Modèles mathématiques de forçage génétique

Dynamiques stochastiques à l'arrière d'un front d'éradication.

Léna Kläy
lena.klay@univ-brest.fr

en collaboration avec

Léo Girardin, Florence Débarre, Vincent Calvez.

19 juin 2023

Introduction

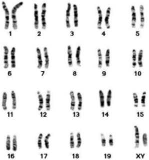
Forçage génétique et caryotypes

Forçage génétique : espèces sexuées, principalement diploïdes.

Drosophile

Grenouille

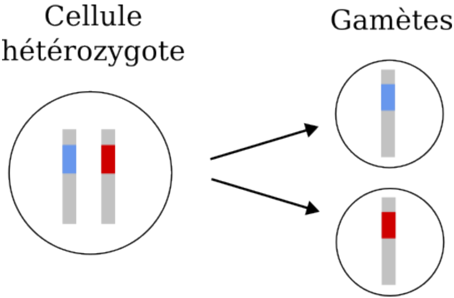
Souris


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
XY

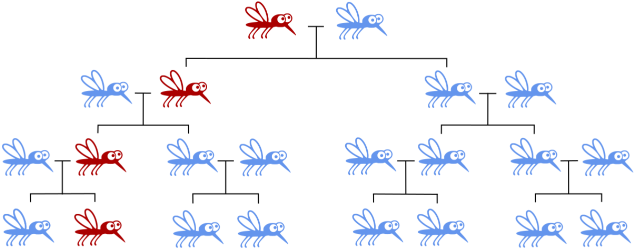
Arabidopsis

Propagation classique (probabilité Mendélienne)

Si le parent est hétérozygote, il a 50% de chance de transmettre l'allèle rouge à sa descendance.


Cellule hétérozygote
Gamètes

Dans une population qui ne contient que des allèles bleus:



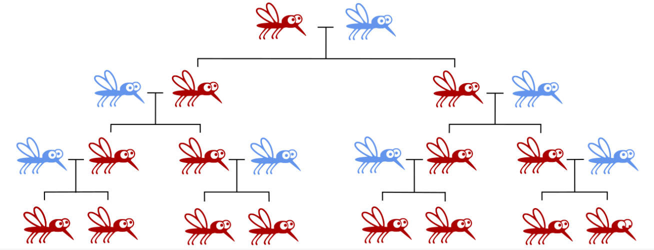
En rouge, les individus ayant au moins un allèle rouge.

Introduction

Forçage génétique

Allèle forcé génétiquement :

Allèle ayant un taux de transmission plus élevé que le taux Mendelien.



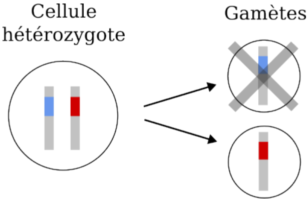
En rouge, les individus ayant au moins un allèle rouge.

Possibilité nouvelle de modifier des populations sauvages, en propageant un trait d'intérêt.

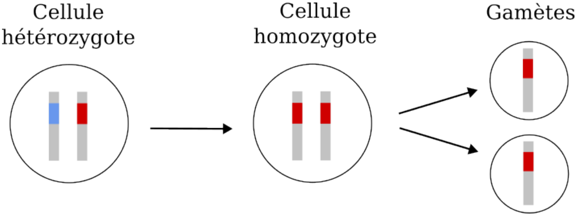
Introduction
Stratégies de forçage génétique

Deux méthodes de forçage génétique *(Gene drive)*.

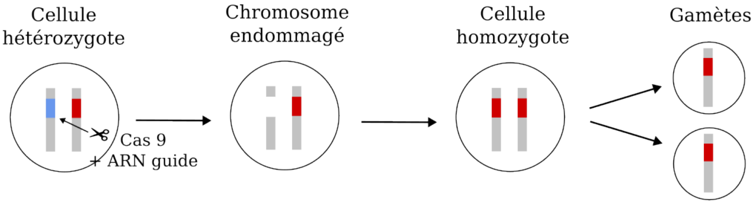
Détruire les allèles bleus *(Toxin-Antidote)* :


Cellule hétérozygote
Gamètes

Convertir l'allèle bleu en un allèle rouge *(Homing)* :

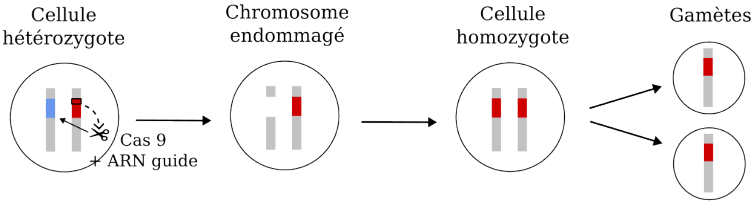

Cellule hétérozygote
Cellule homozygote
Gamètes

Forçage génétique par conversion (CRISPR Cas9)


Cellule hétérozygote
Cas 9
+ ARN guide
Chromosome endommagé
Cellule homozygote
Gamètes

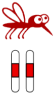
Recombinaison Homogologue (HDR).

Forçage génétique par conversion (CRISPR Cas9)

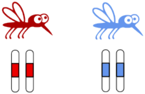

Cellule hétérozygote
Cas 9
+ ARN guide
Chromosome endommagé
Cellule homozygote
Gamètes

Recombinaison Homogologue (HDR).

