

Projet : Génétique des populations

Madeleine DE SOUSA & Sandrine FARTEK

Introduction

Par ce projet, nous souhaitons étudier par simulation, l'influence de différents facteurs sur la vitesse d'évolution de l'hétérozygotie moyenne $\bar{H}(t)$ au sein d'une population. Les facteurs étudiés un à un sont des entorses aux hypothèses du modèle de Wright-Fisher :

- Proportion de mâles et femelles inégale
- Autofécondation
- Variance du succès reproducteur au sein de la population
- Taille de la population croissante

L'espèce considérée est supposée diploïde. On s'intéresse à un locus autosomal, neutre, sans mutation et présentant une diversité initiale maximale, les générations sont non-chevauchantes et il n'y a pas de migration.

Par ces simulations, nous étudierons également l'influence sur l'effectif efficace N_e , ce dernier étant défini comme la taille N qu'une population idéale de Wright-Fisher devrait avoir afin de dériver à la même vitesse qu'une population réelle.

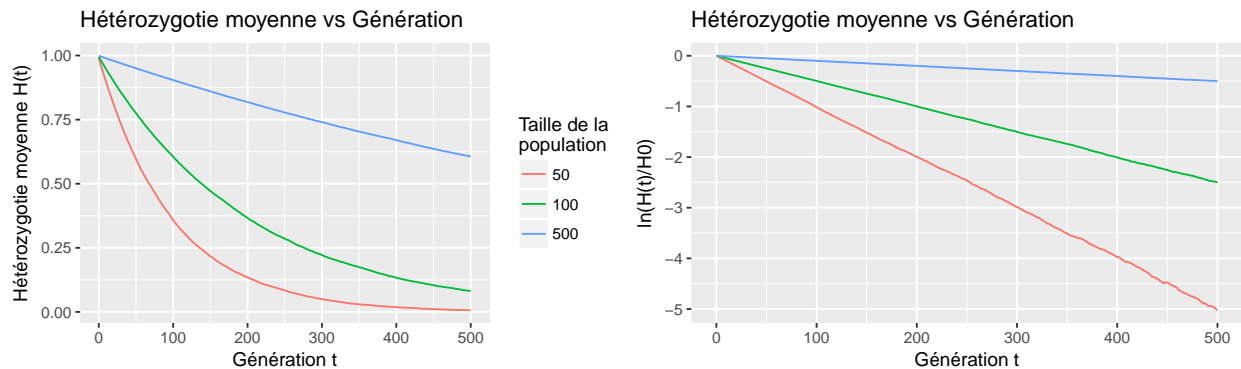


Figure 1: Evolution de l'hétérozygotie moyenne $H(t)$ en fonction des générations t pour une population de Wright-Fisher et de taille constante $N = 100$. La moyenne de l'hétérozygotie est calculée sur 5000 simulations indépendantes.

Matériel et méthode

Code général

Le programme est découpé en 8 fichiers :

- main.c, qui permet de lancer le programme
- d_util, qui permet de gérer l'allocation de mémoire et les variables aléatoires (notamment pour choisir un individu)

- genpop.c, qui permet de gérer les allèles, les locus, le génome, les gamètes et les individus
- model_WrightFisher.c, model_Autofecondation.c, model_CycleDemographique.c, model_SexRatio.c et model_Variance.c qui permettent de gérer les différentes simulations.

L'utilisateur choisit une simulation avec ses paramètres (nombre de générations, taille de la population etc) et le programme va créer un génome, les locus (à partir des allèles) et les tableaux qui contiendront les données de sorties. Puis, le programme lancera la simulation.

