

# Génétique Procaryote : Introduction, généralités.

L'arbre phylogénétique des procaryotes donne deux groupes :

- Les eubactéries ou bactéries vraies.
- Les archéobactéries, qui vivent en conditions extrêmes.

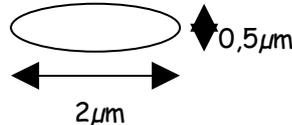
On va surtout s'intéresser aux eubactéries et en particulier, à *Escherichia coli* (Gram -, entérobactérie vivant dans notre intestin).

L'intérêt de la génétique procaryote : actuellement, on a identifié moins de 1% des bactéries vivantes, il y aurait  $5.10^{30}$  bactéries au total. Dans les océans, il y a  $10^6$  bactéries par mL d'eau de mer. On ne sait cultiver qu'un petit nombre de ces bactéries, on ne sait donc pas identifier les autres.

Les bactéries sont :

- responsables de maladies,
- utilisées en biotechnologie,
- des organismes « simples », faciles à étudier.

## I\ Le cycle cellulaire de *E. coli*.



La dimension d'*E. coli* varie en fonction du milieu de culture : si le milieu est riche, la bactérie sera plus grosse.

*E. coli* se multiplie par scissiparité.

T est le temps de génération (au mieux, 20 minutes).

Pour se multiplier pendant la période T, la longueur va être multipliée par 2 et le diamètre va varier. A la fin de cette période (T), on obtient deux cellules filles par division qui ont la longueur initiale.

Dans la bactérie, l'ADN met 40 minutes pour se répliquer : c'est la période C (constante). Il faut 20 minutes entre la fin de la réplication et la division : c'est la période D (constante).

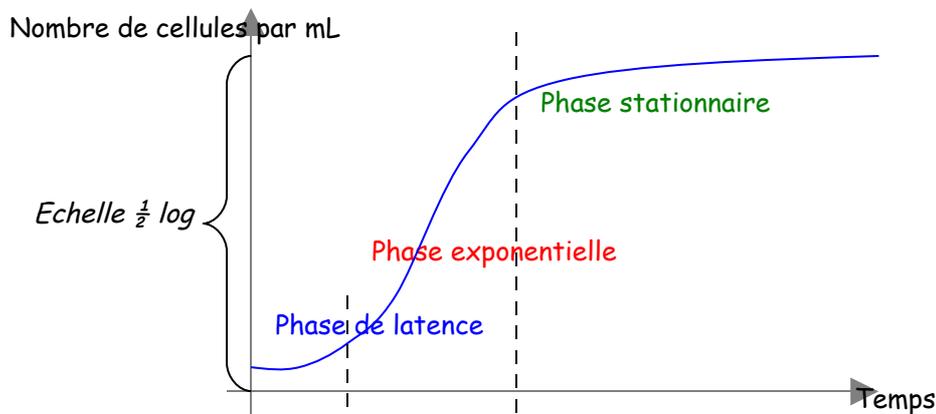
Comment T=20 minutes est possible alors que  $40\text{minutes} + 20\text{minutes} = 60\text{minutes} = \text{constante}$ ?

Concept :  $I + C + D$  ; I est le temps nécessaire à la cellule pour se préparer à une nouvelle initiation de réplication. En général,  $I = T$ . La division cellulaire est instantanée (pas de temps de mitose).

→ Dans la cellule mère, il y a plusieurs cycles de réplication en même temps, donc la période I d'un cycle peut avoir lieu dans une bactérie arrière-grand-mère.

Une bactérie peut croître en milieu liquide ou en milieu solide. Les bactéries sont conservées en laboratoire dans des congélateurs à  $-70^\circ\text{C}$  et à 15% de glycérol pour éviter l'éclatement.

Pour faire pousser, on prend un peu de culture en réserve que l'on met dans un milieu de culture à 37°C et on agite → Mesure de culture.



- La phase de latence.

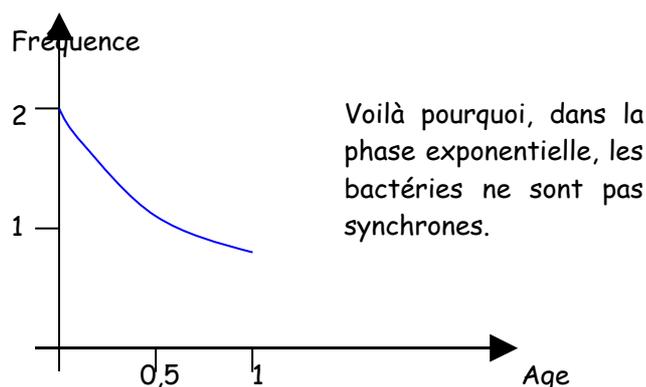
Cette phase correspond au temps nécessaire à la bactérie pour s'adapter à son nouveau milieu pour pouvoir se diviser activement.

- La phase exponentielle.

Toutes les bactéries poussent avec le même T mais toutes les bactéries ne sont pas synchronisées. En anglais, on dit « steady state » (état d'équilibre). Quelque soit le moment de la phase exponentielle, pour tous les paramètres, la concentration ne variera pas : on a un équilibre entre les voies de synthèse et de dégradation. En steady state, on dit que les bactéries sont « physiologiques ». On travaille toujours avec ce type de bactéries.

- La phase stationnaire.

Les bactéries ne poussent plus car il n'y a plus rien à « manger » → ce sont elles que l'on stocke. Ce sont les facteurs environnement et trophique qui limitent la croissance d'une culture de bactéries. Une bactérie se prépare à la phase stationnaire car elle a des capteurs : elle accumule du stock, elle se fait une membrane plus épaisse et ne synthétise que les protéines réellement utiles.



Pour construire une courbe de croissance bactérienne, on mesure le nombre de cellules par mL par spectrométrie car il existe une relation entre DO et nombre de cellules par mL.

→ DO = 1 ⇔  $4 \cdot 10^8$  cellules/mL.

Pour savoir si les bactéries sont mortes ou vivantes, on étale l'échantillon sur un milieu de croissance à différentes dilutions. On fait ensuite la DO des seules bactéries qui ont poussé.

Pour un diamètre de 1mm, on a  $10^7$  cellules par colonie, soit, 23 générations.

$$N = N_0 \times 2^{t/Z} \text{ et } N = d^2 \times 10^7$$

N est le nombre de bactéries,  $N_0$  est le nombre initial de bactéries, d est le diamètre.

