

Génomique évolutive

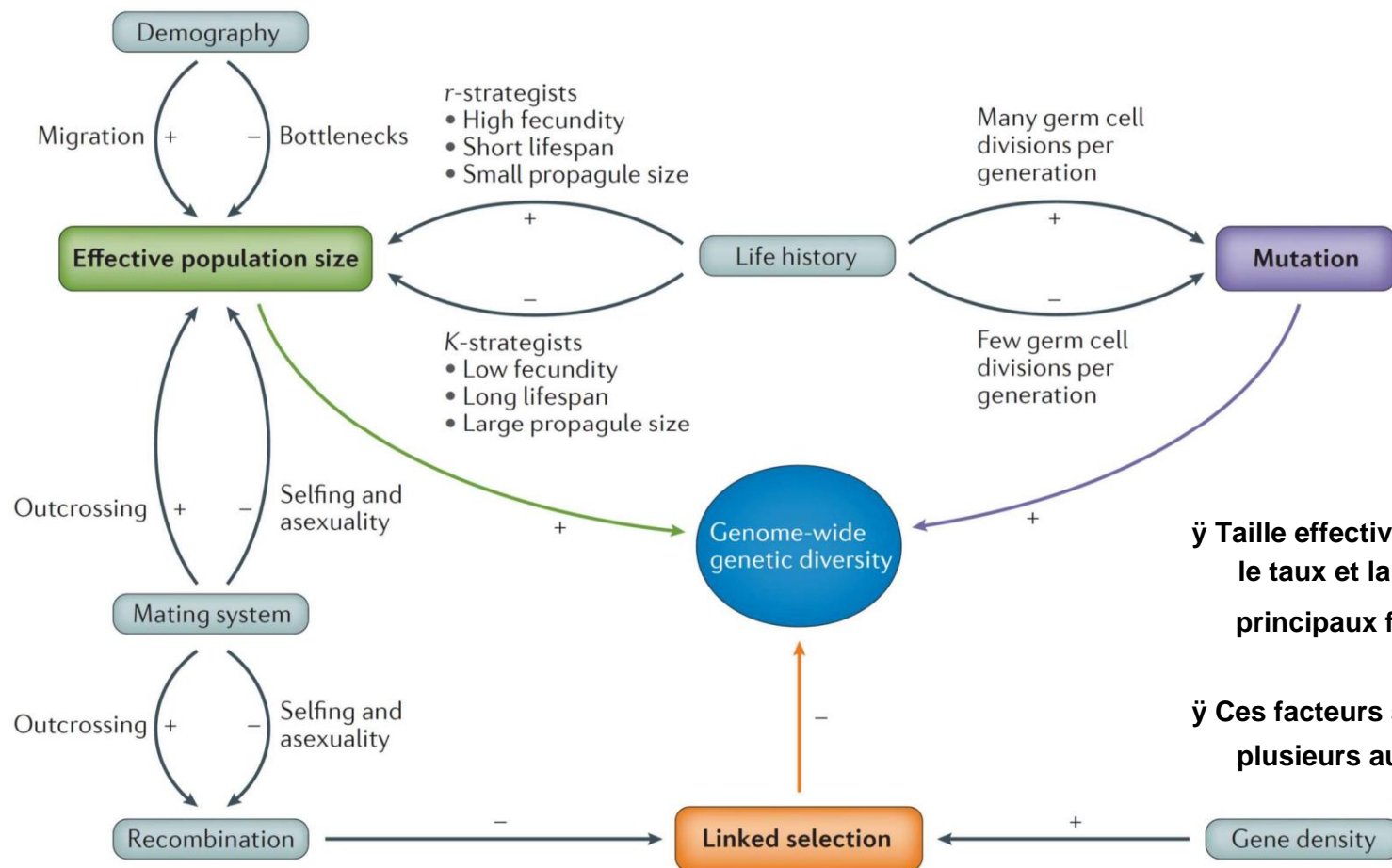
M2 Graduate School « Biologie évolutive »

M2 Méthodes informatiques et statistiques pour les omiques (MISO)

Mutations : taux d'émergence, effet et évolution

Eléonore Durand (eleonore.durand@univ-lille.fr)

Aperçu des déterminants de la diversité génétique



• Taille effective de la population, mutation le taux et la sélection liée sont les principaux facteurs influant sur la diversité.

• Ces facteurs sont à leur tour régis par plusieurs autres paramètres.

Les mutations : la source ultime de toute variation génétique

I. Comment estimer les taux de mutation ?

II. Variation du taux de mutation

III. Mutations et évolution adaptative : exemples avec des expériences de sélection artificielle

Types de mutations

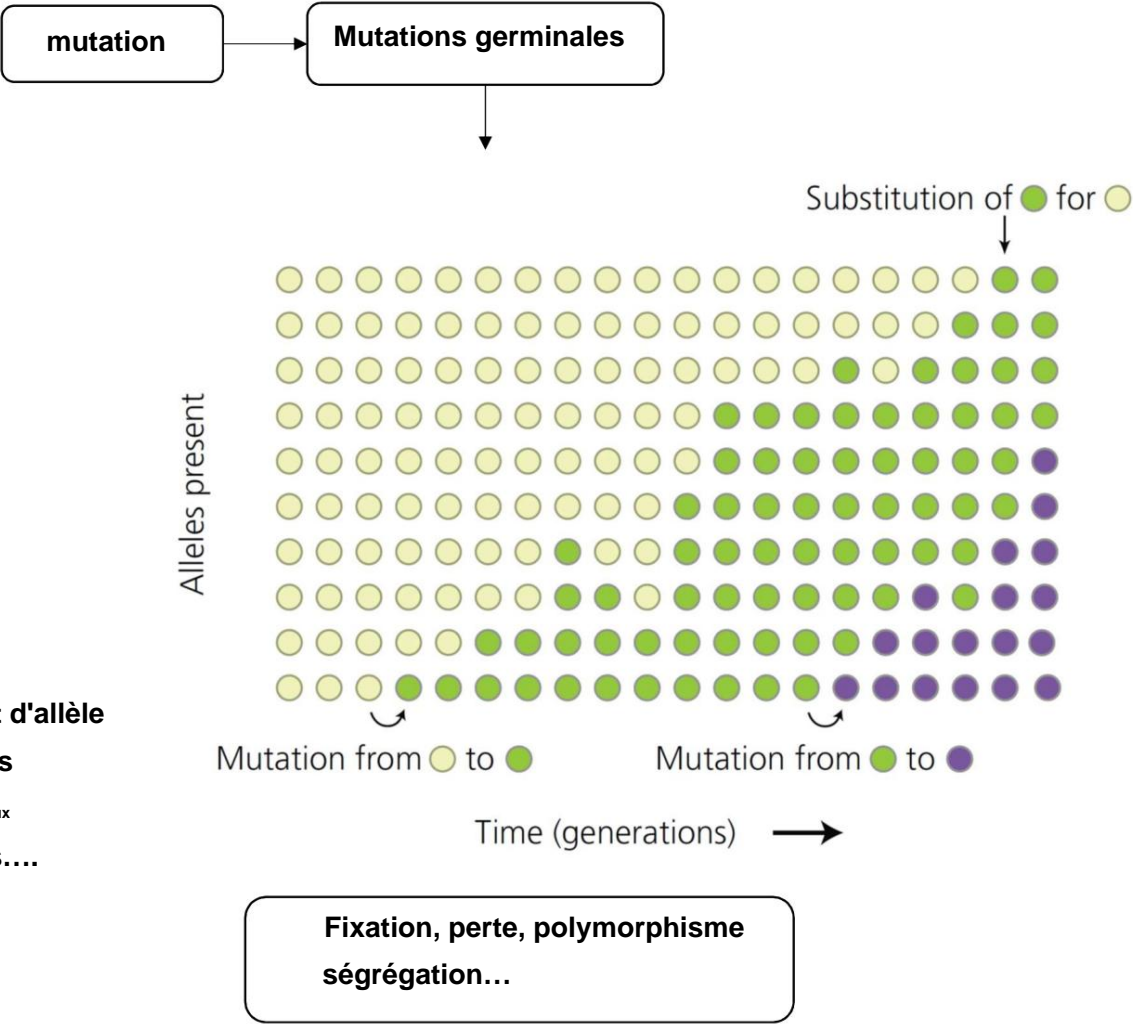
• Mutations à petite échelle (modifications restreintes de l'ADN)

- mutation d'un seul nucléotide ou mutation ponctuelle
- insertion ou délétion•: indel

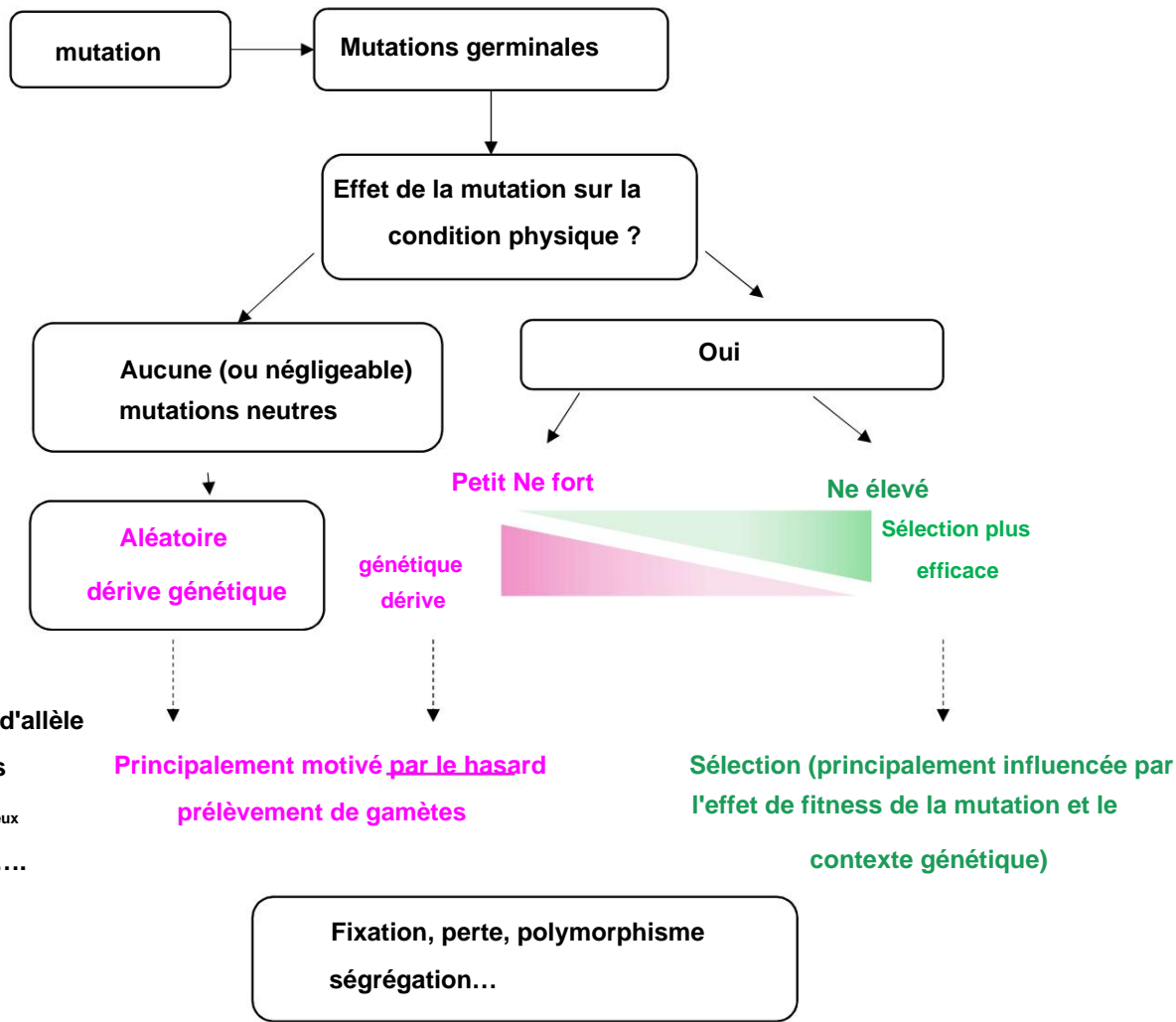
• Mutations à grande échelle

- réarrangements chromosomiques
- gain ou perte de chromosomes
- duplication du génome entier