Les simulations peuvent être expliquées comme ceci :

allocation mémoire de toutes les variables nécessaires

Pour chaque individu :

attribution des gènes paternels et maternels

attribution du sexe #de base, les individus sont hermaphrodites

ajouter cet individu dans la population

Pour chaque générations :

Calcul de l'hétérozygotie et stockage des résultats

Pour i allant de 0 au nombre d'individus :

Choix du père et de la mère aléatoirement dans le tableau des individus

Choix du sexe

Création de l'enfant

Remise en cause du modèle de Wright-Fisher : le sex-ratio

La remise en cause du modèle de Wright-Fisher du sexe ratio permet de simuler une population avec deux sexes différents. Les individus, au lieu d'être hermaphrodites, ont un sexe, et ne peuvent s'accoupler qu'avec le sexe opposé. Pour cela, les individus qui sont stockés dans la population ont un pointeur global__male et global__female qui permettent de différencier les sexes. Pour garder le même ratio mâle-femelle que l'utilisateur a choisi, les enfants n'ont pas un sexe aléatoire.

Remise en cause du modèle de Wright-Fisher : l'autofécondation

La remise en cause du modèle de Wright-Fisher de l'autofécondation permet de simuler une population où certains individus n'ont pas besoin d'un autre individu pour se reproduire. Les individus sont toujours hermaphrodites. Pour cela, lorsque les individus se reproduisent, nous choisissons ceux qui font faire de l'autogamie et ceux qui vont faire de l'allogamie (à partir du taux d'autofécondation que l'utilisateur va rentrer). Les individus qui font de l'allogamie vont faire un enfant à partir de deux individus choisis aléatoirement. Les individus qui font de l'autogamie sont des mères (comme les individus sont hermaphrodites, les individus femelles sont les mêmes que les individus mâles) qui sont choisis aléatoirement pour se reproduire avec elle-mêmes.

Remise en cause du modèle de Wright-Fisher : le succès reproducteur

La remise en cause du modèle de Wright-Fisher du succès reproducteur permet de simuler une population où certains individus n'auront aucun enfant pendant que d'autres en auront plus pour garder un nombre d'individus constant. Les individus peuvent produire 0 gamètes, 2 gamètes (un enfant) ou 4 gamètes (deux enfants) selon la formule :

$$N2 = N - (2 * N0)$$

$$N4 = N0$$

Avec $N2$ qui correspond au nombre d'individus qui produiront 2 gamètes, $N4$ le nombre d'individus qui produiront 4 gamètes et $N0$ le nombre d'individus ne produiront 0 gamètes. L'utilisateur ne pourra pas choisir

un N_0 plus grand que $N - 2$. Pour simuler cela, nous avons choisis que les individus N_2 auront le même comportement que celui décrit dans le modèle de Wright-Fisher (deux individus vont faire un enfant), et les individus N_4 vont faire dans un premier temps un enfant, puis vont choisir un autre individu aléatoirement pour faire le deuxième enfant (puisque un individu peut choisir un autre individu pour faire son enfant). Les individus étant hermaphrodite, nous avons choisis de prendre aléatoirement une deuxième mère pour faire le deuxième enfant.

Remise en cause du modèle de Wright-Fisher : le cycle démographique

La remise en cause du modèle de Wright-Fisher du cycle démographique permet de simuler une population croissante avec des goulots d'étranglements. Afin d'éviter une perte de temps en ré-allocation mémoire, la mémoire des populations ParentPop et OffspringPop est allouée avec le nombre d'individus maximal en fin de cycle, ce dernier étant décrit par l'équation ci-dessous. Comme les individus sont tirés aléatoirement dans des tableaux de pointeurs sur les individus mâles ou femelles et non pas directement dans le tableau contenant les individus, cela ne pose pas de problème.

$$N(t + n) = N(t) * (1 + r)^n$$

Nous avons choisi d'interpréter le goulot d'étranglement comme étant l'avant dernière génération d'un cycle ne produisant que $N(0)$ descendants plutôt que générer l'ensemble des descendants et en supprimer une partie.