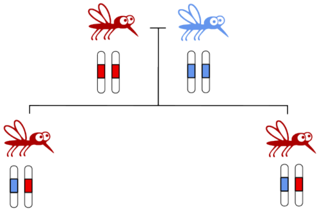
Conversion à chaque génération



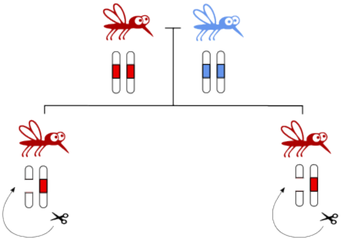
Conversion à chaque génération



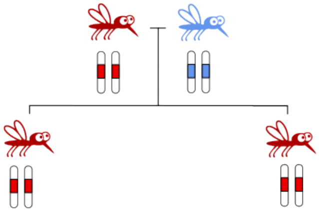
Conversion à chaque génération



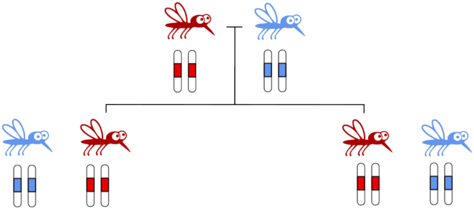
Conversion à chaque génération



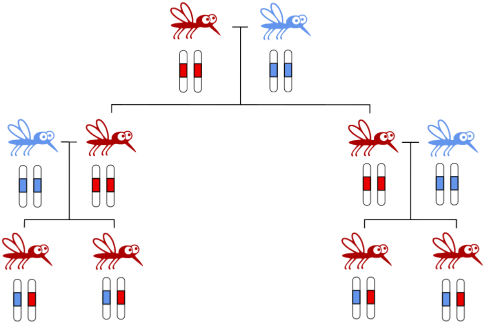
Conversion à chaque génération



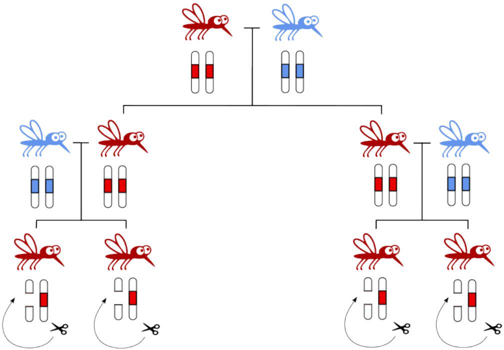
Conversion à chaque génération



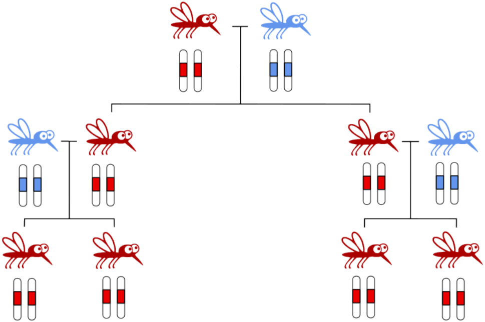
Conversion à chaque génération



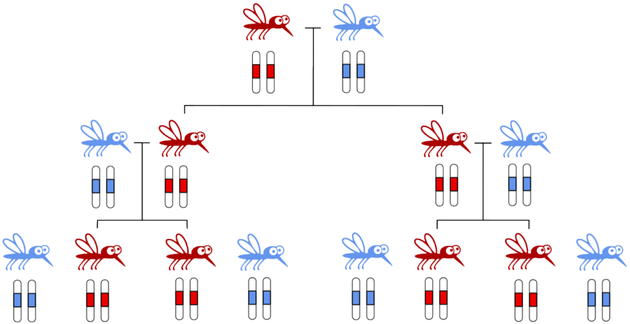
Conversion à chaque génération



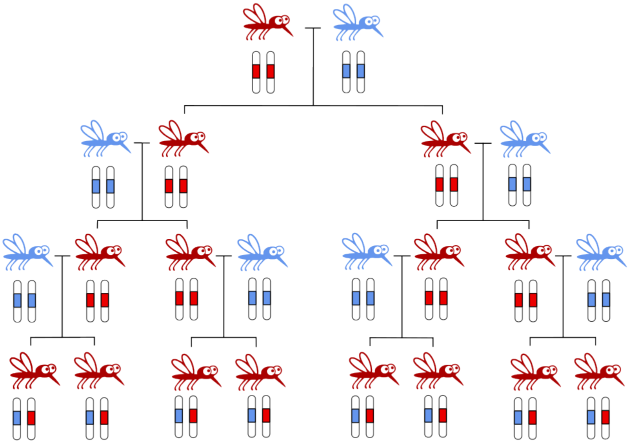
Conversion à chaque génération



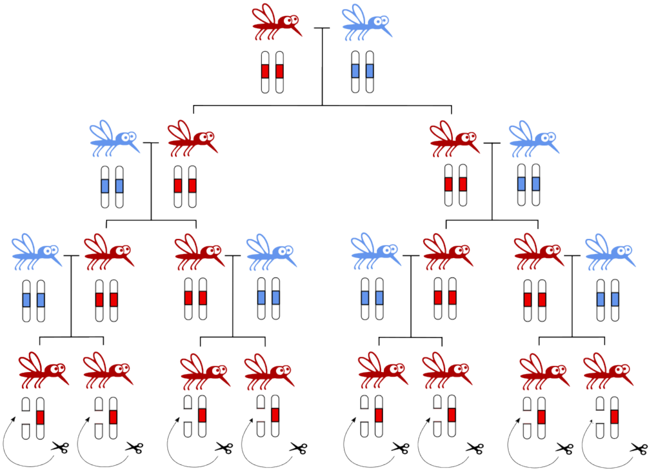
Conversion à chaque génération



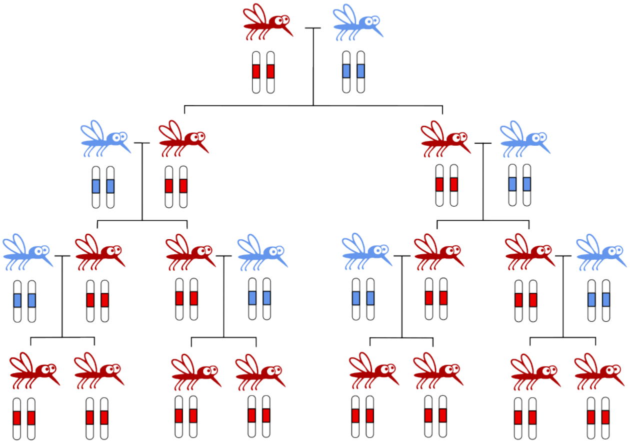
Conversion à chaque génération



Conversion à chaque génération



Conversion à chaque génération

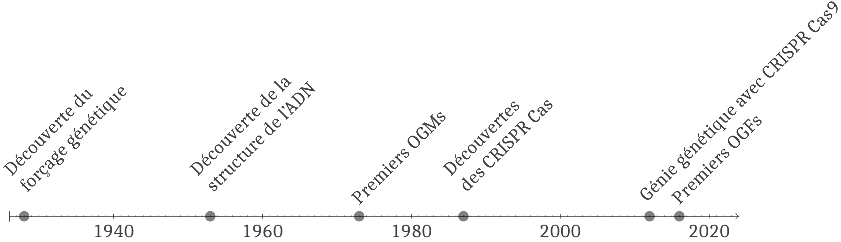


Contexte

Chronologie

OGM : Organisme génétiquement modifié (GMO)

OGF : Organisme génétiquement forcé (GDO)



Découverte du forçage génétique

Découverte de la structure de l'ADN

Premiers OGMs

Découvertes des CRISPR Cas

Génie génétique avec CRISPR Cas9

Premiers OGFs

Précisions

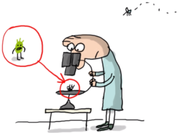
Aucun lâcher dans la nature aujourd'hui

Temps de génération courts (quelques mois)

Applications
Modification

Modification d’une population sauvage :

- Rendre les moustiques résistants à certains pathogènes,
- Éliminer la résistance aux pesticides chez des espèces "nuisibles".



Santé publique

Conservation

Agriculture

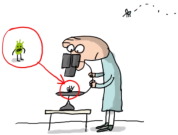
Dessins de Jacques Azam, `1jour1actu.com`

Applications

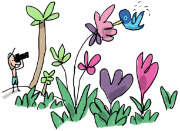
Éradication

Éradication d'une population sauvage :

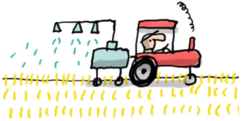
- Éliminer les espèces vectrices de maladie,
- Éliminer les espèces invasives,
- Éliminer les ravageurs.



Santé publique



Conservation



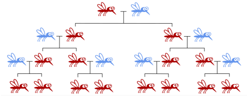
Agriculture

Dessins de Jacques Azam, `1jour1actu.com`

Désavantage de fitness du drive

Avantage de conversion $c \in [0,1]$

Une grande part des descendants héritent de la modification.



Désavantage de fitness

Les individus modifiés ont en moyenne moins de descendants.

$+$ les individus sont modifiés dans la population,

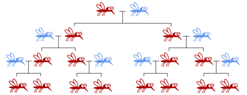
$+$ plus le taux de natalité va chuter,

$\rightarrow$ peut mener à l'extinction.

Désavantage de fitness du drive

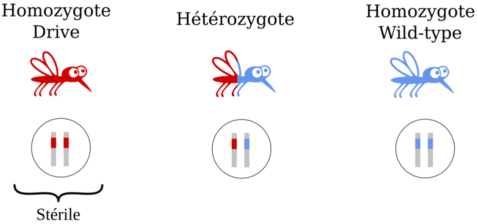
Avantage de conversion $c \in [0, 1]$

Une grande part des descendants héritent de la modification.



Désavantage de fitness

Extinction observée en laboratoire (gène *doublesex*)



Homozygote Drive

Hétérozygote

Homozygote Wild-type

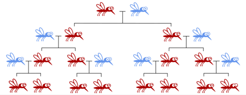
Stérile

600 moustiques dont 150 modifiés → extinction en 8 à 12 générations.

Désavantage de fitness du drive

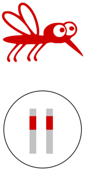
Avantage de conversion $c \in [0, 1]$

Une grande part des descendants héritent de la modification.



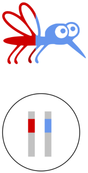
Désavantage de fitness $s \in [0, 1], h \in [0, 1]$

Homozygote
Drive



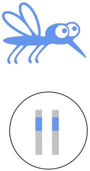
Fitness : 1-s

Hétérozygote



Fitness : 1-sh

Homozygote
Wild-type



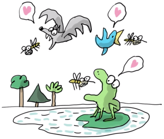
Fitness : 1

Bénéfices-risques

Bénéfices

En santé publique, environnementaux et économiques.

Risques identifiés



Conséquences mal connues sur le long terme,

Déséquilibre des réseaux d'interaction,

Propagation à d'autres régions ou d'autres espèces.

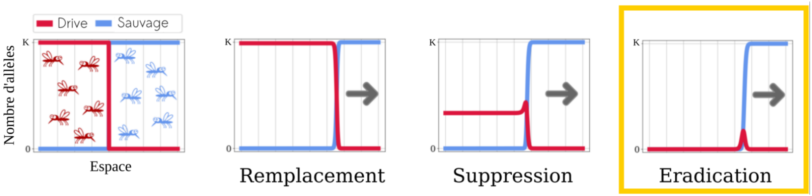
Considérations morales comme...

Dispose t-on trop facilement des autres êtres vivants ?

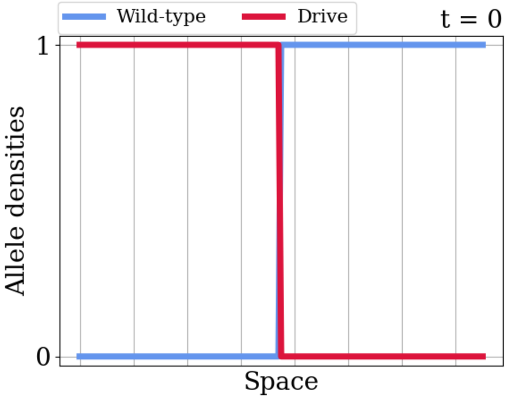
Peut-on moralement planifier l'éradication d'une espèce ?

Modélisation mathématique

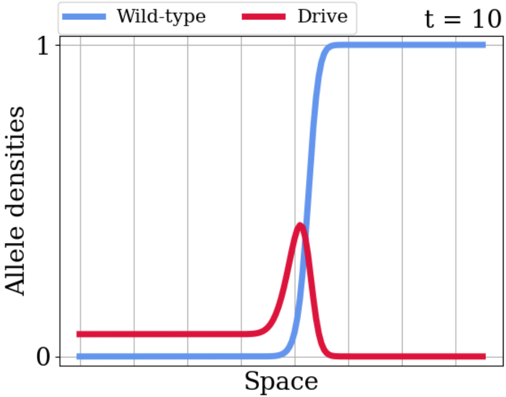
Dans cet exposé, on va se concentrer sur le drive d’éradication.



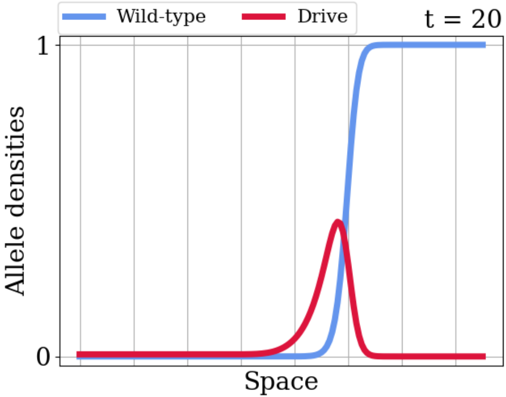
Front d’éradication (modèle déterministe)



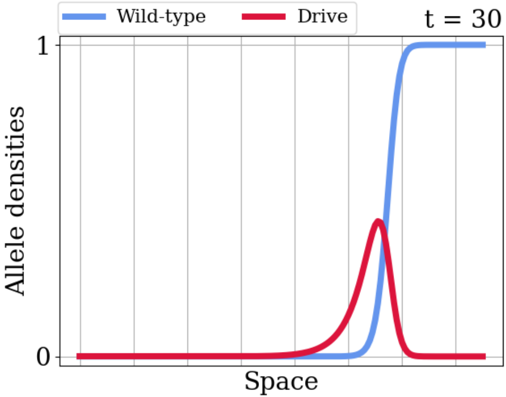
Front d’éradication (modèle déterministe)



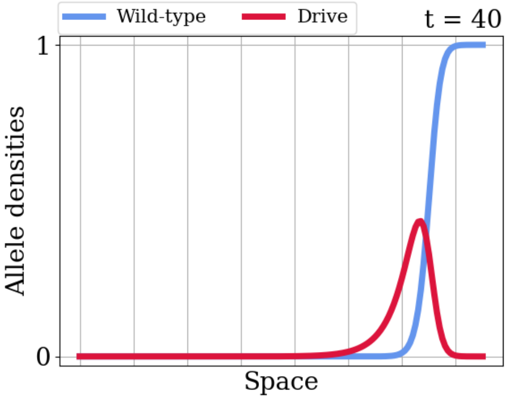
Front d’éradication (modèle déterministe)



