***Disentangling how selection, demography, and sex-biased behavior influence genetic diversity using field study–informed genetic simulations***

Genetic Diversity can be thought of as the collection of all genetic differences found in genomes of a species or a population. This diversity is shaped by behavioral, demographic and selective processes within the population, and some parts of the genome might be more conserved than others thought these forces. For example, some sex-biased behaviors, such as males-only migrations, increases the diversity of the autosomes compared to the X chromosomes, because males carry one X chromosome for every two autosomes. Male dispersal is the main pattern of migration in primate populations and particularly in the Amboseli baboon population[1]. On the contrary, the reproductive skew due to male-rank in primate populations, might increase the X chromosome diversity compared to the autosome, since only few males contribute to the next generations. Those behavioral forces are considered neutral, as opposed to selection, which reduces the diversity of chromosomes depending on mutations selection coefficient. In some cases, such as during hybridization, selection can act preferentially on the X chromosome than in autosomes, as predicted by Haldane's rule, reducing the diversity of the X.

While the genetic diversity of a population reflects past and present behavioral and selective processes, disentangling the effects of each process is challenging in practice, especially in natural populations. Different processes can have the same or opposing effects on diversity, and some processes might interact with one another.

As a solution, we propose using computer simulations, in which we can program the exact processes we want to study. These simulations are made with SLiM. SLiM is an evolutionary simulation framework that combines a powerful engine for population genetic simulations with the capability of modeling arbitrarily complex evolutionary scenarios[2]. With SLiM, we can implement all the demographic and behavior forces to our genetic models in forward simulations. We focus our study on sex-biased behaviors, because they change the diversity of the X and the autosomes differentially and we can then evaluate their influence on genetic diversity. Because selection is also expected to change the X and autosomes diversity differently, we aim to distinguish one force from another. To do so, we use several estimators for the genetic diversity $\theta$[3]:

- $\pi$ (Pi): Measure of the average number of nucleotide differences per site within a sequence (gene, chromosome, genome…). In our case, we calculate the diversity for sexual chromosomes and for one autosome.
- Watterson's $\theta$ (Wa): Measure of the total number of segregating sites in the sample, adjusted by the size of the sample.
- Heterozygosity (h): Calculate from the proportion of all the individuals who would be heterozygotes if there was random union of gametes, for every mutation in the sample and adjusted by the size of the sample.

To compare genetic diversities of the X chromosome and the autosome, we calculate the X-to-A ratio, which is equal to $\theta_X$ / $\theta_A$. In a panmictic population with a sex-ratio (proportion of males) of 0.5, the expected value of the X-to-A ratio is 0.75. In fact, the diversity of a chromosome

is directly proportional to its effective population size, and since in primates there are 3 X chromosomes for each 4 autosomes, the effective population size of the X is ¾ of that of autosomes. In the same way, the expected Y-to-A ratio is 0.25. Deviations from those expected values point to sex-biased processes (which prevent panmixia) or selection acting differently on the X versus autosomes in the population.

The simulations aim to be informative about real-life populations of mammals in general, and primates in particular. To achieve so, we will focus on the Amboseli population of baboons, in which there exists extensive behavioral and demographic data collected for the last fifty years[4]. Such data are very useful to assess what forces are more important in shaping genetic diversity of a primate population. We will also soon have access to the genetic data of the population. We could then compare the results of our simulations with the real X-to-A ratio observed in the population.

As a summary, the goal of the project is to answer these two questions:

- What is the strength of each sex-biased behavior on the X-to-A diversity ratio?
- Are these processes enough for explaining the difference in genetic diversity between the X and the autosomes, or whether instead we need to invoke natural selection?

## I. Influence of behavioral and selection forces on the nucleotide diversity ratio X/A

This project will focus on three sex-biased forces: reproductive skew, male migration, and hybridization, and compare them to Selection. This part aims to present those forces and to assess their expected influence on the X-to-A ratio.

### a. Males Reproductive Skew

In Amboseli, mating success variability in male baboons is very important[5]. Some males, high ranked in the hierarchy, reproduce more than other, low-ranked males. In contrast, the influence of rank on mating success is less significant for females. The high variance in reproductive success in males must decrease the diversity of both the X chromosome and the autosomes, but it will affect more strongly the autosomes than the X chromosome. Indeed, males only carry a single copy of the Xand two copies of each autosome. In other words, if a male does not reproduce, the population will lose the diversity of two copy of any Autosome, but only one copy of X. This effect is even stronger the more the reproductive skew of males and females are different.

**The more the male-biased reproductive skew increases, the more the diversity ratio increases.**

### b. Only-Males Migration

The studied population of baboons at Amboseli is divided into many subpopulations of around 50 baboons. Reproduction is only possible within a subgroup and panmixia is no longer achieved, which must decrease the genetic diversity of the whole population. However,

migrations between subpopulation reduces this effect. In Amboseli, male is the dispersing sex, and male can disperse several times during their adulthood[1]. When a male migrates from a subpopulation to another, it takes two copies of A and one copy of X with him. Thus, with male-biased migration, the A diversity raises more rapidly than the X diversity.

**The higher the probability of migration is, the more the diversity ratio decreases.**

### c. Hybridization

The ancestry of baboons living in Amboseli is derived mainly from Yellow Baboons (*Papio cynocephalus*). However, they live in a hybrid zone with Anubis Baboon (*Papio anubis*), so all individuals carry varying proportions of anubis ancestry. The hybridization can be seen as a migration between two diverged populations. Because only the Anubis males are hybridizing with the Amboseli population, we expect a similar change of the ratio as the only male migration, with a strength even more important because of the divergence between Amboseli and Anubis Baboons.

**The more diverged populations are, the more Hybridization decreases the diversity ratio.**

### d. Selection

Selection tends to remove deleterious alleles and fix advantageous ones of nucleotide sequences which result in reducing their genetic diversity. Some mutations might partly escape selection if they are non-dominant because heterozygote individuals do not express them. But, for non-dominant mutations that are on the X chromosomes, males are haploid for X and selection acts on them. This increases the strength of selection on the X chromosome compared to the Autosomes.

**The more non neutral mutations are, the more the diversity ratio decreases with selection.**

Moreover, when hybridizing some combinations of genomes are very deleterious (or rarely advantageous) for the individuals, and many nucleotide sequences lose their diversity for one genome. We can thus expect a reduction of nucleotide diversity in those sequences. The phenomenon is stronger on X chromosome than on Autosomes because linked to Haldane's rule, the presence of sex chromosome contribute to faster evolution of reproductive incompatibility, and the X chromosomes tends to accumulate more hybridization incompatibilities than others[6].

**The more hybridization increases selection on X than on Autosomes, the more the diversity ratio will decrease.**

## II. SLiM simulations and type of results

### a. Global parameters and possible outputs

In this project, the models simulate a population of **750 individuals**. Each one of them has **one pair of Autosomes**, and **one pair of sexual chromosomes (XX in females, XY in males)**. The chromosomes are **1e8 nucleotides** long and we fix **the mutation rate at 5e-7**, and **the recombination rate at 4e-7**. All mutations are **neutral** and can substitute a nucleotide to another

one with equal probabilities for each possibility. The population has **as many males as females** by default, but the sex-ratio (proportion of males) can be tuned. Every simulation lasts **2e4 ticks**, and a tick equals a generation for Wright-Fisher models, and a year for Non-Wright-Fisher models. **Each reproduction creates 1 offspring**.

For each model we build, we gather useful statistics to monitor several processes. First, we can assess the genetic diversity and the X-to-A ratio of the population with the three estimators: $\pi$ , **Watterson 's Theta** and **Heterozygosity.** The **expected theorical genetic diversity**, equal to **4*Ne*μ**, can serve as a comparison, which we can obtain directly because we know the simulation parameters of the model. Indeed, we choose our population and genetic parameters to match the genetic diversity usually observed in primates, between 4e-4 and 7e-3[7]. The effective population Ne change with the sex-ratio (sr = males effective / females effective):

$$Ne = 4 * \mathrm{N} * \mathrm{sr} * (1 - \mathrm{sr}) \quad (1)$$

In the same vein, the **theorical X-to-A ratio** change with the sex-ratio:

$$\frac{\theta_X}{\theta_A} = 1 - \frac{1}{2} * sr \quad (2)$$

Secondly, the **pedigree** of individuals is also recorded. We can then have access to **the age distribution** of the population, **the age of death,** and **the number of offsprings** for each individual. Base on those data, we can calculate the reproductive skew of males (and females) following this formula [8]:

$$M(r,t) = \tilde{M}(r,t)\, E[\tilde{M}(X,t)] \quad (3a)$$

where,

$$\tilde{M}(r,t) = \frac{N}{R^2} \sum_{i=1}^{N} (ri - \bar{ri})^2 \quad (3b)$$

$$X \sim Multinomial\ (R, \frac{t}{T})$$

With r the reproduction realization, t the time of exposure of all the N sampled individuals. R is the total number of offsprings from the sample, and T the total exposure time. $\bar{ri}$ is the expected number of offspring that individual i would have produced at its age if reproductive rates were perfectly equal within the group[8]. $M(r,t)$ measures the deviation in the number of offspring per individual from a multinomial distribution parameterized by the age structure of the population. Positive values of $M(r,t)$ signify that reproduction is more spread (or skewed) than expected, with more individuals at the extremes (either too many or too few reproductions) than the multinomial expectation. Negative values of $M(r,t)$ mean that reproduction is less spread than the multinomial, with individuals having a more similar number of offspring than expected.

Every model of the project is represented in a mind map in *annex 1*.

### b. The Wright-Fisher model

We first simulate the population with a Wright-Fisher (WF) model, i.e., without overlapping generations. In this case, a tick is equal to a generation, and each offspring has its mother and

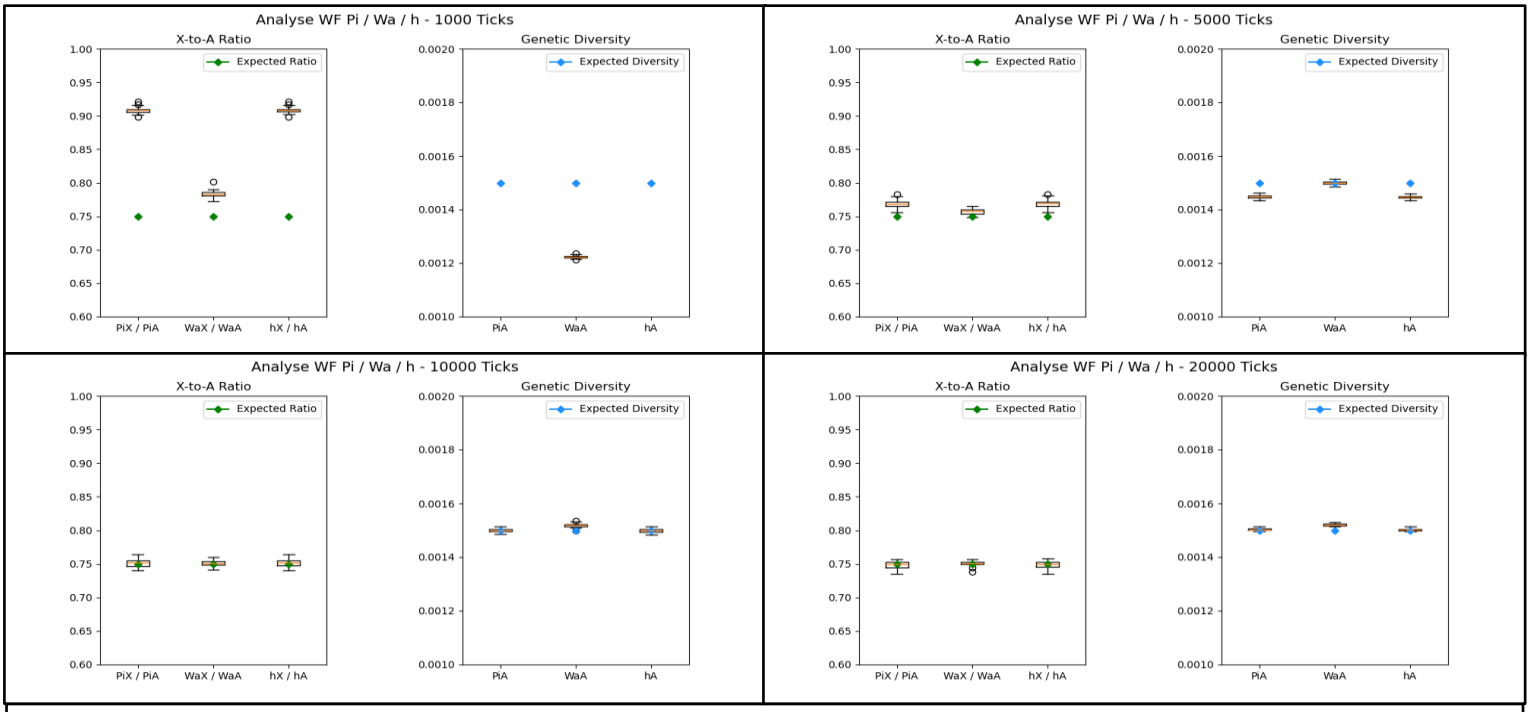



***Figure 1:*** *Boxplots of 100 WF simulations of each total number of Ticks (1000, 5000, 10000, 20000). Left Boxplots represent the X-to-A ratio for π, Wa and h. Green diamonds are the theorical ratio = 0.75. Right Boxplots represent the genetic diversity of the Autosomes for π, Wa and h. Blue diamonds are the theorical diversity = 1.5e-3.*

father chosen randomly within the population. At the beginning of every simulation, all individuals are genetically the same. Consequently, simulations need to last enough time (called a burn-in period) to reach the equilibrium of genetic diversity.

The WF model is used for three objectives: 1) check if π (Pi), Watterson 's Teta (Wa) and heterozygosity (h) are good estimators of genetic diversity, 2) find which one is the best, and 3) find the time of future simulations (beyond the burn-in period). To do so, we simulate populations for many total numbers of Ticks (1000, 5000, 10000 and 20000) and calculate the π, Wa and h values for each chromosome. Each simulation is repeated 100 times. The resulting graphs are shown in *figure 1*.

For 1000 and 5000 Ticks, π, Wa and h greatly underestimate the expected values of ratio and diversity. After 10000 Ticks, **they are good estimators** for both the X-to-A ratio and the genetic diversity. In order to be sure of going beyond the burn-in period, **we will take 20000 Ticks for the future simulations**.

As expected, the genetic diversity, around 1.5e-3, is between 4e-4 and 7e-3. All the three estimators seem to have the same variance, and the same capacity to assess both the X-to-A ratio and the genetic diversity. However, Waterson's estimator might be less efficient when adding selections in our models[9]. In the next simulations, **we will prioritize π has an estimator of genetic diversity and X-to-A ratio**.

