starting heterozygosity (x-axis of graph B) to end heterozygosity (y-axis) for 1000 repetitions. The black dots represent the data points from the single repetitions and the blue line shows their correlation. Left side: Pearson correlation coefficient $r=0.0312$, $p=0.3246$, right side: Pearson correlation coefficient $r=0.2475$, $p=1.999e-15$. The recombination rate is set to $1e-8$.

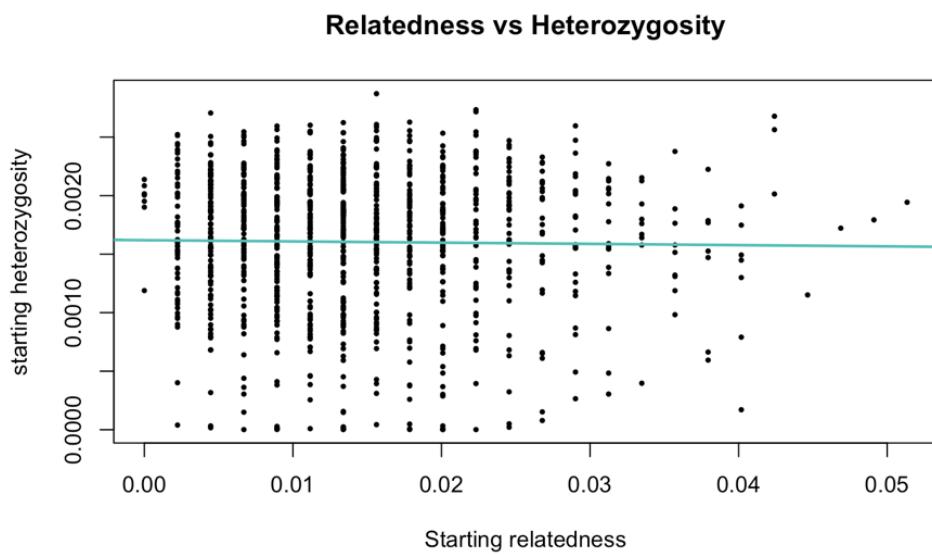


Figure S-17: There is no correlation between starting heterozygosity and starting relatedness (Pearson correlation coefficient $r=-0.0155$, $p=0.6249$). This Graph was run with Model 2 and visualizes the correlation between starting relatedness (8 introduced individuals) and starting heterozygosity. The points represent the data of 1000 repetitions and the blue line the correlation between them. The recombination rate is $1e-8$.

Recombination rate 1e-9

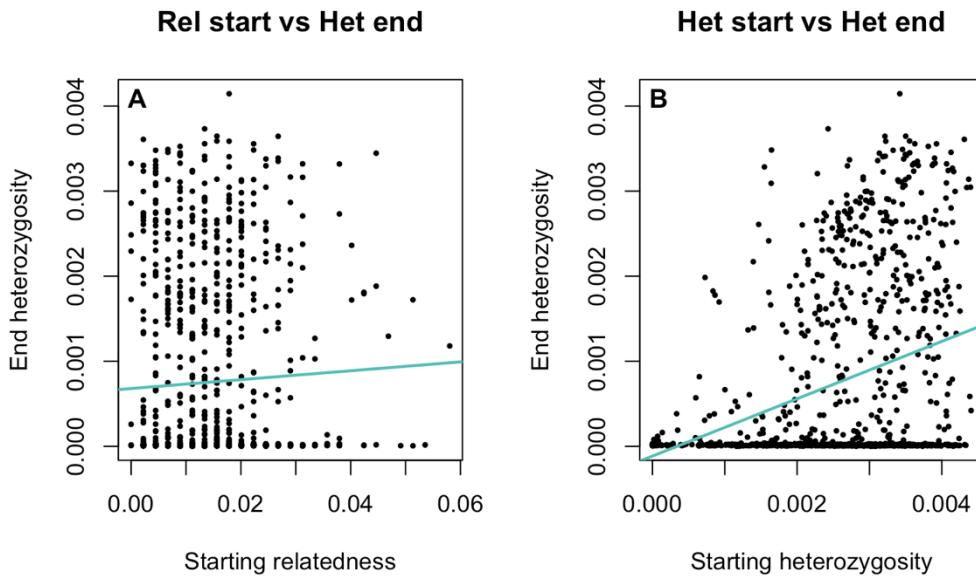


Figure S-18: Starting heterozygosity shows a statistically significant correlation to end heterozygosity (B), but for starting relatedness to end heterozygosity there was no correlation found. The graphs show the correlation analysis between starting relatedness (x-axis of graph A) and starting heterozygosity (x-axis of graph B) to end heterozygosity (y-axis) for 1000 repetitions. The black dots represent the data points from the single repetitions and the blue line shows their correlation. Left side: Pearson correlation coefficient $r=0.0418$, $p=0.1866$, right side: Pearson correlation coefficient $r=0.3275$, $p=2.2e-16$. The recombination rate is set to $1e-9$.