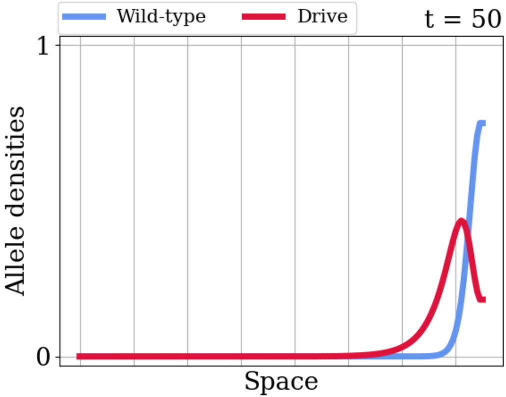
Front d'éradication (modèle déterministe)



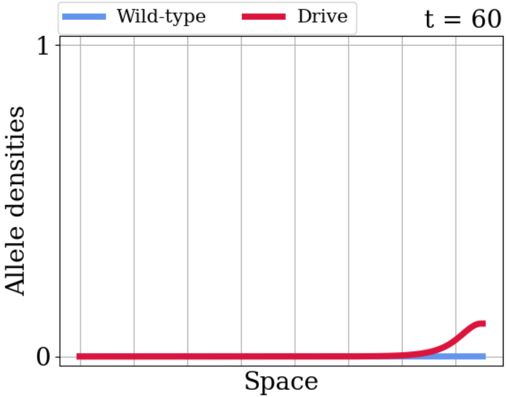
Front d’éradication (modèle déterministe)



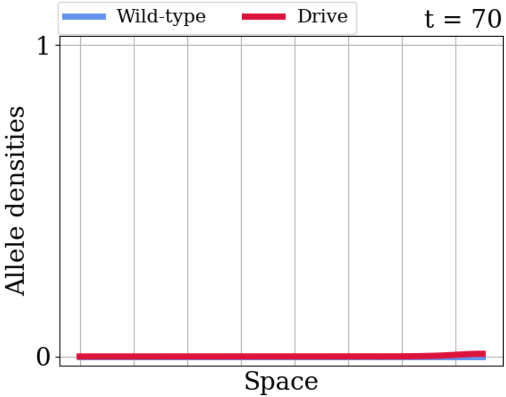
Front d’éradication (modèle déterministe)



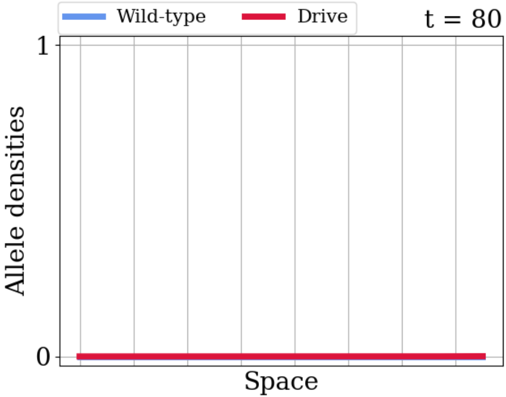
Front d’éradication (modèle déterministe)



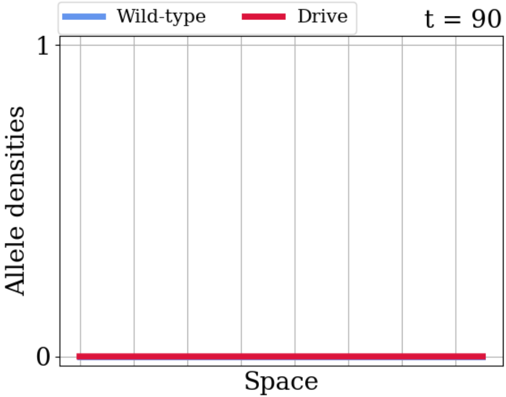
Front d'éradication (modèle déterministe)



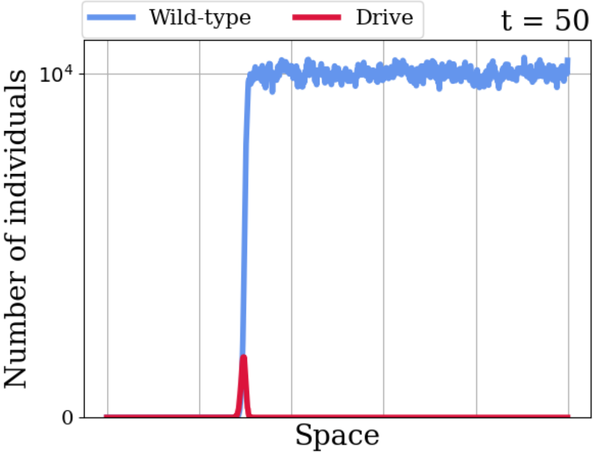
Front d’éradication (modèle déterministe)



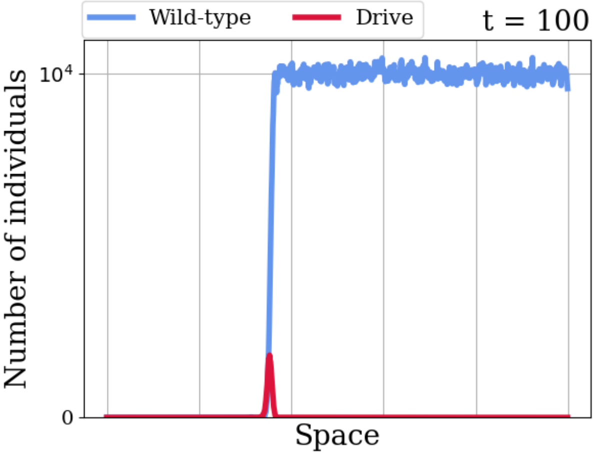
Front d'éradication (modèle déterministe)



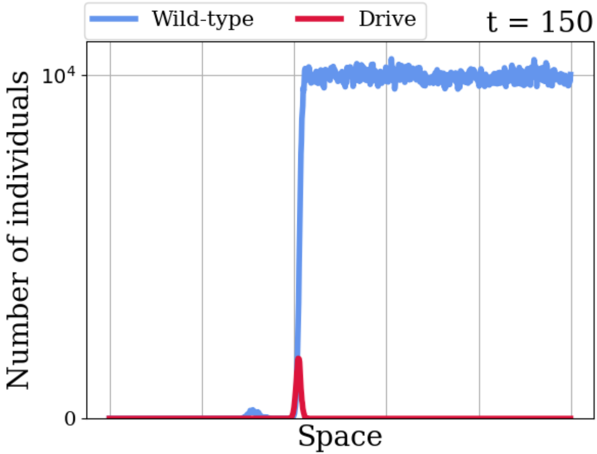
Front d'éradication (modèle stochastique)



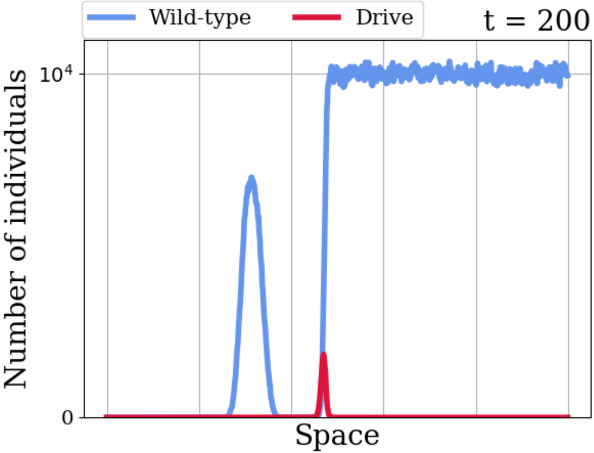
Front d'éradication (modèle stochastique)



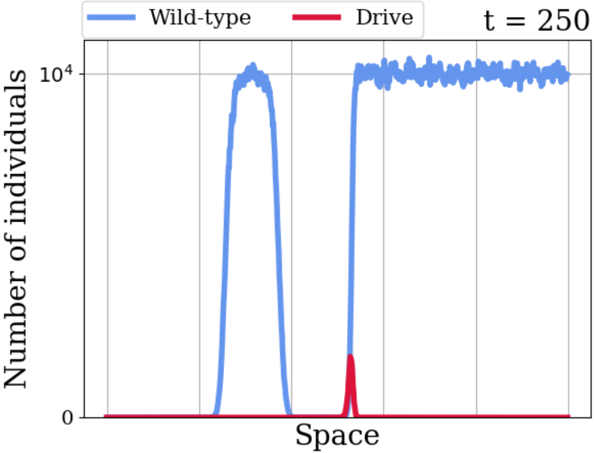
Front d'éradication (modèle stochastique)



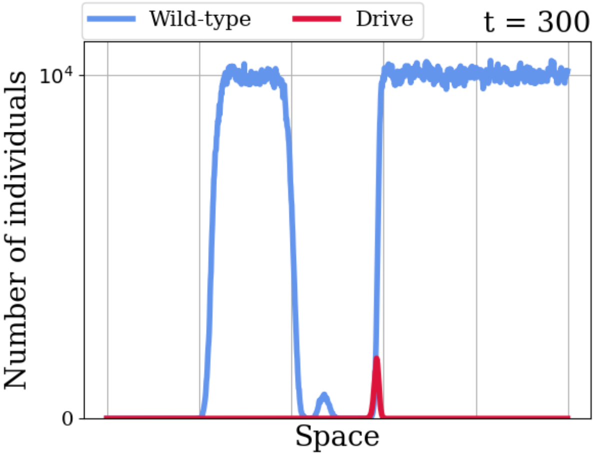
Front d’éradication (modèle stochastique)



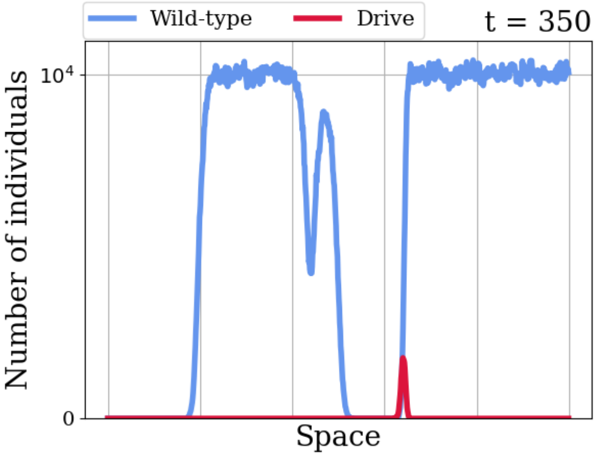
Front d'éradication (modèle stochastique)