### c. The Non-Wright-Fisher Vanilla model

With the non-Wright-Fisher (NWF) vanilla model (NWF_V), we add overlapping generations and a new reproduction system to the former WF model. In this model, a Tick is equal

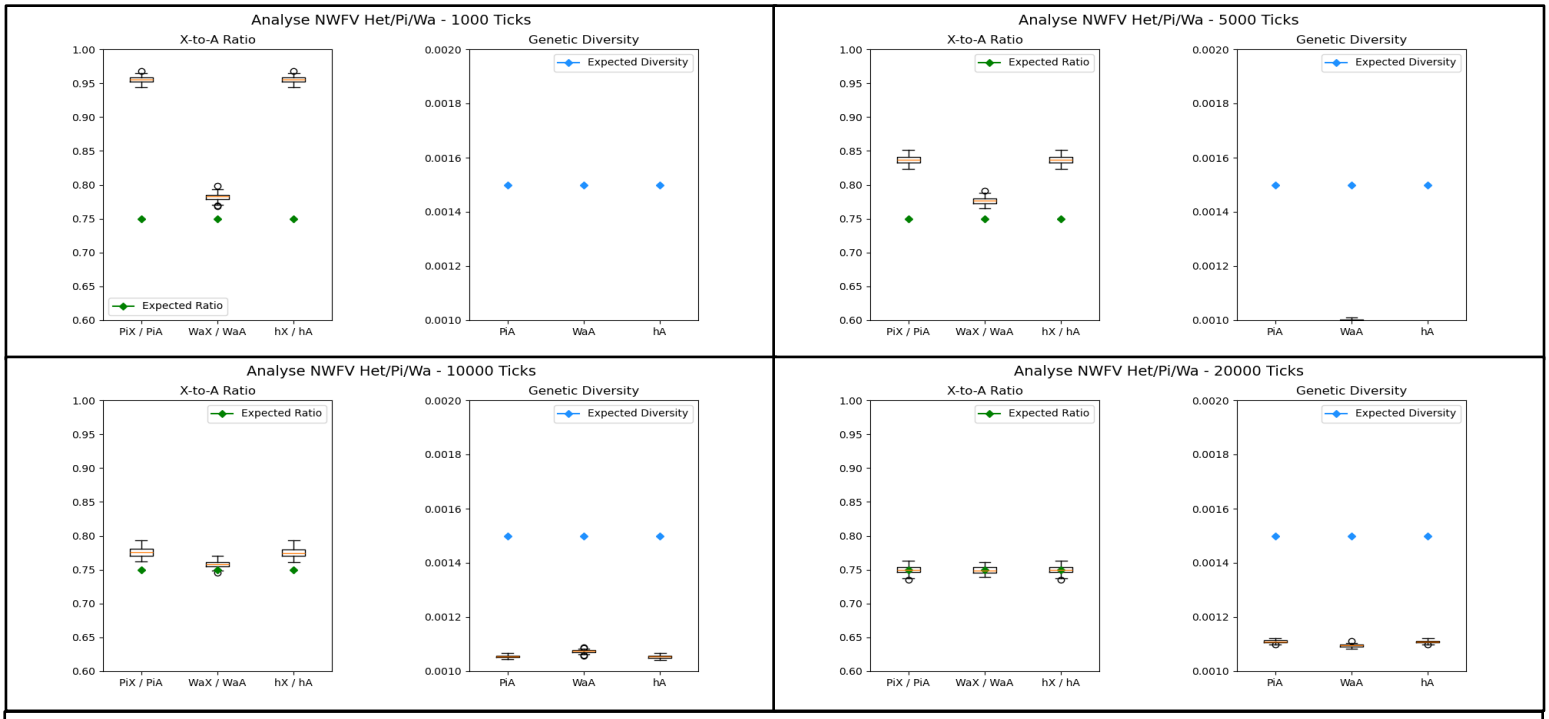



***Figure 2:*** *Boxplots of the 100 NWF_V simulations of each total number of Ticks (1000, 5000, 10000, 20000). Left Boxplots represent the X-to-A ratio for π, Wa and h. Green diamonds are the theorical ratio = 0.75. Right Boxplots represent the genetic diversity of the Autosomes for π, Wa and h. Blue diamonds are the theorical diversity = 1.5e-3.*

to a year. The reproduction is now based on couples. Each Tick, we will form couples until there is no female or male left. In this situation, **every adult reproduces one time (or zero times) a Tick,** which leads to the same number of offsprings for every adult and thus no reproductive skew. The size of population is still maintained at around 750 individuals, thanks to a **carrying capacity regulation**.

We can make the same graph as the WF model, in order to assess the burning period time and the better estimator between h, Wa and π (*figure 2*).

With the NWF vanilla model, the simulations seem to stabilize only after **20000 Ticks**. We will then keep this value for future simulations. Each π, Wa and h are still good estimators of the X-to-A ratio. However, the estimated genetic diversities for the three are below the expected one (around 1.1e-3). This might be explained by the randomness of the population size each Tick, due to the carrying capacity, which increases the genetic drift compared to the WF model[10]. This model is still sufficient because this project is interested in the X-to-A ratio, for which we recover values around the expectation. Moreover, the genetic diversity is still between 4e-4 and 7e-3. **We will keep the π estimator**.

### d. Age structure and reproduction systems

To more accurately reproduce the Amboseli population, we have created two new NWF models, changing the reproduction system:

- The first one is the NWF female choice model (NWF_FC). In each Tick, all females reproduce once, and they choose their mate randomly from the male population. Thus, a

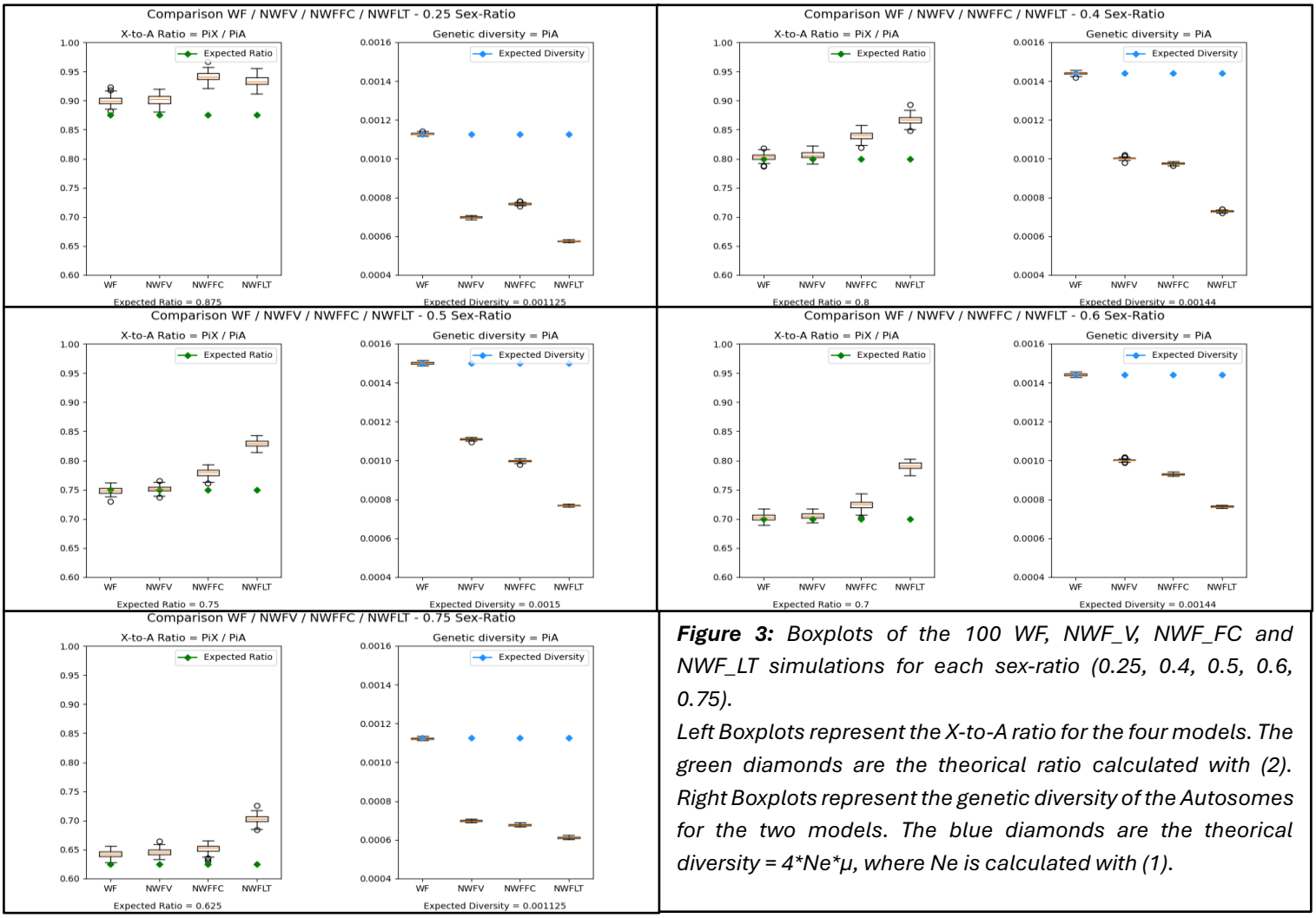



***Figure 3:*** *Boxplots of the 100 WF, NWF_V, NWF_FC and NWF_LT simulations for each sex-ratio (0.25, 0.4, 0.5, 0.6, 0.75).*
*Left Boxplots represent the X-to-A ratio for the four models. The green diamonds are the theorical ratio calculated with (2). Right Boxplots represent the genetic diversity of the Autosomes for the two models. The blue diamonds are the theorical diversity = 4*Ne*μ, where Ne is calculated with (1).*

male can mate several times in a Tick, while others will not reproduce. This mimics the reproductive system of primates, where females may only reproduce once per year.

- The second one is the NWF Life Table model (NWF_LT) which adds an age structure to the population. It uses mortality tables (given by Fernando) for males and females, and a fecundity table only for females[11]. Females live a maximum of 26 and start reproducing at 4, males live a maximum of 23 years and start reproducing at 5 (*annex 2*). Each Tick, all females reproduce one or zero times depending on their fecundity. They choose their mate randomly within the adult male population (age ≥ 5), all with the same probability.

We want to assess the capacity of these models to be a reference for the subsequent simulations. To do so, we want to see how accurately they assess the expected X-to-A ratio, and the expected diversity, compared to the WF and the NWF_V models. We will make different simulations with different values of sex ratio (0.25, 0.4, 0.5, 0.6, 0.75). The results of the 100 simulations for every model are shown in *figure 3*.

For a sex-ratio of 0.5, we obtain the same results as before for both WF_V and NWF_V, the estimations of X-to-A ratio are good. However, when moving away from the symmetrical ratio, estimations become higher than expected even bad for the 0.25 and 0.75 sex-ratios. Because we will later keep a sex-ratio of 0.5, the skewing from the expected for the extreme ratios should not affect us for the next simulations. Even if the sex-ratio in NWF models is not fixed through the

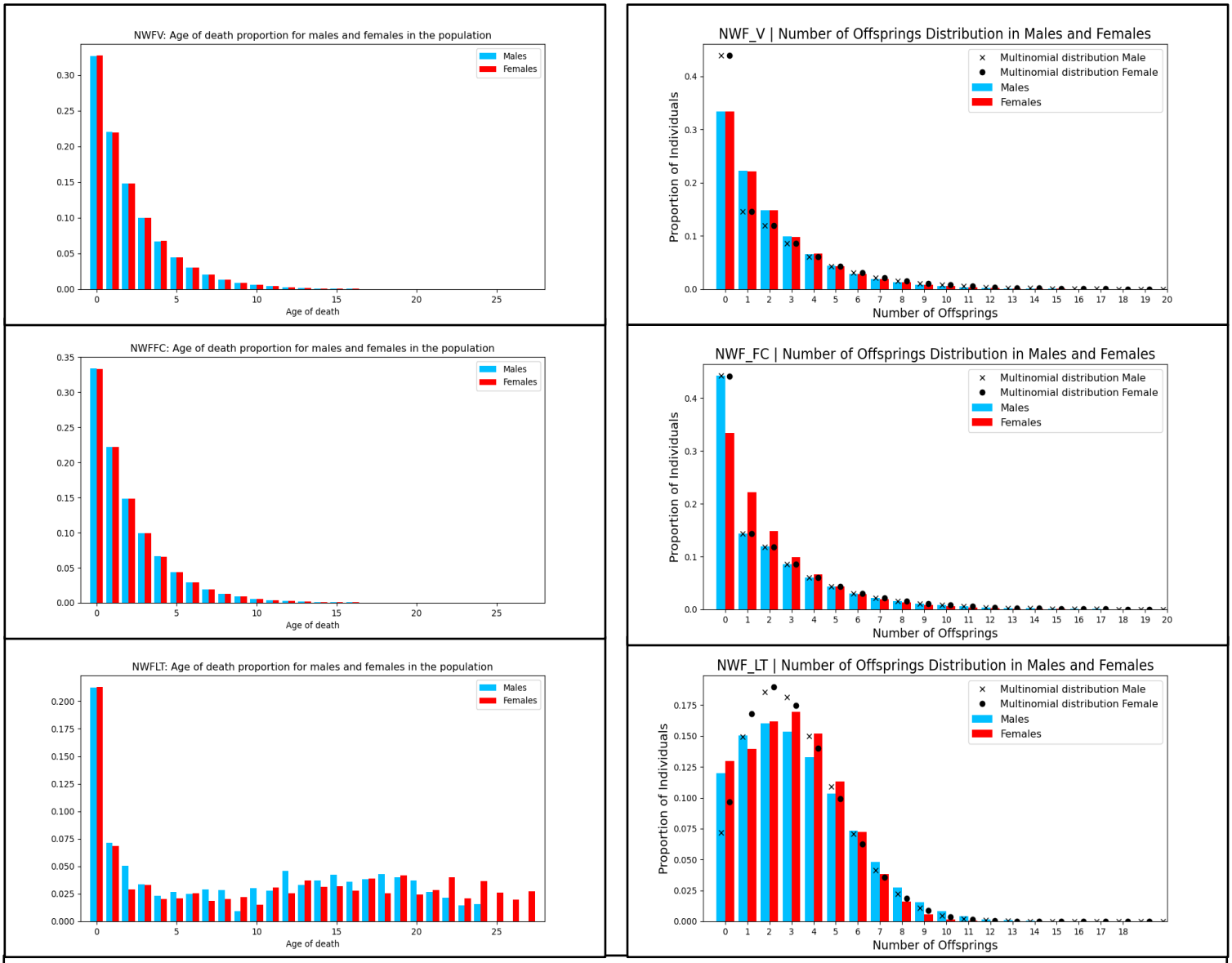



***Figure 4:*** *Bar graphs of Age of Death (left) and Number of offsprings (right) proportions for simulations of the three NWF models: NWFV, NWFFC and NWFLT. The simulation last 10000 Ticks with a burning period of 1000 Ticks.*

years, it varies around 0.1 of the given parameters, so $\pi$ is still a good estimator of the X-to-A ratio. Increasing the sex-ratio reduces the expected X-to-A ratio. In fact, a high sex-ratio will decrease the proportion of X chromosomes in the population. The sex ratio doesn't influence the accuracy of the genetic diversity estimations. WF has always a good estimation of the diversity, whereas NWF_V is always below the expected.

The estimations of the X-to-A ratio are worth for the NWF_FC and the NWF_LT models. They are always higher than the ones of the two previous models and the expected ratio. NWF_FC become slightly better with high sex ratio, which is not the case for NWF_LT model. There are also differences between the NWF models in assessing the genetic diversity of the population. In fact, the diversity of the NWF_V models is higher than the NWF_FC one (except for a sex ratio of 0.25), which is higher than the NWF_LT.

To understand what can explain the differences between the NWF_V model and the two new NWF ones, we studied the demography of the populations. We analyzed the age of death and the

| Reproductive Skew of males and females for each NWF models | | | | |
|---|---|---|---|---|
| | **Males** | **Females** | **Maximum** | **Minimum** |
| **NWFV** | -0.487533 | -0.465266 | 5044 | -0.506629 |
| **NWFFC** | -0.000159 | -0.496887 | 4937 | -0.499141 |
| **NWFLT** | -0.000880 | -0.077242 | 3271 | -0.325602 |

***Figure 5:*** *Table of measure of Reproductive Skew of every NWF model for males and females, and for the maximum and maximum cases. Values are calculated using (3a).*

number of offsprings of every individual, for both males and females, during the last 9000 Ticks of the simulation. For the NWF_LT models, we count the number of offsprings only when an individual reaches adulthood (5 years for males and 4 for females). The results are showcased in *figure 4* (previous page).

As expected, the age of death between males and females for the NWF_V and the NWF_FC are the same. In fact, the simulation model does not include differences of mortality between males and females, or mortality depending on the age. Every individual has the same probability of dying each Tick, and around 33% of individuals die during their first year. This is not the case for the NWF_LT, because of the sex and age-depending mortality. Females live longer than males, and male mortality in the first year is higher than female one. More interestingly, we can see differences in the number of offsprings between males and females for the two new models. This is due to the reproduction system, where a female can reproduce one time per year at most, whereas males may reproduce several times per year. Thus, compared to females, males have a higher variance in reproductive success than females, which creates an unannounced reproductive skew. These can be easily seen on the NWF_FC graph, where more males than females have 0 offsprings, but males also have high number of offsprings (15,16,17…). The difference in age structure and in reproduction system explains the differences observed for the X-to-A ratio between the vanilla model and the two others. Therefore, even in the absence of explicit modelling the reproductive skew, these simulations show that the reproductive system commonly seen in primates and in mammals more generally is sufficient to create a reproductive skew that can be quantified using the genetic diversity of the population.

To numerically quantify the reproductive skew of the different models, we use the equation (3a). We calculate $M(r,t)$ for males and females, for the three NWF models. Results are shown in *figure 5* (next page). To understand the ranges of values of $M(r,t)$, the minimum and maximum $M(r,t)$ were also computed, based, respectively, on a case where every adult has the same number of offsprings and a case where only one adult produces all the offsprings. They use the same parameters R, N, T and t of the NWF models, with only varying r. A reproductive skew equal to zero means that reproduction follows a multinomial distribution (everyone has the same probability to reproduce). A value below zero means that there is less variance in the number of reproductions in individuals compared to the multinomial distribution, and that everyone reproduces the same. Finally, a higher value than zero means that reproduction is positively skewed, i.e. some individuals have more chances to reproduce than others.