Évolution d'une nouvelle mutation



Évolution d'une nouvelle mutation

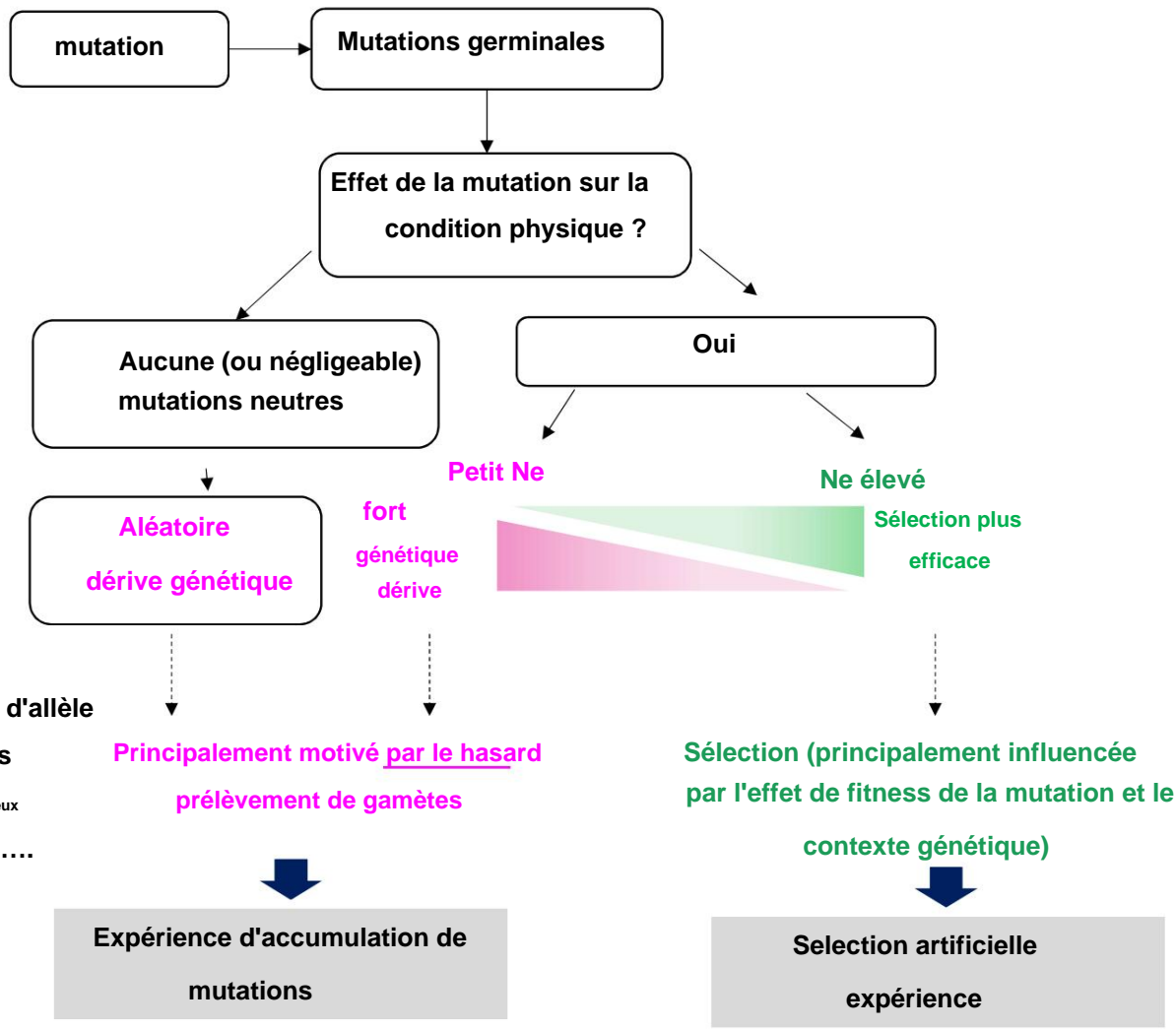


Aptitude : La capacité d'un organisme à survivre et à se reproduire, par rapport aux autres organismes de la même population, constitue

La dérive génétique est le changement des fréquences alléliques dans une population de génération en génération qui se produit en raison d'un échantillonnage aléatoire d'organismes.

Taille effective de la population (N_e) : un concept visant à mesurer la quantité de dérive génétique se produisant dans une population à un moment donné

Évolution d'une nouvelle mutation



Aptitude : La capacité d'un organisme à survivre et à se reproduire, par rapport aux autres organismes de la même population, constitue

La dérive génétique est le changement des fréquences alléliques dans une population de génération en génération qui se produit en raison d'un échantillonnage aléatoire d'organismes.

Taille effective de la population (N_e) : un concept visant à mesurer la quantité de dérive génétique se produisant dans une population à un moment donné

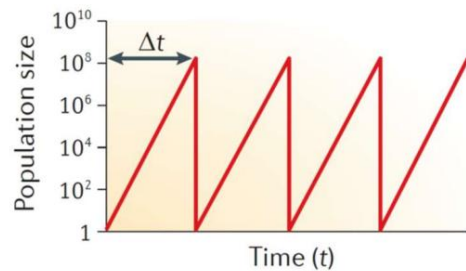
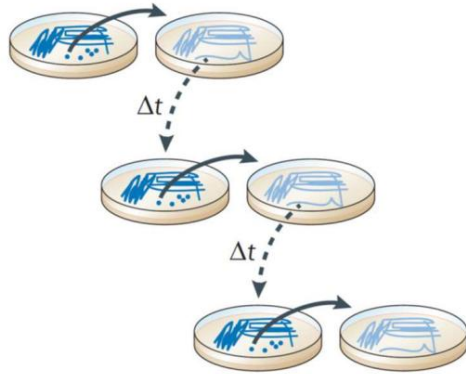
Types d'évolution
expériences en
laboratoire

Types d'expériences d'évolution en laboratoire

Accumulation de mutations

Mutation accumulation

a Single-cell bottlenecks

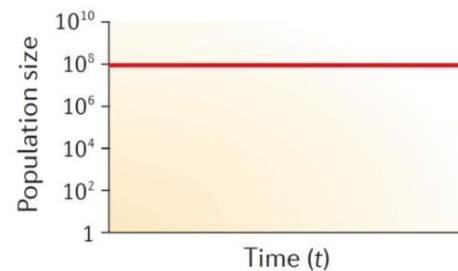
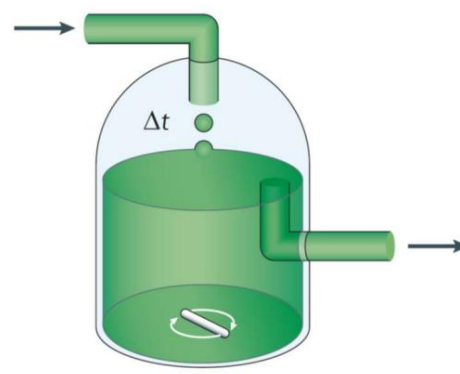


**Forte dérive pour éviter/limiter la sélection
contre les mutations délétères**

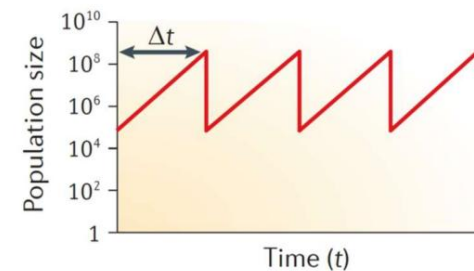
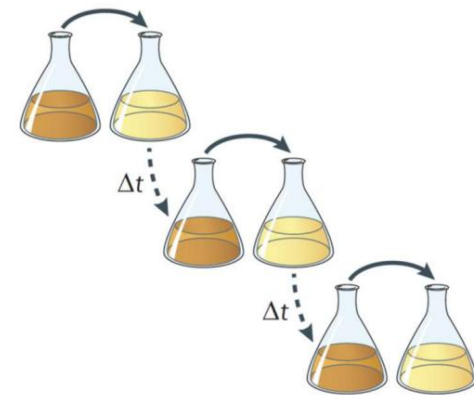
Expérience de sélection artificielle

Adaptive evolution

b Continuous culture



c Serial transfer

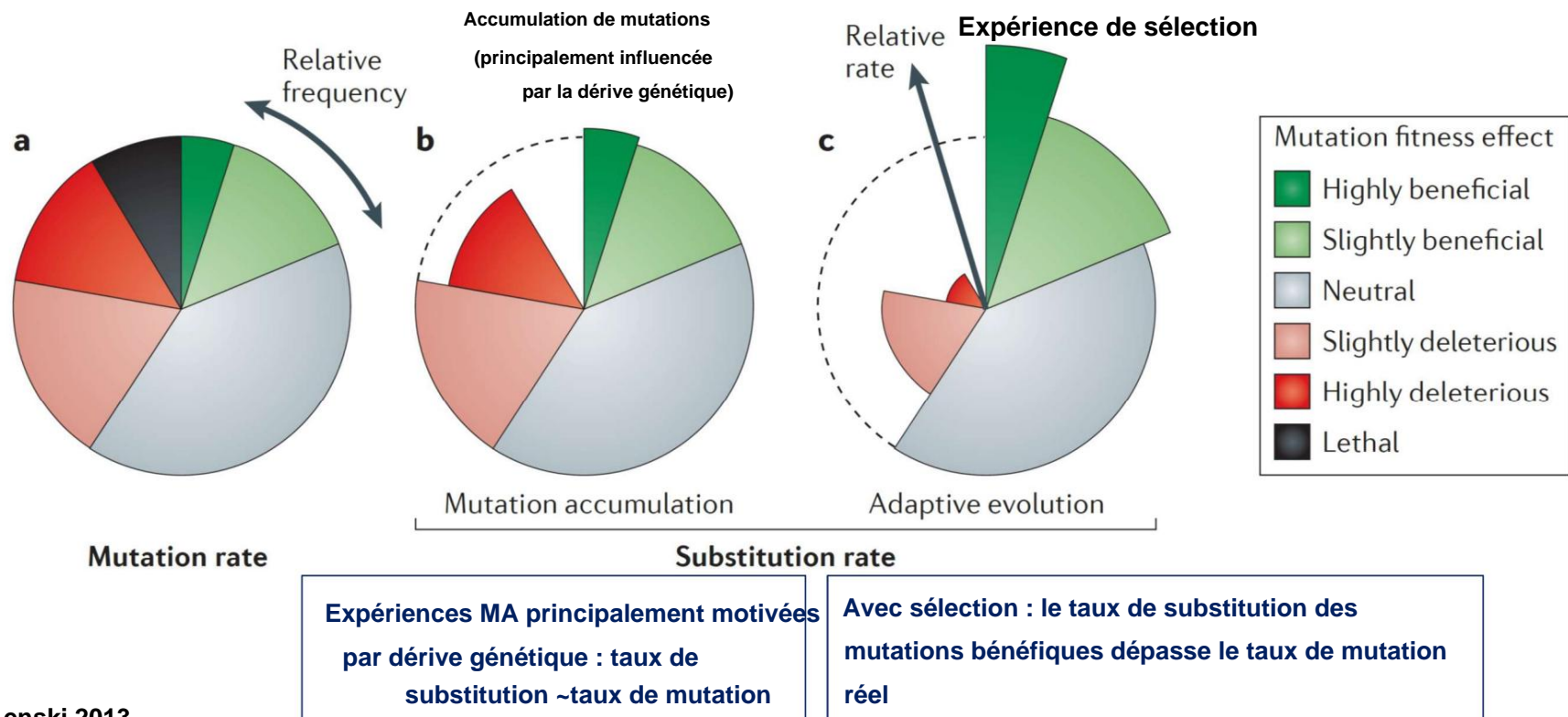


Avec sélection (c'est-à-dire pour un taux de croissance plus élevé sur un milieu particulier) -> grandes populations pour favoriser l'efficacité de la sélection et réduire l'effet de la dérive génétique⁸

Taux de mutation et taux de substitution

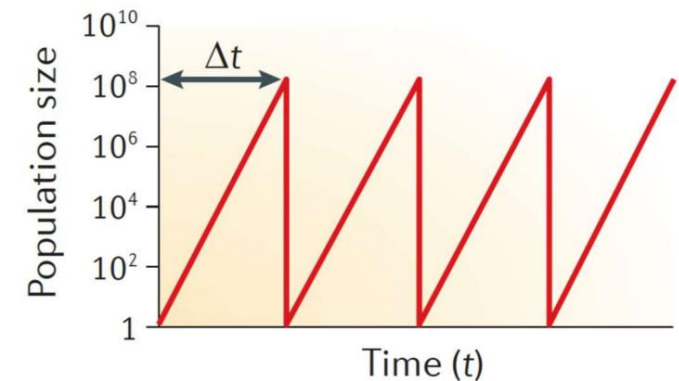
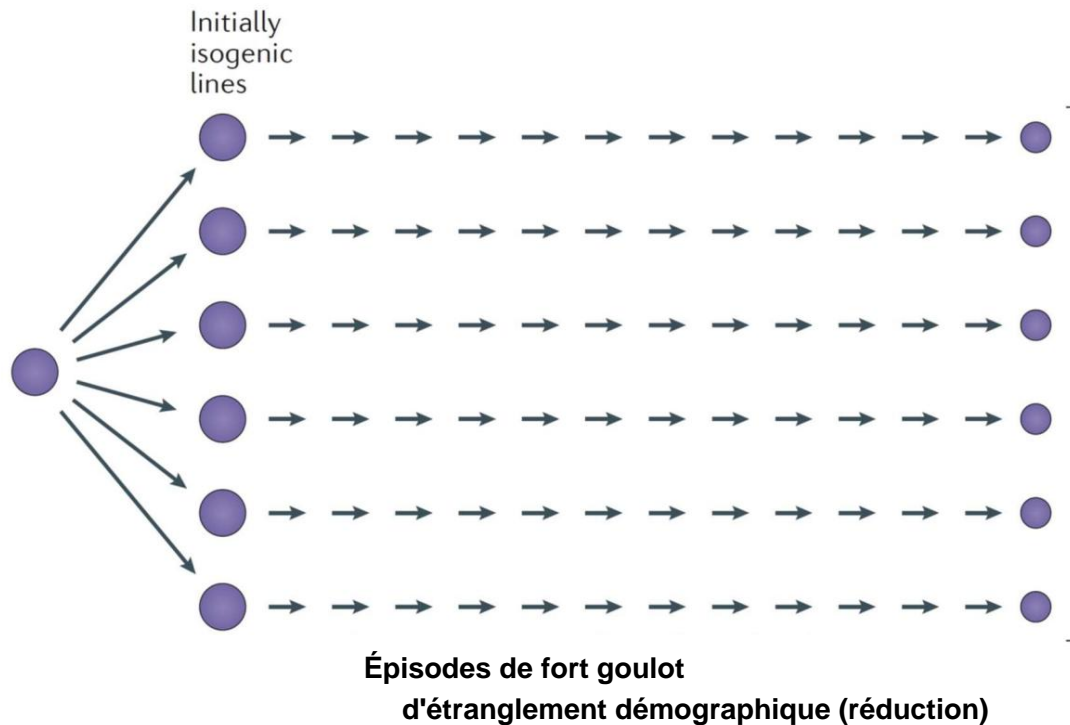
Taux de mutation : reflète la probabilité d'un changement de séquence du génome entre un parent et sa progéniture

Taux de substitution : taux auquel de nouvelles mutations s'accumulent dans une lignée en évolution au fil du temps, qui dépend généralement à la fois du taux de mutation et des effets de la sélection naturelle.



