

TPK1: MODÉLISATIONS NUMÉRIQUES DE LA SÉLECTION NATURELLE ET DE LA DÉRIVE GÉNÉTIQUE

Objectifs méthodologiques :

- *Modéliser, à l'aide d'un logiciel, des cas de sélection ou de dérive afin d'expliquer l'évolution d'une population.*

Objectifs cognitifs :

- *Savoir identifier les mécanismes de l'évolution mis en œuvre à partir de résultats modélisés et discuter des facteurs et de la valeur sélective d'un trait de caractère considéré.*

Exemples de sujets proposés à l'épreuve de travaux pratiques sur cette partie du programme:

- Étude quantitative d'une population en équilibre de Hardy-Weinberg ou non → « calcul de la fréquence allélique et génotypique, hypothèse de HW et vérifier si une population est à l'équilibre et pourquoi »
- Identification et discussion de facteurs de sélection, de la valeur sélective (fitness)
- Modélisations numériques (dérive génétique, sélection naturelle, dynamique de populations...) et leur interprétation ⇒ sur le logiciel Populus, simuler l'évolution des populations de 2 mares isolées (relever N après 50 générations etc) puis si les migrations sont possibles entre ces deux populations, comparer et proposer des mesures pour protéger les tritons, avec le mode ?, modéliser la chaîne juvénile, jeune reproducteur, adulte avec N_0 et probas de survie / passage d'un stade à l'autre, voir si la population est en extinction, changer certains paramètres et voir si c'est mieux, conclure sur les mesures pour protéger les tritons
- Étude de cas de coévolution (amenant ou non sur le modèle de la Reine Rouge)
- Identification d'une homologie ou d'une convergence évolutive (supports et échelles divers)

L'évolution est à l'origine des modifications de la biodiversité au cours du temps.

Elle repose sur l'existence d'une **diversité génétique des individus au sein des populations** et de **modifications de cette diversité de manière orientée ou non**. En cas de rupture de flux génique entre des populations, des espèces émergent, d'autres disparaissant soit par hybridation avec une population, soit par extinction.

Les mécanismes de l'évolution sont au nombre de 4 :

Les mutations sont **aléatoires** et **engendrent un polymorphisme**.

La sélection est un processus majeur de **reproduction différentielle** où la valeur sélective (fitness) se mesure au nombre de descendants viables et fertiles produits, la valeur sélective d'un trait dépend de l'environnement et peut diminuer la valeur sélective d'autres traits (trade-off).

La dérive génétique exerce un **tri aléatoire** dont l'intensité dépend de l'effectif efficace.

Les migrations peuvent tendre à **diminuer l'effet de la sélection naturelle** OU **provoquer une dérive génétique** par effet fondateur.

L'objectif ici est de manipuler les concepts et les mécanismes évolutifs en se rappelant que ces derniers s'appuient sur une approche expérimentale, des observations, des prédictions testables et des modélisations numériques

I. Modélisation de la sélection naturelle et de la dérive génétique

De nombreuses simulations ont été vues et manipulées au collège et au lycée pour modéliser « simplement » les effets de la sélection naturelle et de la dérive génétique sur les variations des fréquences alléliques au sein d'une population.

1. utilisation d'une application facile d'accès pour mettre les idées bien en place : l'application Géné'pop

Lancer tout d'abord sur votre ordinateur l'application Géné'pop à partir du lien suivant :

<https://cosphiloq.fr/genepop/>

Cette application permet d'étudier les variations de la fréquence de deux allèles d'un locus donné au sein d'une population soumises à la dérive génétique et à la sélection.

Plusieurs paramètres sont modifiables :

- l'effectif de la population (0 signifie infini)
- le nom des deux allèles peuvent être choisis
- la fréquence initiale de l'allèle 1 peut être choisie également
- les valeurs sélectives des trois génotypes possibles, fondamental pour étudier la sélection naturelle, et notamment illustrer les sélections directionnelle, stabilisatrice ou balancée
- la fréquence des mutations : un paramètre pour enrichir la simulation et introduire de la variabilité dans la population initiale
- Le nombre de générations

1.1. reprise de l'exemple de la phalène du bouleau ou comment expliquer que la fréquence d'un allèle dans la population peut varier en fonction de la pression de sélection elle-même variable au cours du temps.

L'allèle carbonaria C est apparu par mutation vers 1815, on considère qu'en 1830, la fréquence de cet allèle C était de 0,01 %.

Dans les r , la valeur sélective relative des formes claires typica a été estimée autour de 0,4.