As expected, the reproductive skew for both males and females in the NWF_V model are quite the same. They are below zero and near the minimum values. Reproduction has low variance, which is coherent with the reproduction system. The same result is observed for the females of

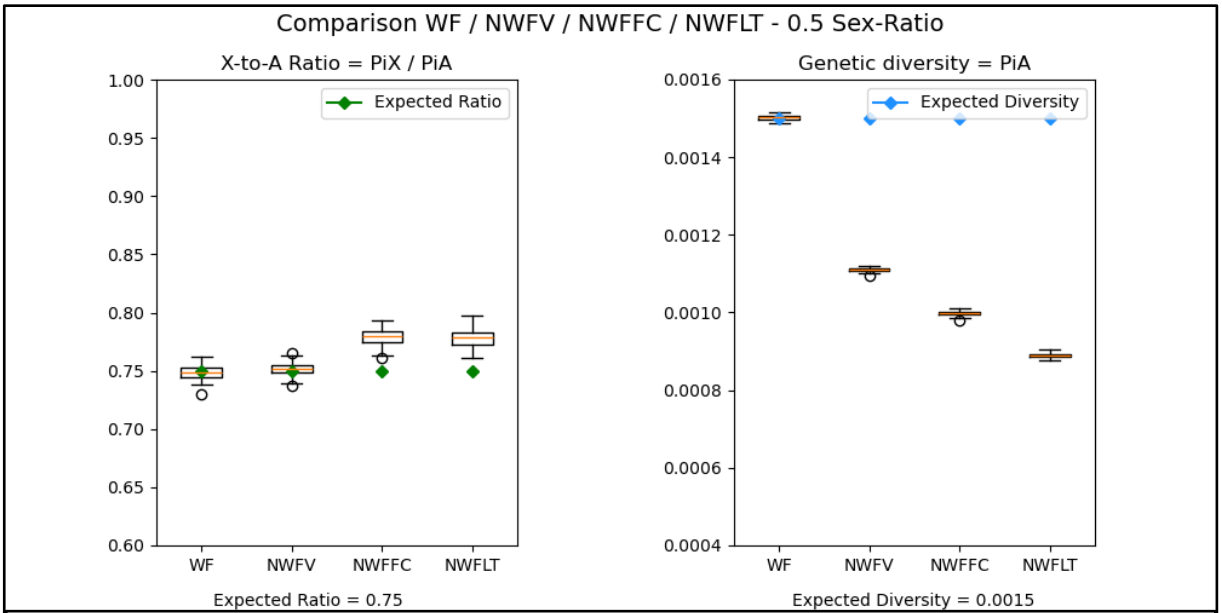



***Figure 6:*** *Boxplots of the 100 WF, NWF_V, NWF_FC and NWF_LT with corrected burn-in period simulations for sex-ratio = 0.5*
*Left Boxplots represent the X-to-A ratio for the four models. The green diamonds are the theorical ratio calculated with (2).*
*Right Boxplots represent the genetic diversity of the Autosomes for the two models. The blue diamonds are the theorical diversity = 4*Ne*μ, where Ne is calculated with (1).*

the NWF_FC model. In fact, in this model females all reproduce once. This is not the case for the males, who are chosen randomly to mate with a female with equal probabilities, which explains the value near zero of the reproductive skew. The conclusion is the same for the males in NWF_LT. More surprisingly, females in NWF_LT have a higher value of reproductive skew than females in NWF_FC. This might be explained using the fertility for females in NWF_LT, which makes some females randomly not reproduce each year. Reproductive skew of Females in NWF_LT is in between a model with low variance and a multinomial model.

At the end of the internship, we discovered a problem in the NWF_LT model (that we couldn't solve for all our simulations because of time). In fact, the burn-in period for NWF_LT should be 60000 Ticks and not 20000 Ticks. The previous diversity values we see for this model is not at equilibrium, and the diversity is higher than it should be. A new graph which the corrected NWF_LT model is shown in *Figure 6*.

We see that the equilibrium X-to-A ratio in NWF_LT is below the one with a burn-in period of 20000 Ticks. The value matches the one with NWF_FC.

To put it in a nutshell, the differences observed between the two new NWF models and the NWF_V model might be explained by passing from an all-equal reproduction to a random reproduction. NWF_FC and NWF_LT were created to add more complexity to NWF_V, with the aim of more realistically modeling the observed demography and behaviors of the Amboseli baboon population.

In the next parts, we will add the forces that influence the X-to-A ratio. As a reference, we use the NWF_LT model. However, some of the results presented (for migration and selection) are done with the uncorrected model (low burn-in period). In those cases, the values of genetic diversity are wrong because we are not at equilibrium, but we can still look at the tendency

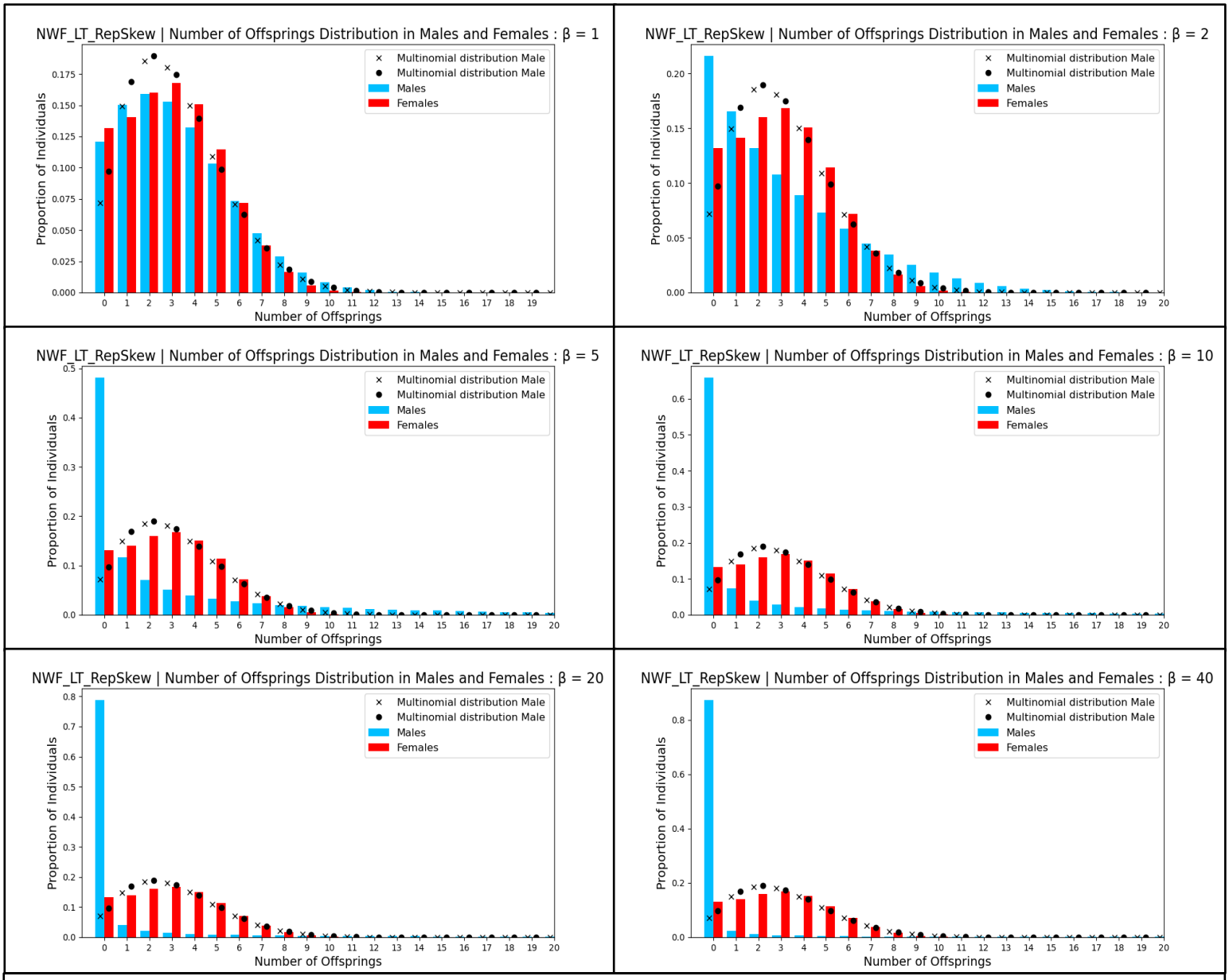



***Figure 7:*** *Bar graphs of Number of offsprings distribution for reproductive skew simulations with β equal 1, 2, 5, 10, 20 and 40. The simulation last 10000 Ticks with a burning period of 10000 Ticks. The distribution when matings follow a multinomial distribution (with the same parameters N, R, $t_i$ and T [eq 3] as the reproductive skew models) is shown.*

caused by the forces. For reproductive skew and hybridization, the corrected NWF_LT (60000 Ticks burn-in period) is used.

### e. Reproductive skew

The first sex-biased behavior we add to our model is a positive reproductive skew in males. We use the **NWF_LT model as a reference**. We model the probability of a male to be chosen by a female following a **Beta-distribution** (**α = 1** and β is tunable) depending on its rank. The rank of a male is a quantile between 0 and 1 (the number of quantiles is equal to the number of males in the population). Males are ranked according to a number (Tag) drawn between 0 and 1 at their birth. **Males keep their tag value during all their life**. So, a male with a high tag value will always

| Reproductive Skew mean of males and females for each value of Beta | | | | |
|---|---|---|---|---|
| | **Males** | **Females** | **Maximum** | **Minimum** |
| **β = 1** | 0.009701 | -0.073462 | 3283.122449 | -0.328291 |
| **β = 2** | 0.424480 | -0.072683 | 3295.061224 | -0.329395 |
| **β = 5** | 2.223916 | -0.075689 | 3284.612245 | -0.328614 |
| **β = 10** | 5.278182 | -0.074994 | 3285.530612 | -0.328832 |
| **β = 20** | 11.292585 | -0.073673 | 3288.551020 | -0.328719 |
| **β = 40** | 23.155331 | -0.074022 | 3289.489796 | -0.329085 |

***Figure 8:*** *Table of measure of Reproductive Skew of every reproductive skew model for males and females, and for the maximum and maximum cases, with β equal 1, 2, 5, 10, 20 and 40. Values are calculated using (3a).*

be in the highest rank and will reproduce more than the average of males. Females reproduce depending on their fertility and choose a male only if the reproduction is successful.

The first thing to check is if a positive reproduction appears in males. To do so, we output the Number of offsprings graphs for males in females for each value of β (1,2,5,10,20 and 40) in *Figure 7* (previous page).

We observe that when β = 1, the distribution is very similar to the NWF_LT (without skew) distribution. This is expected because, when β = 1 every male has the same probability to be chosen as a mate, as in NWF_LT. Then, the more β increases, the more male number of offsprings is skewed. The proportion of males with 0 offsprings and the maximum number of offspring for a male increase (*Annex 3*). On the opposite, female reproduction seems to approach the multinomial values as expected. To validate these observations, we can look at the reproductive skew estimator, shown in *Figure 8*.

As before, the result for β = 1 show no positive reproductive skew in males because M is around zero. M is positive which differs from the negative value of M for NWF_LT, but NWF_LT is still on the standard deviation of β = 1. When β increases, the reproductive skew increases in males and not in females as expected. So, our modulization of reproductive skew in males is efficient! Another thing to notice is that, even with β = 40 and a maximum number of offspring for a male of 294 (*Annex 3*), the estimator is still far below the maximum value. This can be explained by the fact that M is calculated over many generations (around 10000 generations) in our simulations. But the maximum value is calculated as if a male has all the offsprings of all the generations, which is not possible in our work.

Now that we know our model good to simulate a positive skew in male reproduction, we can look at the influence of this skew to the X-to-A ratio and the diversity. Graphs are shown in *Figure 9 (next page)*.

As expected, the genetic diversity is reduced when the reproductive skew increases, because we lose the diversity of males that do not reproduce. Moreover, the diversity decreases slower than the autosome diversity as the increasing X-to-A ratio shows. In Amboseli, we expect the maximum number of offspring for a male to be below 50, and so the reproductive skew might be modeled with a value of β between 2 and 5. Even for these values, the X-to-A ratio is increased compared to the NWF_LT model.

To better match the Amboseli population, we try two other models of reproductive skew which give us the same result as before (increase of the ratio and decrease of the diversity with β). In the first one, we add a decreasing tag value with age because in Amboseli the ranking of males decreases with time[4]. To do so, males are given a tag value between 0.5 and 1 at their birth,

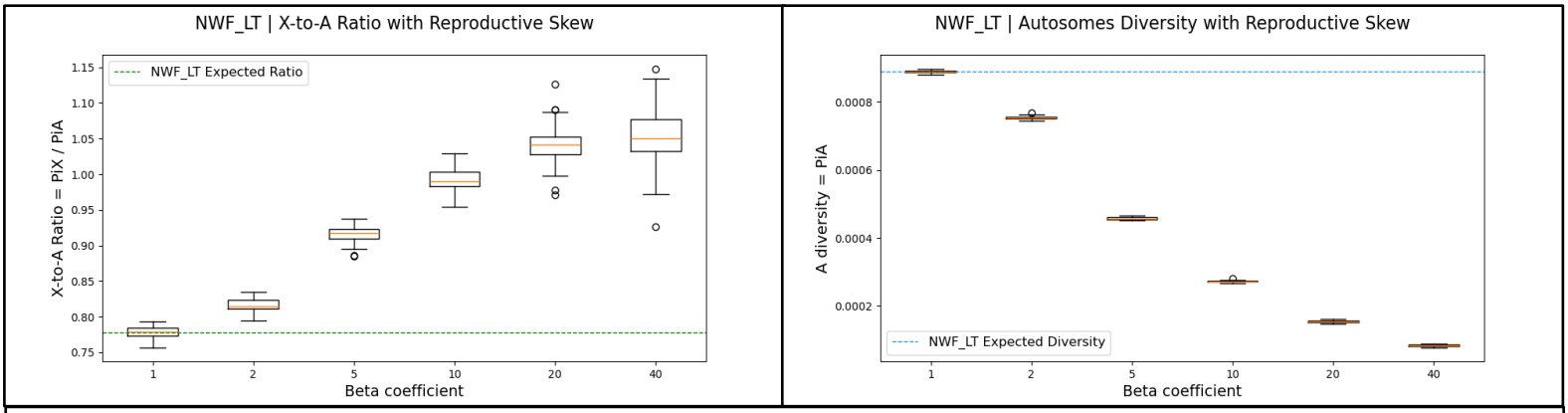



***Figure 9:*** *Boxplots of the 50 Reproductive skew simulations for each β (1, 2, 5, 10, 20, 40).*
*Left Boxplots represent the X-to-A ratio, and the green line is the theorical ratio for the NWF_LT model.*
*Right Boxplots represent the genetic diversity of the Autosomes, and the blue line is the theorical diversity for the NWF_LT model.*

and this value decrease of 10 percent each year when adult. In the second one, we divided the populations in 15 subpopulations with a circle migration system (see f. Migration) and a probability of migrations for males when adult of 1.0. This aims to study smaller populations of 50 individuals.

### f. Migrations

The second sex-biased behavior we study in our simulations is only male migration. We start from the NWF_LT model (without positive reproductive skew). The whole population is divided into **subpopulations** with a **carrying capacity of 50 individuals**, which is the average size of groups encountered in Amboseli[12] (before the decline in the 60's). To keep **the size of the metapopulation equal to 750 individuals**, we use **15 subpopulations**. **Only the males migrate, always with the same probability** every year. The number of migrants in the whole metapopulation follows a **Poisson distribution** on all males, where λ equals the probability of migration. This probability is the tunable parameter to assess the strength of migration. The selected migrants will "choose" their future subpopulations depending on the migration model. Three migration models are coded (see *figure 10 next page*):

- **Star migration**: Migrant can go to any subpopulation with the same probability except the one it was in before.
- **Circle migration**: Migrant can only go in one of the two adjacent subpopulations of the one it was in before, with the same probability.
- **Steppingstone migration**: Migrant can only go in one of the two adjacent subpopulations of the one it was in before, with the same probability. The first and last subpopulations of the metapopulation are not linked.