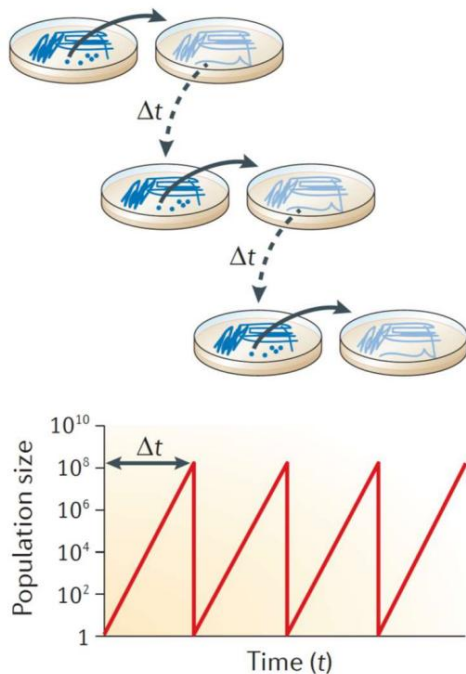
I. Estimation du taux de mutation

Expériences d'accumulation de mutations (AM) : les clones (ou individus proches) sont soumis à un grand nombre d'épisodes séquentiels de fort goulot d'étranglement de population (pour réduire l'effet de la sélection agissant sur les mutations délétères)



I. Estimation du taux de mutation

Expériences d'accumulation de mutations (AM) : les clones (ou individus proches) sont soumis à un grand nombre d'épisodes séquentiels de fort goulot d'étranglement de population (pour réduire l'effet de la sélection agissant sur les mutations délétères)



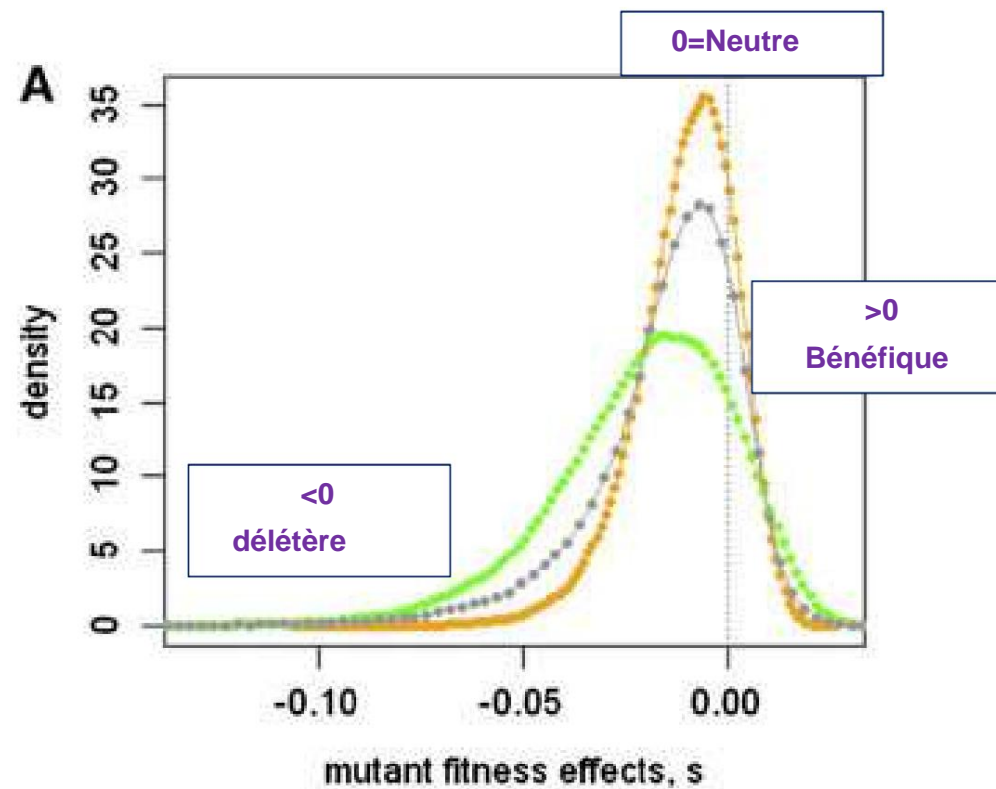
Estimations du taux de

mutation : $\bar{\mu}$ indirectes : basées sur la génétique accumulée, la variance pour les traits liés à la condition physique et la distribution des effets sur la condition physique des nouvelles mutations
(Bateman 1959, Mukai 1972, Keightley 2004)

$\bar{\mu}$ direct : séquençage complet du génome des lignées MA
(plus récent)

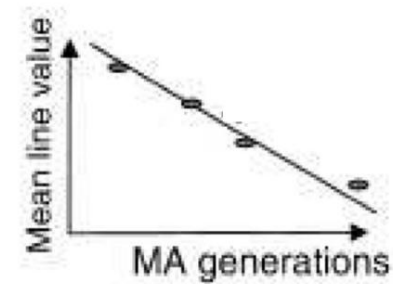
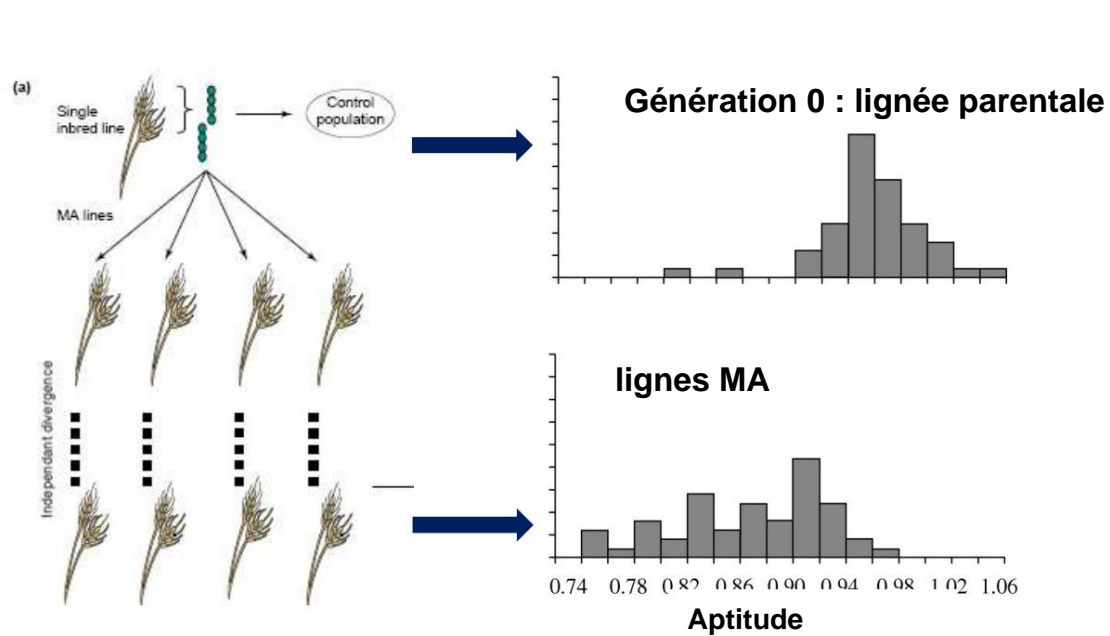
I. Estimation du taux de mutation : utilisation de la distribution de l'effet fitness des mutations

Exemple de distribution théorique de l'effet de fitness (DFE) de nouvelles mutations

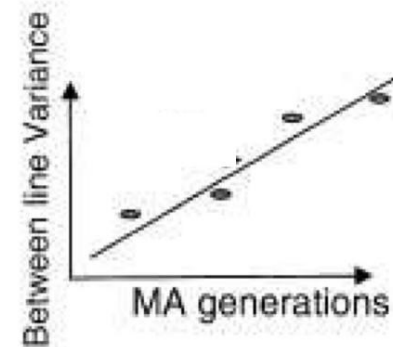


I. Estimation du taux de mutation : utilisation de la distribution de l'effet fitness des mutations

Expériences d'accumulation de mutations (MA) : principes de l'estimation indirecte du taux de mutation



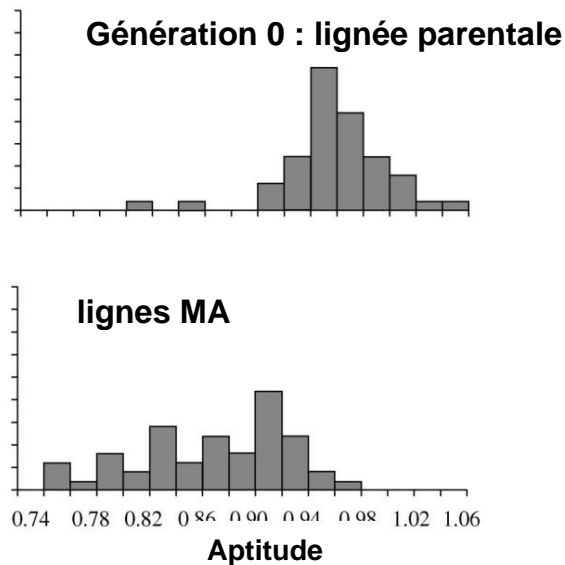
La valeur moyenne
du trait lié à la
condition physique diminue



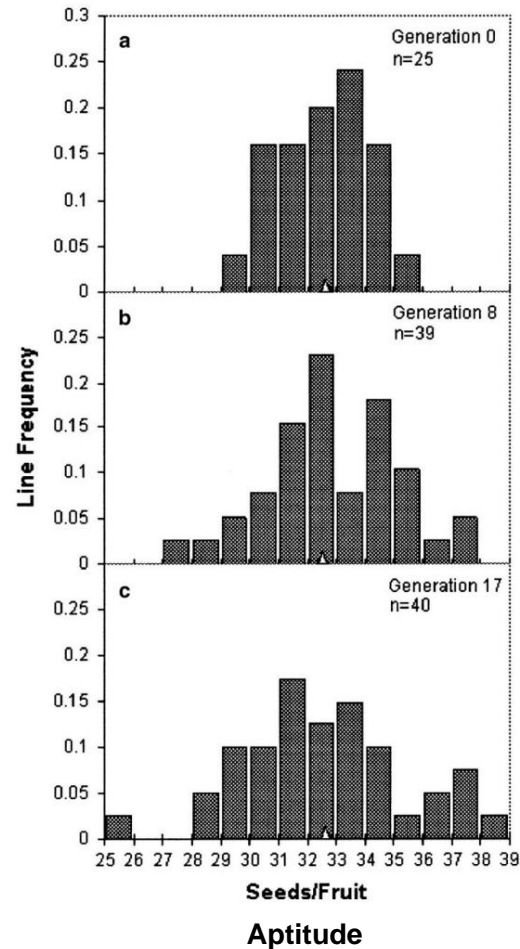
La variance
de fitness augmente

I. Estimation du taux de mutation : utilisation de la distribution de l'effet fitness des mutations

Exemple 1



Exemple 2 (moins fréquent, cas particulier ?)



-> Dans les deux exemples: la variance de fitness augmente au fil des générations

-> Pas toujours clair pour la moyenne aptitude

Shaw et al. 2000

I. Estimation du taux de mutation : séquençage du génome entier des lignées MA

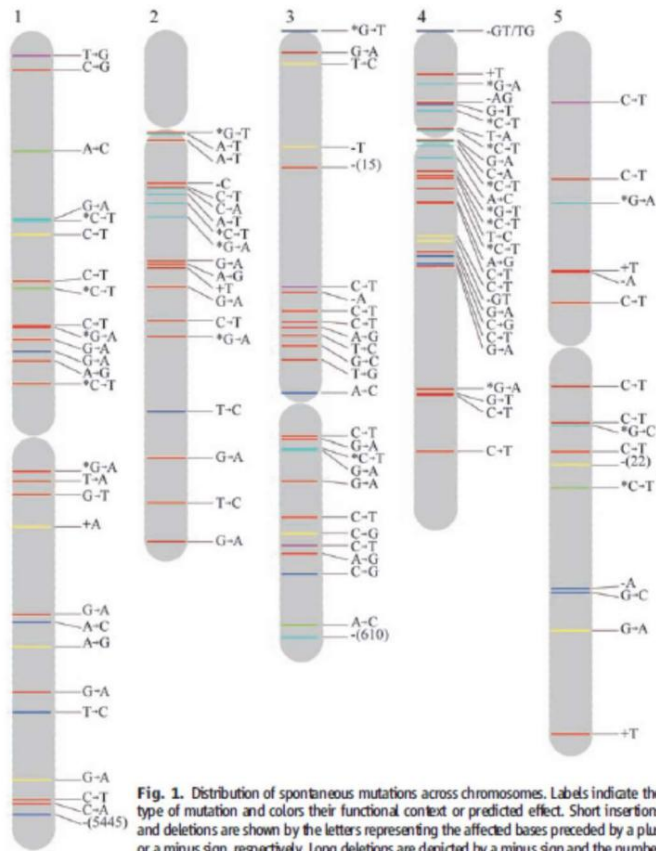
Taux et spectre des mutations à l'aide du séquençage du génome entier des lignées MA



A. thaliana

5 lignes MA, 30 générations

• 99 substitutions de bases, 17 petites et grandes insertions et délétions. • 7×10^9 substitutions de bases/base/génération • la majorité sont des transitions G:C>A:T



- intergénique
- intron

- **substitution non synonyme** • , petits indels avec décalage de cadre dans la séquence codante, délétion de gène
- substitution **synonyme** • UTR

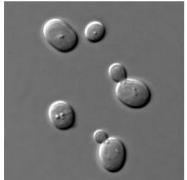
- **élément transposable**
- * **C ou G méthylé**

Ossowski et al 2010

I. Estimation du taux de mutation : séquençage du génome entier des lignées MA

Taux et spectre des mutations à l'aide du séquençage du génome entier des lignées MA

145 lignées diploïdes d'accumulation de mutations (MA) (311 000 générations)



S. cerevisiae

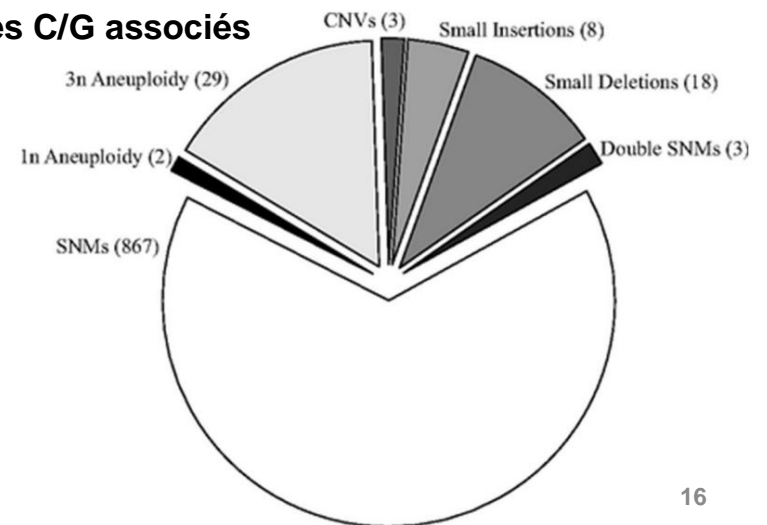
~ $0,17 \times 10^9$ substitutions de bases/par base/génération