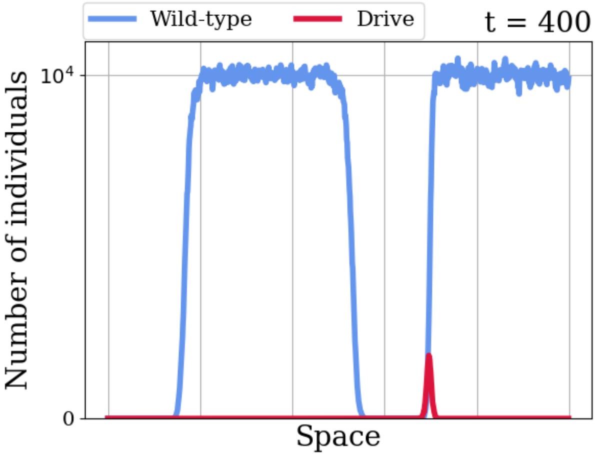
Front d’éradication (modèle stochastique)



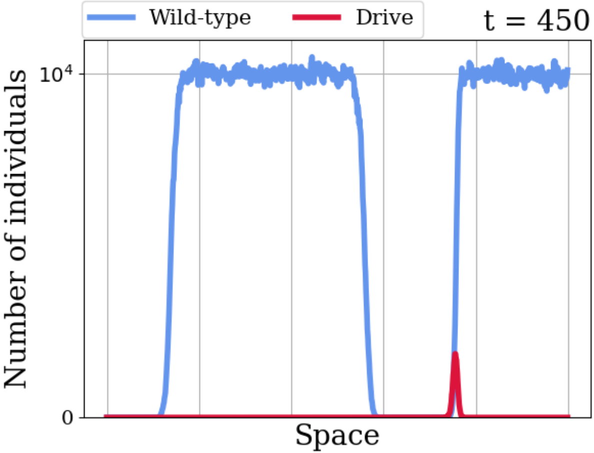
Front d’éradication (modèle stochastique)



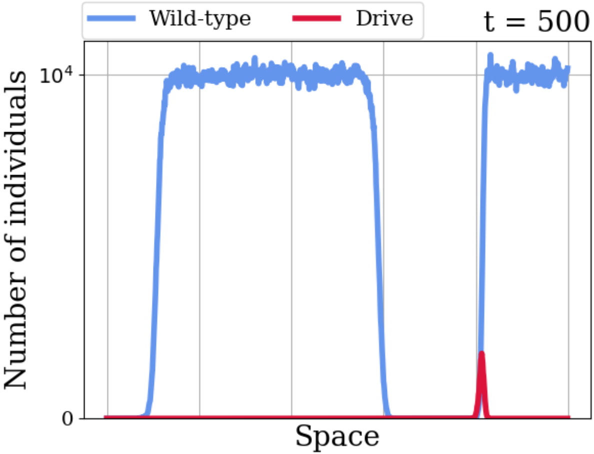
Front d’éradication (modèle stochastique)



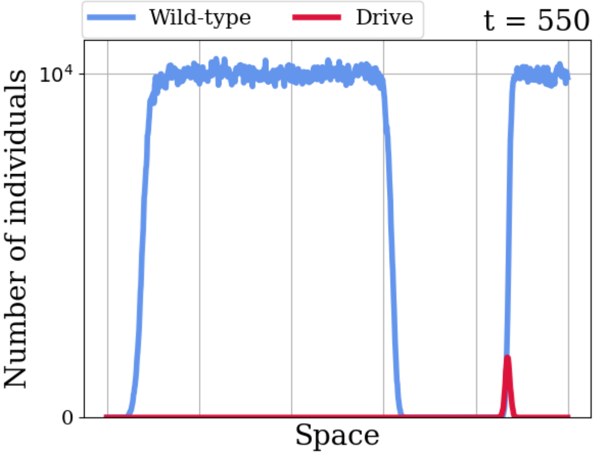
Front d’éradication (modèle stochastique)



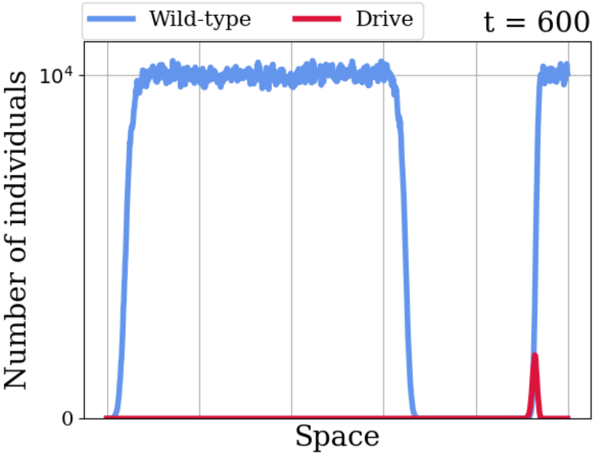
Front d’éradication (modèle stochastique)



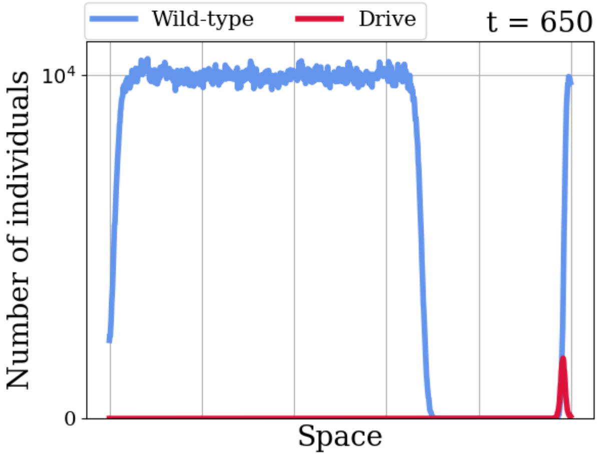
Front d'éradication (modèle stochastique)



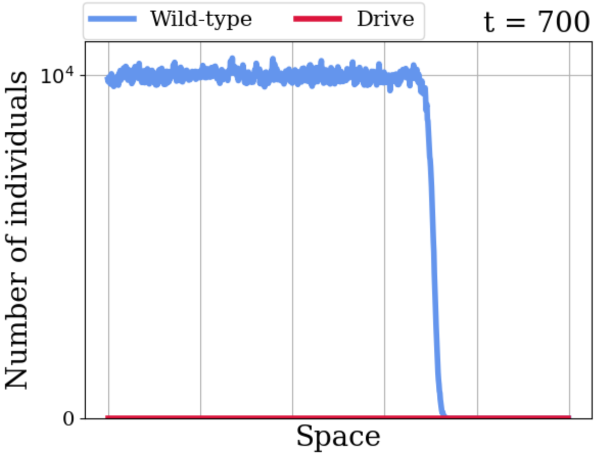
Front d’éradication (modèle stochastique)



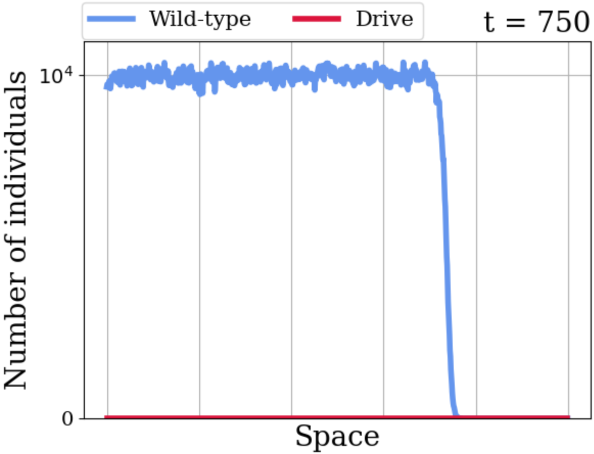
Front d'éradication (modèle stochastique)



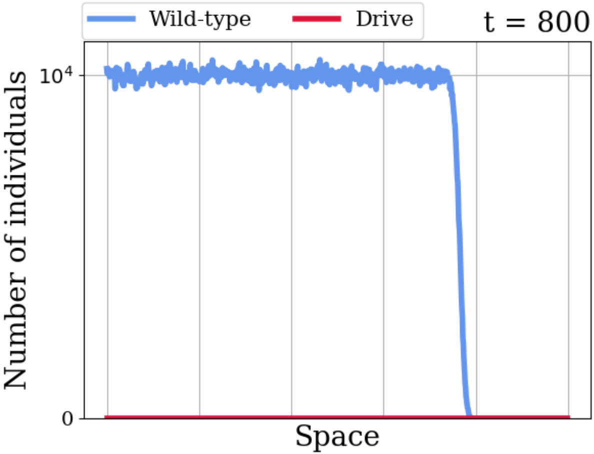
Front d’éradication (modèle stochastique)



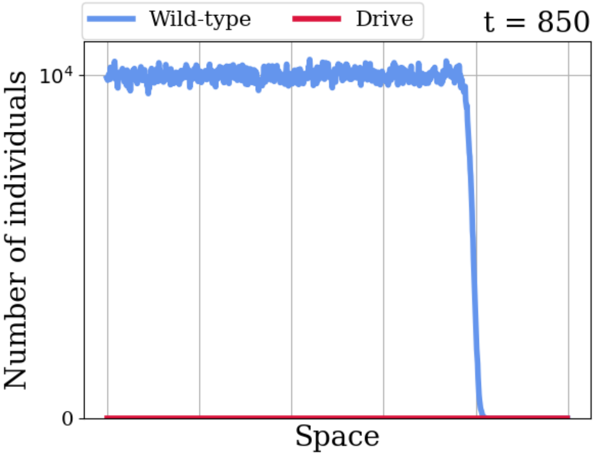
Front d'éradication (modèle stochastique)



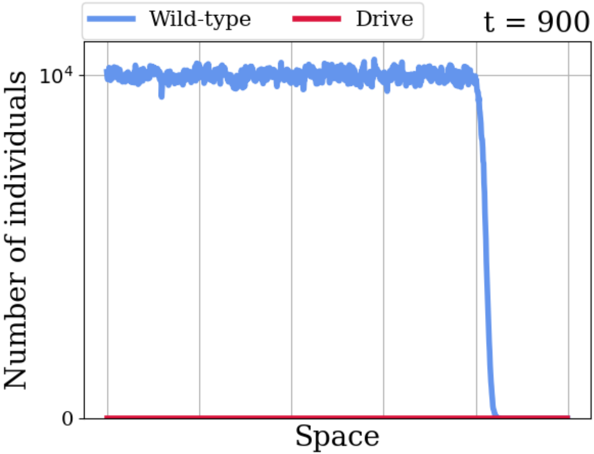
Front d’éradication (modèle stochastique)