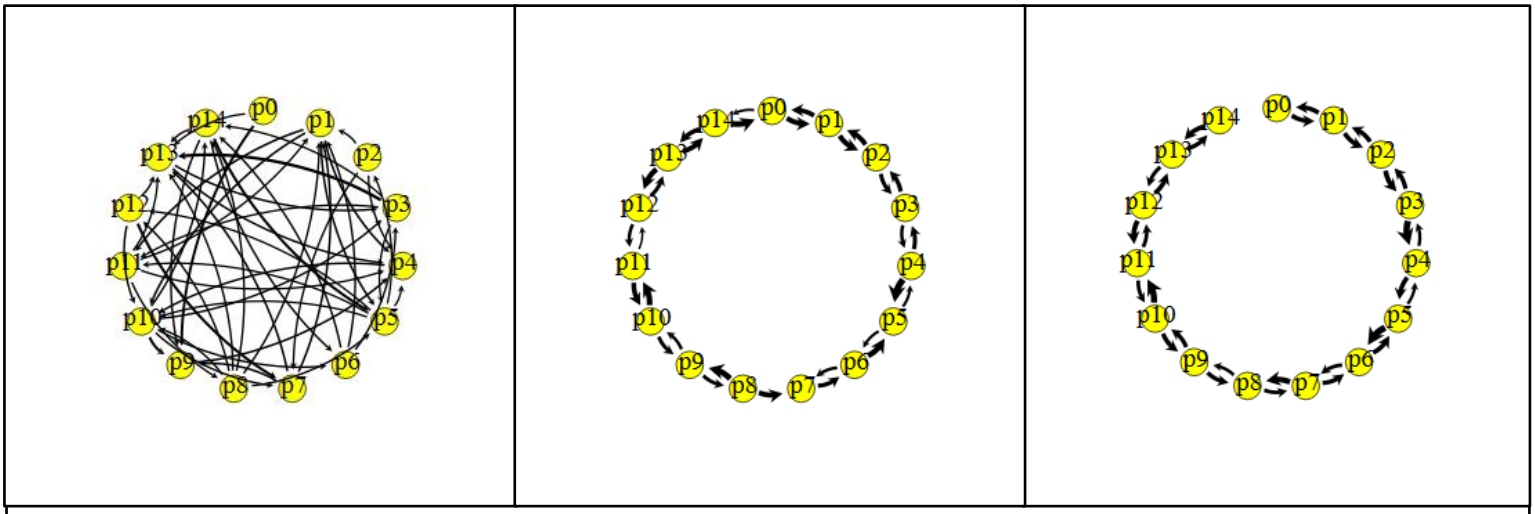



***Figure 10:*** *Visualization of the three migration models: star migration (left), circle migration (middle) and steppingstone migration (right).*

As the population is now divided into many subpopulations, the former value of genetic diversity 4*Ne*μ is no longer true. The value of diversity will be migration dependent. We can define a formula of π for one subpopulation ($\pi_{SG}$) and one for the whole population ($\pi_{TG}$). The two formulas are based on *Laporte, Charlesworth 2002*[13].

$$\pi_{SG} = 4 * n * Ne_G * \mu$$

$$\pi_{TG} = \pi_{SG} * (1 + \frac{(n-1)^2}{n^2} \frac{1}{4 * m_g * Ne_G}) \quad (1b)$$

Where G refers to the studied chromosome (A or X), n is the number of subpopulations, $Ne_G$ the effective population of G in one subpopulation and $m_G$ the migration factor depending on G. $m_G$ follows those formulas for A and X:

$$m_A = m_Z + \frac{1}{2}(m_M + m_F) \quad (2a)$$

$$m_X = m_Z + \frac{1}{3}(m_M + 2 * m_F) \quad (2b)$$

Where $m_Z$ is the zygote migration rate, $m_F$ the female migration rate (both are equal to zero in our simulations) and $m_M$ the male migration rate which is our tunable factor.

We can evaluate the X-to-A ratio of the whole population for a range of migration probabilities between 0.05 and 1 (0.05 steps) and compare it with the expected $\pi_{TG}$. Similarly, we can compare the X-to-A ratio of one subpopulation with the expected $\pi_{SG}$. To check if we correctly simulate migration, we first look at NWF_V X-to-A ratio. Graphs are shown in *figure 11(next page)*.

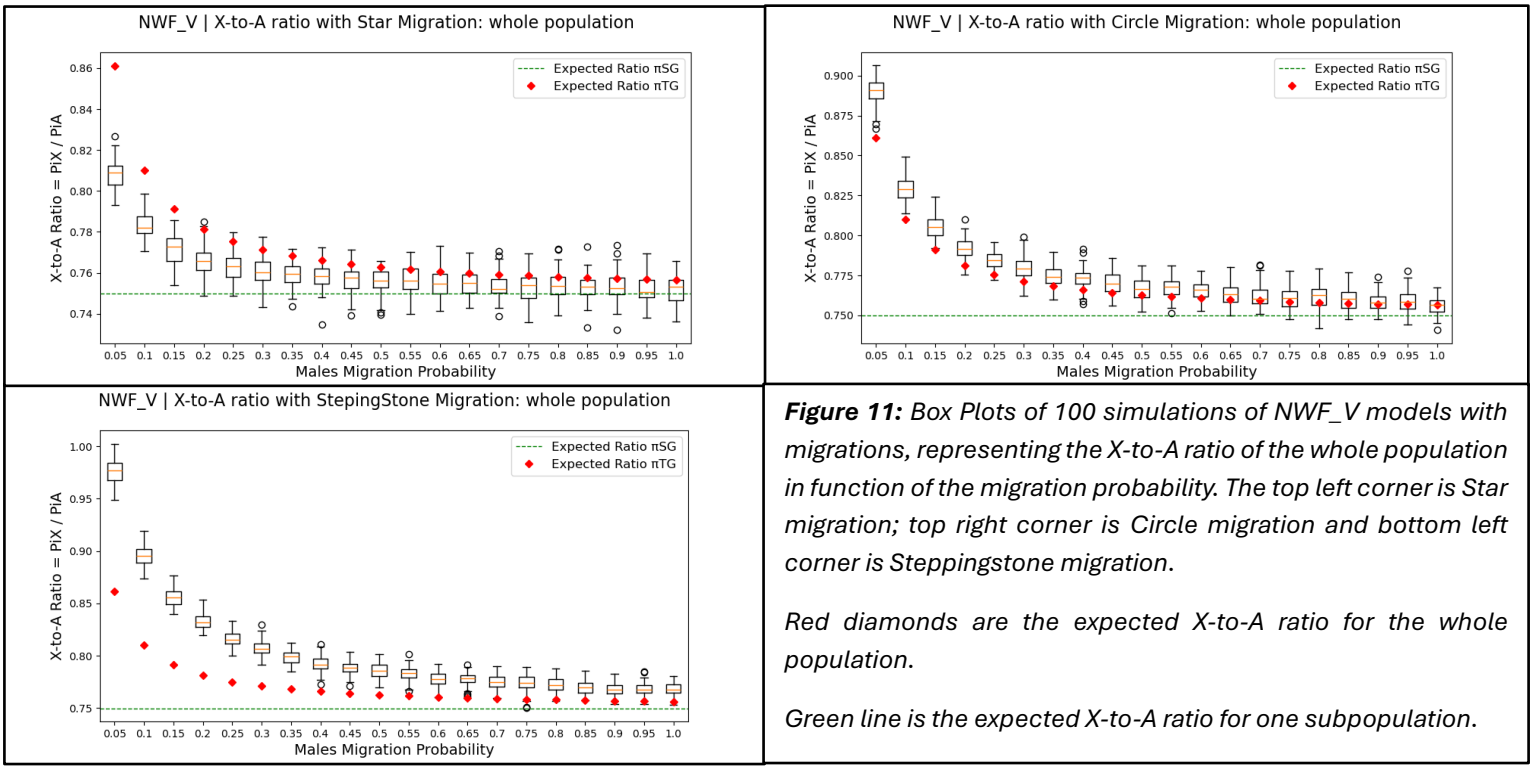



***Figure 11:*** *Box Plots of 100 simulations of NWF_V models with migrations, representing the X-to-A ratio of the whole population in function of the migration probability. The top left corner is Star migration; top right corner is Circle migration and bottom left corner is Steppingstone migration.*

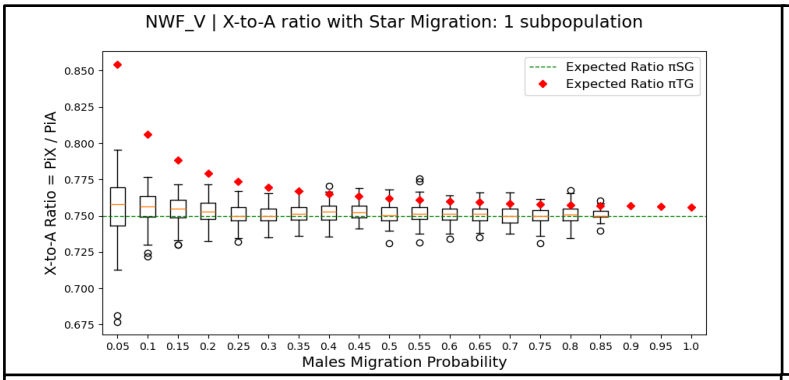
*Red diamonds are the expected X-to-A ratio for the whole population.*

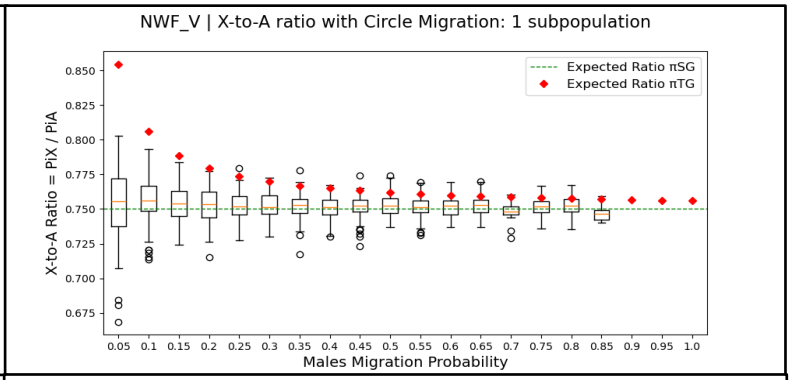
*Green line is the expected X-to-A ratio for one subpopulation.*

We observe that the X-to-A for all three migration models match well the dynamic of the expected ratio in the whole population. In circle migration, the ratio is even close to the values of the expectation. We see a decrease in the X-to-A ratio when the probability of migration increases but not how we expected it. In fact, with high migration probability, the ratio is equal to 0.75 without migration and not underneath. A subdivided population with a high probability of migration is equivalent to one population with the same number of individuals. When the migration rate is high, all the diversity is shared among all the populations, but the subpopulations are not diverging. This is why we observe an increase in the X-to-A ratio when the migration rate is low, because subpopulations are more diverging. In this case, only male migration is still enough to mix all the Autosomes of all the subpopulations and so the diversity is shared across all the metapopulation. This is not the X chromosomes because males carry only one copy of X and subpopulations diverges more on X than on Autosome and the ratio increase.

We can then look at the graphs of the X-to-A ratio for one population only, to see if it matches the expected values for the whole population. Graphs are shown in *Figure 12* (next page).

For Star and Circle migrations, X-to-A ratios match the expectation. There is more variance when migration probabilities are low. However, we observe a decrease in the ratio in Steppingstone migration when migration rate decreases, and variance is very high. In this case, we are looking at subpopulation 1 which is at one subpopulation from extremity. Because of that, this subpopulation is more diverging than others in the center of the line, because few migrants are coming from the other extremity. With low probabilities, we are then approaching the hybridization models when migrations append between diverging populations which is expected to decrease the X-to-A ratio. If we look at more “centered” subpopulations we observe no decrease of the ratio with low probability.



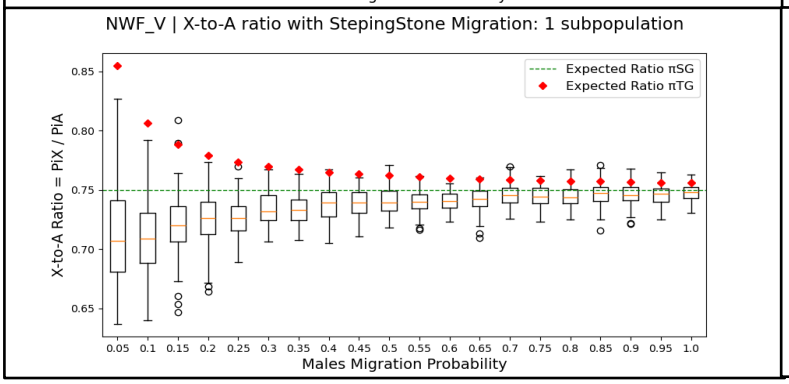
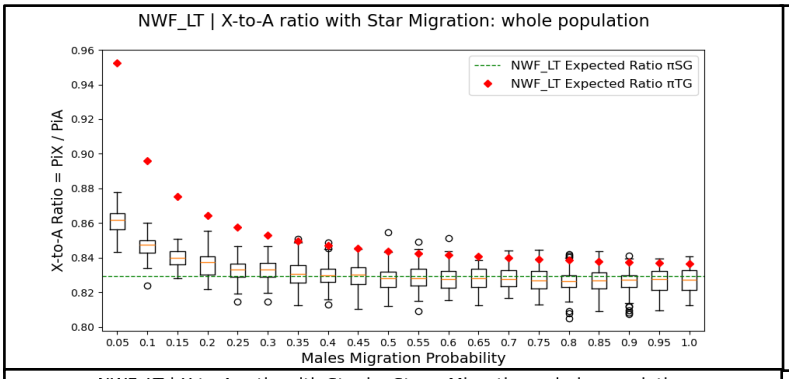
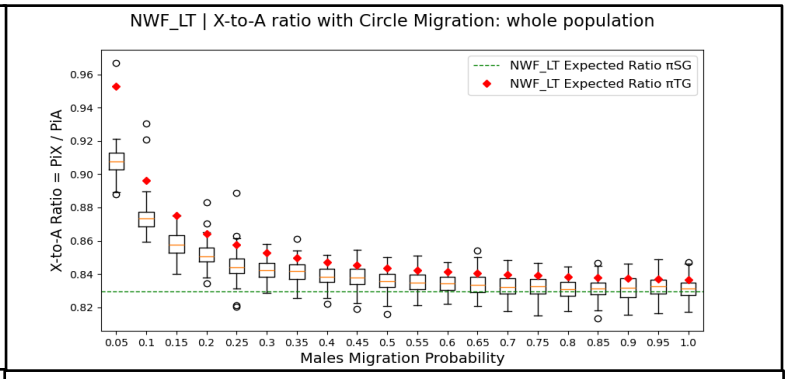




***Figure 12:*** *Box Plots of 100 simulations of NWF_V models with migrations, representing the X-to-A ratio of one population in function of the migration probability. The top left corner is Star migration; top right corner is Circle migration and bottom left corner is Steppingstone migration.*

*Red diamonds are the expected X-to-A ratio for the whole population.*

*Green line is the expected X-to-A ratio for one subpopulation.*



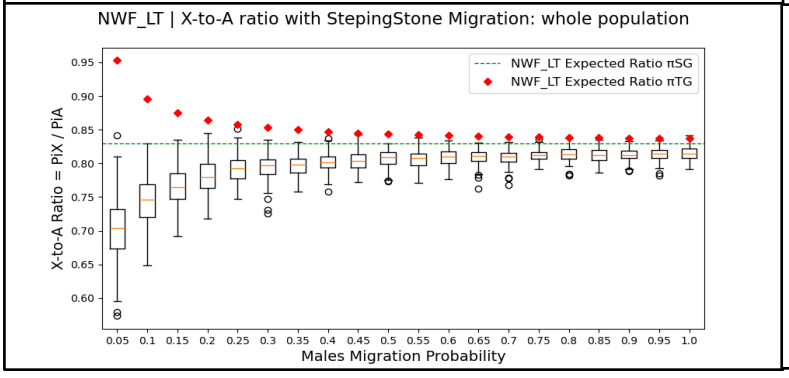




***Figure 13:*** *Box Plots of 100 simulations of NWF_LT models with migrations, representing the X-to-A ratio of the whole population in function of the migration probability. The top left corner is Star migration; top right corner is Circle migration and bottom left corner is Steppingstone migration.*

*Red diamonds are the expected X-to-A ratio for the whole population.*

*Green line is the expected X-to-A ratio for one subpopulation.*

Other verifications of our migration models have been made. We simulated migration with only females or with males and females, or with only one migration in a life of an individual. In all cases, the results were satisfying (*annexes 4, 5 and 6*).