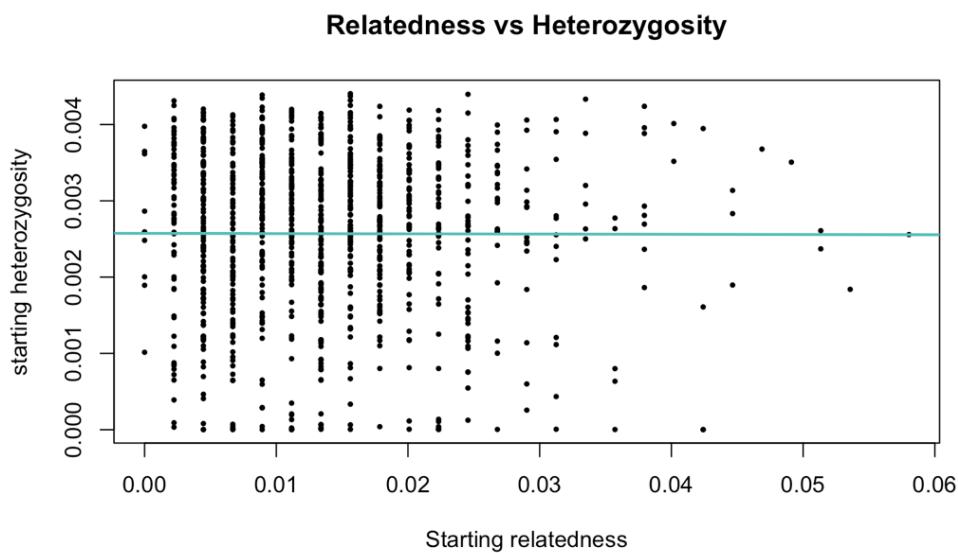


Figure S-19: There is no correlation between starting heterozygosity and starting relatedness (Pearson correlation coefficient $r= -0.0028$, $p=0.9305$). This Graph was run with Model 2 and visualizes the correlation between starting relatedness (8 introduced individuals) and starting heterozygosity. The points represent the data of 1000 repetitions and the blue line the correlation between them. The recombination rate is $1e-9$.

IV. Model 2.1

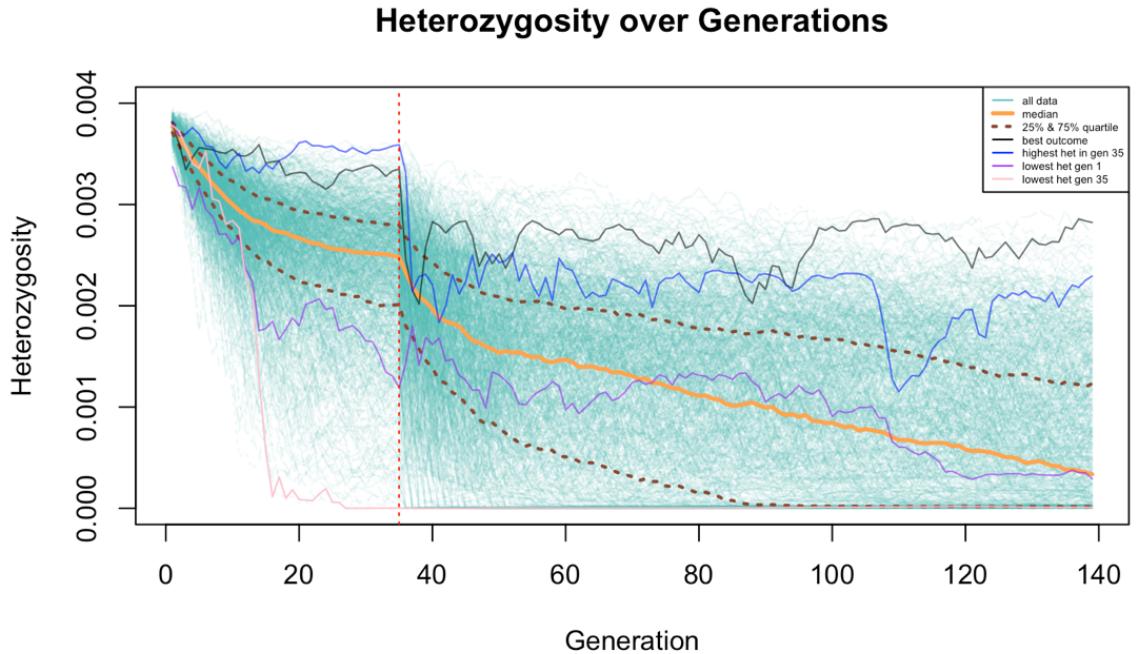


Figure S-20: The overall trend of heterozygosity is declining over the generations, but shows an especially steep fall off in small population sizes (start of the simulation and after the bottleneck in generation 35). Heterozygosity across generations for Model 2 is visualized, where the history of the sihek is implemented. On the x-axis are the generations and on the y-axis the heterozygosity over the whole population. The single lines in turquoise represent single simulations from one of the 1000 repetitions, the orange line the median and the dashed lines show the 25% and 75% quartiles. Some specific lines are indicated in different colors to better visualize the dynamics, namely the one with the best outcome after 140 generations (black), the simulation with the highest (blue) and lowest (pink) heterozygosity values at generation 35 and the one with the lowest heterozygosity at the start of the simulation (purple). The dashed, red line shows the timepoint of the bottleneck (generation 35), which set the population size from 140 to 8 individuals.

