

Un modèle stochastique de démo-génétique

Camille Coron

1 Motivations

Ce travail de recherche, qui est présenté plus en détail dans les articles [3], [4] et [1], a pour but d'étudier la dynamique de la composition génétique d'une population. Les modèles classiques dans ce domaine sont le modèle de Wright-Fisher (introduit par [6], [7] et étudié en détail dans [5]) et ses extensions (diploïdie, ajout de sélection et de mutations, etc...), qui supposent des générations discrètes, non-chevauchantes, mais surtout une taille de population constante ou en tout cas définie indépendamment de la composition génétique de la population. Pour éviter cette hypothèse, qui rend difficile l'étude de certains phénomènes liés à l'extinction des populations (comportement spécifique des individus vivant dans des populations de petite taille, consanguinité, vortex d'extinction, etc...), notre approche consiste à modéliser cette population à l'aide d'un modèle individu-centré, c'est-à-dire un modèle dans lequel le comportement de la population est le résultat du comportement de chaque individu qui la compose.

2 Modèle et changement d'échelle

Modèle On considère une population d'individus diploïdes, caractérisés par leur génome à un locus présentant 2 allèles, A et a . Les individus ont donc pour génotype AA , Aa ou aa , et on modélise la population par un processus de naissance et mort à 3 types $(N_{AA}(t), N_{Aa}(t), N_{aa}(t))_{t \geq 0}$, avec compétition et reproduction mendélienne. Plus précisément, si la population est dans l'état $(n_{AA}, n_{Aa}, n_{aa}) \in \mathbb{N}^3$ et si l'on note $n = n_{AA} + n_{Aa} + n_{aa}$ sa taille, alors le taux $\lambda_{AA}(n_{AA}, n_{Aa}, n_{aa})$ (resp. $\mu_{AA}(n_{AA}, n_{Aa}, n_{aa})$) auquel un individu de génotype AA naît (resp. meurt), vaut :

$$\begin{aligned}\lambda_{AA}(n_{AA}, n_{Aa}, n_{aa}) &= bn \left(\frac{2n_{AA} + n_{Aa}}{2n} \right)^2, \\ \mu_{AA}(n_{AA}, n_{Aa}, n_{aa}) &= n_{AA}(d_{AA} + cn).\end{aligned}$$

Ces taux de naissance et mort sont le résultat de la modélisation suivante : chaque individu meurt naturellement au taux d , se reproduit au taux b et fait mourir par compétition chaque autre individu au taux c . Un individu qui se reproduit choisit un partenaire uniformément au hasard et cette rencontre donne lieu à la naissance d'un nouvel individu. Chacun des deux parents transmet l'un de ses deux allèles, choisi indépendamment et uniformément au hasard, parmi les deux qu'il possède, à ce nouvel individu. Les quatre autres taux de saut sont définis de façon analogue.

Changement d'échelle Ce processus de naissance et mort (du fait notamment de sa dimension et de la non-linéarité de ses taux de saut) est généralement difficile à étudier. Une stratégie classique et très importante en dynamique des populations consiste à étudier un changement d'échelle de ce processus (voir [2] pour plus de motivations et d'exemples). Dans ce cas précis, supposons que pour tout $i \in \{AA, Aa, aa\}$:

$$b = \gamma K + \beta, \quad d_i = \gamma K + \delta_i, \quad \text{et} \quad c = \alpha/K, \quad (1)$$

où K est un paramètre d'échelle qui quantifie la taille initiale de population, et qui tend vers l'infini. Ce changement d'échelle signifie que l'on s'intéresse à une population constituée d'un très grand nombre de petits individus (la compétition totale générée par un individu étant de l'ordre de $1/K$) qui se reproduisent et meurent très rapidement.

Processus limite Notons $(N_{AA}^K(t), N_{Aa}^K(t), N_{aa}^K(t))_{t \geq 0}$ le processus de naissance et mort introduit précédemment, qui dépend maintenant du paramètre d'échelle K . Lorsque K tend vers l'infini, nous observons la convergence du processus de dimension 3 $(N_t^K, X_t^K, Y_t^K)_{t \geq 0}$ tel que

$$\begin{aligned} N_t^K &= \frac{N_{AA}^K(t) + N_{Aa}^K(t) + N_{aa}^K(t)}{K} \quad \text{est la taille de population} \\ X_t^K &= \frac{N_{Aa}^K(t) + 2N_{aa}^K(t)}{2N_{AA}^K(t) + 2N_{Aa}^K(t) + 2N_{aa}^K(t)} \quad \text{la proportion d'allèle } a, \\ Y_t^K &= \frac{4N_{AA}^K(t)N_{aa}^K(t) - (N_{Aa}^K(t))^2}{(N_{AA}^K(t) + N_{Aa}^K(t) + N_{aa}^K(t))^2} \quad \text{la déviation par rapport à la structure de Hardy-Weinberg,} \end{aligned}$$