Now that our migration models are satisfying, we can look at the NWF_LT model. Graphs are shown in *Figure 13*. Here, we use the uncorrected model (with low burn-in period), so we should look at the tendency but not the effective value of the X-to-A ratio.

We observe the same results as with the NWF_V model. The X-to-A ratio increases when the migration probability decreases. We observe a slighter increase than with NWF_V. Moreover, the Steppingstone X-to-A ratio decreases with migration probability which is not expected.

In Amboseli, all the males migrate when reaching adulthood, and they can migrate several times during their life. So, the migration probability between subpopulations is high, near one. We also expect the migration process to be as the star one, there is many possible subpopulations for a male to migrate in. Our simulations show that for these values the X-to-A ratio is not expected to change compared to one population with no migration, and so only male migration cannot explain a decrease on the ratio in Amboseli. But we also see that if populations are enough diverged, the only male migration may have an impact. This is what we will look at in the next part, about hybridization.

### g. Hybridization (without selection)

In this part, we simulate the arrival of Anubis male baboons in Amboseli. Simulations of hybridization are done in 3 periods, based on the NWF_LT model:

- We start with an **ancestral population (P0) of 750 individuals** (no positive skew, no migration), for a **burn-in period time of 60000 Ticks**.
- Then, **half of the population migrates to a new one (P1)**. This is the **Separation** event. The carrying capacity of P0 and P1 are still 750 individuals. The two populations **diverge for a tunable number of ticks**, with no migration or reproduction between them.
- Then, a **tunable number of males migrate from P1 to P0**. This is the Hybridization event. After this, P0 **drifts for 30000 Ticks.**

We tracked the diversity from the Separation to the end of the simulation. We can then follow the evolution of the X-to-A genetic ratio through the events. Graphs are shown in *figures 14, 15 and 16 (*next pages) respectively for a divergence time of 10000,20000 and 30000 Ticks. Other results can be found in *annexes 7 and 8*.

When looking at the differents graphs of one divergence time, we can compare the effect of different number of migrants on the X-to-A ratio. With high number of migrants, we observe a sharp decrease of the ratio compare to low number of migrant. Because the more migrant there is, the more new diversity is commming in the Autosome than on the X. One migrant has a very low influence on the ratio, and there is no decrease (exxcept a small one with 30000 years of divergence). More over, we observe a positive spicke the exact moment of hybridization with 300 migrants. This might be explain by the change of sexratio in the P0 (300 only males are arriving), and the removal of individuals with the carrying capacity.

Moreover, the divergence time also influence the X-to-A diversity. In fact, the more diverged two population are, the less diversity they share. So A migrant will bring lot more new mutations if coming from a diverged population than a no-diverged one. And because, migrants are males they bring more diversity on the autosome than on the X. We observe this results in the graphs. In fact, the drop down of the ratio is sharper with a 30000 years of divergence than a 10000 one. With a 10000 years one, the ratio never go below 0.75, but with 30000 years it does.

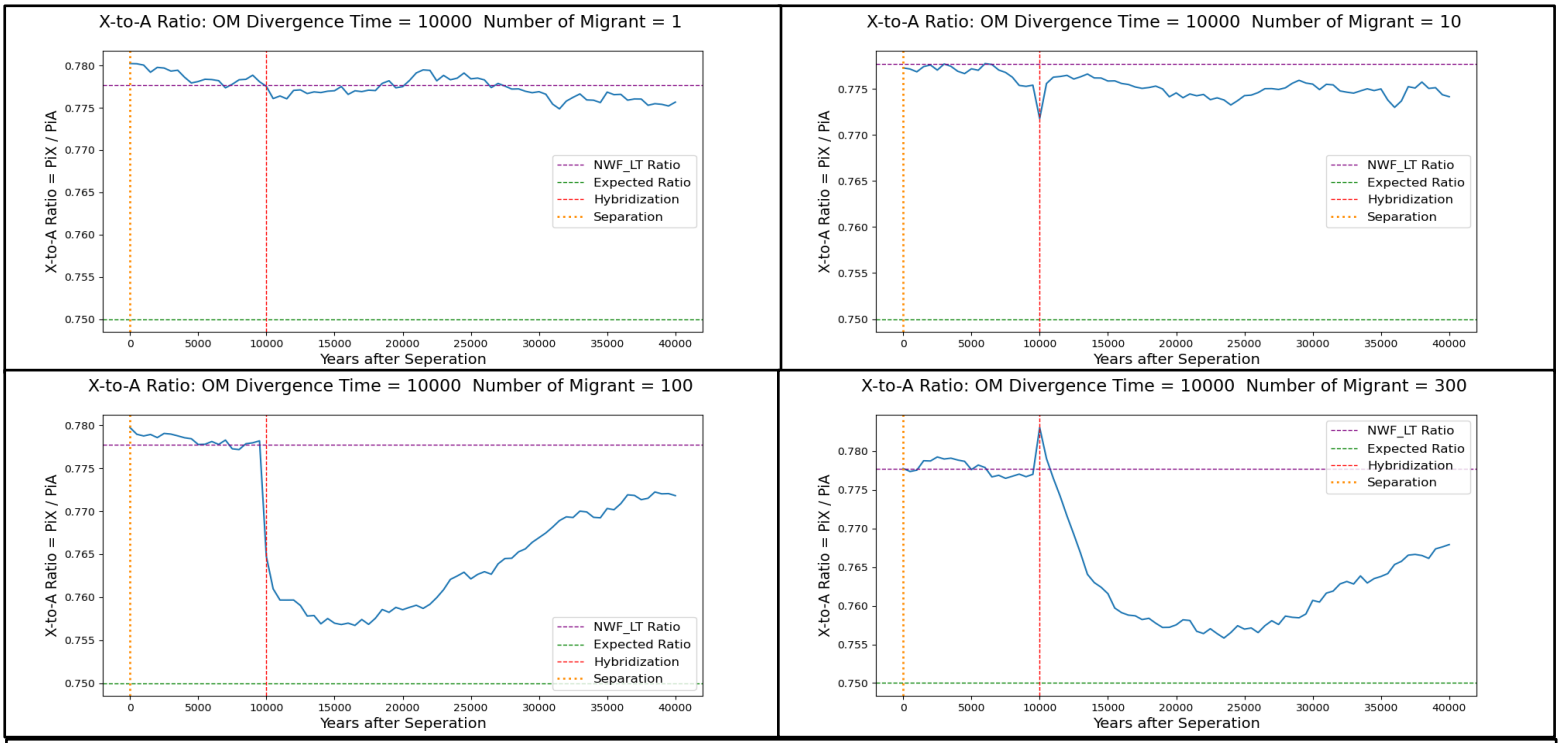



***Figure 14:*** *X-to-A genetic ratio mean of 20 NWF_LT (corrected) simulations with hybridization from the separation to the end of simulations (a value every 500 Ticks). Simulations are done with a time of divergence equal 10000, and four different numbers of migrants: 1, 10, 100 and 300. The yellow and red lines are respectively the separation event and the hybridization event. The purple line is NWF_LT expected ratio at equilibrium with a corrected Burn-in period. The green line is the expected ratio equal 0.75.*

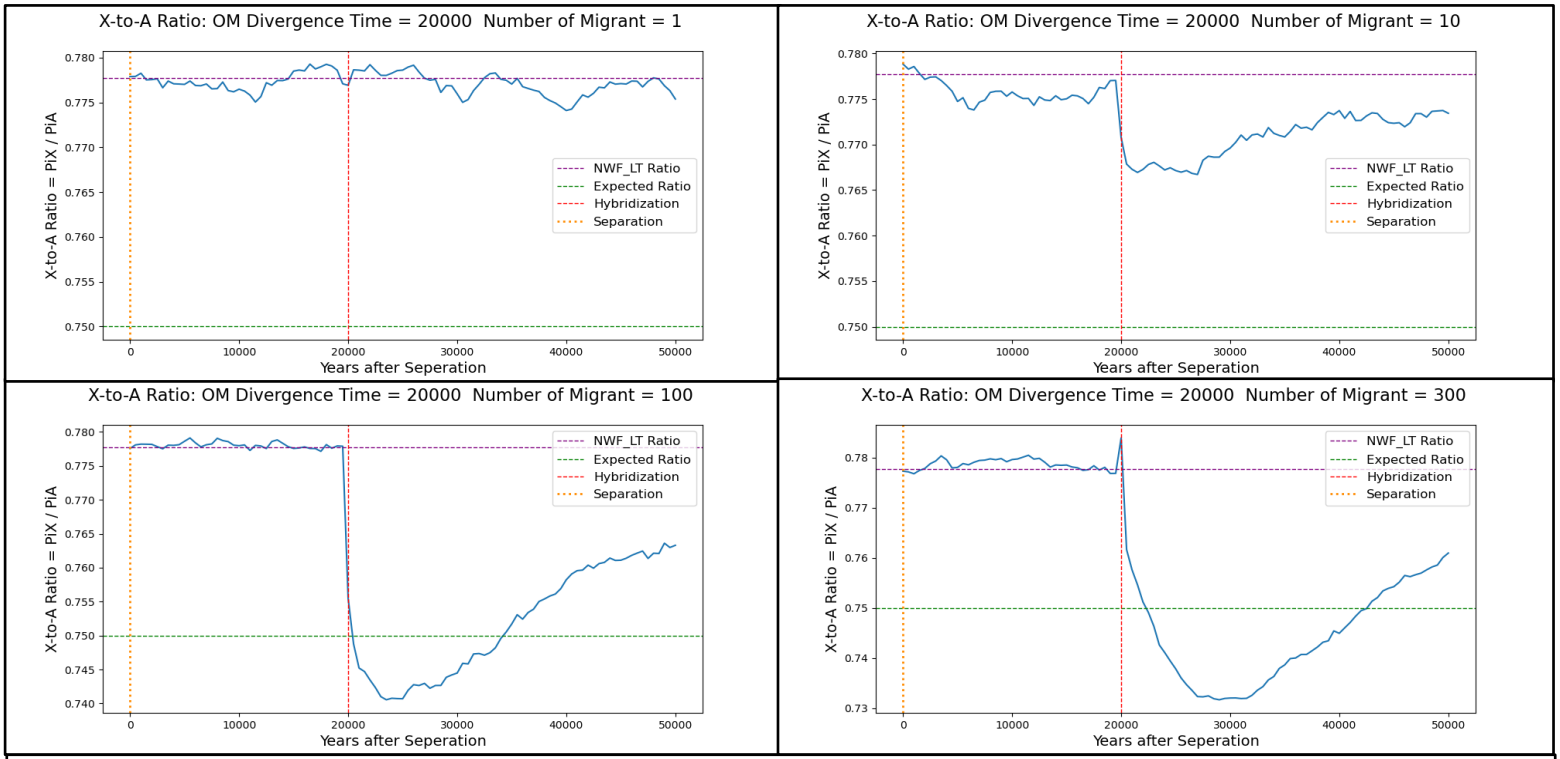



***Figure 15:*** *X-to-A genetic ratio mean of 20 NWF_LT (corrected) simulations with hybridization from the separation to the end of simulations (a value every 500 Ticks). Simulations are done with a time of divergence equal 20000, and four different numbers of migrants: 1, 10, 100 and 300. The yellow and red lines are respectively the separation event and the hybridization event. The purple line is NWF_LT expected ratio at equilibrium with a corrected Burn-in period. The green line is the expected ratio equal 0.75.*

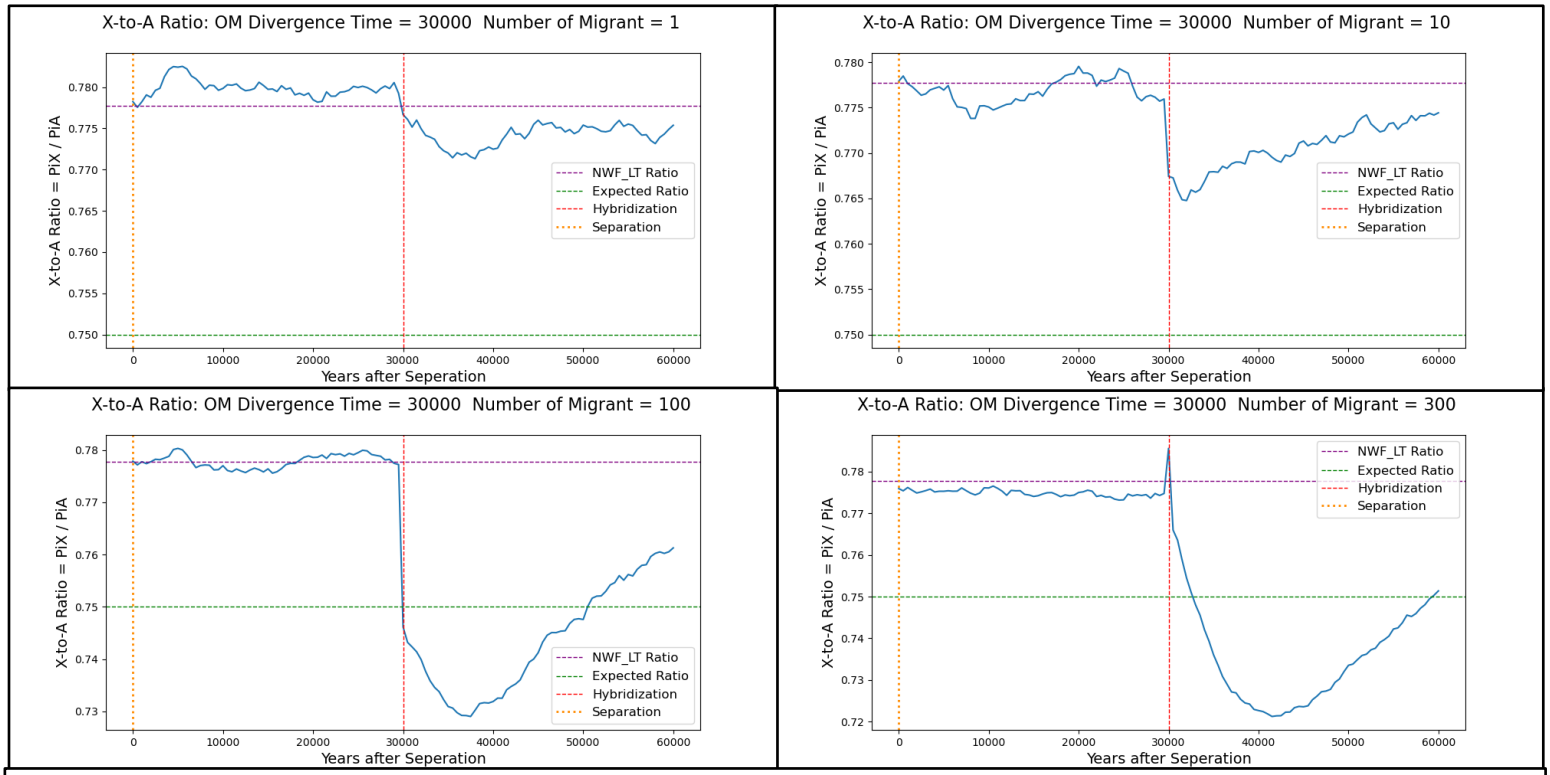



***Figure 16:*** *X-to-A genetic ratio mean of 20 NWF_LT (corrected) simulations with hybridization from the separation to the end of simulations (a value every 500 Ticks). Simulations are done with a time of divergence equal 30000, and four different numbers of migrants: 1, 10, 100 and 300. The yellow and red lines are respectively the separation event and the hybridization event. The purple line is NWF_LT expected ratio at equilibrium with a corrected Burn-in period. The green line is the expected ratio equal 0.75.*

The X-to-A ratio drop down is immediate, then it increases slowly to reach again the equilibrium value. But when the hybridization event is important (many migrants and diverged populations), the equilibrium is reached after a long time. Sometimes more than 30000 years in our simulations, which is equivalent to around 1.5 Millions years because of the population size rescaling (compared to Amboseli).

The separation of Yellow and Anubis baboons is expected to have happened 1.5 Millions years ago (30000 Ticks in our simulations)[14]. Moreover, in Amboseli, migration from Anubis populations have happened several times with low number of males since 1980[15]. So if we look at the graph with 30000 times of divergence and 100 migrants, we observe a decrease of the ratio, even below the 0.75 referent value. Hybridization is a sex-biased processes that may reduce the X-to-A ratio in Amboseli.

### h. Selection

The final force we studied is the stronger selection on the X chromosome than on the Autosomes. To simulate this, we make **deleterious mutations appear only on the X chromosome**. In the X, the **mutation rate is still 5e-7**. Each mutation can be deleterious or neutral. **The proportion of deleterious mutations can be tuned**. Moreover, **selection coefficients of deleterious mutations are drawn following a gamma distribution**. We keep **γ = 1**, but **we change the mean value** of the distribution.