## II\ Le génome bactérien.

### A\ Le génome chromosomique.

L'ADN de *E. coli* est formé d'ADN double brin circulaire, d'une origine *C*, d'un terminus *C* et fonctionne en répllication en mode  $\theta$  bidirectionnelle : l'initiation commence en ori *C*, il y a ensuite séparation des deux brins et allongement du brin complémentaire par une polymérase.

L'ADN a été séquencé dans la souche sauvage de référence MG16SS et pèse 4936kpb. On peut connaître le nombre de plages ouvertes de lecture (ORF) afin de connaître le nombre de gènes (à peu près 5000 gènes).

Remarque : informations au [www.pasteur.fr/bio/colibri.html](http://www.pasteur.fr/bio/colibri.html) ou [htm](#).

Sur une population bactérienne, il y a toujours moins de 0,1% de particules sans ADN, donc, la partition du chromosome entre les deux cellules filles est régulée.

Le génome minimal est le nombre minimal de gènes nécessaires à la vie (250 gènes pour *E. coli*). Les gènes nécessaires à la vie sont toujours très conservés. Pour *E. coli*, 100 gènes servent à la traduction et 28 ont un rôle inconnu.

### B\ Le génome extra chromosomique : plasmide, virus et transposons (Elts genetic Mobils).

# Gène et mutation.

## I\ Le gène.

### Définitions :

- Selon **Mendel** : Le gène est l'unité structurale et fonctionnelle de l'hérédité, porteuse de l'information d'une génération à l'autre. Cette définition est basée expérimentalement sur l'expérience de la complémentation.
- Le gène est un segment d'ADN qui spécifie un polypeptide ou un ARN actif. On exclu les séquences consensus. C'est la définition minimaliste de **Watson**.

Dans une bactérie, les gènes sont souvent organisés en un même opéron s'ils sont impliqués dans une même fonction.

Nomenclature : par exemple, l'arginine peut être écrite argA ; soit, trois lettres en minuscules et une lettre en majuscule et le tout est souligné.

Si l'on a plusieurs allèles, on ajoute le numéro de l'allèle : argA1. Le gène argA code pour la protéine ArgA : une lettre majuscule, deux lettres minuscules et une lettre majuscule ; rien n'est souligné.

## II\ Recherche des mutants.

Toute analyse génétique débute par l'isolement d'un mutant. On identifie ensuite le gène muté.

On peut aussi prendre un gène, le faire muter, le « remettre » dans la cellule et comprendre son rôle : c'est de la génétique inverse.

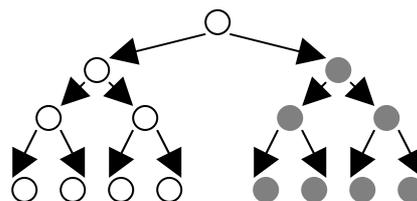
### A\ définitions.

Mutation : c'est le changement héréditaire dans la séquence nucléotidique qui ne se traduit pas forcément par un phénotype observable.

Mutant : un mutant est un organisme ou une cellule qui porte une mutation (pas forcément observable).

En pratique, on parle de mutant quand celui-ci présente un phénotype différent de celui du sauvage et on considère que le nombre de mutants est égale au nombre de mutations.

○ Sauvage (+) ● Mutant



Fréquence de mutants : c'est le nombre de cellules mutantes sur le nombre total de bactéries.

Taux de mutation : c'est la probabilité qu'un événement mutationnel se produise en un laps de temps donné. Le taux de mutation est obtenu en mesurant la fréquence de mutant par unité de temps choisi.

## B\ Différents types de mutants et les stratégies d'isolement (repérage et isolement).

Une souche **auxotrophe** est incapable de synthétiser un constituant cellulaire.

Une souche **prototrophe** est capable de synthétiser tous ses constituants cellulaires sur un milieu minimum (eau, sels minéraux, sucre).

Les **mutants de sucre** : ces mutants sont incapables de d'utiliser un sucre donné comme source d'énergie (C et N).

Les **mutants de résistance** (à un composé toxique) : c'est un mutant qui pousse même en présence d'un composé toxique.

Les **mutants de fonctions essentielles** ou **Mutants conditionnels** : en conditions permissives, la mutation ne s'exprime pas et la bactérie pousse. En conditions non permissives, la mutation s'exprime et la bactérie meure. On trouve dans ce cas, les mutants thermosensibles. Ce sont des mutations qui changent un acide aminé, donc il y a variation dans la structure tridimensionnelle de la protéine.

Isolement :

- Mutant de résistance : sélection.
- Mutant auxotrophe : repérage.
- Mutant de sucre : repérage.
- Mutant conditionnel : repérage.

### 1\ Sélection.

Si la fréquence de mutation de résistance à l'ampicilline est de  $10^{-9}$ , on fait une structure d'au moins  $10^9$  cellules ; on centrifuge puis n étale sur une boîte contenant de l'ampicilline. Sur les  $10^9$  cellules, une seule poussera → La souche est isolée.

### 2\ Repérage.

Repérage de mutants d'auxotrophie, de sucre et conditionnels.

Si la fréquence Arg résistance est égale à  $10^{-6}$ , on étale sur une première boîte de milieu complet : tout pousse.

On repique ensuite sur une seconde boîte sans arginine et l'on regarde ce qui ne pousse pas et par corrélation avec la boîte 1, on repère la souche.