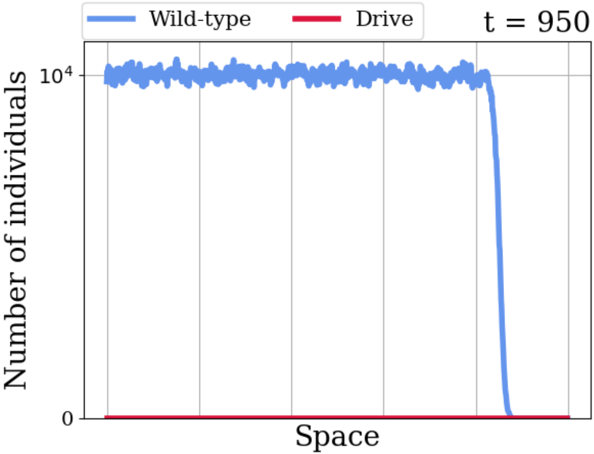
Front d’éradication (modèle stochastique)



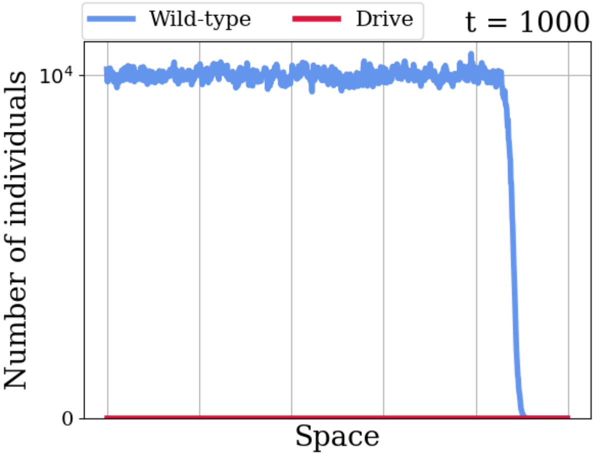
Front d'éradication (modèle stochastique)



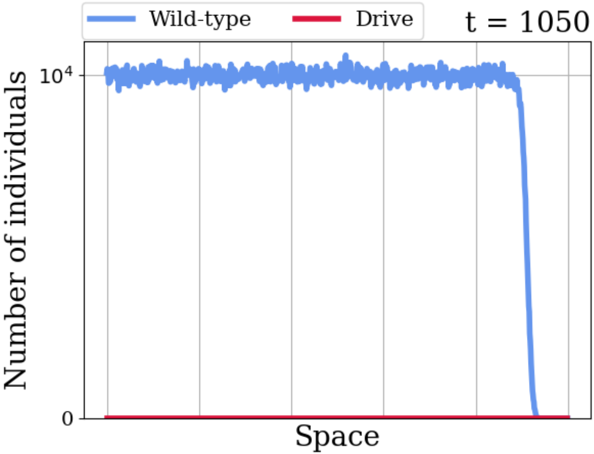
Front d'éradication (modèle stochastique)



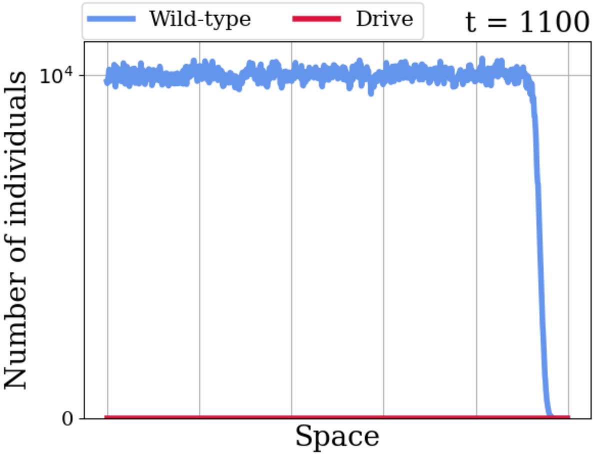
Front d’éradication (modèle stochastique)



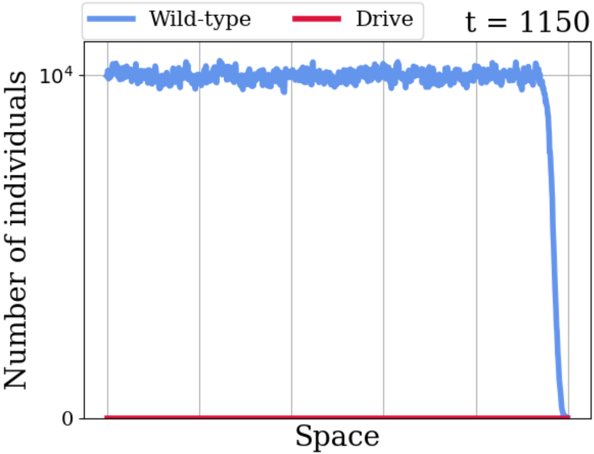
Front d’éradication (modèle stochastique)



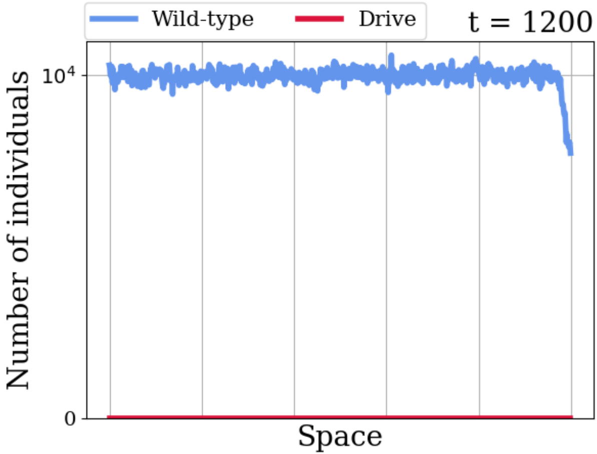
Front d’éradication (modèle stochastique)



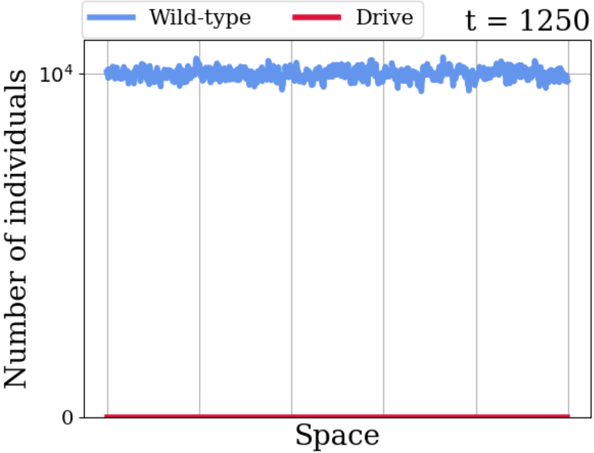
Front d’éradication (modèle stochastique)



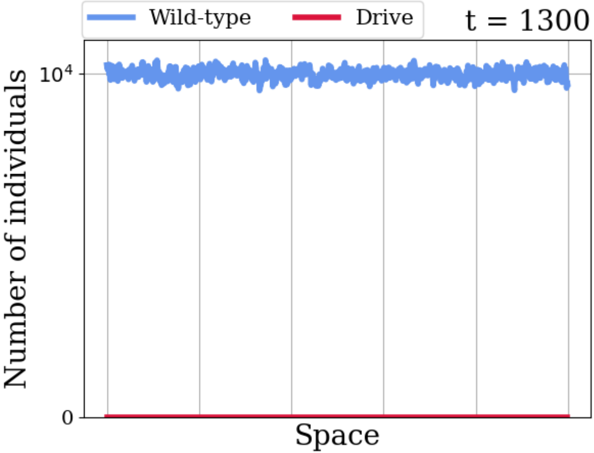
Front d'éradication (modèle stochastique)



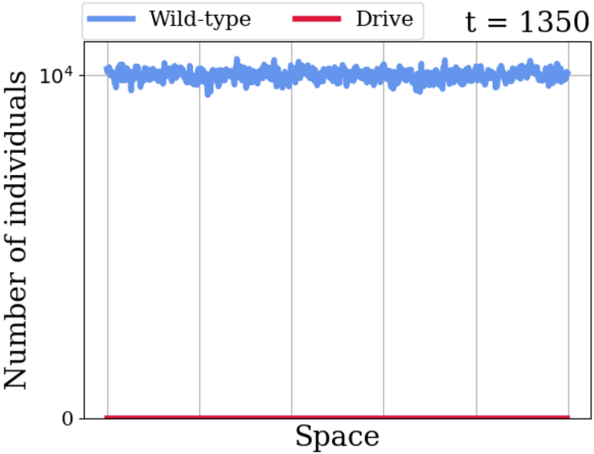
Front d'éradication (modèle stochastique)



Front d'éradication (modèle stochastique)



Front d'éradication (modèle stochastique)



Modèle déterministe

continu en temps et en espace

Modèle déterministe, continu en temps et en espace

$$\begin{cases} \overset{\text{Variations in time}}{\boxed{\partial_t n_{\mathrm{D}}}} - \overset{\text{Diffusion}}{\boxed{\sigma^2 \partial^2_{xx} n_{\mathrm{D}}}} = \overset{\text{Growth term}}{\boxed{\left(r\left(1-\frac{n}{K}\right)+1\right)}} \left[\overset{\text{Fitness}}{\boxed{(1-s)}} \overset{\text{Genotype}}{\boxed{\frac{n_{\mathrm{D}}^2}{n}}} + \overset{\text{Fitness}}{\boxed{(1-sh)}} \overset{\text{Genotype}}{\boxed{\frac{2\, n_{\mathrm{W}} n_{\mathrm{D}}}{n}}} \overset{\text{Conversion}}{\boxed{\frac{1+c}{2}}} \right] - \overset{\text{Mortality}}{\boxed{n_{\mathrm{D}}}}, \\ \boxed{\partial_t n_{\mathrm{W}}} - \boxed{\sigma^2 \partial^2_{xx} n_{\mathrm{W}}} = \boxed{\left(r\left(1-\frac{n}{K}\right)+1\right)} \left[\boxed{1} \boxed{\frac{n_{\mathrm{W}}^2}{n}} + \boxed{(1-sh)} \boxed{\frac{2\, n_{\mathrm{W}} n_{\mathrm{D}}}{n}} \boxed{\frac{1-c}{2}} \right] - \boxed{n_{\mathrm{W}}}. \end{cases}$$

n_{D} : nombre d'allèles drive (rouge),

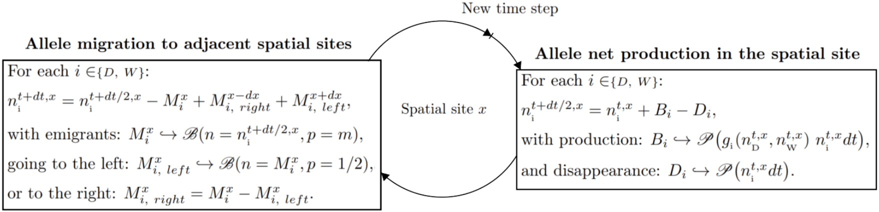
n_{W} : nombre d'allèles wild-type (bleu).