vers une dynamique lente-rapide. Plus précisément, la variable Y^K tend vers 0, tandis que le couple $(N_t^K, X_t^K)_{t \geq 0}$ converge vers un processus de diffusion de dimension 2 $(N_t, X_t)_{t \geq 0}$, qui satisfait l'équation suivante :

$$\begin{aligned} dN_t &= \sqrt{2\gamma N_t} dB_t^1 + N_t \left[\rho - \xi N_t + \sigma X_t (2h + X_t(1 - 2h)) \right] dt, \\ dX_t &= \sqrt{\frac{2\gamma X_t(1 - X_t)}{2N_t}} dB_t^2 + \sigma X_t(1 - X_t) \left[h + X_t(1 - 2h) \right] dt, \end{aligned} \tag{2}$$

où $(B_t^1, B_t^2)_{t \geq 0}$ est un mouvement brownien standard de dimension 2, et les paramètres ρ , σ et h dépendent des paramètres β , et δ_i introduits dans le changement d'échelle (??).

3 Comportement en temps long

3.1 Comportement quasi-stationnaire

Le comportement en temps long du processus de diffusion donné par (??) est simple :

Théorème 3.1. *Presque sûrement, $T_0 = \inf\{t > 0 : N_t = 0\} < \infty$, et $\inf\{t > 0 : X_t \in \{0, 1\}\} < T_0^N$.*

Ce résultat signifie d'une part que la population finira par s'éteindre et d'autre part qu'avant cette extinction, l'un des allèles initialement présents aura envahi la population (on dit que la fixation a lieu avant l'extinction). Il est donc plus intéressant d'étudier le comportement de cette population conditionnée à la survie, qui est observé dans les populations réelles, et donné par le théorème suivant :

Théorème 3.2. *Il existe une mesure de probabilité ν sur $[0, 1] \times \mathbb{R}_+^*$ telle que pour tout point de départ (n, x) , la loi de (N_t, X_t) sachant que $N_t > 0$ converge vers ν lorsque le temps t tend vers l'infini.*

En particulier la loi de la proportion X_t sachant que $N_t > 0$ converge lorsque t tend vers l'infini, et la loi limite est représentée pour différentes valeurs de paramètres, en figure ???. On remarquera que selon les valeurs de paramètres utilisées, on pourra observer une coexistence en temps long des allèles, ou non. Ce comportement en temps long montre une stratégie qui facilite la survie de la population.

3.2 Impact de la démographie sur la fixation des mutations

Notons que le Théorème ??, dont le résultat paraît plutôt intuitif, n'est pas vrai pour n'importe quelle dynamique de taille de population, comme nous le montrons dans la Figure ???.

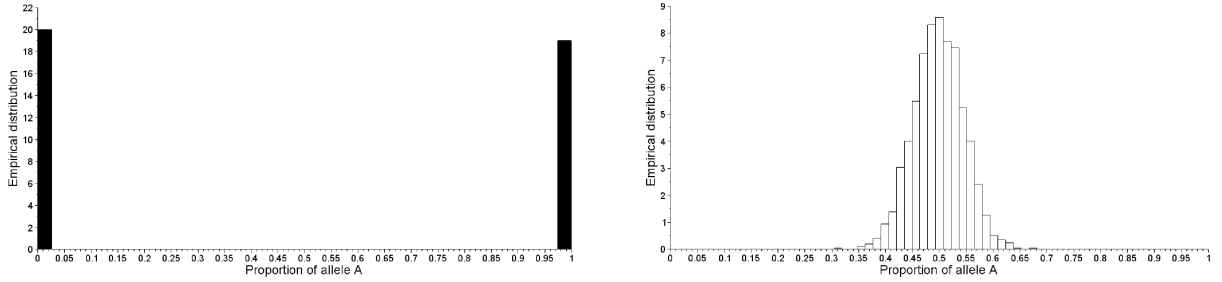


FIGURE 1 – Loi limite de la proportion d’allèle a , dans un cas neutre ($\delta_i = \delta$ pour tout i), et un cas de surdominance ($\delta_2 < \delta_1 = \delta_3$).

FIGURE 2 – Pour cet exemple, la taille de population ($N_t, t \geq 0$) et la proportion d’allèle a ($X_t, t \geq 0$) sont supposées suivre l’équation suivante :

$$dN_t = (N_t)^{1/2-\epsilon} dB_t^1 + N_t[\rho - \xi N_t]dt,$$

$$dX_t = \sqrt{\frac{X_t(1-X_t)}{2N_t}} dB_t^2,$$

où $(B_t^1, B_t^2)_{t \geq 0}$ est un mouvement brownien standard de dimension 2, et $\epsilon \in \mathbb{R}$. Nous montrons que tant que $\epsilon > 0$, il existe une probabilité strictement positive pour que l’extinction ait lieu avant la fixation. Ce résultat est illustré dans cette figure pour $\epsilon = 0.3$, $\rho = -1$ et $\xi = 0.1$.