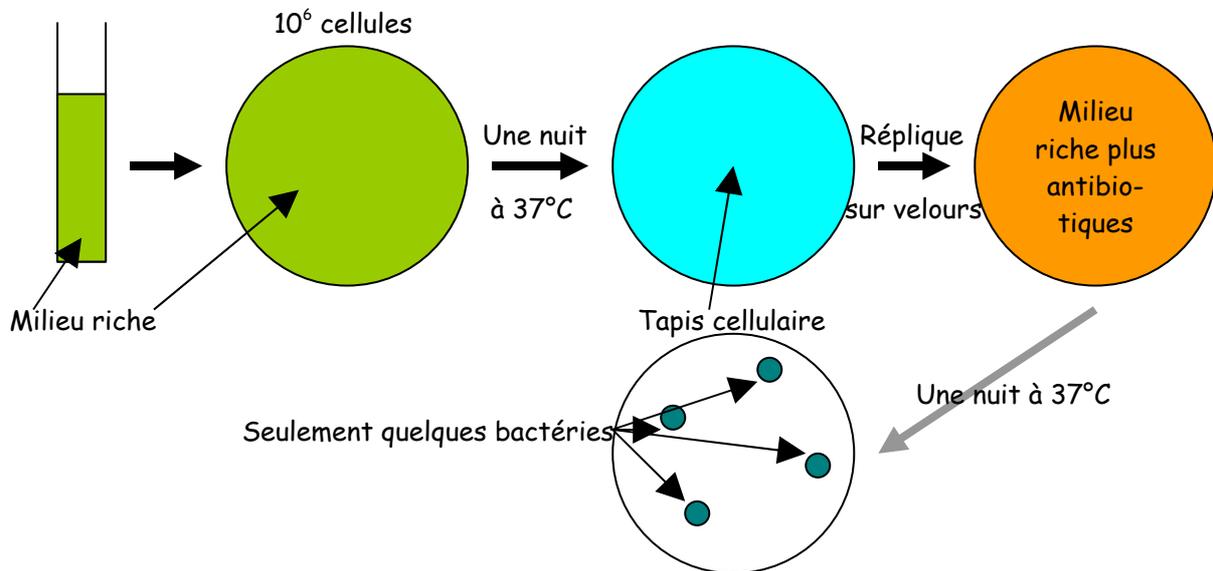
Pour espérer repère un mutant, on doit faire un million de boîtes !!! Pour diminuer le nombre de boîtes, on peut augmenter la fréquence de mutation par mutagenèse.

## III\ Mutation spontanée - Nature des mutations - Suppression.

On a deux hypothèses quant à l'apparition des mutations :

- Au hasard : ce sont des mutations spontanées ou directes.
- Pour s'adapter : on parle dans ce cas de mutations adaptatives.

## A\ Expérience de démonstration (Lederberg 1952).



On prend des échantillons mutants déterminés par l'expérience I sur le tapis cellulaire et l'on recommence autant d'expériences que d'échantillons.

Si l'on obtient la même fréquence de mutation qu'en I, la mutation est dite adaptative.

Si la fréquence augmente à chaque expérimentation, on parle alors de mutations spontanées.

On note que si la fréquence augmente jusqu'à 100% pour de nombreux cycles, c'est que l'on est en présence de mutation spontanée indépendante de l'agent sélectif (c'est un dogme).

## B\ Nature des mutations.

### 1\ Mutations ponctuelles.

Les mutations ponctuelles concernent un seul nucléotide. Il existe différents types de mutations ponctuelles :

- Conservatives.
- Faux sens.
- Non sens.
- Décalantes.

Mutation conservative : cette mutation ne modifie pas l'acide aminé intégré dans la chaîne polypeptidique. En effet, le code génétique est dégénéré.

Mutation faux sens : cette mutation entraîne l'incorporation d'un mauvais acide aminé dans la chaîne polypeptidique. Si cet acide aminé est important, on obtient un phénotype mutant.

Mutation non sens : il y a remplacement d'un nucléotide par un autre et engendre ainsi un codon stop. C'est aussi un cas d'apparition de phénotype mutant.

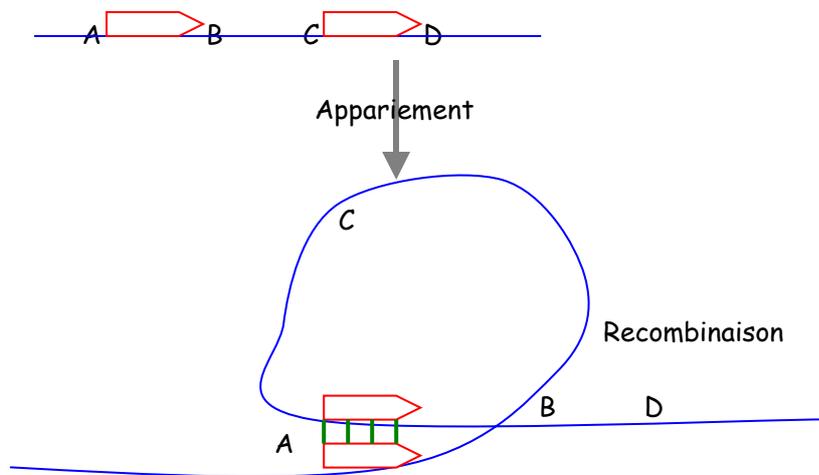
Mutation décalante : il s'agit ici de la délétion ou de l'insertion d'un nucléotide qui décale le cadre de lecture. Il y a souvent engendrement d'une mutation faux sens et apparition d'un peptide incorrect et donc un phénotype mutant.

## 2\ Remaniements génétiques.

Les remaniements génétiques concernent diverses centaines de nucléotides. Il existe différents types de remaniements :

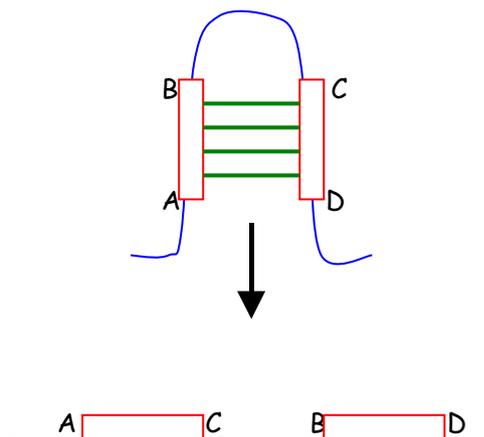
- inversion - délétion - duplication : ces trois phénomènes impliquent un processus de recombinaison homologue.

### $\alpha$ \ Délétion.



On a déléte la séquence entre B et C : on obtient alors A  D.

### $\beta$ \ Inversion.

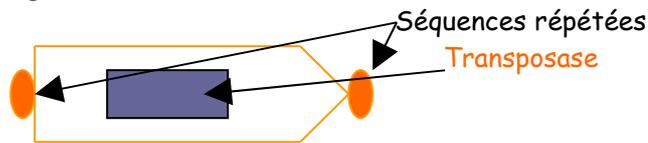


Ces remaniements permettent, au cours de l'évolution, de faire évoluer les gènes essentiels.

### $\gamma$ \ Les séquences homologues.

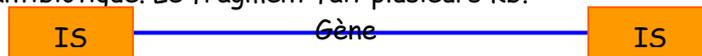
EGM : Element Genetic Mobil ; IS : Insertion sequence ; Tn : Transposon.

- IS : c'est un fragment d'à peu près 700pb. On trouve plusieurs copies de IS sur le chromosome. Ce sont des séquences qui permettent des recombinaisons homologues.