Results will be shown on heatmaps depending on the selection coefficient and the deleterious mutations proportion. Those two parameters represent the strength of selection. The

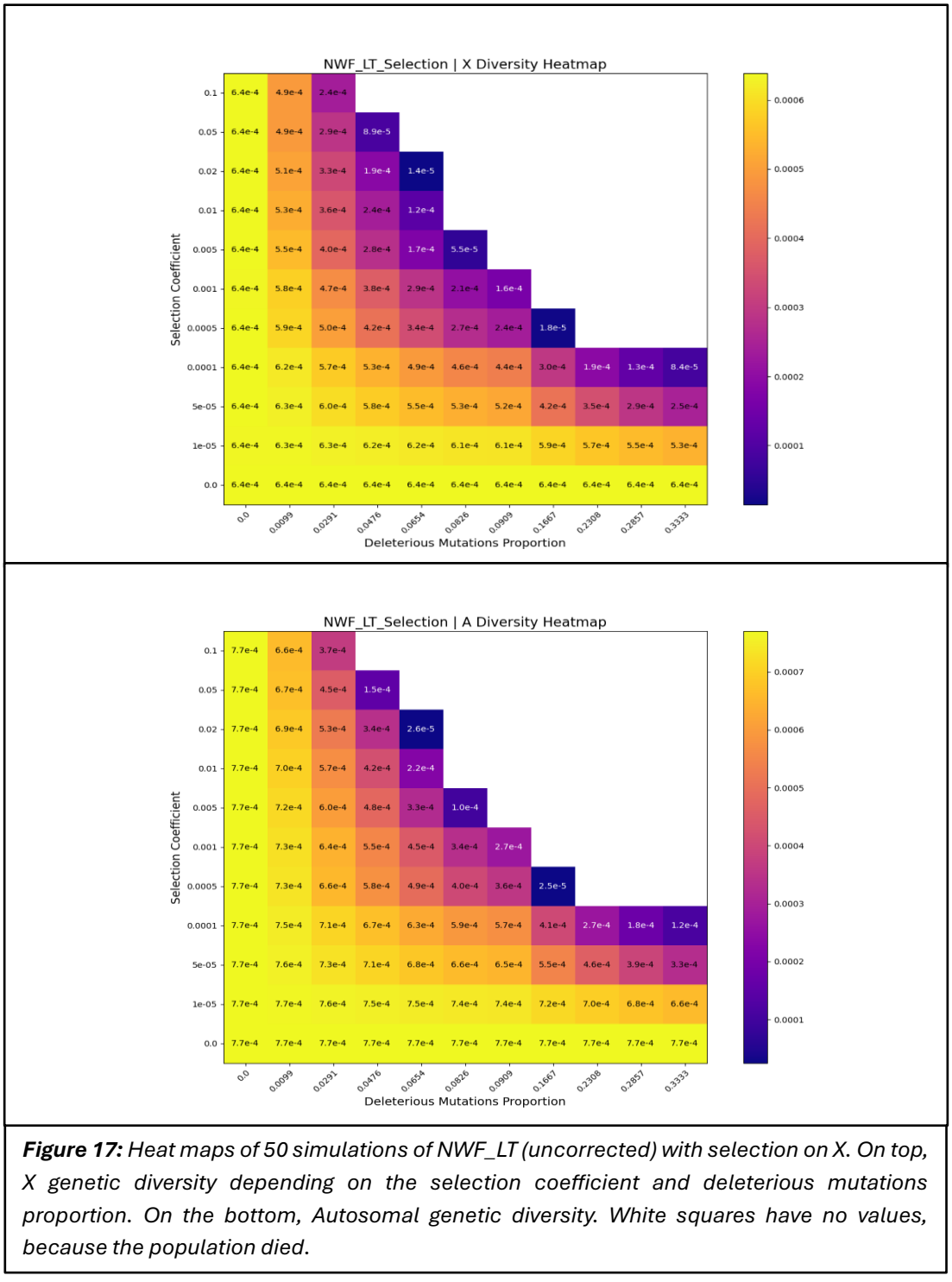



***Figure 17:*** *Heat maps of 50 simulations of NWF_LT (uncorrected) with selection on X. On top, X genetic diversity depending on the selection coefficient and deleterious mutations proportion. On the bottom, Autosomal genetic diversity. White squares have no values, because the population died.*

higher they are, the stronger selection on the X is. To see if selection simulations work, we can look at the X diversity and autosome diversities heat map, shown in *figure 17*. The referent model used here is the NWF_LT uncorrected one. So, we should not focus on the raw genetic values but on the tendency we observe.

When the selection coefficient increases, the X diversity decreases. We have the same result when the deleterious mutation proportion increases. As expected, selection acts negatively on the X diversity (even by a factor of ten). Selection removes deleterious mutations, but also the diversity of adjacent sequences. More surprisingly, we observe a decrease in the autosome diversity with selection. This is explained by the fact that selection reduces the populations size

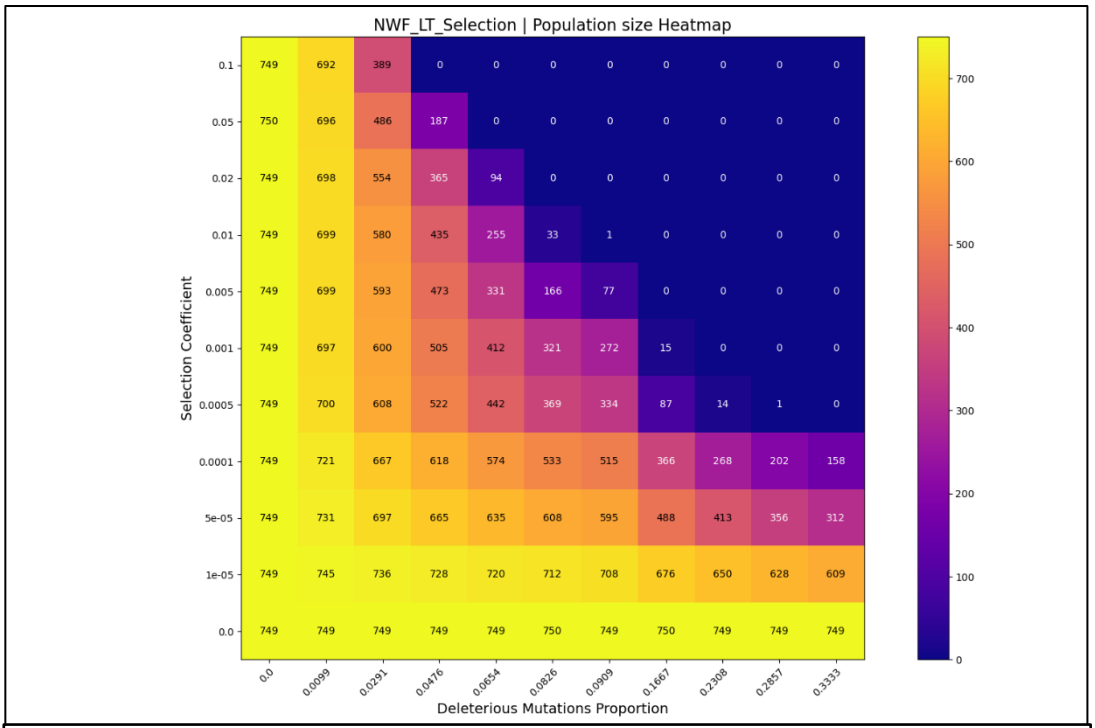



***Figure 18:*** *Heat maps of 50 simulations of NWF_LT (uncorrected) with selection on X. Mean population size of the 10000 last ticks of the simulations depending on the selection coefficient and deleterious mutations proportion.*

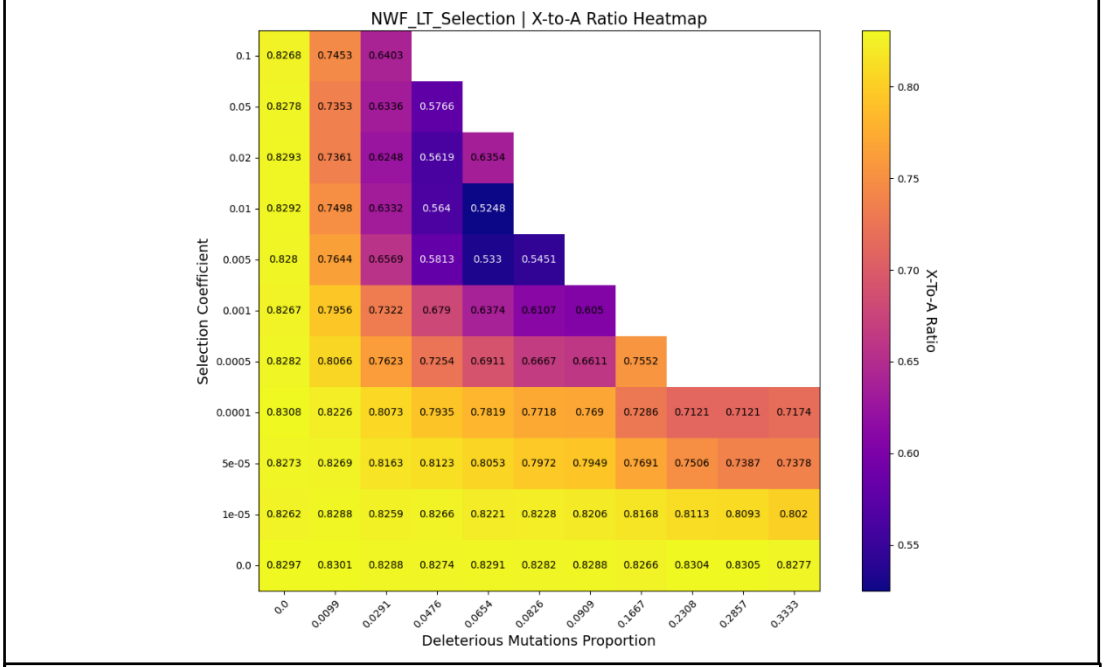



***Figure 19:*** *Heat maps of 50 simulations of NWF_LT (uncorrected) with selection on X. X-to-A genetic ratio depending on the selection coefficient and deleterious mutations proportion.*

in our simulation (see *figure 18*) which obviously changes the equilibrium diversity. The population can be less than 100 individuals (or even die) with a high selection. This might be seen as a limit of our simulations, because we would want the population to be around 750 individuals.

We can now look at the X-to-A ratio depending on the selection coefficient and the deleterious mutations proportion, in *figure 19*. The X-to-A ratio decreases with the strength of selection. We can even say that selection seems to be the strongest of the four forces on the ratio. The decrease can be more than 60%. However, because of the reduced population size and the burn-in period problem, results might be hard to interpret.

Selection on X appears to be the strongest of the four forces. Sex-biased processes may only increase (reproductive Skew) or decrease (Hybridization) the strength of selection.

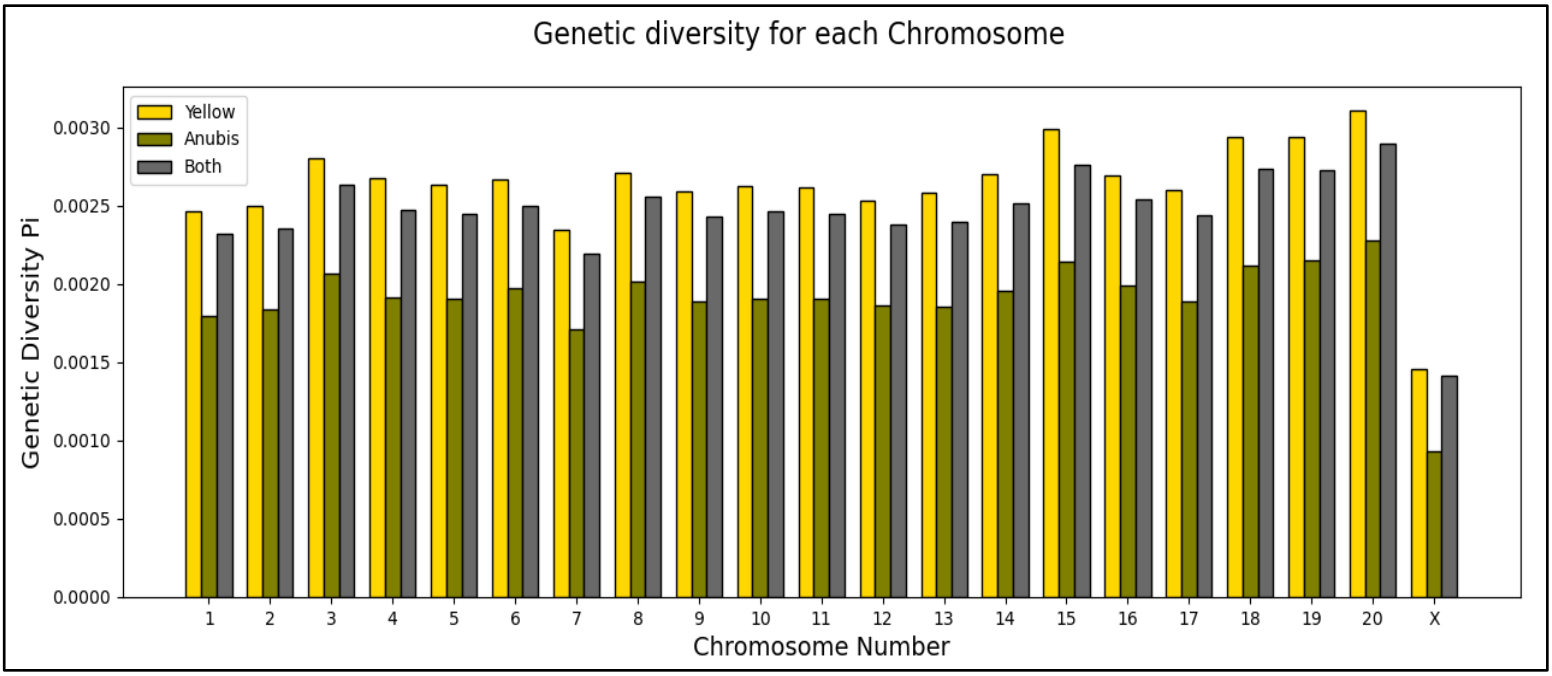
## III. Reel Baboon Data

In this part, I will present some results we had on reel diversity we observe in baboons. To do so, Iker created BCF files of every chromosome we found in baboons, with yellow and anubis individuals from the genetic reference panels. A bcf file is the binary version of a Variant Call Format (VCF) file. This format keeps the nucleotide diversity we can observe in the sampled individuals. At each nucleotide site, the format gives us the different nucleotides that exist in the population, and the nucleotide each individual wears.

SLiM can generate a population based on VCF files. To do so, it also needs Fasta files of the reference sequences we use for the VCF files. So, for every chromosome we convert the BCF file into a VCF file to create a population in SLiM equivalent to the sampled individuals. We can then measure the genetic diversity using π as in the previous part. This is done for yellow baboons, anubis baboons and both at the same time. Results are shown in *Figure 20*.

First, we can see that genetic diversity in both Yellow and Anubis baboons is in the high range of primate values (between 3e-4 and 3e-3). For the Yellow ones, the autosomal diversity is around 2.5e-3, and for the Anubis ones around 1.8e-3. The diversity is higher in Yellow baboons because the referent sequence for VCF files is an Anubis sequence. As expected, the X chromosome diversity is reduced compared to autosomes diversity in both species, but by a half and not a quarter. This result is confirmed with the X-to-Autosome measure. In fact, for every autosome the X-to-A ratio is below the expected value 0.75. For both species, the ratio is around 0.5. It is lower in Anubis than in Yellow baboons.