→ *Le cycle de développement de la phalène du Bouleau étant annuel, déterminer à l'aide de l'application, combien d'années sont nécessaires à la fixation de l'allèle C en fonction de l'effectif initial de la population.*

Pour cela :

- *partir d'un effectif de la population de 1000 et d'un nombre de générations égale à 500.*
- *effectuer plusieurs simulations pour les mêmes paramètres*
- *décrire vos observations*

- *recommencer les simulations pour un effectif de population de 100 et à l'inverse pour un effectif infini.*
- *décrire vos observations et proposer une explication*

Dans la région de Manchester, la fréquence maximale observée de l'allèle C a été entre 1957 et 1964, elle était alors de 0,99.

Les dernières mines de charbon ont été fermées en 1983, la fréquence de l'allèle C est tombée au dessous de 0,8 à partir de 1990 :

- en 1991 : $F(C) = 0,6$
- en 2004 : $F(C) = 0,14$
- en 2013 : $F(C) = 0,01$

→ *Estimer le coefficient de sélection s'étant exercé sur l'allèle C entre 1983 et 2004 pour expliquer une telle variation de fréquence. Considérer le cas avec $N = 1000$ et le cas où la population est infinie.*

→ *Faire de même entre 2004 et 2014.*

→ *Interpréter ces résultats d'après vos connaissances*

1.2. exemple de du maintien de l'allèle d HbS dans les populations ou comment expliquer que la fréquence d'un allèle dans la population peut varier en fonction de la pression de sélection elle-même variable dans l'espace

Aujourd'hui, les individus souffrant de la drépanocytose en Europe ont une espérance de vie autour de 45 ans grâce aux progrès de la médecine.

→ *Réinitialiser les valeurs par défaut de Géné'pop, prendre $N = 1000$*

→ *Partir d'une fréquence initiale de 0,5 et prendre pour valeurs sélectives : $w_{A/A} = 1$, $w_{A/S} = 1$ et $w_{S/S} = 0,8$,*

→ *Déterminer ainsi $F(A)$ et préciser si ce résultat est en accord avec les fréquence de l'allèle Hbs observée sur le document 11 du cours.*

→ *A présent, changer les valeurs sélectives pour celles observées au Nigéria à savoir*

$$w_{A/A} = 0,88, w_{A/S} = 1 \text{ et } w_{S/S} = 0,13$$

→ *Partir d'une fréquence initiale de 0,5 et prendre pour valeurs sélectives : $w_{A/A} = 0,88, w_{A/S} = 1$ et $w_{S/S} = 0,13$*

→ *Déterminer ainsi $F(A)$, préciser si ce résultat est en accord avec les fréquence de l'allèle Hbs observée sur le document 11 du cours.*

2. utilisation d'un autre logiciel un peu plus complet: PopG

Le logiciel PopG permet en plus de Génepop de modifier :

- Les effets de migration : la migration induit des brassages et peut modifier les résultats au regard des résultats attendus compte tenu des simulations précédentes. Peut aider à expliquer le maintien de génotypes à faible fitness dans une population
- Le nombre de populations qui évoluent en parallèle : entre 0 et 1000. Permet de définir une reproductibilité ou non des résultats.

On s'appuiera une nouvelle fois sur l'étude des variations de fréquence de l'allèle C dans le cas du mélanisme industriel en zone polluée !

→ *Lancer dans les applications le logiciel PopG*

→ *Pour lancer une simulation, il suffit de sélectionner « New run » dans le graphique vierge qui s'est ouvre.*

→ *Réaliser plusieurs modélisations en utilisant les paramètres indiqués ci dessous :*

	Cas numéro 1	Cas numéro 2 (théorique)	Cas numéro 3 (théorique)	Cas numéro 4 (théorique)	Cas numéro 5 cas actuel sans migration	Cas numéro 6 cas actuel avec migration
Effectif de la population	1000	1000	1000	10	1000	1000
Fitness C/C	1	1	1	1	0,9	0,9
Fitness C/c	1	1	1	1	0,9	0,9
Fitness c/c	0,4	1	1	1	1	1
Fréquence initiale de C	0,01	0,01	0,9	0,9	0,01	0,01
Nombre de générations	100	100	100	100	20	20
migration						10 %

→ *Discuter des résultats observés*

3. utilisation d'un autre logiciel : Populus

On restera ici purement théorique

→ *Lancer dans les applications le logiciel Populus*

→ *Dans le menu « Model », choisir « Mendelian Genetics /Drift and selection », cela calcule les fréquences alléliques en tenant compte de la dérive et de la sélection.*

w_{AA} indique la valeur sélective du génotype AA. Cela permet de calculer les fréquences alléliques à la prochaine génération $t+1$, avec $p(t)$ la fréquence de A et q_t celle de a à la génération t :

$$p_{t+1} = (p_t) \frac{p_t w_{AA} + q_t w_{Aa}}{p_t^2 w_{AA} + 2 p_t q_t w_{Aa} + q_t^2 w_{aa}}$$