La transposase et les séquences répétées sont nécessaires à la transposition.

- Tn : le transposon est un élément composite constitué de deux IS en répétition inversée, encadrant un morceau d'ADN portant un gène de résistance  $\zeta$  un antibiotique. Le fragment fait plusieurs Kb.



Le transposon « transpose » car IS contient la transposase et les séquences aux extrémités des IS. Le gène entre les IS n'intervient pas.

Tout peut être transposé, soit, seulement un IS, soit, tout le chromosome ( $\rightarrow$  Transposition inverse) à l'extérieur des IS  $\rightarrow$  On a soit Tn, soit le chromosome, soit un IS qui transpose.

### 3\ Suppression.

Une seconde mutation va supprimer l'effet de la première. Une seconde mutation intragénique se situe dans le même gène que la première. Une seconde mutation extragénique se déroule dans un autre gène que la première.

Les mutations intragéniques sont rares. Elles suppriment les mutations non sens, faux sens et décalantes.

Les mutations extragéniques peuvent être :

- Indirectes.

Ces mutations permettent d'identifier les différents partenaires d'un même processus de biosynthèse. Exemple du métabolisme du glucose :

Galactose  $\rightarrow$  (GalK)  $\rightarrow$  Galactose-1P  $\rightarrow$  (GalE)  $\rightarrow$  UDP-Glucose.

- Si tout va bien, on obtient du glucose.
- S'il y a mutation de GalE, il y aura accumulation de Galactose-1P toxique, donc, mort de l'individu.
- Si la mutation a lieu sur GalK, on observe une accumulation de galactose, non toxique. Il y a mise en évidence d'un nouveau gène GalK qui appartient à la chaîne du métabolisme.
- Directes informationnelles : elles concernent les gènes des ARNt.

La seconde mutation a lieu dans le gène de l'ARNt et implique la mutation de l'anticodon de cet ARN. Ce nouvel ARNt va pouvoir insérer un acide aminé dans la chaîne polypeptidique, même s'il lit un codon stop, la chaîne continue.

Exemple de mutation :

- Mutation supE : elle inclut Glu en face du codon stop AMBRE (UAG).
- Mutation supF : elle inclut Tyr en face du codon AMBRE.
- Mutation supL : elle inclut Lys en face des codons AMBRE et OCHRE (UGA).
- Remarques : AMBRE  $\rightarrow$  UAG ; OCHRE  $\rightarrow$  UGA ; OPALE  $\rightarrow$  UAA.
- Indirect informationnelles : elles concernent les gènes impliqués dans la traduction des ARNm en affectant l'efficacité.

## C\ Nature des mutations.

### 1\ Fidélité de la réplication.

Les mutations apparaissent essentiellement au moment de la réplication. Cette mutation va être fixée au cycle de réplication suivant.

Si, pendant la réplication, il y a erreur, la cellule peut réaliser une correction. Cela explique pourquoi, in vivo, le taux de mutation est très faible.

### 2\ Modification de base.

La modification de base se déroule après la réplication.

Par exemple, une cytosine qui perd son NH<sub>2</sub> devient une tyrosine. La mutation sera fixée par la réplication.

### 3\ Réarrangements.

Les réarrangements ont lieu pendant la synthèse ou après la réplication. Ce sont des phénomènes très rares.

## IV\ Mutagenèse.

On peut réaliser une mutagenèse par des agents chimiques (NitrosoGuanidine), par des agents physiques (UV qui créent des dimères de T), avec des transposons (très pratique).

On augmente la fréquence de transposition en accroissant l'expression de la transposase. Si la transposition porte un gène de plasmide juste à côté d'un gène chromosomique, cela les lie. Les deux phénotypes seront toujours retrouvés ensemble. Si un des deux gènes n'est pas sélectionnable et qu'on le lie à un gène sélectionnable, cela le fait passer d'une situation de repérage à une situation de sélection → C'est un avantage.

Comme inconvénient, on ne pourra pas isoler de mutant de ce genre essentiel car, quand un transposon entre dans un gène, il l'inactive.

Mutagenèse dirigée.

# Echanges génétiques.

Chez les eucaryotes, les parents participent autant l'un que l'autre dans les échanges réciproques. Chez les procaryotes, les échanges sont inégaux et non-réciproques. On trouve comme mécanismes d'échanges :

- La conjugaison : une souche donneuse, une souche réceptrice.
- La transduction : une souche donneuse, une souche réceptrice, vecteur (phage).
- La transformation : souche réceptrice, ADN.

Pour transférer, un ADN doit s'insérer dans un réplicon pour une recombinaison homologe (dépend de la protéine *RecA*).

Réplicon : chromosome ou plasmide selon qu'il y ait homologie ou non. Lors de la recombinaison, le produit entrant sera répliqué, le produit sortant sera perdu car il sera seul dans le cytoplasme et donc détruit par la nucléase.

## I\ La conjugaison.

La conjugaison est processus par lequel un ADN est transféré d'une bactérie donneuse à une bactérie réceptrice. Il faut qu'il y ait un contact physique.

La capacité d'une bactérie à donner son ADN dépend de la présence du plasmide F (facteur de fertilité).

Le plasmide F est composé de 100Kb, est présent en une seule copie par bactérie et est extrêmement stable, donc jamais perdu pour la population, car il a développé des systèmes :

- Système de Partition : il va se répartir pareillement dans les deux cellules filles.
- Système de Killing : si le système de partition ne fonctionne pas et qu'une cellule fille ne porte pas ce plasmide, elle sera tuée.

Le facteur F peut s'intégrer au chromosome bactérien (avec une fréquence de  $10^{-4}$ ). Comme F peut être intra ou extra chromosomique, on le nomme « épisode ». Ce facteur s'intègre par recombinaison homologe, car il possède des : -IS2 -IS3 -Transposons. Il possède aussi des gènes *tra*, nécessaires au mécanisme de transfert de la conjugaison (1/3 du génome de F), une origine de transfert -*OriT*- et bien sûr, une ou plusieurs origines de réplication.

*tra* code pour les pili sexuels qui ont pour rôle d'aller chercher un partenaire. Un individu avec pili est dit : mâle ou F+ ; un individu sans pili est dit femelle ou F-.

Pour que la conjugaison soit possible, il faut F+ et F-.

Les pili, en se dépolymérisant, vont rapprocher les partenaires → contact.

Une fois le contact réalisé, F de F+ passe progressivement vers F- à partir d'*OriT*. Ce qui entre en premier est le bout 5'.