Résolution numérique par différences finies (Crank-Nicolson).

Modèle stochastique

discret en temps et en espace

Modèle stochastique, discret en temps et en espace



New time step

Allele migration to adjacent spatial sites

For each $i \in \{D, W\}$:

$n_{\text{i}}^{t+dt,x} = n_{\text{i}}^{t+dt/2,x} - M_i^x + M_{i,\ right}^{x-dx} + M_{i,\ left}^{x+dx}$,

with emigrants: $M_i^x \hookrightarrow \mathscr{B}(n = n_{\text{i}}^{t+dt/2,x}, p = m)$,

going to the left: $M_{i,\ left}^x \hookrightarrow \mathscr{B}(n = M_i^x, p = 1/2)$,

or to the right: $M_{i,\ right}^x = M_i^x - M_{i,\ left}^x$.

Spatial site x

Allele net production in the spatial site

For each $i \in \{D, W\}$:

$n_{\text{i}}^{t+dt/2,x} = n_{\text{i}}^{t,x} + B_i - D_i$,

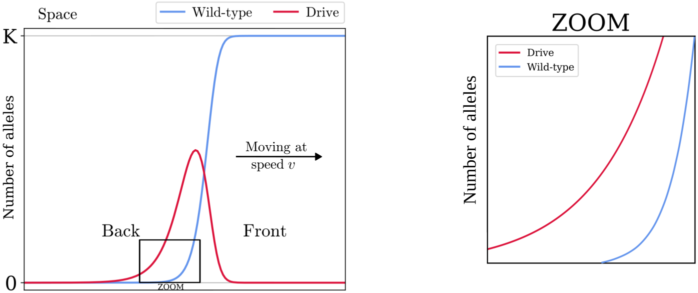
with production: $B_i \hookrightarrow \mathscr{P}\big(g_{\text{i}}(n_{\text{D}}^{t,x}, n_{\text{W}}^{t,x})\ n_{\text{i}}^{t,x} dt\big)$,

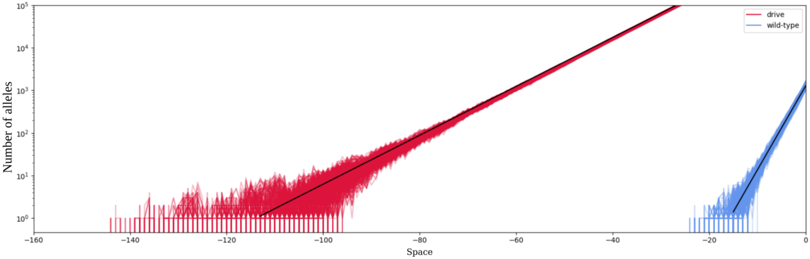
and disappearance: $D_i \hookrightarrow \mathscr{P}\big(n_{\text{i}}^{t,x} dt\big)$.

- une étape de migration,
- une étape de naissance-mort.

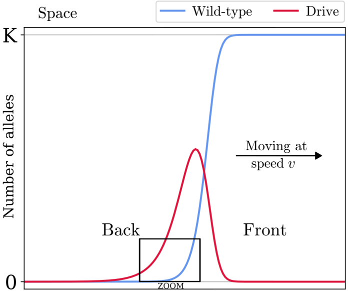
Modèle population-centré (en opposition à individu-centré).

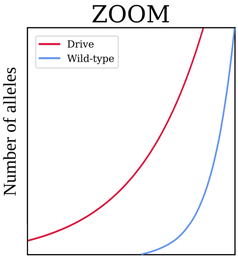
A l’arrière d’un front d’éradication

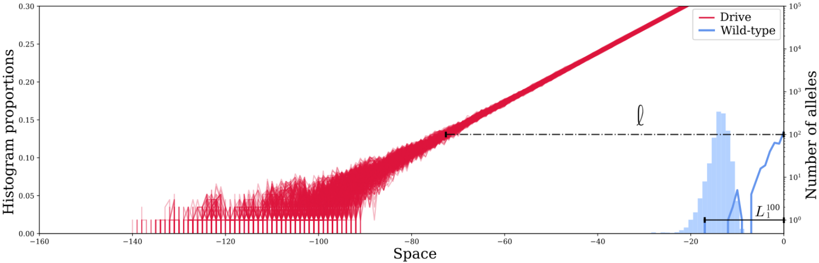




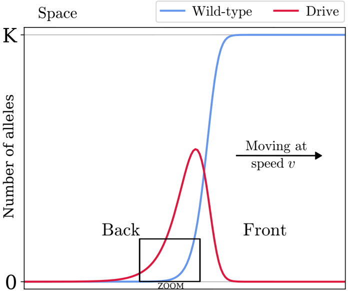
A l’arrière d’un front d’éradication

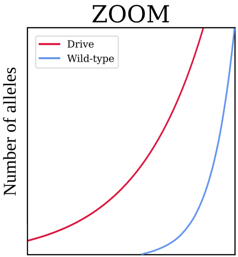


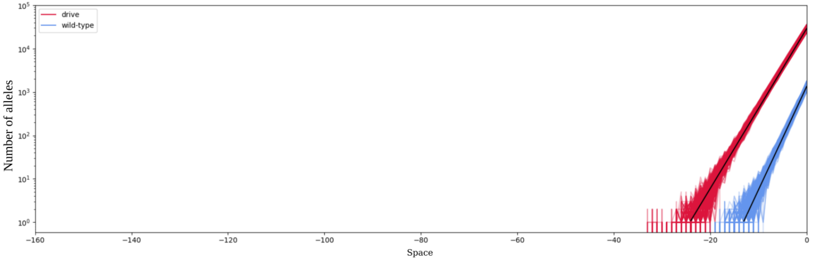




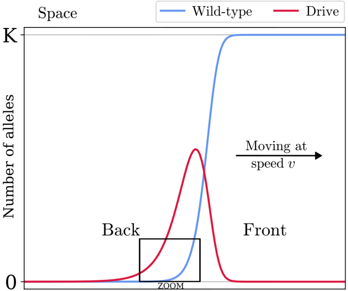
A l’arrière d’un front d’éradication

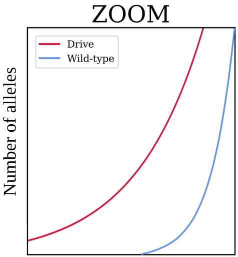


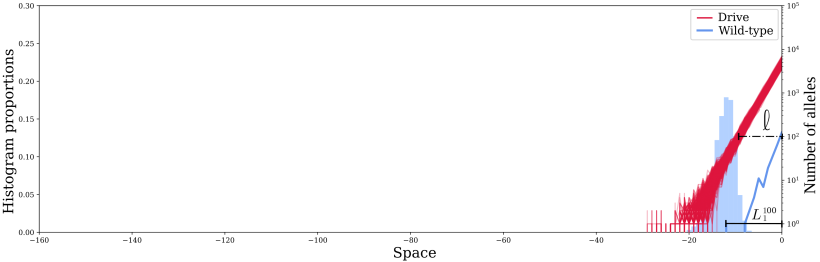




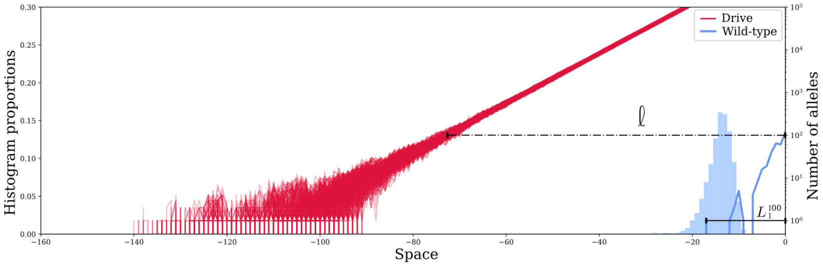
A l’arrière d’un front d’éradication







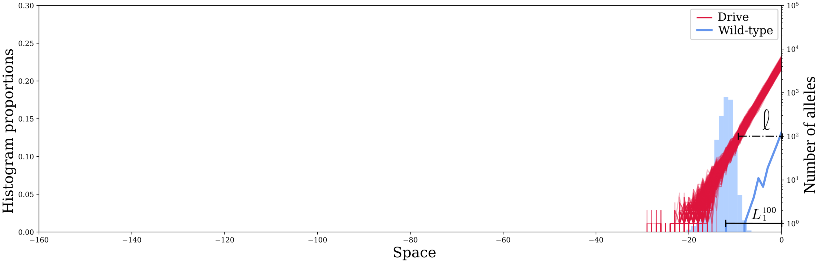
Critère de faible probabilité de recolonisation



- Si le dernier allèle bleu est entouré de plus de 100 allèles rouge, la recolonisation est très peu probable.*
- De manière équivalente : si $\ell > L_1^{\mathcal{N}}$, la recolonisation est très peu probable.*

*(dans un laps de temps raisonnable)

Critère de faible probabilité de recolonisation



- Si le dernier allèle bleu se situe sur un site avec moins de 100 allèles rouge, alors on ne peut rien dire.
- De manière équivalente : si $\ell < L_1^{\mathcal{N}}$, on ne peut rien assurer.