These results show that, as in many primates, the X-to-Autosome ratio is below 0.75 for baboons. If in Amboseli the ratio is the same



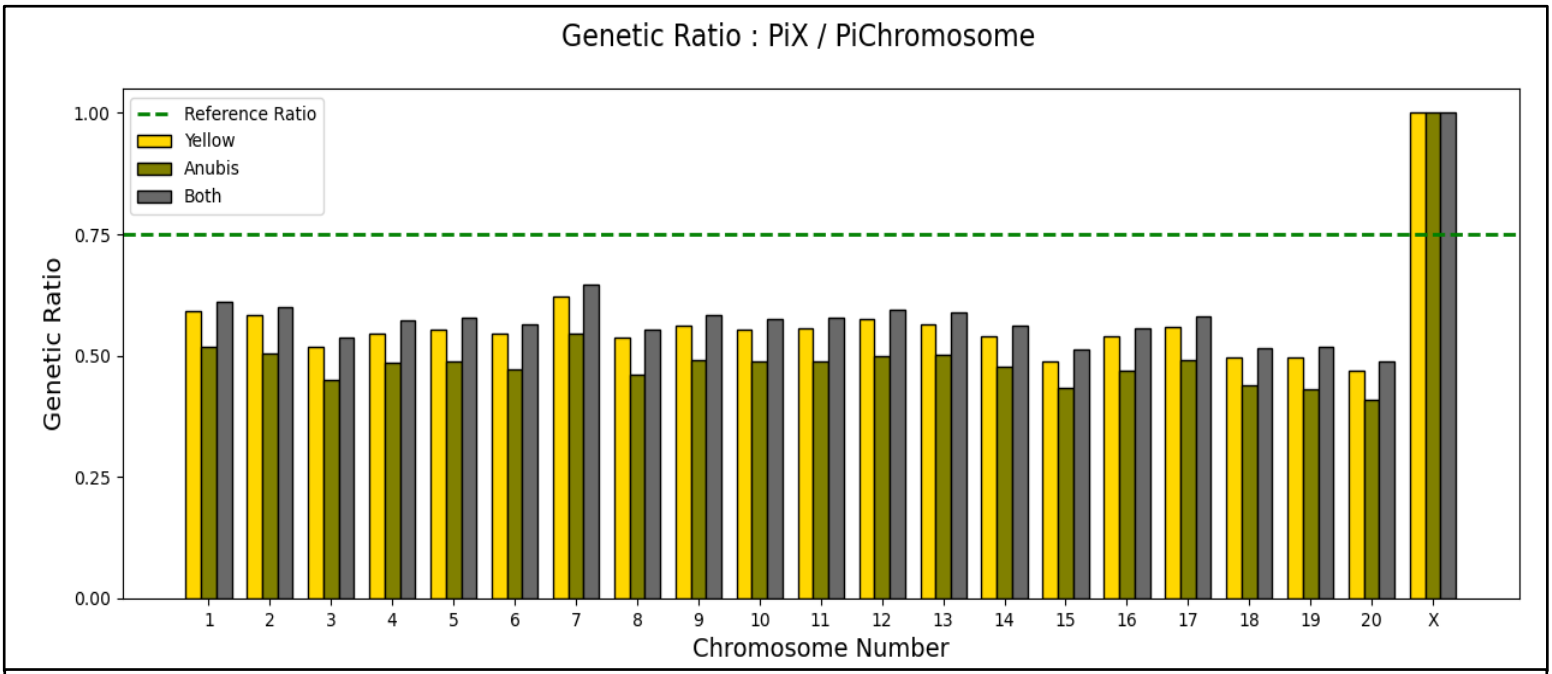




***Figure 20:*** *On top, genetic diversity observed in Yellow baboons, Anubis baboons and both for every autosome and the X chromosome.*

*On the bottom, Genetic ratio (= $\pi_x$ / $\pi_{chromosome}$) observed in Yellow baboons, Anubis baboons and both for every autosome and the X chromosome. The green line is the reference X-to-A genetic ratio of 0.75.*

## IV. Conclusion

This project aimed to find what forces happening in Amboseli is expecting to change the X-to-A genetic ratio from the 0.75 expectation. We firstly find that the demographic structure of the population has an influence on the ratio: the sex ratio, the difference in mortality between males and females, the different modes of reproduction. The NWF_LT has a higher ratio than the

expected value, around 0.77. Then, two sex-biased processes have been proven to change the ratio:

- The positive reproductive skew in males, linked to ranking. This force increases the ratio.
- Hybridization with Anubis baboons, with only male migrants. This force decreases the ratio.

The strength of these two forces is lower than the strength of selection on the X-to-A ratio. In fact, an increased selection on X chromosomes compared to on autosomes can decrease the ratio a lot.

We also calculated the reel value of diversity observed in Yellow and Anubis baboons and found a X-to-A ratio around 0.5. When this ratio will be available for the Amboseli population, we can compare the three ratio and see what forces can explain differences (or not) between them.

The simulations encounter several limits in this project. Because it is impossible to simulate the Amboseli reel population of 10000 baboons, we had to re-scale it to 750 individuals and more mutation rates. However, the rescaling was impossible to continue in migrations model, because populations would have 5 individuals, which is not enough for a simulation. Another issue is the burn-in period. In fact, at the end of the internship we found out that a burn-in period of 20000 Ticks was not enough for the NWF_LT model, but we didn't have time to check if others needed a higher bun-in period. Nevertheless, a lower burn-in period may not change the tendency of the force studied, but only the value of genetic diversity. Moreover, we didn't study all the processes that can change the X-to-A ratio such as bottle neck processes, different chromosome mutations rates, ... This may reduce our possibility of interpreting the reel data in Amboseli. Finally, even though Amboseli is a well-studied population, we still lack some data to complete our model, to better match the reality. We don't have the fertility in males depending on the rank, ... Using new data may also help us to try new ways of simulating the behavior forces.

We also proved SLiM to be a very powerful software to do population genetic computations. SLiM allows us to add all the demographic and behavioral processes while tracking the genetic data of every individual. Because simulations are forward in time, we also can simulate selection efficiently.

Only the first part of the problematic has been answered yet, but we didn't have time to compare the behavior forces to selections and see if selection is stronger than the others. Using SLiM, we can combine models to simulate both the reproductive skew and selection, or hybridization and selection, at the same time.

## V. References

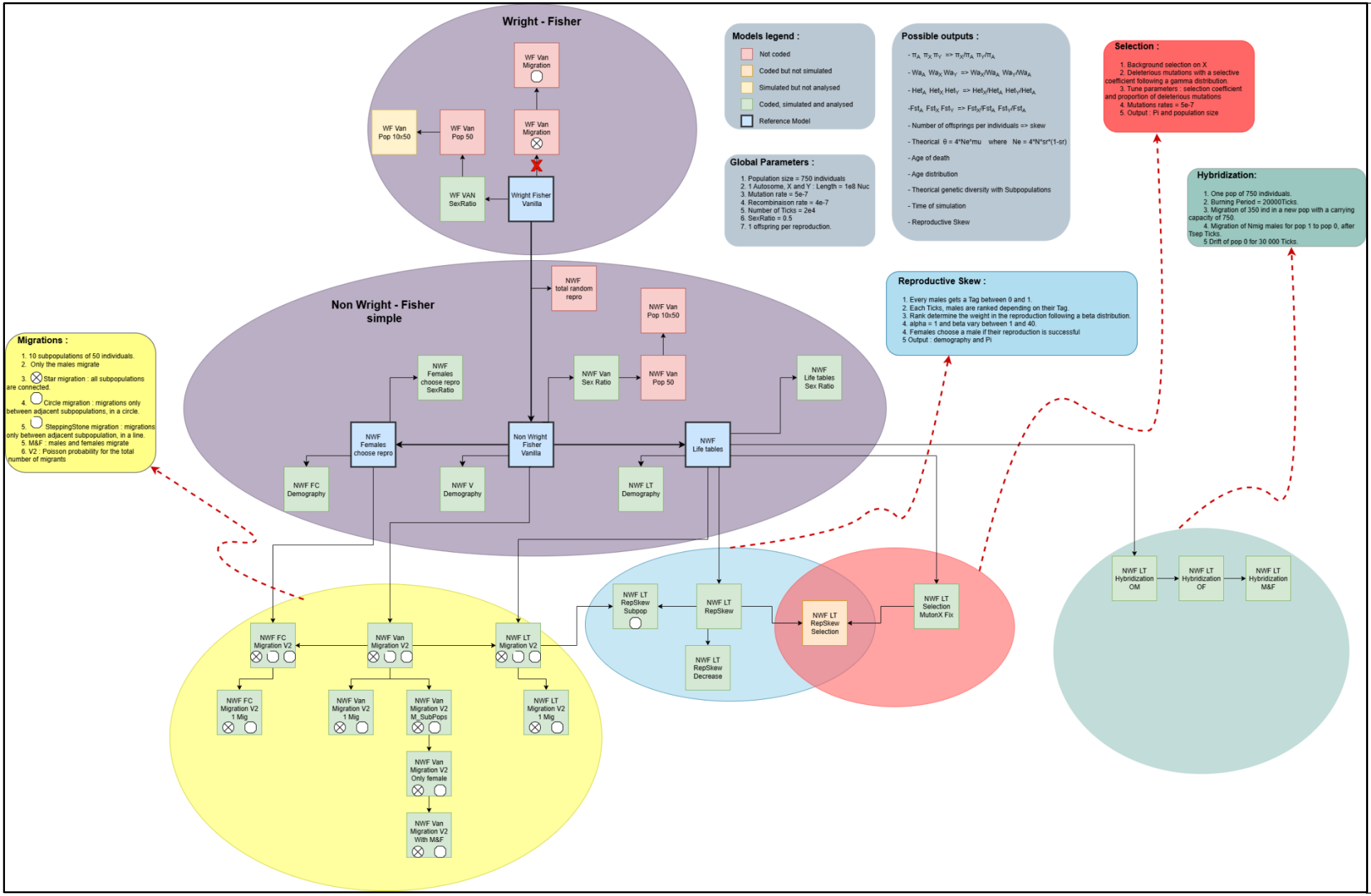
1. Alberts, S. C. & Altmann, J. Balancing Costs and Opportunities: Dispersal in Male Baboons.

*Am. Nat.* **145**, 279–306 (1995).

2. Haller, B. C. & Messer, P. W. SLiM 4: Multispecies Eco-Evolutionary Modeling. *Am. Nat.* **201**, E127–E139 (2023).
3. W.Hahn, M. *Molecular Population Genetics*. (Oxford University Press, 2019).
4. Alberts, S. C. & Altmann, J. The Amboseli Baboon Research Project: 40 Years of Continuity and Change. in *Long-Term Field Studies of Primates* (eds Kappeler, P. M. & Watts, D. P.) 261–287 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012). doi:10.1007/978-3-642-22514-7_12.
5. Noë, R. & Sluijter, A. A. Reproductive Tactics of Male Savanna Baboons. *Behaviour* **113**, 117–169 (1990).
6. Matute, D. R. & Cooper, B. S. Comparative studies on speciation: 30 years since Coyne and Orr. *Evolution* **75**, 764–778 (2021).
7. Kuderna, L. F. K. *et al.* A global catalog of whole-genome diversity from 233 primate species. *Science* **380**, 906–913 (2023).
8. Ross, C. T. *et al.* The multinomial index: a robust measure of reproductive skew. *Proc. R. Soc. B Biol. Sci.* **287**, 20202025 (2020).
9. Charlesworth, B., Morgan, M. T. & Charlesworth, D. The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303 (1993).
10. Rogers, A. & Prügel-Bennett, A. Evolving Populations with Overlapping Generations. *Theor. Popul. Biol.* **57**, 121–129 (2000).
11. Bronikowski, A. M. *et al.* Female and male life tables for seven wild primate species. *Sci. Data* **3**, (2016).
12. Altmann, S. & Altmann, J. *Baboon Ecology*. (The University of Chicago Press, 1970).
13. Laporte, V. & Charlesworth, B. Effective Population Size and Population Subdivision in Demographically Structured Populations. *Genetics* **162**, 501–519 (2002).
14. Vilgalys, T. P. *et al.* Selection against admixture and gene regulatory divergence in a long-term primate field study. *Science* **377**, 635–641 (2022).

15. Charpentier, M. J. E. *et al.* Genetic structure in a dynamic baboon hybrid zone corroborates behavioural observations in a hybrid population. *Mol. Ecol.* **21**, 715–731 (2012).

## VI. Annexes



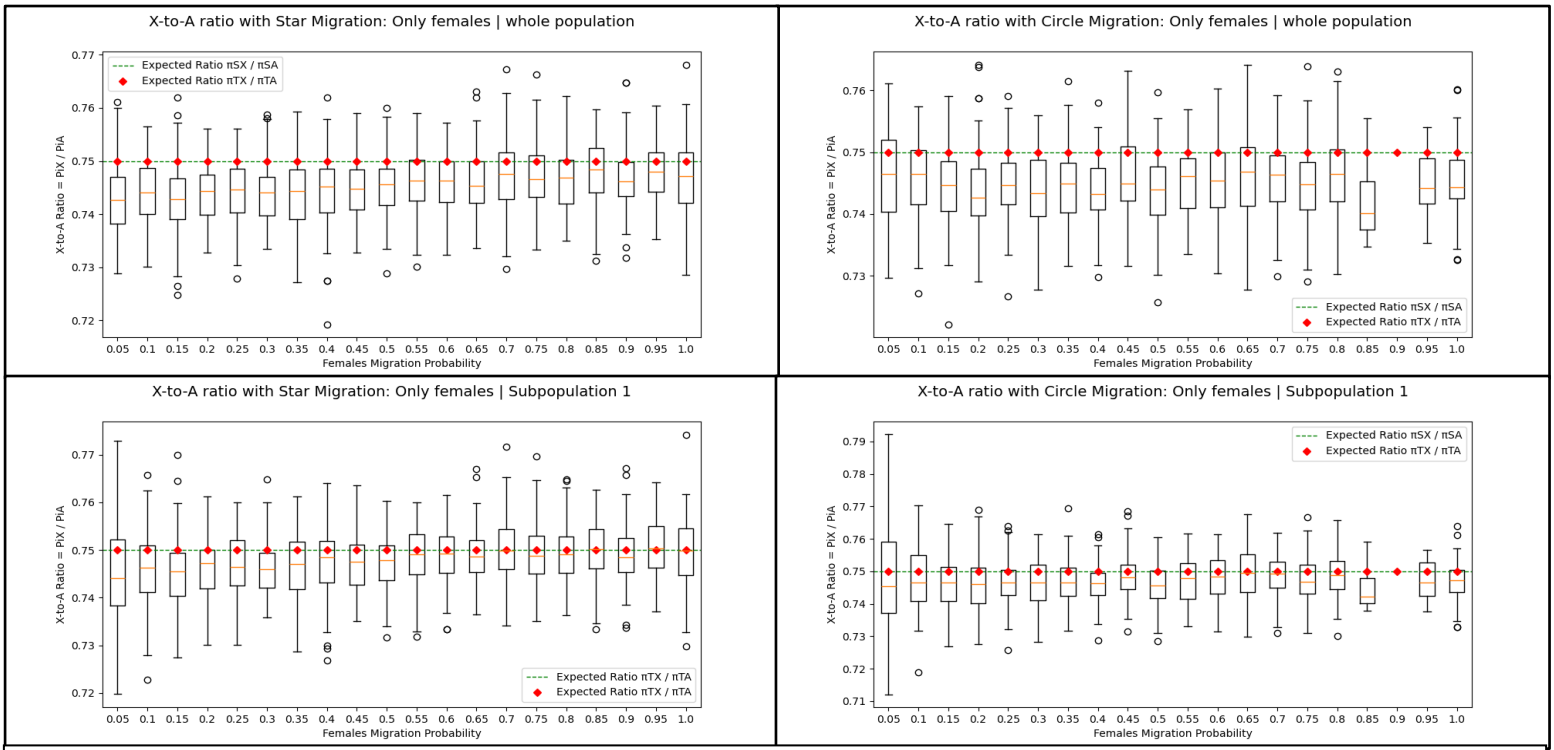

***Annex 1:*** *Mind map of every model done during the project*

| Age | Female_Mortality | Male_Mortality | Female_fecundity |
|---|---|---|---|
| 0.0 | 0.215909091 | 0.215809285 | 0.0 |
| 1.0 | 0.092391304 | 0.096 | 0.0 |
| 2.0 | 0.045908184 | 0.076106195 | 0.0 |
| 3.0 | 0.054393305 | 0.057471264 | 0.0 |
| 4.0 | 0.037610619 | 0.042682927 | 0.005 |
| 5.0 | 0.03908046 | 0.050955414 | 0.211 |
| 6.0 | 0.04784689 | 0.049217002 | 0.264 |
| 7.0 | 0.037688442 | 0.058823529 | 0.272 |
| 8.0 | 0.041775457 | 0.06 | 0.279 |
| 9.0 | 0.046321526 | 0.02393617 | 0.281 |
| 10.0 | 0.034285714 | 0.068119891 | 0.269 |
| 11.0 | 0.065088757 | 0.067251462 | 0.276 |
| 12.0 | 0.060126582 | 0.112852665 | 0.29 |
| 13.0 | 0.087542088 | 0.091872792 | 0.291 |
| 14.0 | 0.081180812 | 0.112840467 | 0.285 |
| 15.0 | 0.088353414 | 0.140350877 | 0.274 |
| 16.0 | 0.083700441 | 0.137755102 | 0.263 |
| 17.0 | 0.125 | 0.165680473 | 0.271 |
| 18.0 | 0.093406593 | 0.219858156 | 0.271 |
| 19.0 | 0.163636364 | 0.254545455 | 0.236 |
| 20.0 | 0.115942029 | 0.329268293 | 0.233 |
| 21.0 | 0.147540984 | 0.345454545 | 0.259 |
| 22.0 | 0.240384615 | 0.416666667 | 0.047 |
| 23.0 | 0.164556962 | 0.476190476 | 0.108 |
| 24.0 | 0.333333333 | 1.0 | 0.0 |
| 25.0 | 0.363636364 | 1.0 | 0.0 |
| 26.0 | 0.428571429 | 1.0 | 0.0 |
| 27.0 | 1.0 | 1.0 | 0.0 |