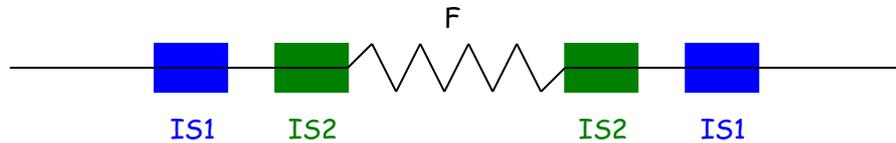
→ F- devient F+

→ F+ reste F+ car au fur et à mesure que F passe dans F-, il est répliqué.

La conjugaison est rapide : autour de deux minutes. Le facteur F est un réplicon donc, tout se fait extrachromosomiquement.

Quand F vient s'intégrer au chromosome bactérien, la bactérie devient Hfr. La recombinaison de F est alors prise en charge par la réplication du chromosome (bactérien). F peut se ré-exciser à la fréquence de  $10^{-4}$ .

Si F s'excise en utilisant la séquence IS1, il y a erreur, normalement, c'est avec IS2.



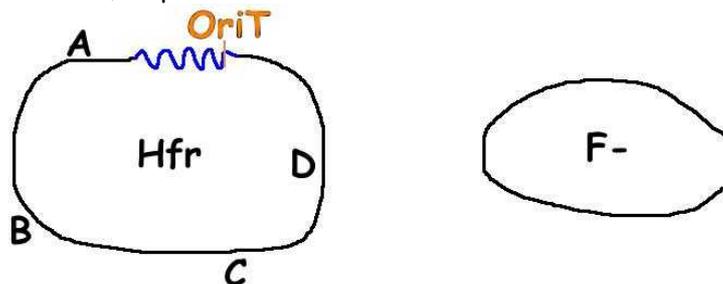
Alors, le produit de l'excision comportera un bout de chromosome bactérien → on parle alors du « facteur F' ».

F' peut conjuguer comme une bactérie F+.

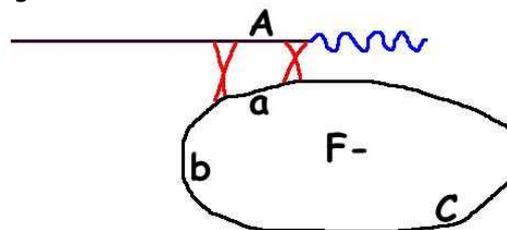
F' × F- :

→ F- devient F' et donc, on parle de diploïdie partielle pour le bout de chromosome bactérien contenu dans le plasmide et dans le chromosome de la nouvelle F'. Cela permet des expériences de complémentation. C'est intéressant d'un point de vue évolutif.

Remarque : A l'arrivée de toutes ces conjugaisons, on n'aura plus que des « mâles ». Pour lutter contre ce phénomène, on prend Hfr comme bactérie donneuse → Hfr × F-.

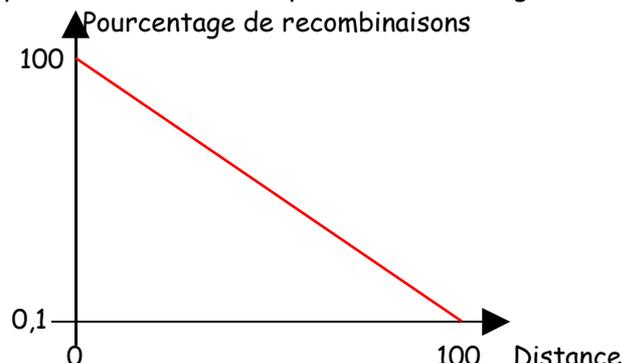


Lors de la conjugaison, avec le plasmide intégré, vont passer A, B, C... Donc, dans F-, il peut y avoir recombinaison homologue.



F- (a) deviendra F-(A) car tout le plasmide ne pénètre pas (il manque la fin).

La conjugaison s'arrête spontanément et au hasard, donc, la probabilité pour qu'un marqueur passe dans la souche réceptrice dépend de la position de celui-ci par rapport à OriT : plus le marqueur est proche d'OriT, plus il a de chances de passer → C'est le gradient de transfert.



C'est grâce à ses expériences de recombinaison que l'on a su que le chromosome de *E. coli* était circulaire avant même de savoir réellement que c'était un chromosome.  
En conjugaison, on travaille sur de grandes parties de chromosomes.

Applications :

- Transformer une bactérie leu<sup>-</sup> en leu<sup>+</sup> par une souche donneuse leu<sup>+</sup> Hfr.
- Mutagenèse régionalisée toujours avec Hfr.

## II\ Transduction.

Il faut un vecteur (e) donneur et un récepteur. Ce vecteur va être un phage. Il existe deux types de transduction :

- Spécialisée avec le phage  $\lambda$ . C'est toujours la même région d'ADN de cellule donneuse qui est transduite.
- Généralisée avec le phage P1. N'importe quel bout d'ADN peut être transduite.

Les phages injectent leur ADN dans les bactéries. Cet ADN se circularise puis peut avoir deux avenir possibles car  $\lambda$  et P1 sont des phages tempérés :

- Le cycle lytique : la réplication du génome phagique est très efficace et tous les gènes sont exprimés. Finalement, il y a lyse des bactéries et libération de 100 à 500 phages.
- Le cycle lysogénique : l'ADN injecté dans la bactérie ne se réplique pas. A partir de là, on distingue deux cas (un pour  $\lambda$  et un pour P1) :
  - Pour  $\lambda$  : l'ADN non-répliqué s'insère sur le chromosome bactérien au locus *attB* (entre les gènes *bio* et *gal*). On a alors un prophage qui n'exprime qu'un gène : celui de rester intégré. Sa réplication est prise en charge par le métabolisme bactérien.
  - Pour P1 : l'ADN non-intégré se circularise et reste extra-chromosomique, mais on parle aussi de prophage. Il va se comporter comme un plasmide. C'est un réplicon qui se suffit à lui-même et possède un système de partition et de killing.

Les prophages  $\lambda$  et P1 peuvent naturellement entrer en cycle lytique. P1 donne alors directement des phages alors que  $\lambda$  doit d'abord s'exciser.

Lors de l'encapsidation, au lieu d'encapsider du génome phagique, le phage P1 encapside de l'ADN chromosomique (dans 0,3% des cas). Ces particules phagiques qui contiennent de l'ADN chromosomique sont transductrices. Chaque région du chromosome a la même probabilité d'être encapsidée dans la particule phagique → C'est la transduction généralisée.