→ Compléter le tableau suivant à l'aide des résultats obtenus avec les paramètres indiqués :

N (effectif)	w_{AA}	w_{Aa}	w_{aa}	Variations de $F(A)$	Mécanisme(s) impliqués et remarques
500	1	0,8	0,8	Fixation de l'allèle A	Sélection directionnelle dans une grande population la sélection directionnelle favorise l'allèle ayant la meilleure valeur sélective
500	0,8	0,8	1	Fixation de l'allèle a	
500	0,8	1	0,8	Pas de fixation, ni de dérive	Sélection balancée Dans le cas où l'hétérozygote a la valeur sélective la plus élevée, les deux allèles sont maintenus
20	1	0,8	0,8	Aléatoire	pas de sélection, uniquement génétique du fait d'un faible effectif Une petite population évolue toujours par dérive, donc les valeurs sélectives utilisées n'ont pas d'importance.
20	0,8	0,8	1	Aléatoire	
20	0,8	1	0,8	Aléatoire	
500	1	1	1	Aléatoire	Dérive génétique Quand il n'y a pas de sélection naturelle, la dérive génétique agit d'autant plus vite que l'effectif de la population est faible.
20	1	1	1		

→ Illustrer la loi de Kimura, qui affirme qu'on observe une sélection ou une dérive si la valeur du paramètre $4Ns$ est nettement supérieure ou nettement inférieure à 1 (où N est l'effectif et s le coefficient de sélection).

N	w_{AA}	w_{Aa}	w_{aa}	s	$4Ns$	Résultats
1000	1	1	0.99975	0.000025	0,01	Aléatoire car dérive génétique seule.
1000	1	1	0.9975	0.0025	10	Fixation de A car sélection de l'allèle A.

II. Modélisation de la dérive génétique seule

1. Modélisation des travaux de Peter Buri (1956)

Il s'agit ici de simuler l'expérience historique de Buri montrant la dérive génétique sur de petites populations.

→ Aller sur <http://virtualbiologylab.org/population-genetics/> et lancez le modèle 3.

Dans des tubes, on place des *Drosophiles* possédant soit deux allèles dominants (+/+) donnant des yeux rouges, soit 2 allèles récessifs (*bw/bw*) donnant des yeux bruns, soit des hétérozygotes (+/*bw*) aux yeux rouges. Dans les conditions de culture, la couleur des yeux n'a aucun effet sur la capacité à trouver leur nourriture ou un partenaire. C'est un tirage aléatoire parmi les mouches présentes dans le tube qui donne la nouvelle population du tube.

Il est possible de choisir le nombre de mouches qui est prélevé à chaque génération et remplacé dans un nouveau tube (paramètre «*Pop-Size*»), sachant que le sex-ratio reste de 50%-50%.

Il est également possible de choisir la fréquence de l'allèle *bw* au début de l'expérience (paramètre «*Init-Prop-bw*»). Le nombre de générations est indiqué, ainsi que le nombre de tubes pour lesquels un allèle a été fixé (paramètre «*#Colonies Fixed*»), avec le nombre moyen de générations nécessaires pour obtenir cette fixation (paramètre «*Mean Fix. Time*»). La fréquence de l'allèle *bw* dans l'ensemble des tubes est donnée dans le paramètre «*Overall Prop. bw*».

→ Lancer une simulation avec les paramètres par défaut, qui ressemblent à ceux de l'expérience historique. Pourquoi Peter Buri en a-t-il conclu que cette expérience démontrait l'existence de la dérive génétique ?

On observe qu'il y a dans tous les tubes une fixation d'allèles, si on attend plusieurs dizaines de générations. Cependant, contrairement à une sélection naturelle directionnelle, ce n'est pas toujours le même allèle qui est fixé. Cela peut aussi bien être l'allèle sauvage que l'allèle récessif *bw*, de façon équiprobable puisque la fréquence de *bw* dans l'ensemble des tubes reste de 0,5. Si on répète la simulation sans changer les paramètres, les résultats obtenus dans chaque tube changent, ainsi que le temps avant fixation. Cela montre qu'il s'agit de phénomènes aléatoires, c'est pourquoi Peter Buri a bien démontré qu'il s'agit de dérive génétique.

→ Que se passe-t-il si on augmente le nombre de mouches par tube ? Discuter des résultats observés.

Si on met 80 mouches par tube, il n'y a plus de fixation d'allèles. La population est de grande taille, mais les allèles étudiés sont neutres. C'est donc la loi d'Hardy-Weinberg qui s'applique. On s'attend donc à ce que, en moyenne, la fréquence de *bw* dans l'ensemble des tubes reste égale à 0,5.