Taux de duplication du génome entier de $9,7 \pm 1,8 \times 10^5$ événements par génome diploïde par génération (29 lignées)

léger biais vers les suppressions

Biais AT de mutation avec un taux élevé de mutations de C/G à T/A

taux de mutation plus élevé dans des contextes de nucléotides C/G associés à la méthylation de la cytosine



I. Estimation du taux de mutation : utiliser la diversité génétique

Estimation indirecte à partir de polymorphismes « neutres » accumulés au fil du temps dans les populations naturelles (substitutions synonymes et intergéniques)

Supposons que les substitutions synonymes sont neutres (pas toujours, c'est-à-dire biais d'utilisation des codons)

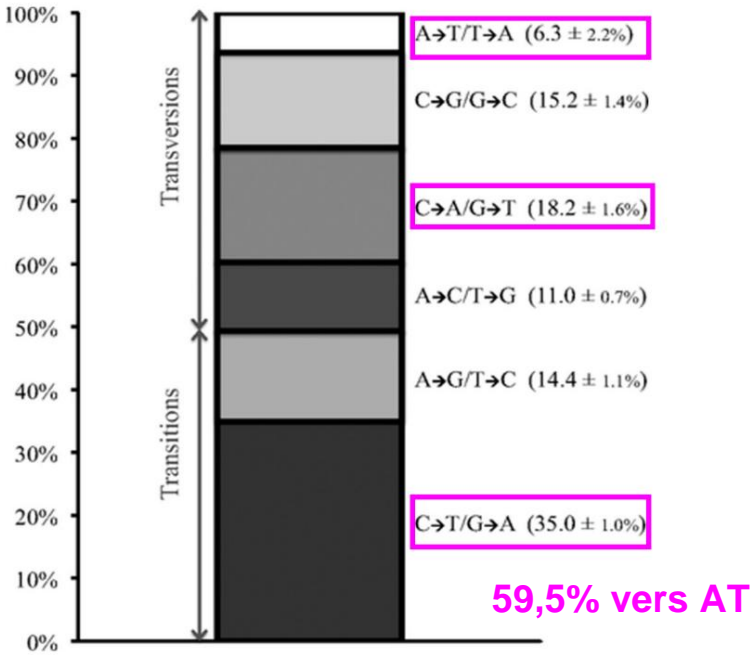
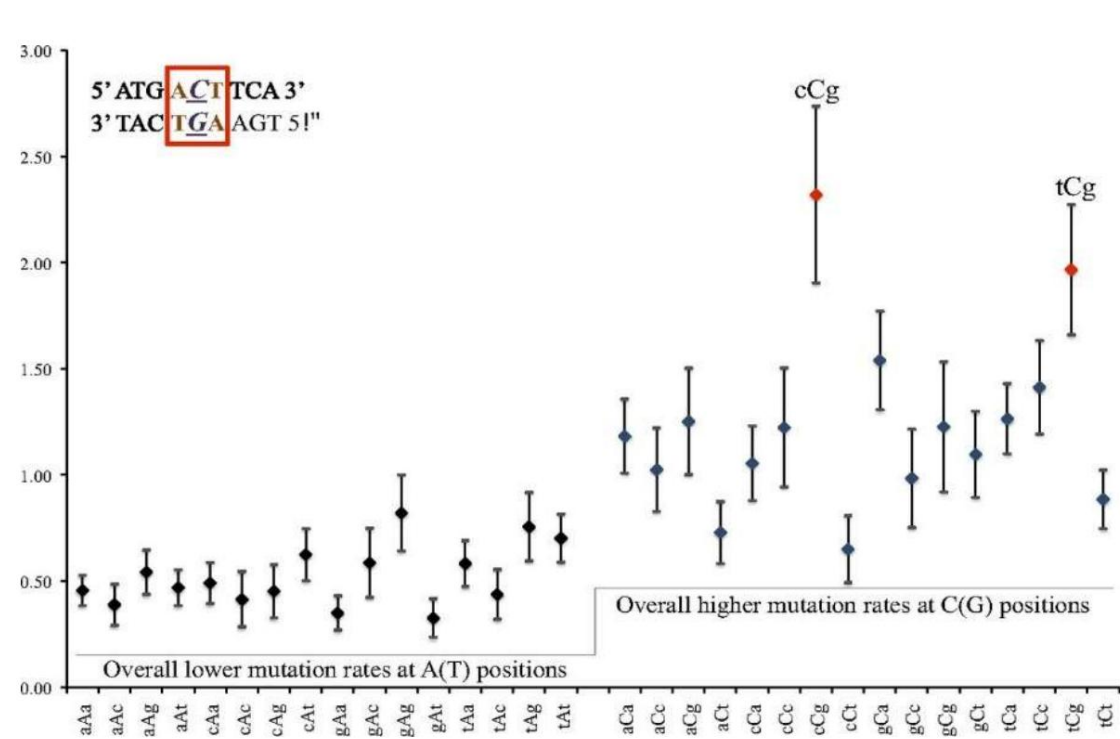
Fortement influencé par la qualité de l'alignement et de l'annotation du génome



| | |
|-----|---------|
| GCU | Alanine |
| GCC | Alanine |
| GCA | Alanine |
| GCG | Alanine |

II. Variation du taux de mutation le long des chromosomes

- Certains contextes génétiques peuvent favoriser des types spécifiques de mutations (mutations d'un seul nucléotide, insertions/délétions)
- Effets environnementaux et types et taux de mutation (stress et dommages à l'ADN)
- Léger biais vers AT (observé dans divers organismes)

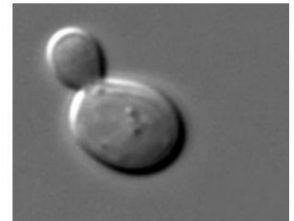


Taux de mutation relatifs de chacun des six changements possibles de nucléotides dans la levure.

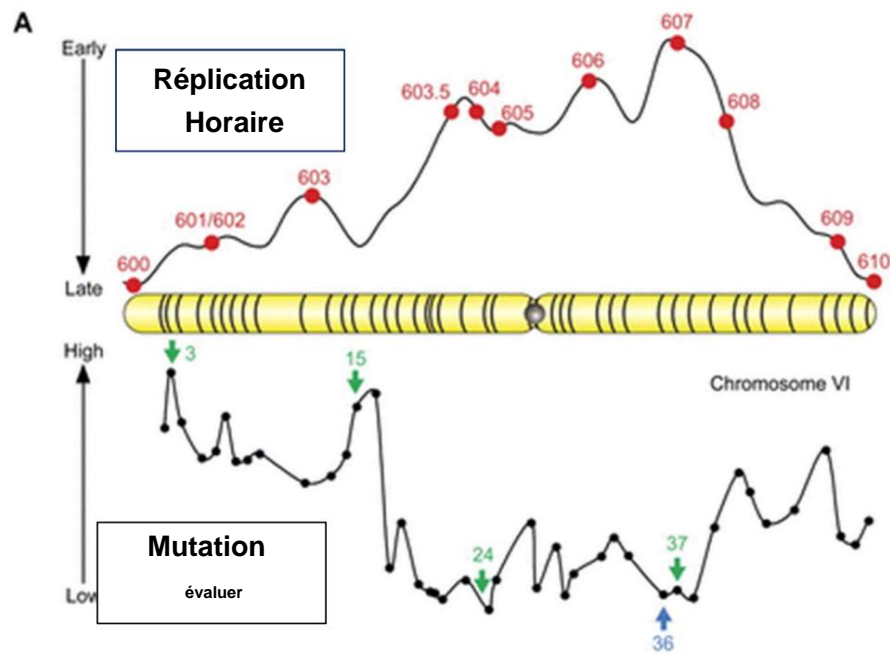
II. Variation du taux de mutation le long des chromosomes

Taux de mutation corrélé à la vitesse de réplication le long des chromosomes:

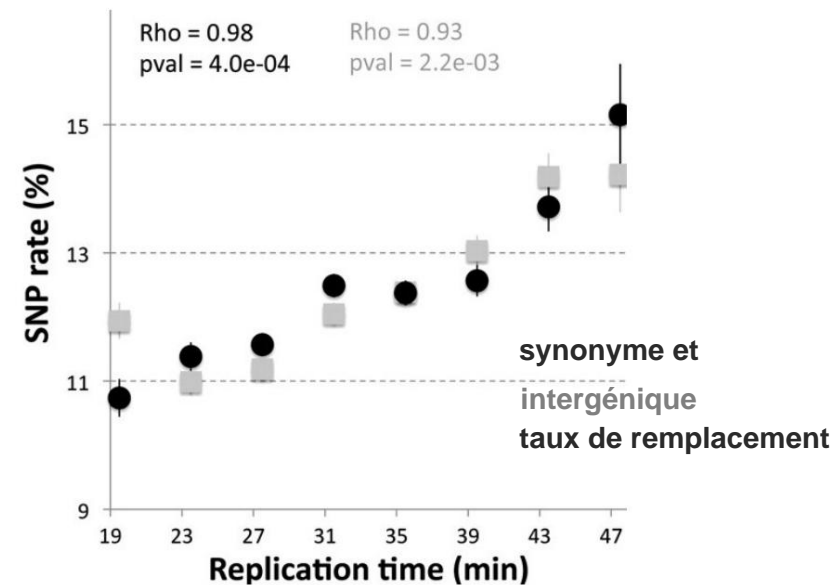
Taux de mutation plus élevé dans les régions à réplication lente/tardive



S. cerevisiae



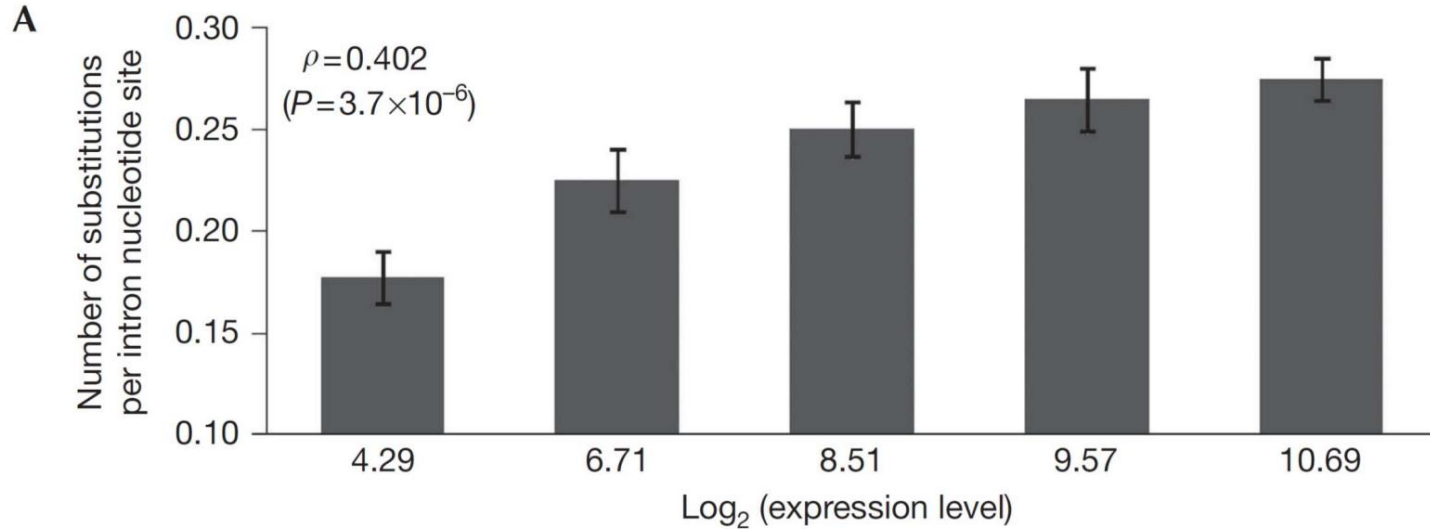
Lang et Murray 2011



Agier & Fisher 2011

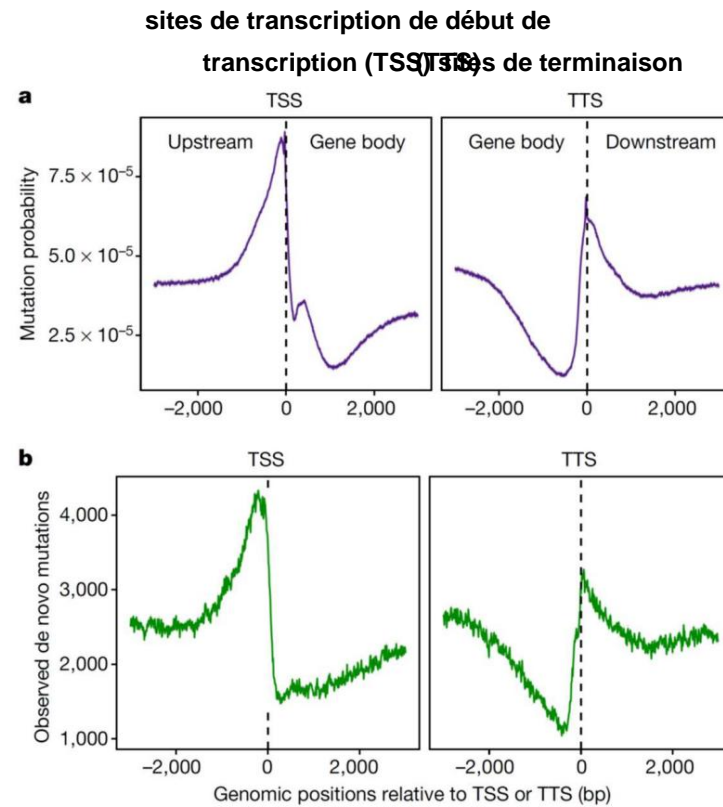
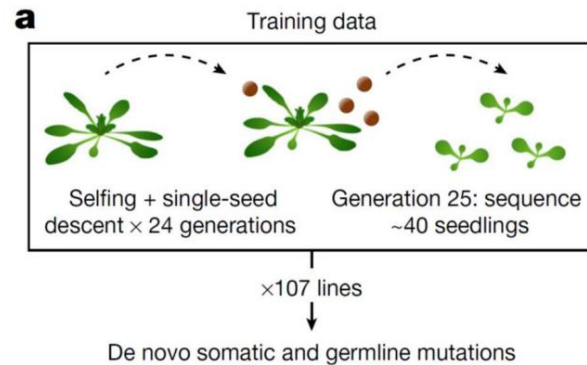
II. Variation du taux de mutation le long des chromosomes