Résultats

Variation du sex-ratio

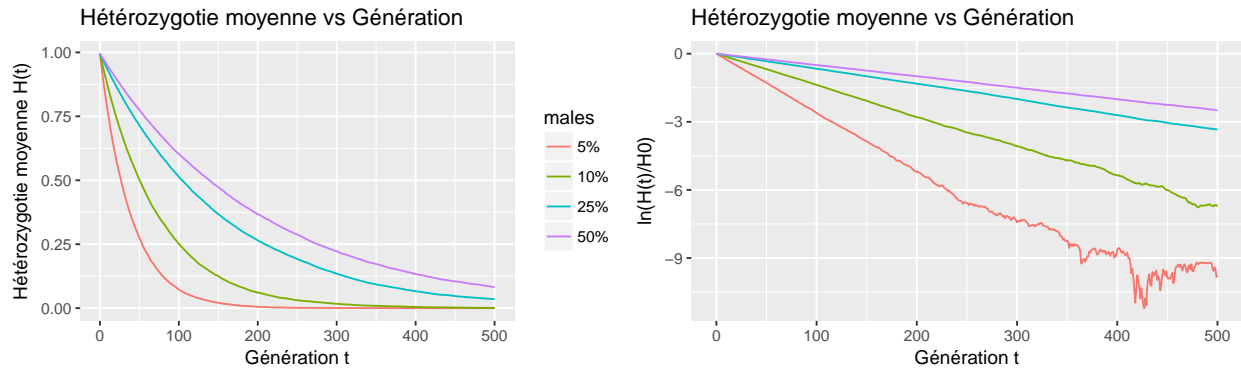


Figure 2: Évolution de l'hétérozygotie moyenne $H(t)$ en fonction des générations t pour des populations variant en proportion de mâles et de taille constante $N = 100$. La moyenne de l'hétérozygotie est calculée sur 5000 simulations indépendantes.

On observe sur les courbes, Figure 2, que la proportion de mâles/femelles joue sur l'hétérozygotie moyenne de la population. En effet, un nombre différent de mâles et de femelles va entraver la diversité génétique puisque certains individus d'un sexe s'accoupleront plus souvent afin de garder une population de taille constante.

Pour le modèle de Wright-Fisher, l'effectif efficace attendu est de $N_e = N$. Pour une population avec reproduction sexuée, l'effectif efficace dépend du ratio mâles/femelles. Chaque parent a une probabilité de $\frac{1}{2}$ de transmettre son gène. Si on regarde pour les mâles, il y a une probabilité de $\frac{1}{4 * N_m}$ pour qu'un gène

provienne du père. Sachant que la probabilité qu'un individu dans une population de Wright-Fisher ait un gène d'un parent est de $\frac{1}{2N}$, l'effectif efficace pour une population de mâle et de femelles est de :

$$\frac{1}{2Ne} = \frac{1}{4Nm} + \frac{1}{4Nf} \Rightarrow Ne = \frac{4Nf * Nm}{Nf + Nm}$$

Pour 50 mâles, nous avons $Ne \approx 100$, ce qui signifie que nous avons une variation du polymorphisme d'une population naturelle de 100 individus. Chaque individus a pu se reproduire. Lorsqu'il y a moins ou plus de femelles, on observe un déséquilibre, puisque les individus ne se reproduisent qu'une fois. Cela veut dire qu'il n'y aura trop ou pas assez de femelles pour tenir l'effectif d'une population de 100 individus.

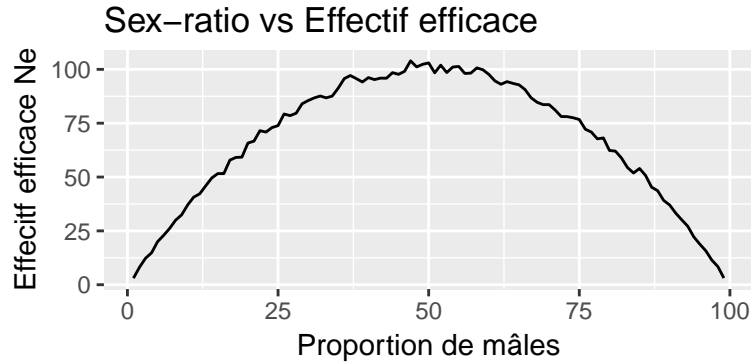


Figure 3: Évolution de l'effectif efficace Ne en fonction de la proportion de mâles dans une population de taille constante $N = 100$. La moyenne de l'hétérozygotie est calculée sur 500 simulations indépendantes.