Résultats antérieurs

Facteurs favorisant les phénomènes de recolonisation :

- un taux de diffusion faible [Paril et al. 2022; Champer et al. 2021],
- un taux de consanguinité élevé [Champer et al. 2021],
- l'absence d'espèces compétitrices ou prédatrices [Liu et al. 2023].

Notre étude porte sur de l'influence de :

- le coût de fitness drive s (résultats numériques contradictoires [Paril et al. 2022; Champer et al. 2021]),
- la capacité de charge de l'environnement K (jamais étudié),
- le taux de diffusion.

Déterministe pour caractériser ℓ grâce aux profils exponentiels,

Stochastique pour approximer L_1^{100} via un processus de Galton-Watson.

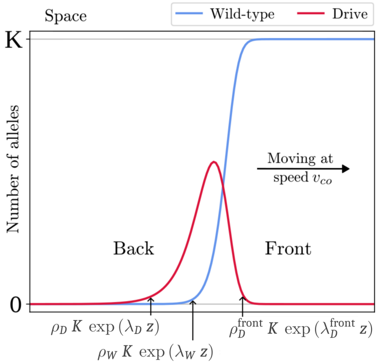
Étude de la distance déterministe ℓ

On se place dans le cas d'un front d'invasion drive tiré :

$$(1 - sh)(1 + c) > 1,$$

et d'éradication :

$$r < \frac{s}{1 - s}.$$



La dynamique du système :

$$\begin{cases} \partial_t n_{\mathrm{D}} - \sigma^2 \partial_{xx}^2 n_{\mathrm{D}} = \left(r\left(1 - \frac{n}{K}\right) + 1 \right) \left[(1 - s)\frac{n_{\mathrm{D}}^2}{n} + (1 - sh)\, \frac{2\, n_{\mathrm{W}} n_{\mathrm{D}}}{n}\, \frac{1 + c}{2} \right] - n_{\mathrm{D}}, \\ \partial_t n_{\mathrm{W}} - \sigma^2 \partial_{xx}^2 n_{\mathrm{W}} = \left(r\left(1 - \frac{n}{K}\right) + 1 \right) \left[\frac{n_{\mathrm{W}}^2}{n} + (1 - sh)\, \frac{2\, n_{\mathrm{W}} n_{\mathrm{D}}}{n}\, \frac{1 - c}{2} \right] - n_{\mathrm{W}}. \end{cases}$$

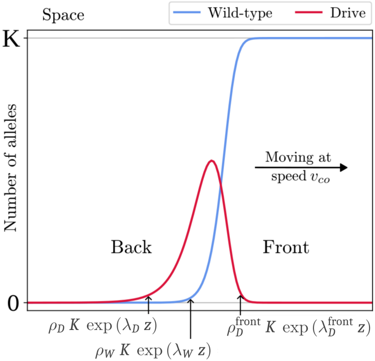
Étude de la distance déterministe ℓ

On se place dans le cas d'un front d'invasion drive tiré :

$$(1 - sh)(1 + c) > 1,$$

et d'éradication :

$$r < \frac{s}{1-s}.$$



A l'arrière du front :

$$\begin{cases} 0 = \sigma^2(\lambda_D)^2 + v\lambda_D + \Big[(r+1)(1-s) - 1\Big], \\ 0 = \sigma^2(\lambda_W)^2 + v\lambda_W + \Big[(r+1)(1-sh)(1-c) - 1\Big]. \end{cases}$$

Étude de la distance déterministe ℓ

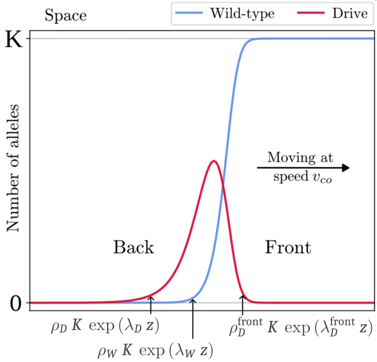
On se place dans le cas d'un front d'invasion drive tiré :

$$(1 - sh)(1 + c) > 1,$$

et d'éradication :

$$r < \frac{s}{1 - s}.$$

On connaît alors :


Space
Wild-type
Drive
K
0
Number of alleles
Moving at speed v_{co}
Back
Front
$\rho_D\ K\ \exp\left(\lambda_D\ z\right)$
$\rho_W\ K\ \exp\left(\lambda_W\ z\right)$
$\rho_D^{\text{front}}\ K\ \exp\left(\lambda_D^{\text{front}}\ z\right)$

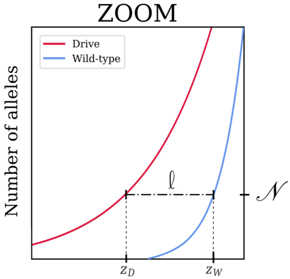
$$v = 2\sigma\sqrt{(1 - sh)(1 + c) - 1},$$

$$\lambda_D^{front} = \frac{1}{\sigma}\left(-\sqrt{(1 - sh)(1 + c) - 1}\right),$$

$$\lambda_D = \frac{1}{\sigma}\left(-\sqrt{(1 - sh)(1 + c) - 1} + \sqrt{(1 - sh)(1 + c) - (r + 1)(1 - s)}\right),$$

$$\lambda_W = \frac{1}{\sigma}\left(-\sqrt{(1 - sh)(1 + c) - 1} + \sqrt{(1 - sh)(1 + c) - (r + 1)(1 - sh)(1 - c)}\right).$$

Relation entre ℓ et les profils exponentiels



Soit z_D et z_W tels que :

$$\rho_D\, K\, \exp(\lambda_D\, z_D) = \mathcal{N},$$

$$\rho_W\, K\, \exp(\lambda_W\, z_W) = \mathcal{N}.$$

Nous pouvons écrire ℓ comme :

$$\ell = z_W - z_D = \frac{1}{\lambda_W} \log\left(\frac{\mathcal{N}}{\rho_W K}\right) - \frac{1}{\lambda_D} \log\left(\frac{\mathcal{N}}{\rho_D K}\right).$$

Comme $\rho_W < 1 \ll K/\mathcal{N}$ et $\rho_D < 1 \ll K/\mathcal{N}$ en échelle logarithmique :

$$\begin{aligned} \ell &= \left(\frac{1}{\lambda_D} - \frac{1}{\lambda_W}\right) \log\left(\frac{K}{\mathcal{N}}\right) - \frac{1}{\lambda_W} \log(\rho_W) + \frac{1}{\lambda_D} \log(\rho_D) \\ &\sim \left(\frac{1}{\lambda_D} - \frac{1}{\lambda_W}\right) \log\left(\frac{K}{\mathcal{N}}\right) \end{aligned}$$

Désavantage de fitness du drive s

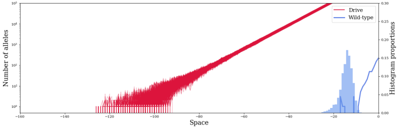
$$\ell \sim \left(\frac{1}{\lambda_D} - \frac{1}{\lambda_W}\right) \log\left(\frac{K}{\mathcal{N}}\right)$$

Grâce au système sur λ_D et λ_W, on prouve que $\frac{d}{ds}\left(\frac{1}{\lambda_D} - \frac{1}{\lambda_W}\right) < 0$.

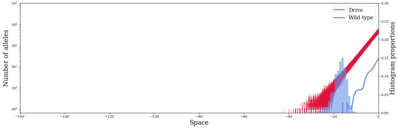
$$\begin{cases} 0 = \sigma^2(\lambda_D)^2 + v\lambda_D + \Big[(r+1)(1-s) - 1\Big], \\ \\ 0 = \sigma^2(\lambda_W)^2 + v\lambda_W + \Big[(r+1)(1-sh)(1-c) - 1\Big]. \end{cases}$$

Ainsi quand s augmente, ℓ décroit.

$s = 0.3$



$s = 0.7$



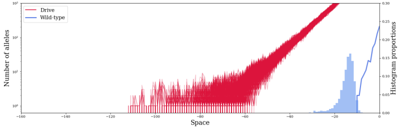
Capacité de charge de l'environnement K

$$\ell \sim \left(\frac{1}{\lambda_D} - \frac{1}{\lambda_W}\right) \log\left(\frac{K}{\mathcal{N}}\right)$$

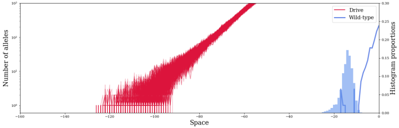
λ_D et λ_W sont indépendantes de K, ainsi ℓ croit avec K.

Si K est multiplié d'un facteur 10, ℓ croit de $\log(10)\left(\frac{1}{\lambda_D} - \frac{1}{\lambda_W}\right)$.

$K = 10^5$



$K = 10^8$



NB : K influence la hauteur du front sans changer la valeur des pentes.

Coefficient de diffusion σ

$$\ell \sim \left(\frac{1}{\lambda_D} - \frac{1}{\lambda_W}\right) \log\left(\frac{K}{\mathcal{N}}\right)$$

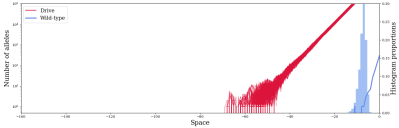
Comme λ_D et λ_W sont inversement proportionnelles à σ,

$$\lambda_D = \frac{1}{\sigma}\left(-\sqrt{(1-sh)(1+c)-1}\sqrt{(1-sh)(1+c)-(r+1)(1-s)}\right),$$

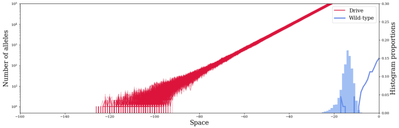
$$\lambda_W = \frac{1}{\sigma}\left(-\sqrt{(1-sh)(1+c)-1}\sqrt{(1-sh)(1+c)-(r+1)(1-sh)(1-c)}\right).$$

Alors ℓ est proportionnelle à σ.