Les gènes hautement transcrits peuvent avoir un taux de mutation élevé (vulnérabilité mutationnelle plus élevée)



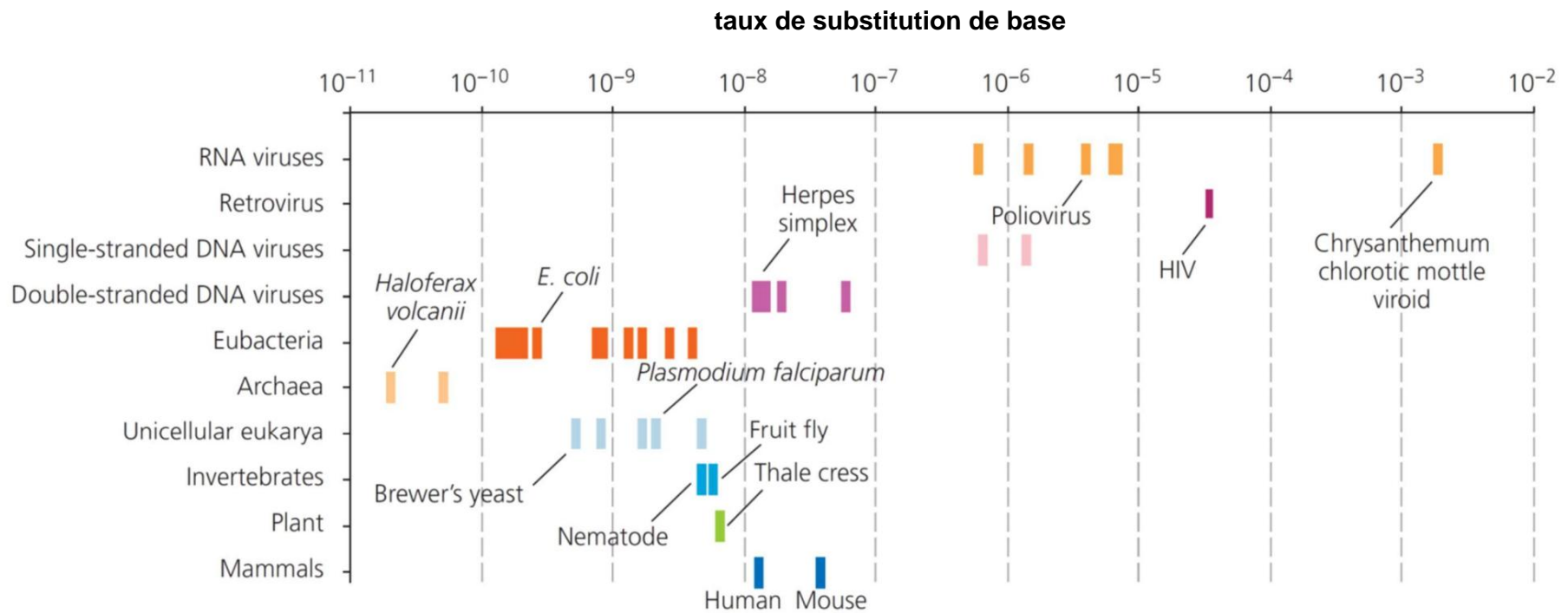
II. Variation du taux de mutation le long des chromosomes

Taux de mutation plus faibles dans les corps de gènes -> caractéristiques épigénétiques
-> biais de mutation et contexte génétique



A. thaliana

II. Variation du taux de mutation dans divers organismes



Taux de substitution de bases dans divers organismes, Figure de Herron & Freeman 5e édition, préparée à l'aide d'estimations compilées par Lynch (2010a) et, pour le VIH, Mansky et Temin (1995).

II. Variation du taux de mutation et conséquences sur la valeur adaptative

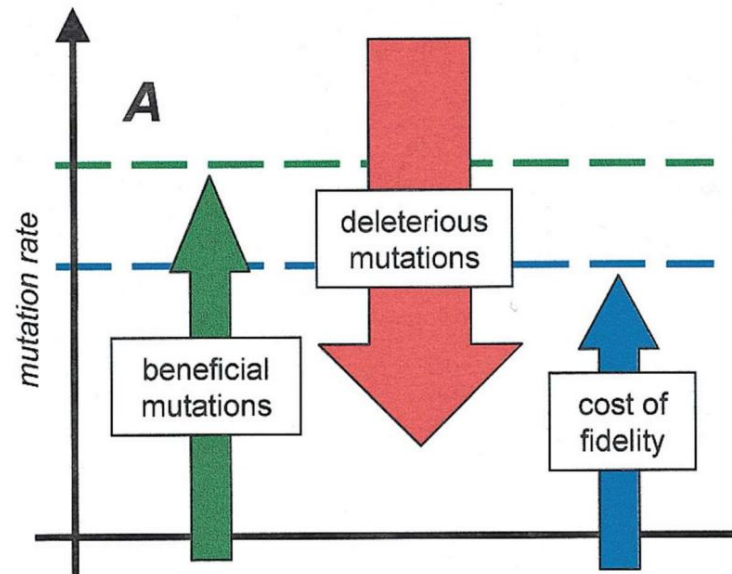
Sélection pour un taux de mutation plus faible : pour limiter l'accumulation de mutations délétères (qui diminuent la fitness)

Deux forces sélectives possibles contraignent le taux de mutation à évoluer vers zéro :

- la probabilité accrue d'acquérir des mutations bénéfiques sous une mutation plus élevée

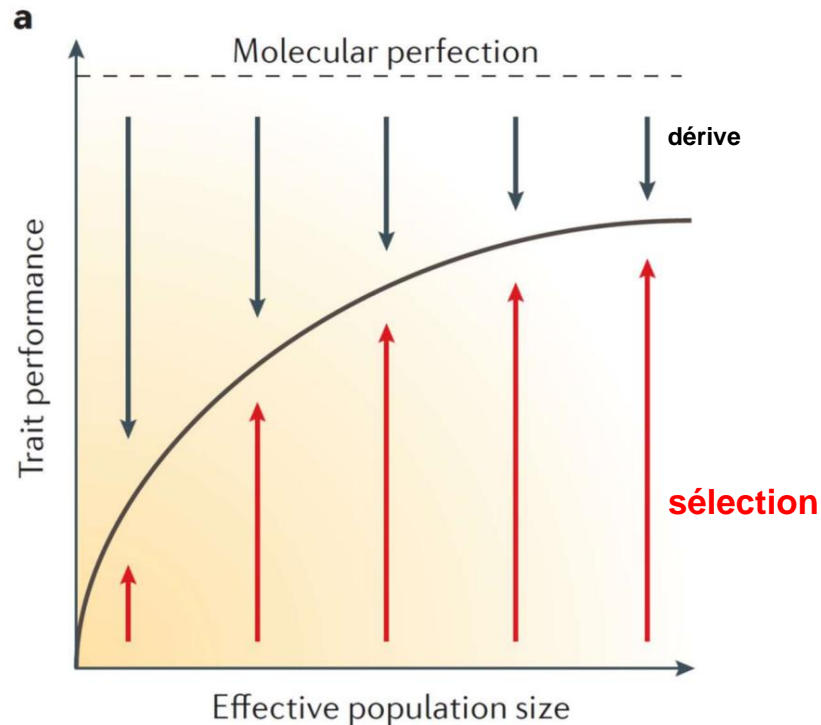
évaluer

- contraintes sur la fidélité de la réplication (coût de la fidélité)