Variation du taux d'autofécondation

On constate que la dérive s'accélère quand le taux d'autofécondation augmente. En effet, quand l'hétérozygotie diminue, le nombre maximal d'individus hétérozygotes diminue également. Donc le fait qu'un individu se reproduise avec lui-même, sachant qu'il a moins de chance d'être hétérozygote, accélère le processus de fixation des allèles.

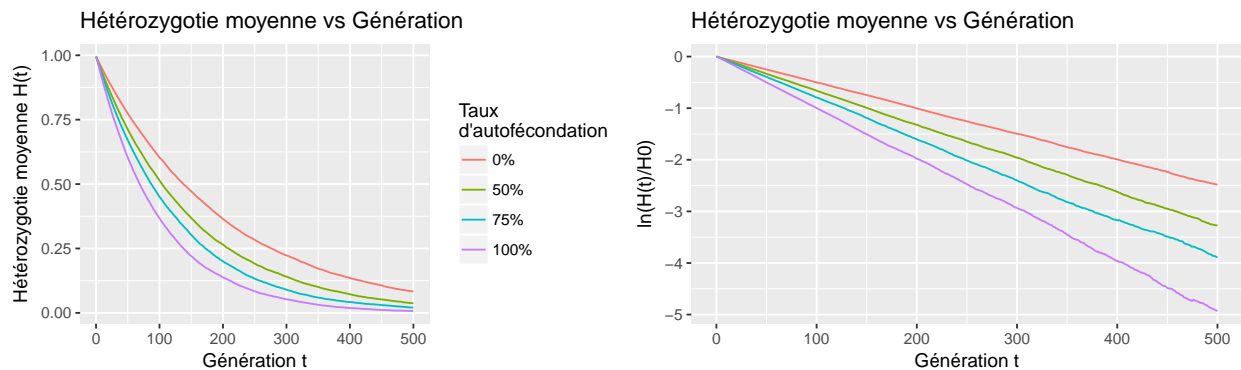


Figure 4: Évolution de l'hétérozygotie moyenne $H(t)$ en fonction des générations t pour des populations variant en taux d'autofécondation et de taille constante $N = 100$. La moyenne de l'hétérozygotie est calculée sur 5000 simulations indépendantes.

```
##          Ne
## 0%    100.59561
## 50%   76.67351
## 75%   64.11731
## 100%  50.90704
```

On observe ici que le Ne diminue lorsque le taux d'autofécondation augmente, mais il diminue moins rapidement que pour le sexe-ratio. Cela est dû au fait que si un individu hétérozygote se reproduit avec lui-même, l'enfant a plus de chance d'être hétérozygote.