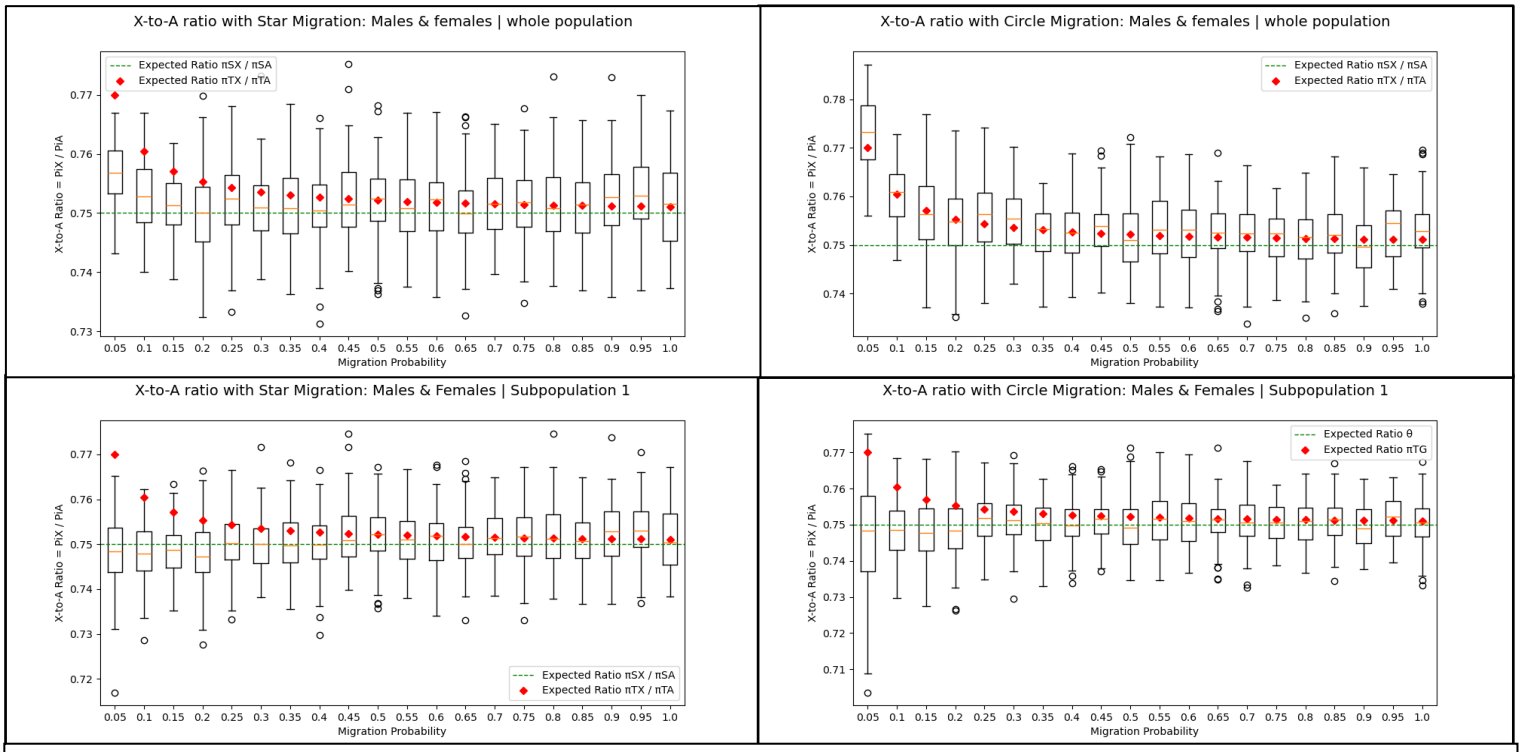
***Annex 2**: Lifetable used for the NWF Life table model. There are female and male mortality, and female fecundity, for each age.*

| Beta | Noff Max |
|---|---|
| 1 | 20 |
| 2 | 29 |
| 5 | 51 |
| 10 | 83 |
| 20 | 162 |
| 40 | 294 |

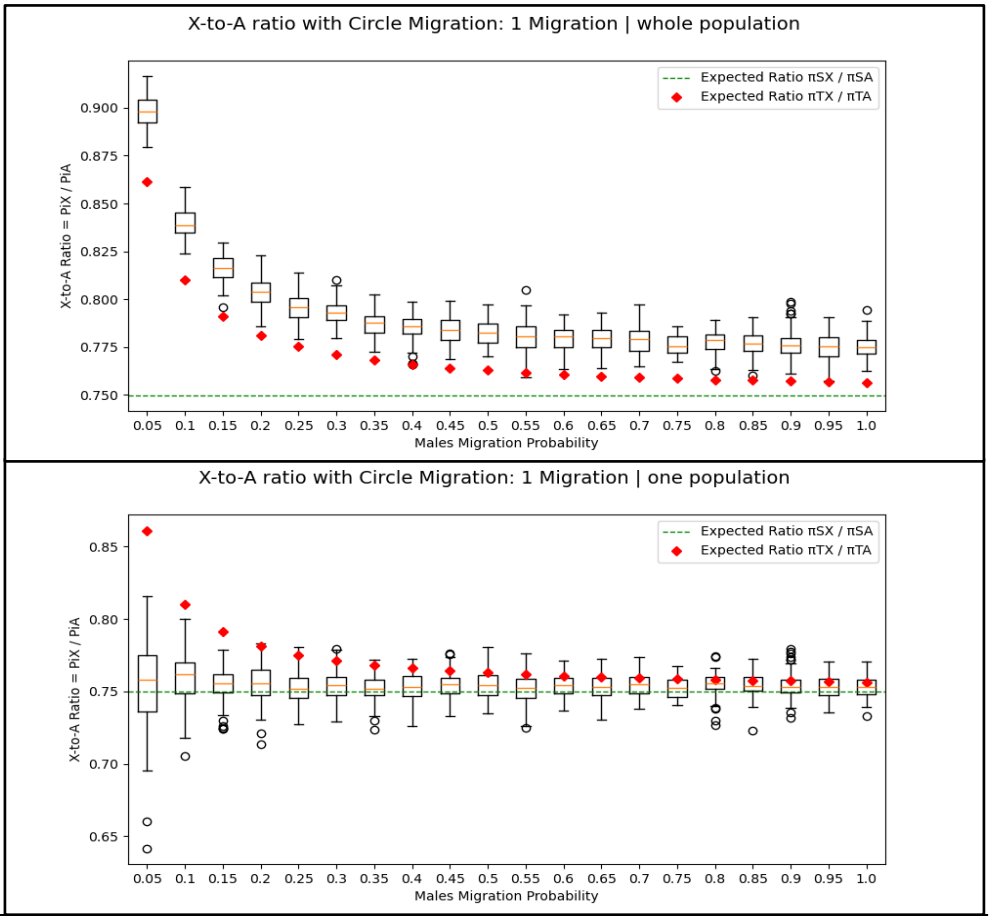
***Annex 3**: Maximal number of offspring for a male in the reproductive skew model for all values of β.*



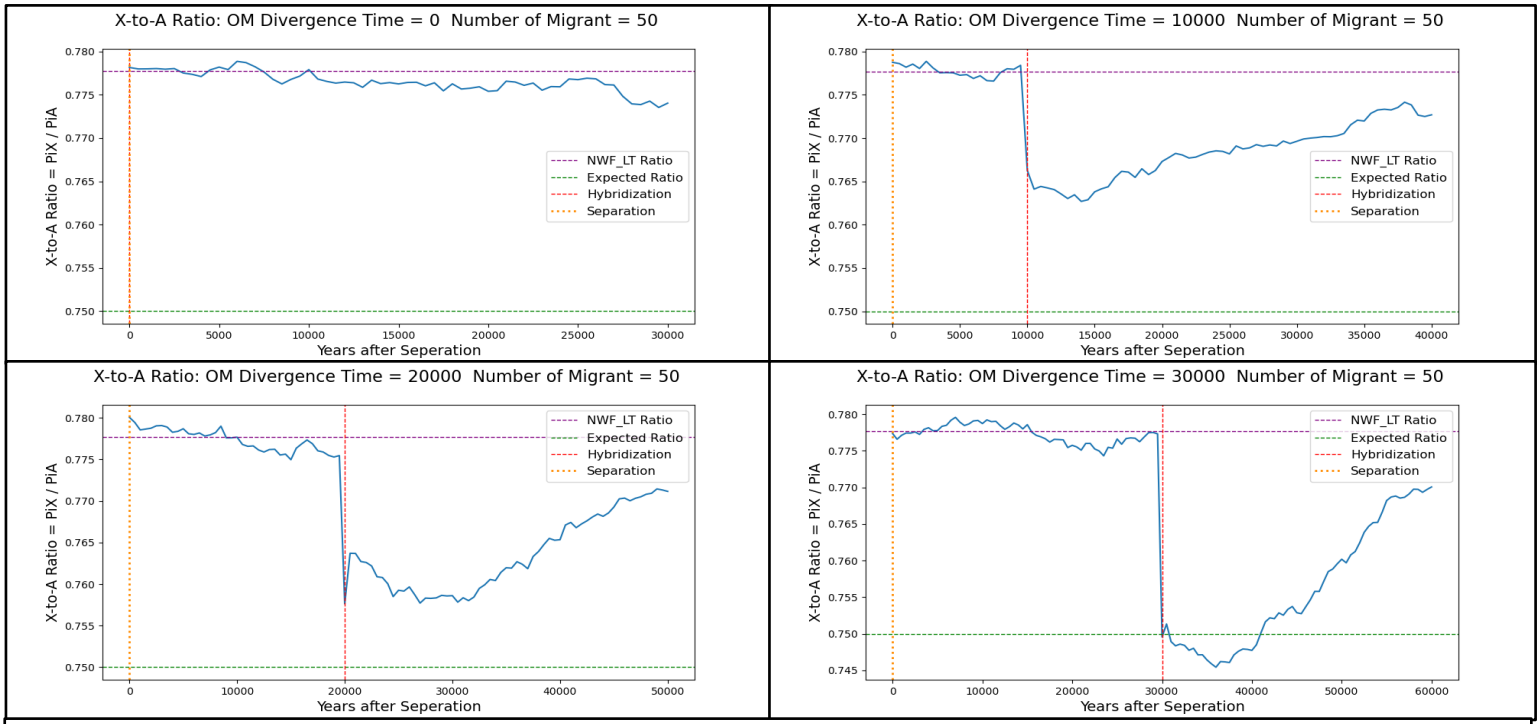

***Annex 4**: Box Plots of 100 simulations of NWF_V models with only female migrations, representing the X-to-A ratio of the whole population in function of the migration probability. The top graphs are for the whole populations, and on the bottom for one subpopulation, on the left for star migrations and on the right for circle migration.*



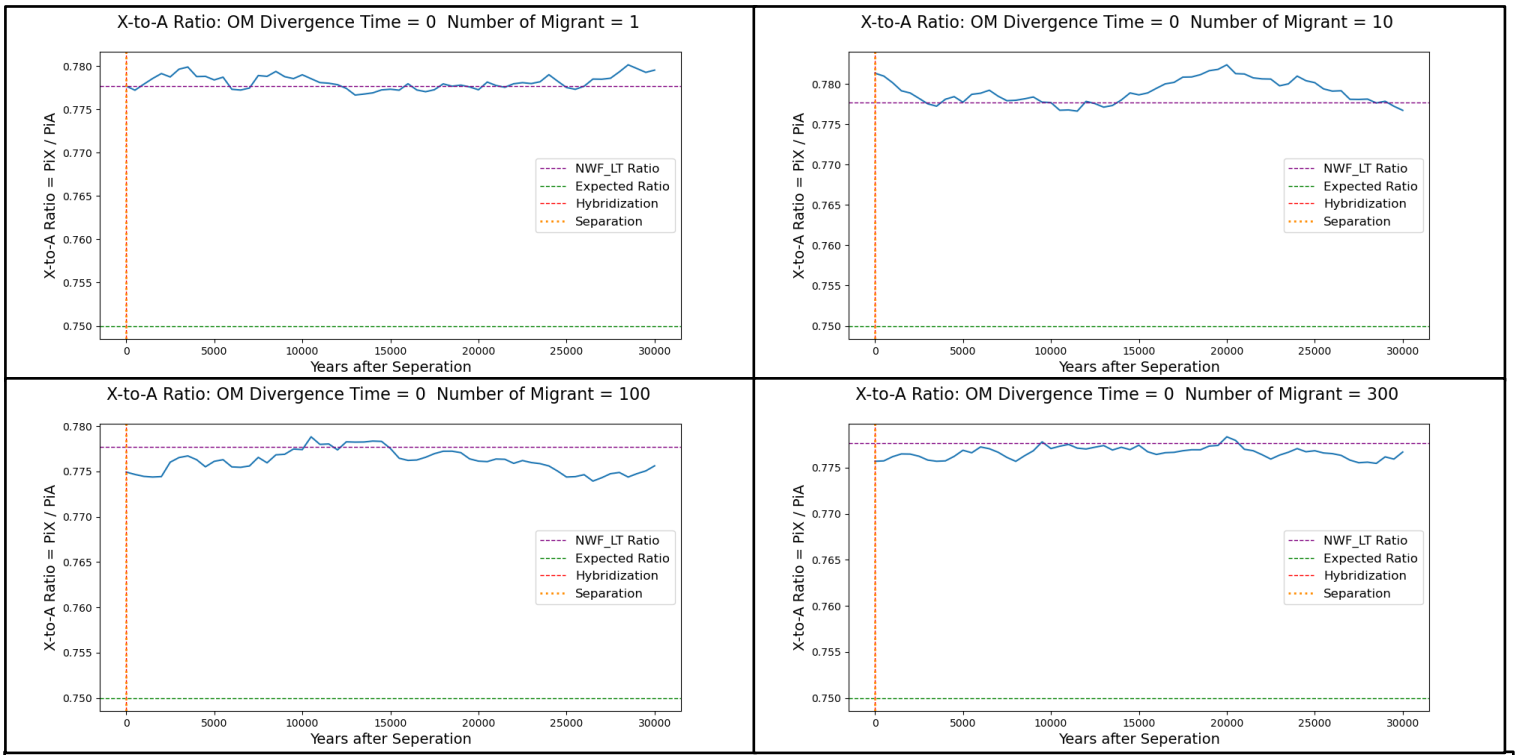

***Annex 5**: Box Plots of 100 simulations of NWF_V models with male and female migrations, representing the X-to-A ratio of the whole population in function of the migration probability. The top graphs are for the whole populations, and on the bottom for one subpopulation, on the left for star migrations and on the right for circle migration.*




***Annex 6**: Box Plots of 100 simulations of NWF_V models with only one migration for males, representing the X-to-A ratio of the whole population in function of the migration probability. Both graphs are for circle model. The top graphs are for the*




***Annex 7:*** *X-to-A genetic ratio mean of 20 NWF_LT (corrected) simulations with hybridization from the separation to the end of simulations (a value every 500 Ticks). Simulations are done with 50 migrants and different divergence time: 0, 10000, 20000 and 30000. The yellow and red lines are respectively the separation event and the hybridization event. The purple line is NWF_LT expected ratio at equilibrium with a corrected Burn-in period. The green line is the expected ratio equal 0.75.*




***Annex 8:*** *X-to-A genetic ratio mean of 20 NWF_LT (corrected) simulations with hybridization from the separation to the end of simulations (a value every 500 Ticks). Simulations are done with no divergence time, and four different numbers of migrants: 1, 10, 100 and 300. The yellow and red lines are respectively the separation event and the hybridization event. The purple line is NWF_LT expected ratio at equilibrium with a corrected Burn-in period. The green line is the expected ratio equal 0.75.*