$\sigma = 0.5$



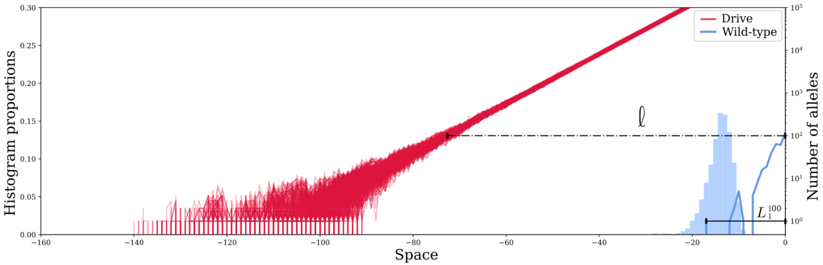
$\sigma = 1$



Conclusion de la partie déterministe

La distance déterministe ℓ décroit lorsque

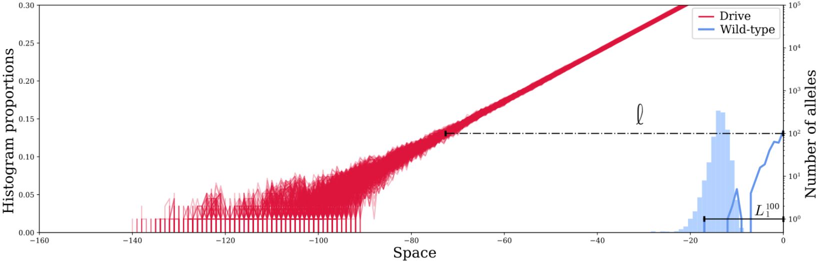
- le drive est moins fit (s croit),
- l'environnement est moins peuplé (K décroit),
- les individus se déplacent moins vite (σ décroit).



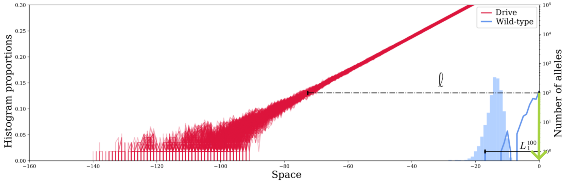
Caractériser les fluctuations stochastiques

Pour caractériser les fluctuations **stochastiques**, on aimerait simplifier notre problème à l’aide d’un processus de Galton-Watson :

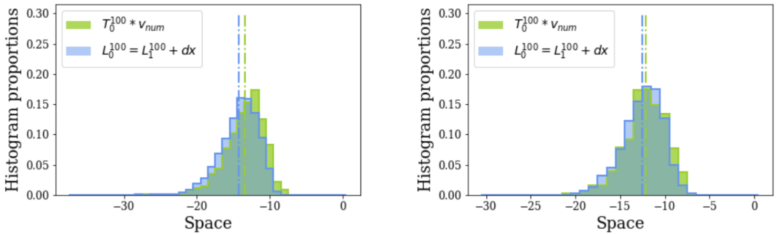
- dont on mesurerait le temps d’extinction,
- composée d’une unique population.



D'une distance à un temps d'extinction



T_0^{100} : temps d'extinction du dernier site contenant plus de 100 allèles sauvages (bleus).



(a) $s = 0.3$

(b) $s = 0.7$

Numériquement T_0^{100} multiplié par la vitesse du front, et L_0^{100} ont la même distribution.

D’un système global à une population isolée

$$\begin{cases} \partial_t n_{\mathrm{D}} - \sigma^2 \partial_{xx}^2 n_{\mathrm{D}} = \left(r \left(1 - \frac{n}{K} \right) + 1 \right) \left[(1-s) \frac{n_{\mathrm{D}}}{n} + (1-sh) \frac{2\, n_{\mathrm{W}}}{n} \frac{1+c}{2} \right] n_{\mathrm{D}} - n_{\mathrm{D}}, \\ \partial_t n_{\mathrm{W}} - \sigma^2 \partial_{xx}^2 n_{\mathrm{W}} = \left(r \left(1 - \frac{n}{K} \right) + 1 \right) \left[\frac{n_{\mathrm{W}}}{n} + (1-sh) \frac{2\, n_{\mathrm{D}}}{n} \frac{1-c}{2} \right] n_{\mathrm{W}} - n_{\mathrm{W}}. \end{cases}$$

Pour approximer la dynamique de la population sauvage à l’arrière du front, on considère une population isolée décrite par :

- la dynamique sauvage lorsque le drive est prédominant ($\frac{n_{\mathrm{D}}}{n} \approx 1$),
- et une petite taille de population (relative, $\frac{n}{K} \approx 0$).

D’un système global à une population isolée

$$\begin{cases} \cancel{\partial_t n_{\mathrm{D}} - \sigma^2 \partial^2_{xx} n_{\mathrm{D}} = \left(r\left(1 - \frac{n}{K}\right) + 1\right) \left[(1-s)\frac{n_{\mathrm{D}}}{n} + (1-sh)\,\frac{2\,n_{\mathrm{W}}}{n}\,\frac{1+c}{2}\right] n_{\mathrm{D}} - n_{\mathrm{D}}}, \\ \boxed{\partial_t n_{\mathrm{W}} - \sigma^2 \partial^2_{xx} n_{\mathrm{W}} = \left(r\left(\cancel{1 - \frac{n}{K}}\right) + 1\right) \left[\cancel{\frac{n_{\mathrm{W}}}{n}} + (1-sh)\,\frac{2\,\cancel{n_{\mathrm{D}}}}{\cancel{n}}\,\frac{1-c}{2}\right] n_{\mathrm{W}} - n_{\mathrm{W}}.} \end{cases}$$

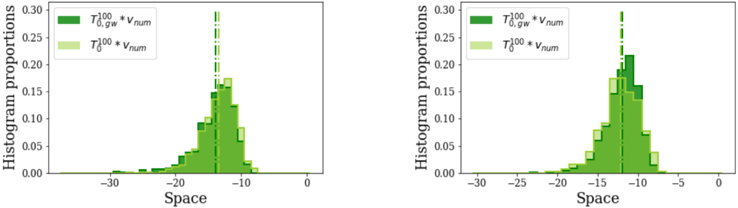
Pour approximer la dynamique de la population sauvage à l’arrière du front, on considère une population isolée décrite par :

- la dynamique sauvage lorsque le drive est prédominant ($\frac{n_{\mathrm{D}}}{n} \approx 1$),
- et une petite taille de population (relative, $\frac{n}{K} \approx 0$).

D’un système global à une population isolée

$$\left\{\begin{array}{l} \cancel{\partial_t n_{\mathrm{D}} - \sigma^2 \partial_{xx}^2 n_{\mathrm{D}} = \left(r\left(1-\frac{n}{K}\right)+1\right)\left[(1-s)\frac{n_{\mathrm{D}}}{n} + (1-sh)\,\frac{2\,n_{\mathrm{W}}}{n}\,\frac{1+c}{2}\right]n_{\mathrm{D}} - n_{\mathrm{D}}}\,, \\ \boxed{\partial_t n_{\mathrm{W}} - \sigma^2 \partial_{xx}^2 n_{\mathrm{W}} = \left(r\,\xcancel{\left(1-\frac{n}{K}\right)}+1\right)\left[\cancel{\frac{n_{\mathrm{W}}}{n}} + (1-sh)\,\frac{2\,\cancel{n_{\mathrm{D}}}}{\cancel{n}}\,\frac{1-c}{2}\right]n_{\mathrm{W}} - n_{\mathrm{W}}\,.} \end{array}\right.$$

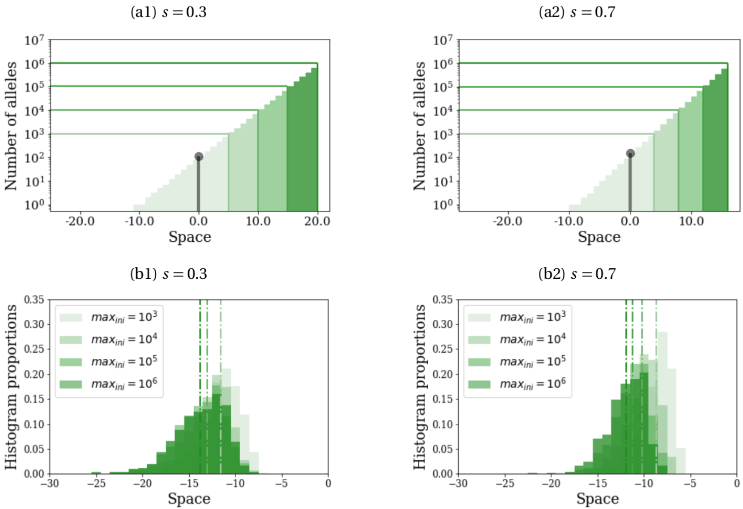
$T_{0,gw}^{100}$: temps d’extinction du processus de Galton-Watson spatial associé.



(a) $s = 0.3$ (b) $s = 0.7$

Numériquement T_0^{100} et $T_{0,gw}^{100}$ ont la même distribution.

… avec une certaine condition initiale



On observe l'importance du "réservoir" d'allèles sauvages (à droite) qui retarde le moment de l'extinction.

Conclusion de la partie stochastique

Comment déterminer jusqu’où peut aller le dernier allèle sauvage ?

On prouve numériquement que l’on peut caractériser $\mathbf{L}_1^{100}$

grâce au temps d’extinction d’un processus spatial de Galton-Watson :

- dont la dynamique est donnée par la population sauvage ($p_{\mathrm{D}} = 1$),
- avec migration,
- en conservant le profil exponentiel comme condition initiale.

Ouvre la voie à une étude probabiliste plus détaillée

Caractériser le temps d’extinction de ce processus de Galton-Watson.

Suite de l’article non présentée ici :

- Étude des dynamiques en 2D,
- Étude des phénomènes de ré-invasion drive.

Merci pour votre attention !

Champer, Jackson et al. (Feb. 2021). “Suppression Gene Drive in Continuous Space Can Result in Unstable Persistence of Both Drive and Wild-Type Alleles”. In: *Molecular Ecology* 30.4, pp. 1086–1101. ISSN: 1365-294X. DOI: `10.1111/mec.15788`.

Liu, Yiran et al. (2023). “Adversarial Interspecies Relationships Facilitate Population Suppression by Gene Drive in Spatially Explicit Models”. In: *Ecology Letters* 26.7, pp. 1174–1185. ISSN: 1461-0248. DOI: `10.1111/ele.14232`.

Paril, Jeff F. and Ben L. Phillips (July 2022). “Slow and Steady Wins the Race: Spatial and Stochastic Processes and the Failure of Suppression Gene Drives”. In: *Molecular Ecology*, mec.16598. ISSN: 0962-1083, 1365-294X. DOI: `10.1111/mec.16598`.