Variation du taux de croissance

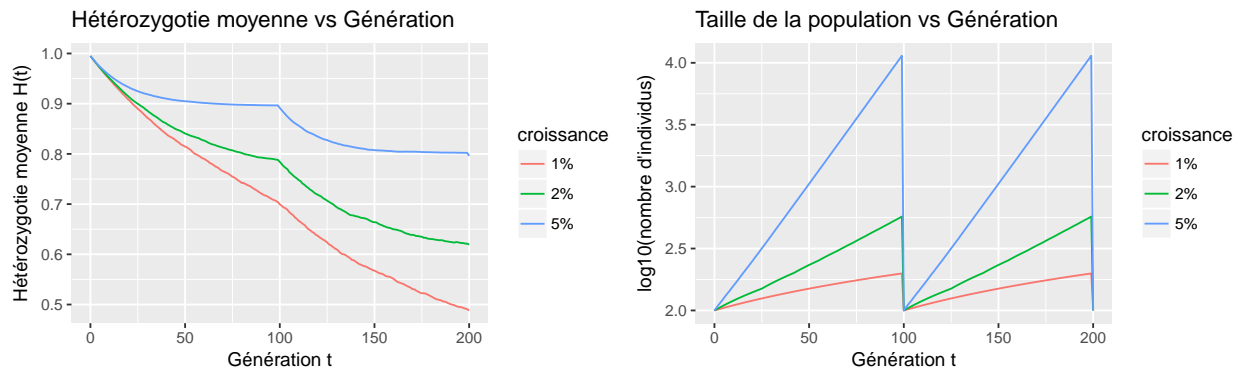


Figure 5: Évolution de l'hétérozygotie moyenne $H(t)$ en fonction des générations t pour des populations variant en taux de croissance et de taille initiale $N = 100$. Un goulot d'étranglement se produisant entre la 99ème et 100ème génération de chaque cycle. La moyenne de l'hétérozygotie est calculée sur 500 simulations indépendantes.

Comme la taille de la population augmente, il est normal d'observer une dérive génétique plus lente. Quand on simule des populations de Wright-Fisher de différentes tailles, les plus grandes dérivent effectivement plus lentement. Cependant l'hétérozygotie semble converger à partir d'un certain nombre de générations, plus ou moins grand selon le taux de croissance. Le goulot d'étranglement à la 100ème génération permet de déclencher un nouveau cycle de diminution de l'hétérozygotie, de la même valeur semble-t-il.

Ici, comme la taille de la population varie, l'effectif efficace devient :

$$\frac{1}{Ne} = \frac{1}{n} \sum_{i=1}^n \frac{1}{N_i}$$

L'effectif efficace Ne (calculé sur les 100er cycles) montre que plus la croissance est importante, plus le nombre d'individus d'une population idéale de Wright-Fisher dérivant à la même vitesse que la population étudiée est grand. Un $Ne = 251$ paraît cependant faible sachant que la taille de la population atteinte en fin de cycle est de l'ordre de 10^4 .

```
##          Ne
## 1%  125.3738
## 2%  150.5753
## 5%  251.1267
```

Variation du succès reproducteur

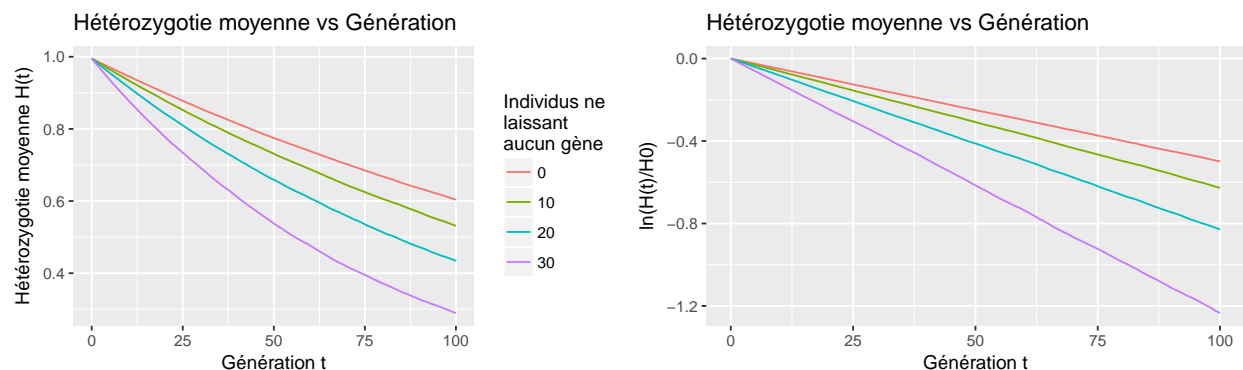


Figure 6: Évolution de l'hétérozygotie moyenne $H(t)$ en fonction des générations t pour des populations variant sur le nombre de gènes transmis par individus et de taille constante $N = 100$. La moyenne de l'hétérozygotie est calculée sur 5000 simulations indépendantes.