Lors de l'excision du phage  $\lambda$ , il peut y avoir des problèmes : l'excision contient un bout de *bio* ou de *gal*. Lors de l'encapsidation, *bio* ou *gal* va l'être aussi → c'est la transduction spécialisée.

Applications.

Un phage P1 virulent entre systématiquement en cycle lytique.

On va réaliser un étalement des bactéries sur une boîte. On y dépose ensuite un phage P1 virulent ; il entre dans une bactérie, s'y multiplie et produit 100 à 500 nouveaux phages. Ces derniers vont infecter d'autres bactéries.

Remarque : Une bactérie lysogène devient immune, c'est à dire, insensible à une surinfection.

→ Au final, on observera une plage de lyse claire : une plage de phage introduits.

Si l'on avait utilisé un phage P1 sauvage, celui-ci aurait pu entrer en cycle lytique ou en cycle lysogénique. Donc, dans la plage de lyse finale, il reste des bactéries lysogènes qui continuent à se multiplier : on a alors une plage de lyse trouble.

Avec P1 :

Dans 90% des cas, un fragment d'ADN transduit est détruit par les nucléases. Dans les 10% restant, s'il peut y avoir une recombinaison homologe, le fragment est inséré dans le chromosome de la bactérie réceptrice qui prendra ce nouveau caractère et perdra son ancien caractère car on est toujours en situation d'haploïdie

Applications :

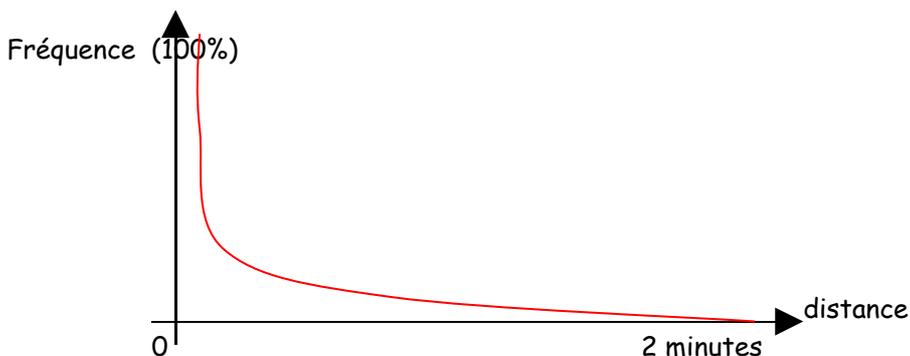
- Mutagenèse régionalisée.
- Cartographie

→ En transduction de ce type, on travaille sur de petites portions de chromosome : jusqu'à deux minutes, soit 90Kb.

On peut affiner une cartographie en localisant la mutation par rapport à un marqueur connu. Si l'on réalise la transduction d'une bactérie  $m^+$ ,  $leu^-$  par une bactérie  $m^-$ ,  $leu^+$  et que l'on sélectionne les bactéries  $leu^+$ , on pourra obtenir deux cas de figure :

- Ces bactéries sont  $leu^+$ ,  $m^-$  : on a cotransduit les marqueurs  $leu$  et  $m$ . →  $m$  est à moins de deux minutes du marqueur  $leu$ .
- Ces bactéries sont  $leu^+$ ,  $m^+$  : on n'a pas cotransduit  $leu$  et  $m$ . → donc  $m$  est à plus de 2 minutes de  $leu$ .

Si l'on est à moins de deux minutes ( $leu$  et  $m$ ), on peut calculer la fréquence de cotransduction qui est absolument corrélée à la distance ( $e$ )  $leu$  et  $m$ .



Cette corrélation suit la loi de Wu

$$F = (1 - d/L)^3$$

D est la distance ( $e$ ) 2 marqueurs.

L : distance maximum ( $e$ )

2 : marqueurs c'est à dire, 2 minutes ou 90Kb.

### III\ La transformation.

Capture d'ADN exogène par une bactérie dite compétente.

Certaines bactéries sont compétentes naturellement. Elles récupèrent l'ADN double brin. Elles dégradent un brin et intègre l'autre grâce à l'énergie due à la dégradation du premier brin. Ensuite, il y a recombinaison homologe.

*E. coli* n'est pas compétente mais on peut la rendre compétente par un choc électrique. Ce qui entre est double brin :

- s'il est linéaire, il est détruit.
- Si c'est un plasmide, il est intégré.

L'intérêt du plasmide peut porter :

- sur le phénomène de résistance.
- Peut être un vecteur de clonage.
- Peut muter un gène.

Un plasmide porter souvent un gène de résistance, donc, pour sélectionner une bactérie transformée, on la fait pousser sur un milieu plus antibiotique spécifique : méthode peut efficace.

# Régulation de l'expression génique.

## I\ Expression génique.

L'ARN nucléotidique n'existe pas dans la cellule. Dès que l'ARN est synthétisé, il est pris en charge par les ribosomes pour éviter l'action des nucléases. Il y a donc un couplage transduction/traduction du fait que la bactérie n'a pas de noyau.

Si l'on a un opéron ABC, et qu'il y ait une mutation non-sens dans B, on n'aura, ni la protéine B, ni la protéine C, car le gène C sera digéré par la nucléase (→ l'effet polaire) et aussi car la transcription ne va pas jusqu'à C.

## II\ Régulation.

Il y a plusieurs niveaux de régulation :

- Initiation de la transcription : activation ou non de promoteur.
- Post-transcriptionnel, qui joue sur la stabilité des messagers.
- Initiation de la traduction : accès ou non aux séquences reconnues par les ribosomes (Shine Dalgarno).
- Post-traductionnel : vitesse de traduction ou stabilité de la protéine.

Le niveau le plus économique est la régulation de l'initiation de la transcription.

### A\ Initiation de la transcription.

On s'intéresse à la division cellulaire. On a cherché des mutants de division. On parle de gènes essentiels donc on cherche des mutants conditionnels : thermosensibles (pousse à 30°C et non à 42°C).