2. Modélisation de l'effet fondateur et du goulot d'étranglement démographique

Afin de d'étudier l'effet fondateur et le goulot d'étranglement démographique, on a modélisé l'évolution d'une population sur un continent et sur des îles colonisées à partir de ce continent.

→ Toujours sur le site web <http://virtualbiologylab.org>, choisir à présent le modèle 2

Les individus sont représentés par des couleurs qui correspondent à leur génotype dans le cas où il y aurait codominance des deux allèles. Les allèles des populations sont symbolisés avec une couleur jaune, bleue ou rouge à l'état homozygote, ou une superposition de deux de ces couleurs s'il s'agit d'un hétérozygote (par exemple, un individu violet a pour génotype bleu/rouge). Tous les génotypes ont même valeur sélective.

On dispose d'un continent (« *Mainland* »), dont on fixe la capacité d'accueil (« *Mainland-K* »). Les fréquences des allèles de la population du continent sont données en bas à gauche. Sur le continent, on peut déclencher un goulot d'étranglement démographique (presser le bouton violet « *Bottleneck* »), en fixant au préalable le nombre de survivants (« *Bottleneck-size* »).

→ **Fixer une population continentale à environ 800 individus et créer un goulot d'étranglement démographique avec dix survivants. Renouveler cinq fois la simulation. Obtient-on toujours les mêmes résultats ? Pourquoi ?**

Après chaque simulation, les résultats sont différents en terme de structure génétique de la population. Les allèles présents et leur fréquence fluctuent d'une simulation à l'autre. On s'y attend car après le goulot d'étranglement, la population est tellement réduite qu'elle évolue au hasard selon la dérive génétique, jusqu'à ce que l'effectif se rétablisse. Durant ce temps, l'évolution est totalement aléatoire.

L'île 1 est colonisée à partir de la population continentale, et l'île 2 à partir de l'île 1. Pour chaque île, il est possible de déterminer le nombre d'individus migrant sur d'île (« *Founding N-x* ») et la capacité d'accueil de l'île (« *Isle-x-K* »).

→ **Choisir des paramètres induisant une évolution de la population continentale selon la loi de Hardy-Weinberg, puis réaliser des colonisations par une dizaine d'individus, pour une capacité d'accueil de 50 individus.**

→ **Comparer les diversités génétiques des populations du continent et des deux îles au cours de la même simulation et d'une simulation à l'autre.**

Pour être dans les conditions de Hardy-Weinberg, il suffit ici de choisir une grande population de plusieurs centaines d'individus, car les traits sont neutres, et la population est fermée. Suite à la simulation proposée, les îles ne possèdent pas tous les allèles de la population source, soit à cause du génotype des fondateurs, soit par élimination d'allèles par dérive génétique après la colonisation. Ces deux effets étant aléatoires, ce ne sont pas toujours les mêmes allèles qui sont fixés ou éliminés. Les trois populations n'ont donc pas les mêmes allèles, et pas les mêmes fréquences d'allèles, que ce soit lors d'une simulation ou d'une simulation à l'autre.

→ **Que se passe-t-il si on augmente la capacité d'accueil des îles à 200 individus ?**

Si on augmente la capacité d'accueil des îles, il reste toujours un échantillonnage au hasard des fondateurs, suivi d'une période de dérive génétique, tant que la population est réduite. Mais ensuite, la population de l'île atteint au moins une centaine d'individus et évolue selon Hardy-Weinberg, et les fréquences alléliques se modifient peu.

→ **D'après cette simulation et vos connaissances, pourquoi les îles proches des continents sont-elles riches en espèces endémiques qu'on ne retrouve nulle part ailleurs ?**

Le hasard intervient dans le tri des allèles des populations des îles, par effet fondateur d'abord, puis par des périodes de dérive génétique chaque fois que la population est peu nombreuse, soit parce que la capacité d'accueil est faible, soit par des effets de goulot d'étranglement climatiques. Il en résulte soit une extinction, soit la

formation d'une population avec un polymorphisme génétique particulier. Si le flux génétique entre cette population et celle de sa population source s'interrompt, alors une nouvelle espèce endémique de l'île se forme.

Conclusion

Les mécanismes de l'évolution découlent de nombreuses observations et peuvent être testés expérimentalement. L'évolution des espèces repose sur des interactions avec les paramètres biotiques et abiotiques.

Les mécanismes de l'évolution modifie la biodiversité, parfois de manière extrêmement rapide, entraînant parfois l'émergence de nouvelles espèces et la disparition d'autres.

La compréhension de ces mécanismes permet l'établissement des relations de parenté entre les organismes, ces derniers gardant des traces de leur histoire évolutive, aussi bien dans leur morphologie, leur anatomie, que dans leur génome.