Lorsque certains individus ont un succès reproducteur nul (c'est-à-dire qu'ils n'ont pas de descendant), les autres individus vont plus se reproduire pour garder la taille de la population constante. Cela entraîne une forte baisse de l'hétérozygotie moyenne de la population, puisque un individu a plus de chance de se reproduire avec un individu ayant les mêmes parents que lui.

```
##          Ne
## 0  100.65384
## 10  80.50353
## 20  60.64848
## 30  40.71429
```

Ici, le Ne diminue très rapidement, ce qui correspond au fait que les individus qui vont se reproduire entre-eux ont de très grandes chances de provenir du même parent, favorisant alors la consanguinité.

Discussion

Avec ces simulations, nous pouvons voir que les différents facteurs entraînent un grand déficit d'hétérozygoties. Nous retrouvons bien l'effet Wahlund puisqu'il y a un excès d'homozygotes et que nous ne subdivisons pas notre population. Si nous regardons l'effectif efficace du Sex-ratio, nous retrouvons bien la représentation en "cloche" avec 50% de mâles = N . Ici la simulation fait en sorte de garder une population de taille constante, mais une grande différence de mâles et de femelles devrait entraîner une décroissance de la population dans une population réelle.

Pour la variation du taux d'autofécondation, lorsque tous les individus sont autogames, l'hétérozygotie moyenne baisse fortement. Si chaque individu fait un enfant lui-même, la diversité entre individus va rester la même sur chaque génération. De ce fait, il ne va pas y avoir une diminution de la diversité. L'hétérozygotie va bien évoluer au cours du temps (formule : $Het(t) = Het(0) * \exp\left(-\frac{t}{2N}\right)$) mais le taux d'autofécondation ne va pas évoluer. Finalement, le résultat ressemblera fortement à celui d'une population consanguine.

Lorsque nous observons les courbes de la variation du taux de croissance, nous pouvons voir que les courbes ont exactement la même allure avant et après le goulot d'étranglement. Le goulot d'étranglement peut donc accélérer la diminution de l'hétérozygotie moyenne dans une population à forte croissance.

La variation du succès reproducteur ressemble beaucoup à une population consanguine. En effet, si certains individus vont faire plus d'enfants que les autres, nous aurons des enfants qui pourront avoir les mêmes allèles.

Que le succès reproducteur joue sur ces nouveaux individus ou non, ces nouveaux individus proviennent pour la plupart de la même descendance, si bien qu'à la fin ils risquent de se reproduire entre-eux. Si nous étudions une population avec un certain pourcentage d'individus consanguins, nous aurons les mêmes résultats à la fin, c'est-à-dire que toute la population aura les mêmes allèles.

Si nous étudions les N_e de chaque simulations, les résultats sont bien vérifiés par les courbes de Wright-Fisher. Par exemple, la courbe du sexe-ratio avec 50% de mâles à la même allure que la courbe de Wright-Fisher avec une population de 100. La courbe de 100% d'autofécondation a la même allure que la courbe de Wright-Fisher avec $N = 50$. La courbe du succès reproducteur entre 20 et 30 individus 0 correspond à peu près à la courbe de Wright-Fisher avec un $N = 50$.

En comparant les N_e entre-eux, on se rends compte que la simulation qui a une hétérozygotie moyenne qui décroît le plus rapidement est le succès reproducteur. Cela semble correct puisque certains individus vont complètement disparaître pendant que d'autres vont faire plusieurs enfants. Plus nous augmentons le nombre d'individus 0, plus nous aurons des individus qui feront beaucoup d'enfants. L'hétérozygotie moyenne va donc baisser très rapidement.

Conclusion

Ces différentes simulations ont permis de mettre en évidence les différents facteurs qui entravent plus ou moins rapidement l'hétérozygotie moyenne d'une population. Cela permet de comprendre l'évolution d'une population et la rapidité de la fixation d'un allèle. Néanmoins, sur une population réelle, ces différents facteurs et d'autres externes (migration par exemple) agissent en même temps. Il faut donc prendre cela en compte sur de réelles simulations.