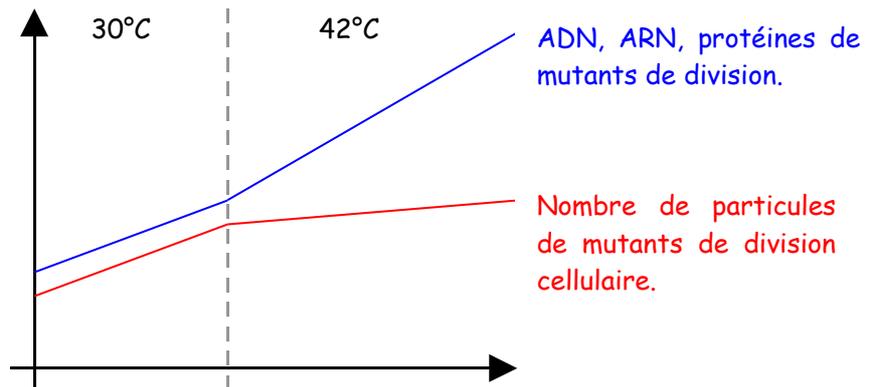
Parmi ceux qui ne poussent pas à 42°C on va chercher les mutants de division cellulaire.

Lorsqu'on les fait pousser à 37°C, on obtient :

Il y a deux cas :

- Les bactéries ne se sont pas divisées et sont sous forme de filament.
- Les bactéries se divisent bien.

On doit confirmer que se sont bien les filaments qui sont mutants de division cellulaire. On les remet alors à 30°C, puis à 42°C.



Le nombre de particules n'augmente plus car il n'y a plus de filamentation : il n'y a plus de division mais le métabolisme fonctionne toujours donc, les quantités d'ADN, d'ARN et de protéines croissent.

On a isolé plusieurs mutants de division dont *ftsZ84* (le gène *ftsZ*) code pour la protéine *ftsZ* qui polymérise pour former un anneau. La constriction de *ftsZ* nécessite le passage  $GTP \rightarrow GDP$  pour finalement cliver la cellule mère.

On a recherché des partenaires de *ftsZ84* par recherche de suppresseurs de la mutation *ftsZ84* qui sont exprimés à forte dose : ce sont les « suppresseurs multicopies ».

On réalise la digestion par *Sau3A* d'ADN de *E. coli* pour obtenir des fragments d'environ 2,5Kb.

On prend une autre souche portant un vecteur de clonage (un plasmide) portant une Ori, un gène de résistance à un antibiotique et un site de clonage multiple.

On fait pousser avec l'antibiotique, on récupère le plasmide et on réalise une digestion complète avec *BamH1* : les coupures sont les mêmes qu'avec *Sau3A*.

On va mettre le plasmide digéré et de nombreux fragments d'ADN dans un eppendorf puis faire agir une ligase pour insertion : on a alors une banque génomique.

Dans le mélange de ligation, on transforme *ftsZ84* puis on étale sur milieu riche avec antibiotique (celui du départ) pour ne sélectionner que les transformants puis on incube à 42°C → C'est une double sélection.

En suite, on séquence l'insert et on l'injecte dans la banque de données et on retrouve à quoi correspond l'insert et quels sont les gènes qui suppriment la mutation *ftsZ84*. On va donc sortir ce gène *ftsZ84* sauvage mais aussi, le gène *rcsB* qui code pour un régulateur transcriptionnel.

*RcsB* supprime la mutation *ftsZ84* en sur exprimant le gène *ftsZ84*. En effet, la quantité de protéines produites est plus importante que l'inactivation de ces protéines.

On va prouver ceci en utilisant les « fusions transcriptionnelles ».

Le gène rapporteur n'a plus son promoteur donc, il est transcrit par le promoteur *ftsZ84*. En présence de *rcsB*, on a 50 fois plus de produit du gène rapporteur, donc, *rcsB* active *ftsZ*.

Comment *rcsB* agit sur *ftsZ* ?

On effectue une analyse par délétion : on va effectuer la délétion partielle du promoteur de *ftsZ84* et doser le produit du gène rapporteur. On trouvera alors où est le site d'action de *rcsB*.

Dans le cas d'une étude par analyse par mutation ponctuelle, on mute systématiquement chaque base.

On peut aussi réaliser une analyse par transcription in vivo pour montrer que rcsB active directement la transcription. Pour cela, il faut une matrice ADN, RNAPolymérase, des NTP marqués et du tampon pour pouvoir synthétiser de l'ARN. Ensuite, on fait migrer les produits obtenus sur un gel d'électrophorèse en conditions dénaturantes.

En présence de rcs (ou rcs+), on voit une beaucoup plus forte activité transcriptionnelle de ftsZ. Rcs est suffisant pour activer la transcription de ftsZ.

RcsB est un régulateur transcriptionnel. Il appartient à un système à deux compartiments : deux gènes rcsB et rcsC. Ici, le premier gène (C) code pour le « senseur » et le second (B) code pour le régulateur.

Le gène rcsC code pour une protéine transmembranaire qui « ressent » un signal extérieur abîmant la membrane. Il y a alors auto-phosphorylation de la partie cytoplasmique de rcsC. Le phosphate est ensuite transmis à rcsB qui est alors seulement activé. Elle va avoir une action sur une cible : dans ce cas, ftsZ.

On sait que rcsC/rcsB régule : la division cellulaire, la réplication de l'ADN (en l'arrêtant), la synthèse de la capsule bactérienne (en l'activant), la synthèse des flagelles pour la mobilité et le chimiotactisme, le système de résistance à des drogues multiples.

Le complexe rcsC/rcsB va intervenir dans la régulation de plus de 50 gènes : c'est un régulateur global de bactérie.

Le fait que rcsB se fixe avant la séquence -35 fait que la fixation de la polymérase est augmentée par rcsB → Augmentation de la transcription.

Si rcsB se fixe sur la séquence -3, il empêche la fixation de la polymérase et réprime donc la transcription.

# Génétique des populations.

## Introduction.

La génétique des populations s'intéresse aux conséquences de la transmission de l'information, de génération en génération pour la structure d'une population.

En 1735, Carl Von Linné crée la *Systema naturae*.

Buffon : pour lui, les ressemblances entre individus sont accessoires par rapport à la transmission de génération en génération. Il y a pré-éminence de l'interfécondité sur les critères morphologiques.

Cuvier : « collection de tous les corps organisés nés les uns des autres ou de parents communs et de ceux qui leur ressemblent autant qu'ils se ressemblent entre eux. » Cuvier était fixiste.

Lamarck : pionniers de l'évolutionnisme, il a mis en doute la notion d'espèce. Pour lui, il n'y a pas de classe constante mais des individus qui se succèdent. Pour lui : « les espèces se fondent les unes dans les autres » au point que l'on ne voit pas les limites.