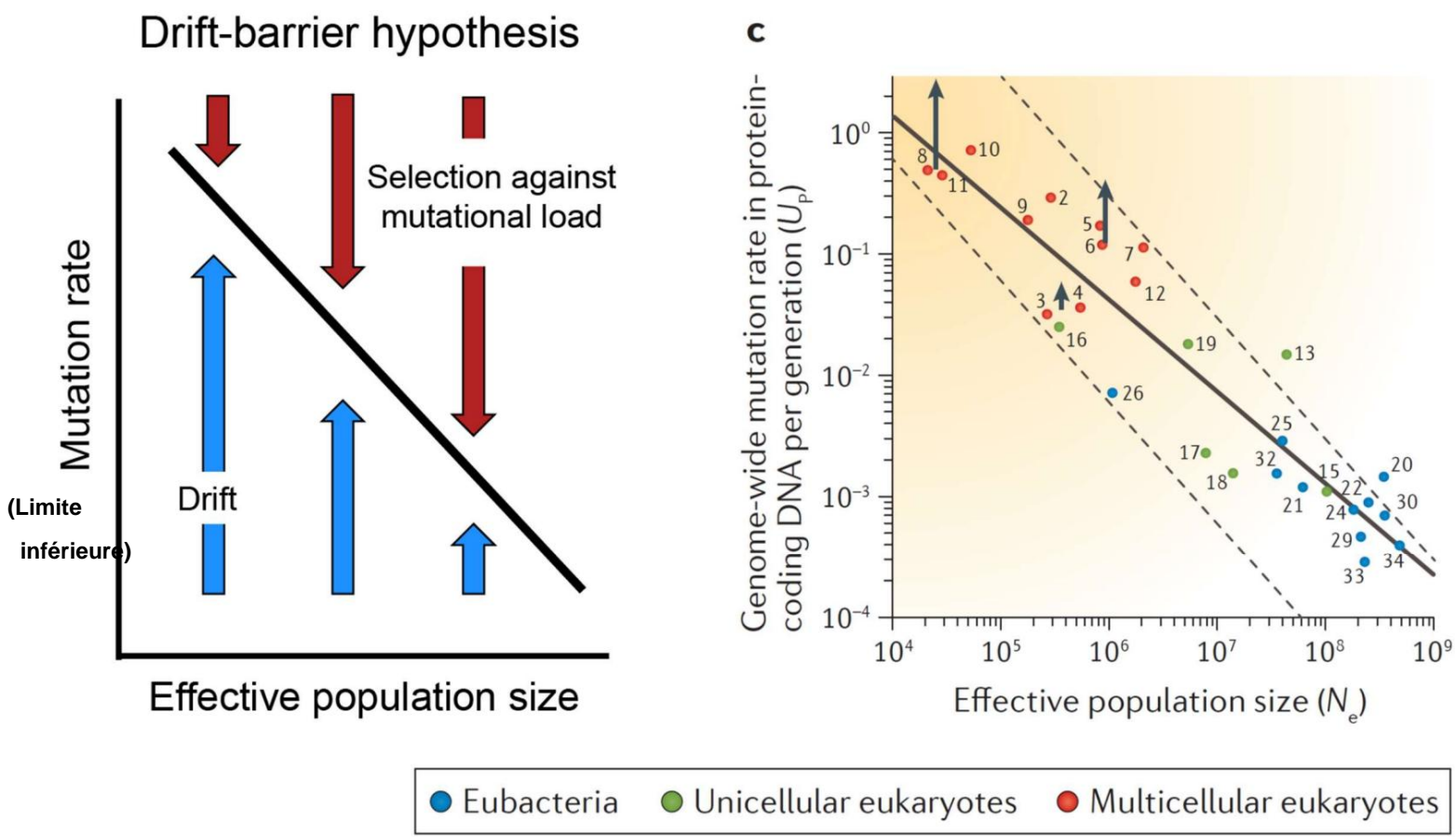
Chiffre de Sniegowski et al. 2010

II. Variation du taux de mutation et conséquences sur la valeur adaptative

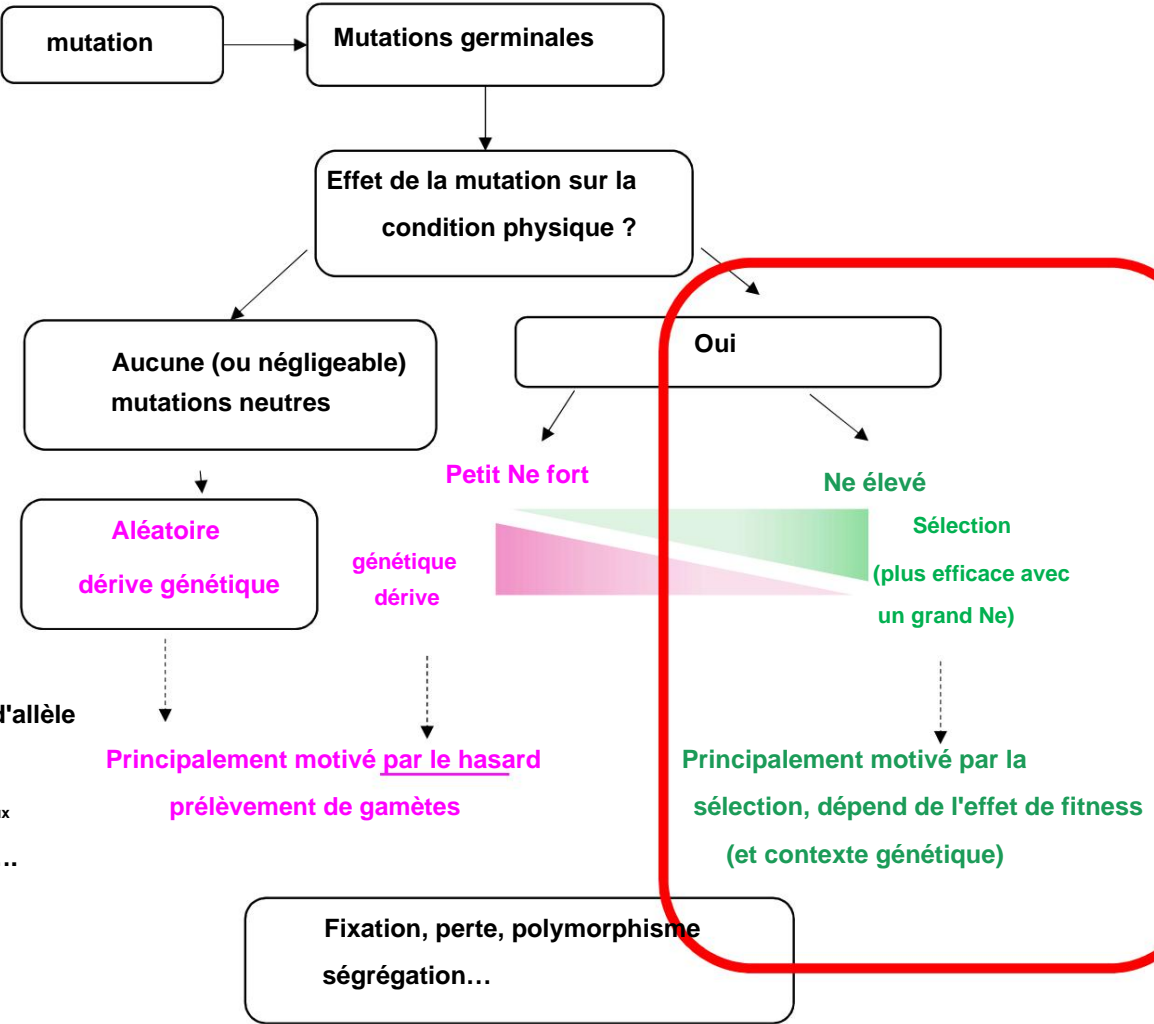


Le niveau de «*hyperperfection*» de tout trait moléculaire devrait s'affiner dans les populations plus importantes (avec un N_e élevé), à mesure que l'efficacité de la sélection augmente et que les effets stochastiques de la dérive génétique diminuent.

II. L'hypothèse de la barrière de dérive pour l'évolution du taux de mutation



Évolution d'une nouvelle mutation



Aptitude : La capacité d'un organisme à survivre et à se reproduire, par rapport aux autres organismes de la même population, constitue

La dérive génétique est le changement des fréquences alléliques dans une population de génération en génération qui se produit en raison d'un échantillonnage aléatoire d'organismes.

Taille effective de la population (N_e) : un concept visant à mesurer la quantité de dérive génétique se produisant dans une population à un moment donné

**Évolution
adaptative**

III. Étudier l'évolution adaptative

Quelles mutations peuvent contribuer à une réponse à la sélection pour les traits liés à la condition physique (c'est-à-dire l'évolution adaptative) ?

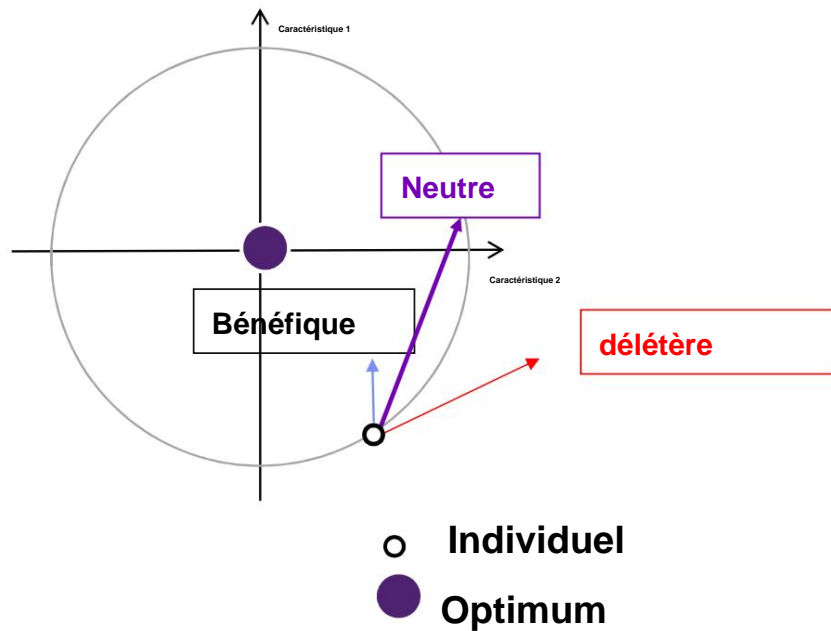
• Modèle de Fisher-Orr (1930, 1998)

• Illustration avec évolution expérimentale à long terme avec des micro-organismes asexués (Lenski expérience)

• Populations sexuées vs asexuées : rôle de la recombinaison

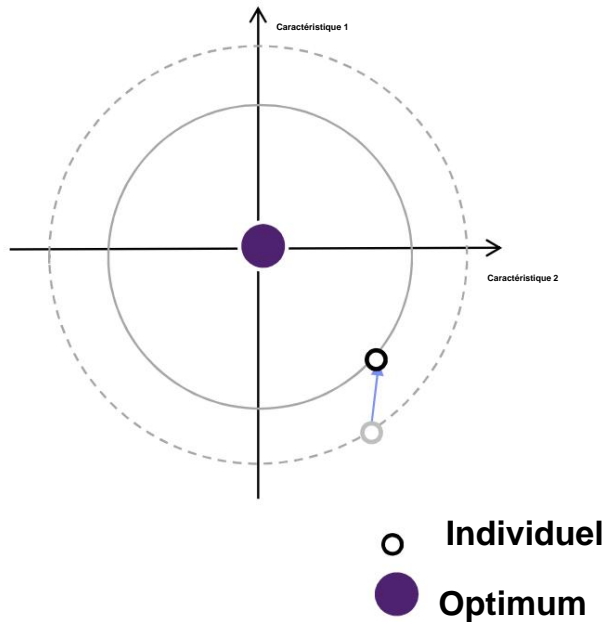
III. Marche adaptative: modèle de Fisher-Orr

• Mutations affectant les valeurs d'un ou deux traits liés à la forme physique



III. Marche adaptative: modèle de Fisher-Orr

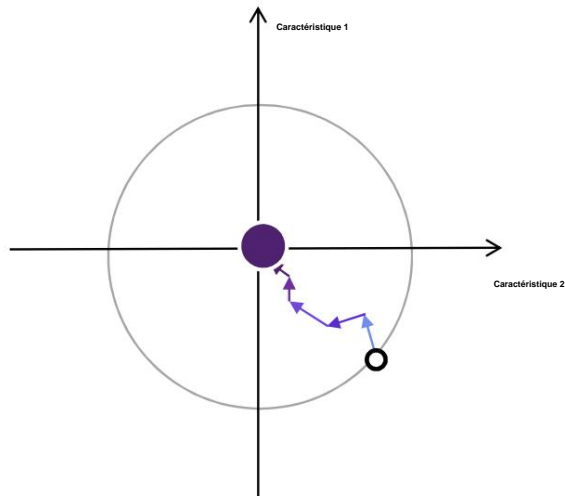
• Mutations affectant les valeurs d'un ou deux traits liés à la fitness • Évolution vers des valeurs de traits qui maximisent localement la capacité d'un organisme dans un environnement (accumulation de mutations qui le rapprochent de son optimum)



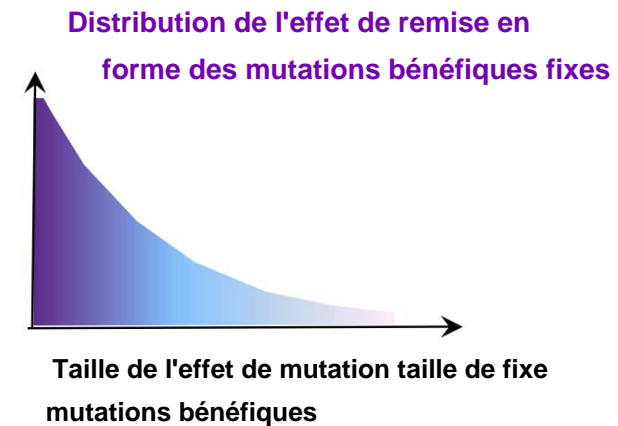
Fisher 1930, Orr 1998

III. Marche adaptative: modèle de Fisher-Orr

• Mutations affectant les valeurs d'un ou deux traits liés à la fitness • Évolution vers des valeurs de traits qui maximisent localement la capacité d'un organisme dans un environnement (accumulation de mutations qui le rapprochent de son optimum) • Évolution par la sélection d'un petit nombre de grandes mutations à effet et un nombre élevé de petites mutations à effet



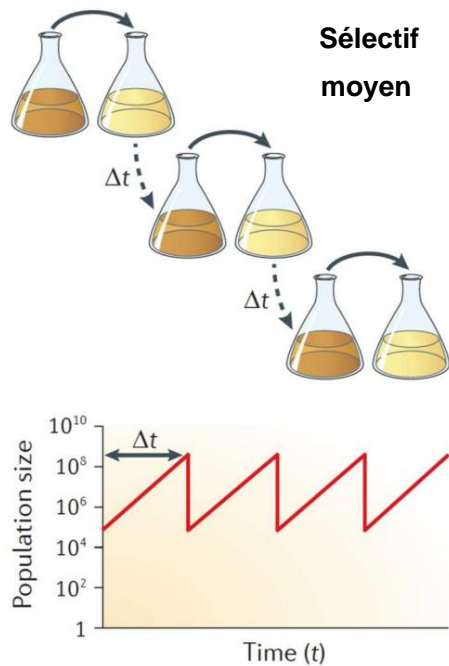
○ Individuel
● Optimum



III. Évolution expérimentale à long terme pour étudier le taux d'évolution adaptative

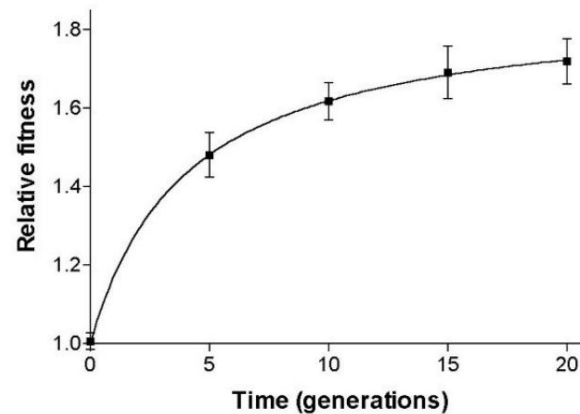
Exemple avec l'expérience de Richard Lenski avec *E. coli*

De nombreuses générations sur sélectif moyen (avec grand N_e)

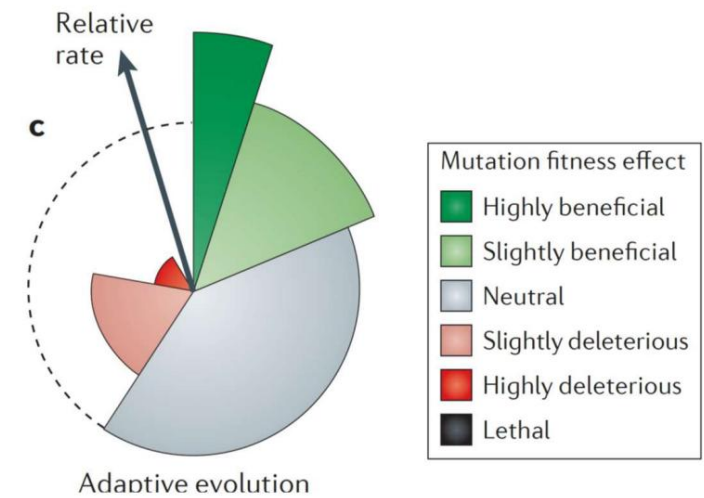


Test de condition physique

c'est-à-dire le taux de
croissance évolué/le gain de forme physique des ancêtres



Mutations accumulées



III. Évolution expérimentale à long terme pour étudier le taux d'évolution adaptative

Exemple avec l'expérience de Richard Lenski avec *E. coli*



février 2010

Célébration des 50 000 générations des lignes à long terme



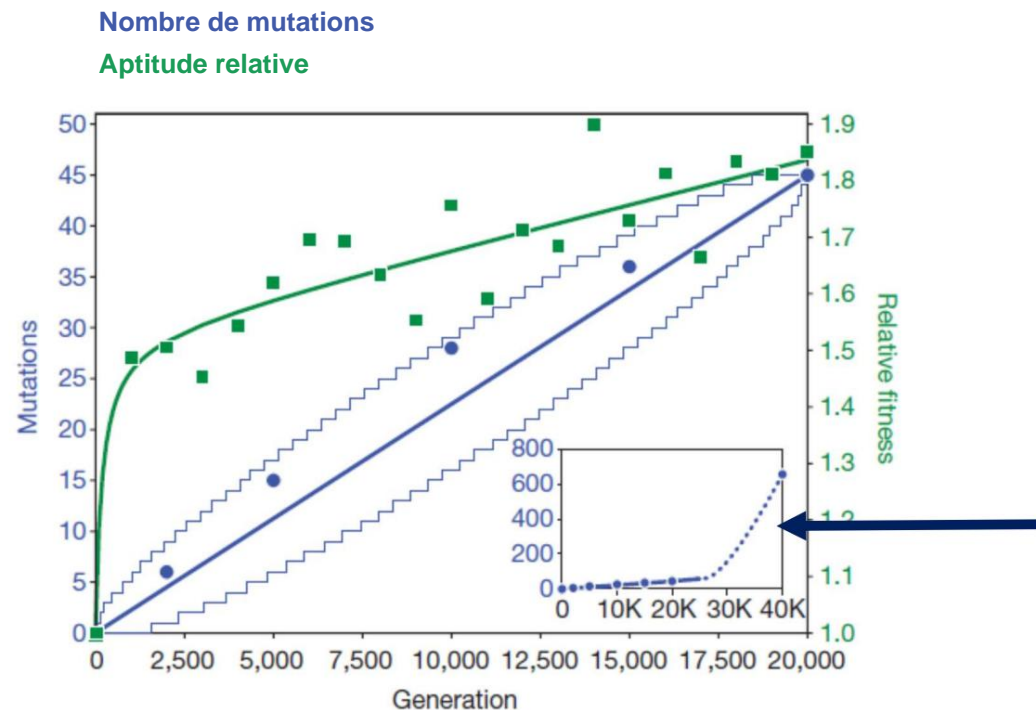
2016 Tour 12 583 assiettes !

Lenski, RE (2000). Le site du projet d'évolution expérimentale à long terme d'*E. coli*. <http://myxo.css.msu.edu/ecoli> <https://www.blount-lab.org/photos.html>

III. Évolution expérimentale à long terme pour étudier le taux d'évolution adaptative

Exemple avec l'expérience de Richard Lenski avec *E. coli*

- 12 populations *E. coli* évolutives issues d'un même clone sur un milieu pauvre en glucose depuis plusieurs générations
- Séquençage du génome entier aux générations 2 000, 5 000, 10 000, 15 000, 20 000 et 40 000

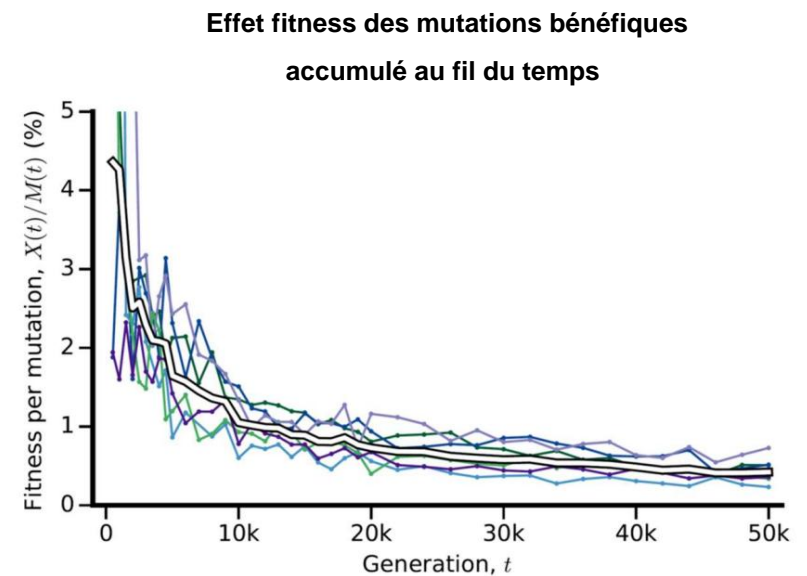
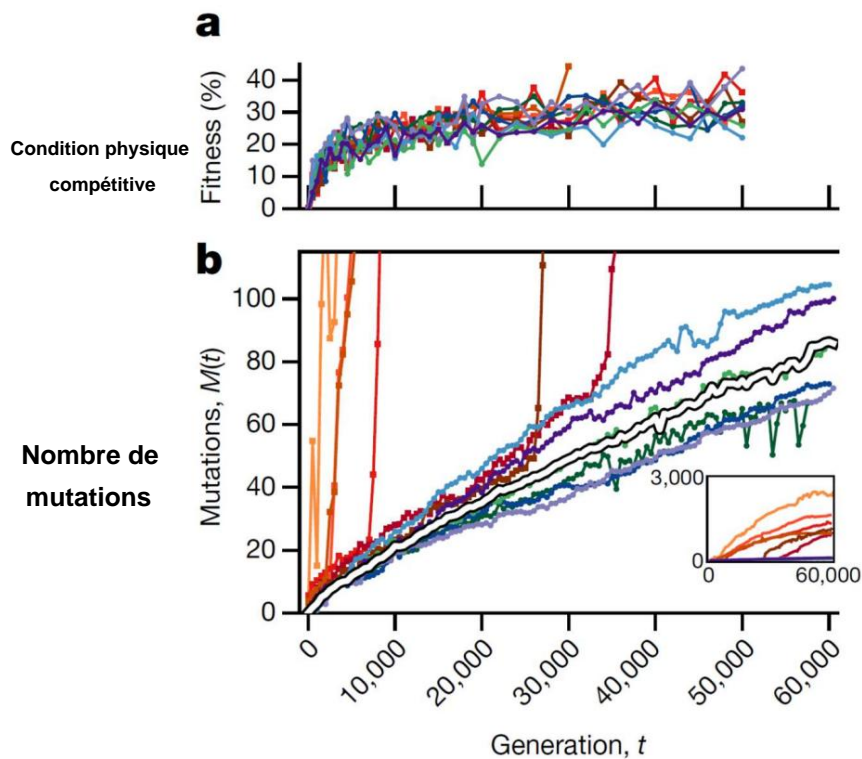


Émergence d'une
lignée hypermutable

III. Évolution expérimentale à long terme pour étudier le taux d'évolution adaptative

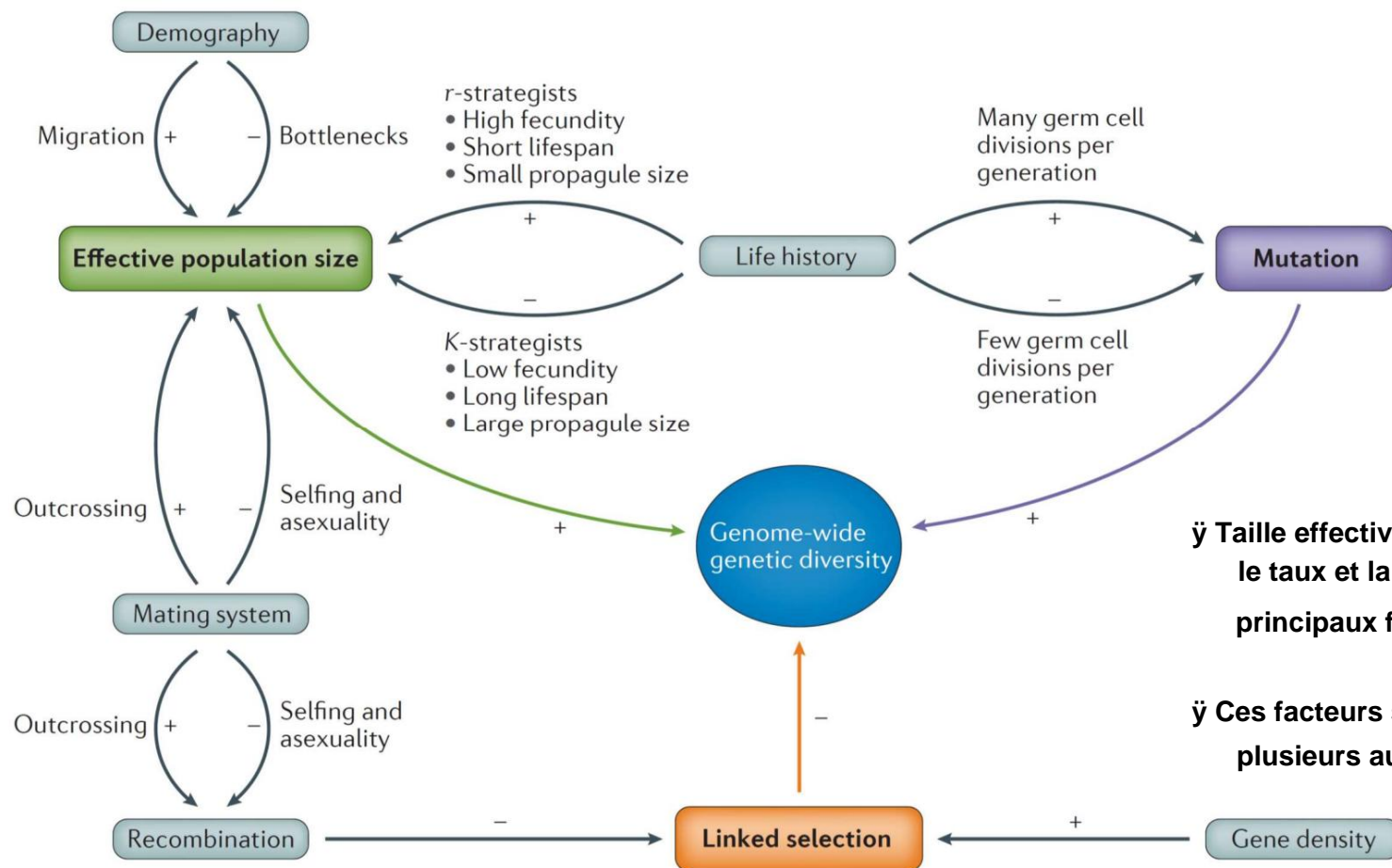
Exemple avec l'expérience de Richard Lenski avec *E. coli*

60 000 générations : séquençage des douze populations toutes les 500 générations (6 mutateurs et 6 non mutateurs)



Bon et al. (2017) La dynamique de l'évolution moléculaire sur 60 000 générations

Aperçu des déterminants de la diversité génétique

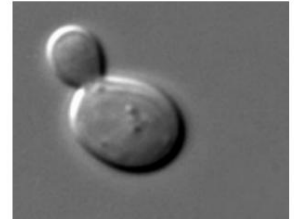
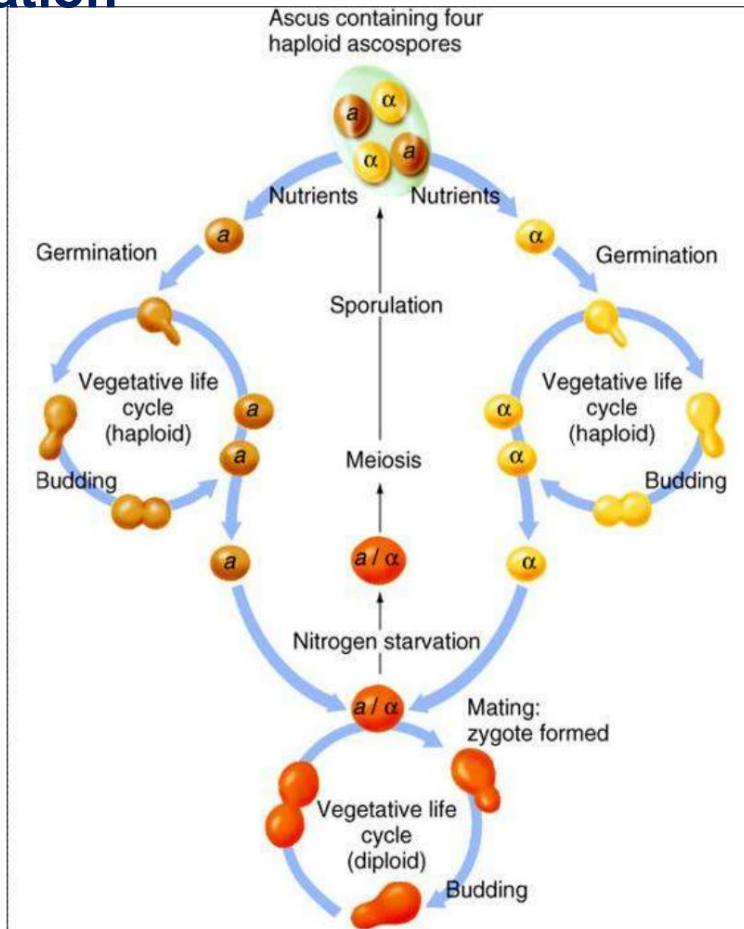


• Taille effective de la population, mutation le taux et la sélection liée sont les principaux facteurs influant sur la diversité.

• Ces facteurs sont à leur tour régis par plusieurs autres paramètres.

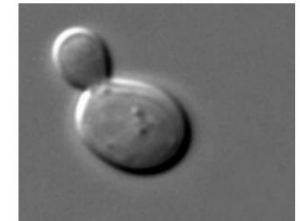
III. Populations sexuées vs asexuées : rôle de la recombinaison dans la vitesse d'adaptation

Cycle de vie de la levure



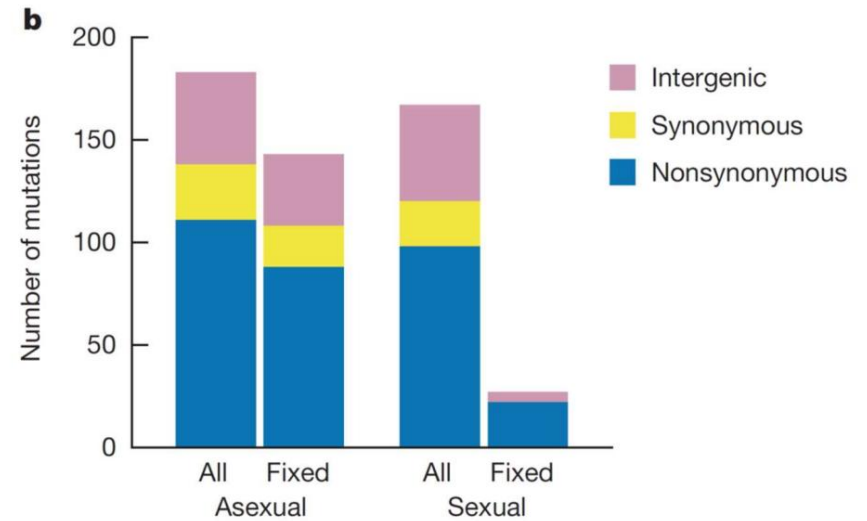
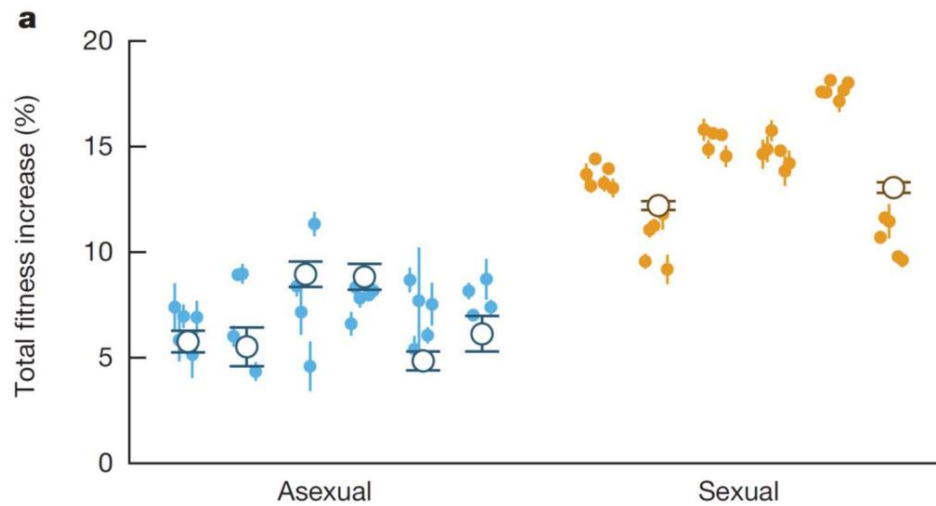
S. cerevisiae

III. Populations sexuées vs asexuées : rôle de la recombinaison dans la vitesse d'adaptation

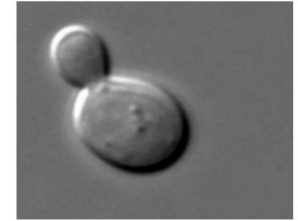


S. cerevisiae

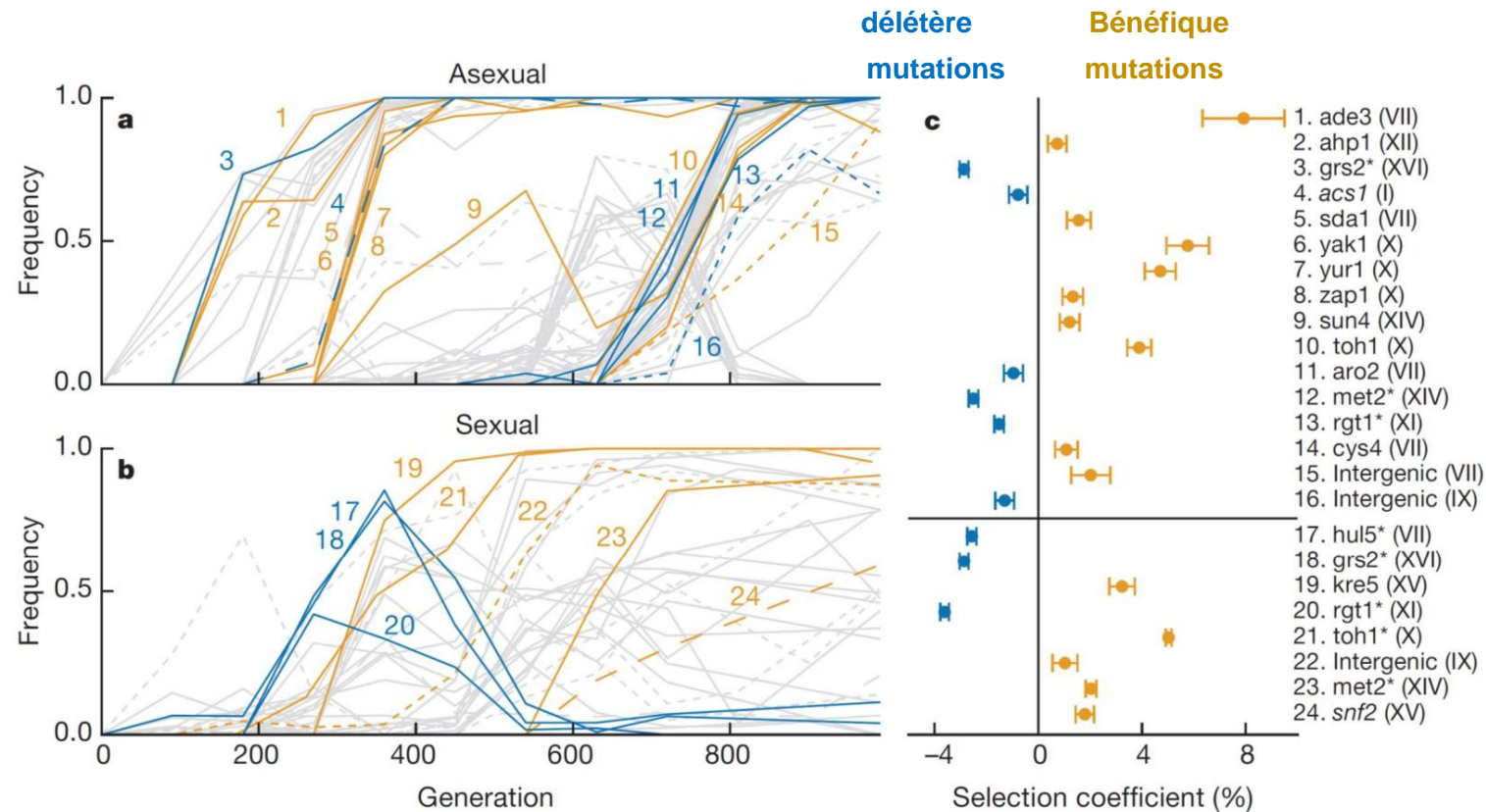
Évolution des populations de *S. cerevisiae* sous ~ 1000 générations



McDonald et al. 2016

**S. cerevisiae**

III. Populations sexuées vs asexuées : rôle de la recombinaison dans la vitesse d'adaptation



McDonald et al. 2016

III. Taux d'évolution adaptative : un sujet complexe !

Taux d'émergence de nouvelles mutations :

- **Elle varie selon les espèces et dans le temps**
- **Influencé par le contexte génomique et épigénétique (et l'environnement)**
- **progresser grâce aux progrès de la génomique**
- **Accès plus facile aux modifications structurelles grâce au séquençage à lecture longue**

Effet fitness et évolution adaptative : influencés par de nombreux facteurs

- **Populations asexuées versus sexuées**
- **N_e et démographie**
- **Types de mutations, effets et changements de ploïdie**
- **Epistasie et pléiotropie contraintes**
- **Effet environnemental, environnements fluctuants vs constants (variation permanente vs nouvelles mutations)**
- **....**

Nouvelles mutations versus variation permanente

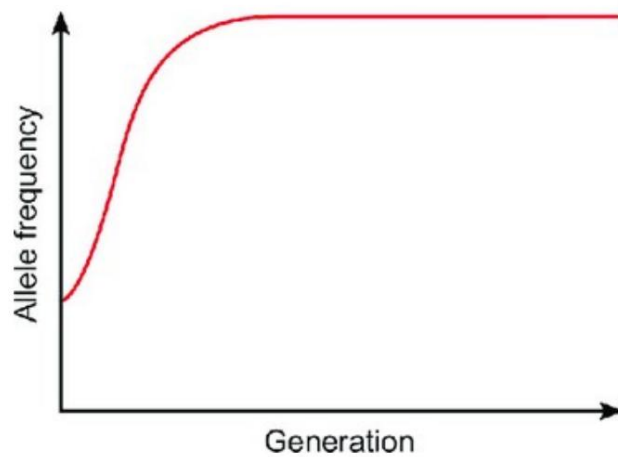
Standing variation

Mutation **present** before selection

Starting allele frequency = $p(\text{resistant allele})$

→ **No** waiting time

High probability of fixation



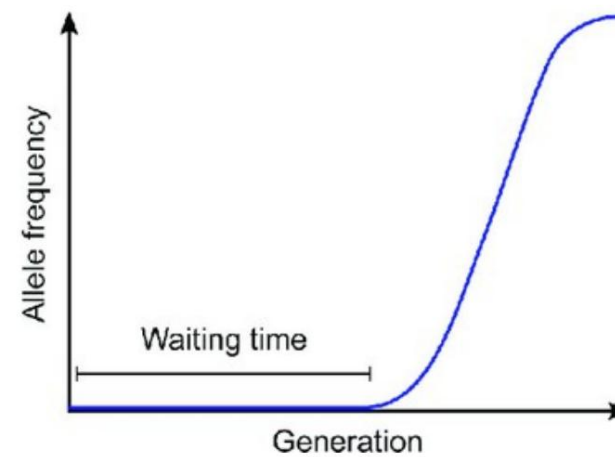
De novo variation

Mutation **not present** before selection

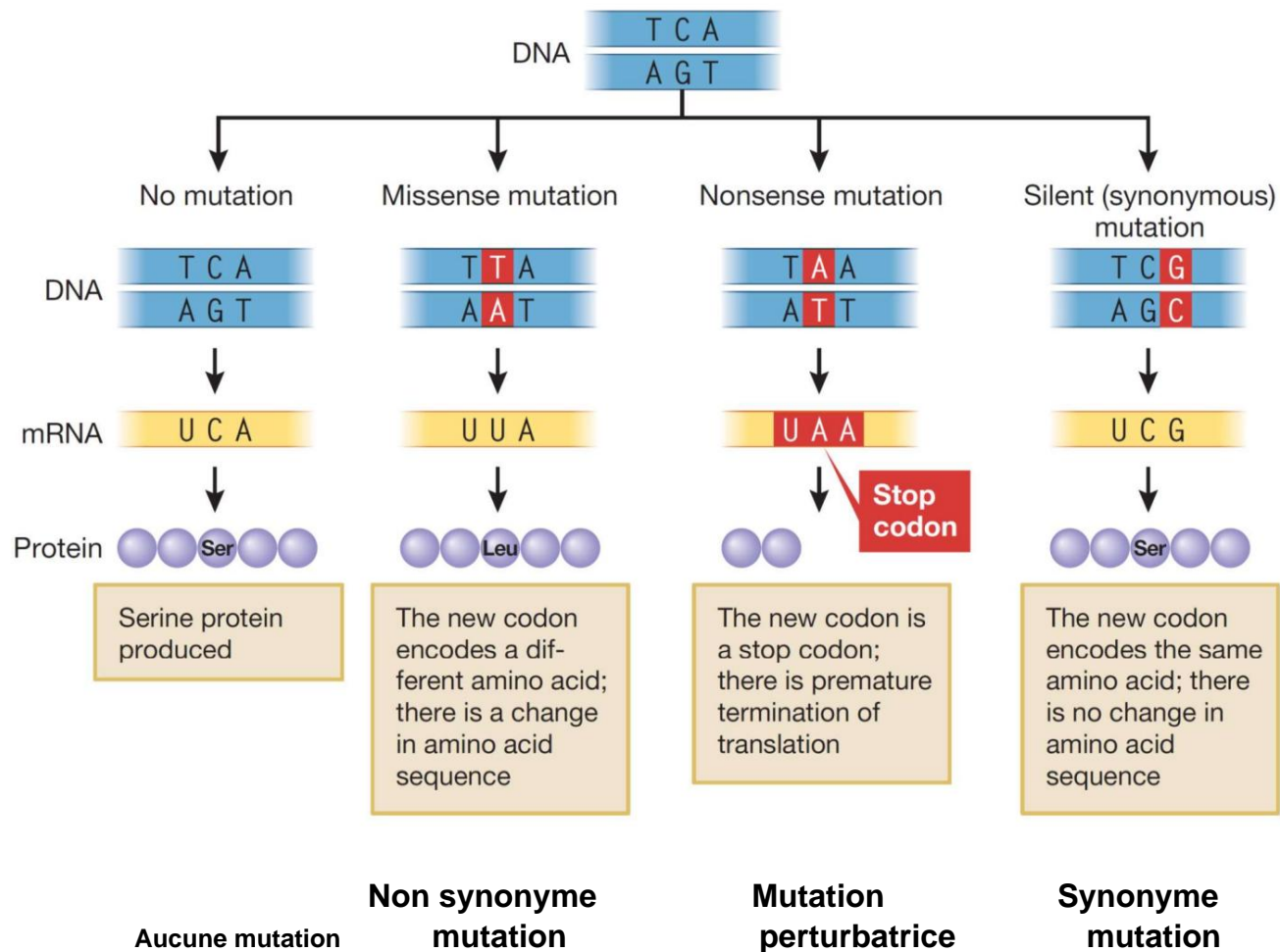
Starting allele frequency = $1/(2N_e)$

→ Waiting time

Low probability of fixation



I. Mutations à petite échelle: mutations d'un seul nucléotide

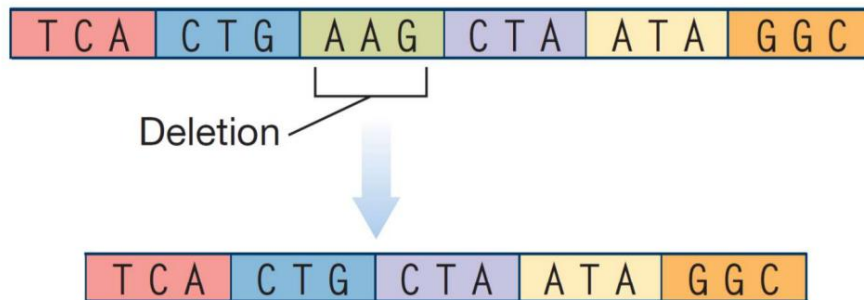


Même un changement de base peut avoir un effet majeur sur le polypeptide résultant!

I. Mutations à petite échelle : insertions et délétions

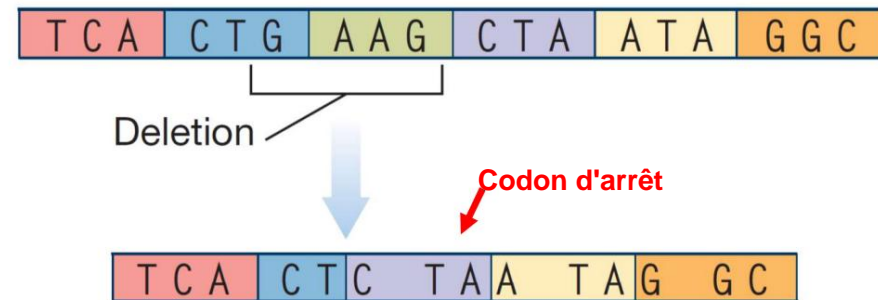
Petites insertions ou suppressions (indels)

A In-frame



Dans le cadre : multiple de 3
 -> Aucun changement dans les
 codons en aval

B Frameshift



Frameshift : non multiple de 3 ->
 Changements dans tous les codons en aval
 et donc dans la séquence d'acides aminés !