Darwin : « Je considère le terme « d'espèces » comme arbitrairement donné par pure commodité à un ensemble d'individus se ressemblant beaucoup entre eux, il n'est pas différent de celui de variété... »

On se rend compte maintenant que l'espèce est antinomique de l'évolution.

Les espèces sont des groupes de populations réellement ou potentiellement capables de se croiser et qui sont reproductivement isolés des autres groupes ayant des propriétés.

L'accent est mis ici sur l'isolement reproductif :

- Espèce : peut se reproduire entre eux.
- Population : se reproduisent effectivement entre eux.

La copie du matériel héréditaire est très fiable mais il y a des mutations qui introduisent une variabilité dans le patrimoine des individus. La génétique des populations a pour but de déterminer les lois qui régissent l'évolution de ces variabilités sous l'influence de différents paramètres possibles.

- Evolution aléatoire à chaque méiose.
- Pression de sélection du milieu qui garde ou refuse les mutations.
- Interaction entre les gènes dans un même génotype.

## I\ La variabilité génétique.

Un gène est une sous-unité du matériel héréditaire responsable de la synthèse d'une chaîne polypeptidique sauf les gènes codant pour des ARNt et ARNr.

Le milieu où l'organisme se développe peut agir sur l'expression du génome, c'est à dire, sur le phénotype.

Un gène est de l'ADN ou un enchaînement de nucléotides recopiés à chaque génération cellulaire. Il peut y avoir des erreurs e recopiage. Il y a donc le gène sauvage, déterminé arbitrairement et les allèles.

Les gènes allèles sont des gènes qui s'excluent mutuellement dans le gamète.

On va considérer deux allèles  $A$  et  $a$ . On a donc trois génotypes possibles :

- $A/A$ .
- $A/a$ .
- $a/a$

S'il y a une relation de dominance ( $A$ ) /récessivité ( $a$ ), on a deux génotypes.  $A/A$  et  $A/a$  donnent le phénotype [ $A$ ] et  $a/a$  donne le phénotype [ $a$ ].

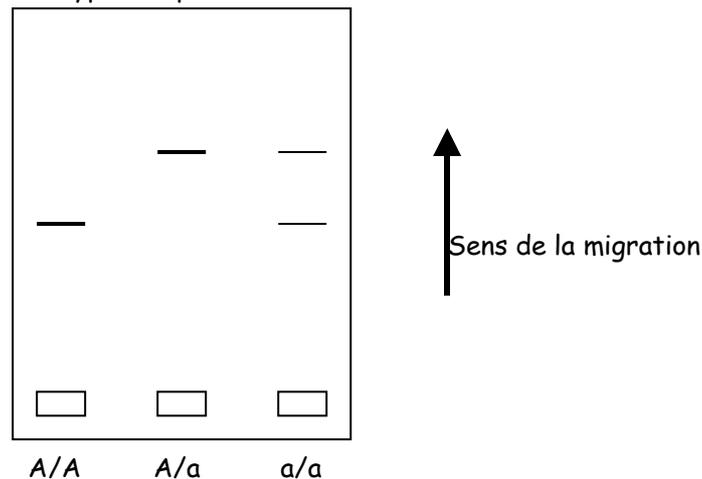
Si, pour un gène, il y a de nombreux allèles différents, on va parler de série allélique.

Plusieurs gènes différents peuvent concourir à l'expression du même phénotype.

L'épistasie est la mutation d'un gène qui masque l'expression d'un autre gène.

On va quantifier la variabilité génétique par électrophorèse des protéines.

On fait migrer les protéines d'une population d'un même gène. Si l'on a des individus  $A/A$ ,  $A/a$  et  $a/a$ , on aura trois types de protéines différentes.



Attention : la migration se fait en fonction de la charge ; en fonction de la mutation, on ne verra pas forcément de différence entre :

- La lysine, l'arginine, l'histidine qui sont chargées positivement.
- L'acide glutamique et l'acide aspartique qui sont chargés positivement.

L'électrophorèse séquentielle.

Cette électrophorèse est réalisée à différents pH. Elle a permis de passer de 6 à 37 allèles connus déterminés pour un gène chez la drosophile.

La fréquence allélique : c'est la proportion, au locus considéré, de tous les allèles considérés.

Exemple :  $n_1$  ind.  $A/A \rightarrow p$

$N_2$  ind.  $a/a$

$N_3$  ind.  $A/a \rightarrow q$

$$\text{Freq}(A) = (2n_1 + n_3) / (2(n_1 + n_2 + n_3)) = \langle P \rangle$$

$$\text{La variance de } \langle P \rangle : \text{var} \langle P \rangle = (\langle P \rangle \cdot [1 - \langle P \rangle]) / 2n = r^2$$

$$\sqrt{\text{var}} = \text{écart type} = \sqrt{r^2} = r$$

L'écart de confiance : si l'on prend la moyenne de la distribution (des allèles dans une population) :

- +/- 1 écart type → 68% des résultats se trouveront dans cet intervalle.
- +/- 2 écart type → 95% des résultats se trouveront dans cet intervalle.
- +/- 3 écart type → 97,7% → Ça ne veut plus rien dire.

Cette notion de fréquence allélique soutend la notion de polymorphisme des populations naturelles.

Un gène est polymorphe si l'allèle le plus fréquent à une fréquence inférieure à 0,95. Un allèle est dit rare quand sa fréquence est inférieure à 0,0005, soit 5 pour mille.

Exemple : 1 à 2 individus sont hétérozygotes pour un allèle rare à n'importe quel locus pour étude sur 43 gènes codant pour des enzymes différentes et ce sur 250 000 européens humains.

On peut déterminer ce polymorphisme par détection directe.

On va couper l'ADN des individus que l'on étudie par des enzymes de restriction puis, on fait une migration par électrophorèse, puis un southern. C'est la technique de RFLP (Restriction Fragment Length Polymorphism).

## II\ Structure de la variabilité génétique.

On va déterminer l'effet des croisements sur la fréquence relative des gènes ou des allèles que l'on a eu au départ.

Il faut déterminer la fréquence génique après la fréquence génotypique après la fréquence phénotypique.

On cherche à définir des modèles : simplification volontaire d'une situation complexe en vue d'éliminer les détails superflus pour s'intéresser seulement à l'essentiel de la question.

Le but final est de comprendre les effets combinés de tous les facteurs qui peuvent agir sur une population.

On a établi des modèles mathématiques. On les a appliqués sur plusieurs générations puis on a comparé les déductions du modèle par rapports aux résultats de la nature afin de savoir si le modèle était valable ou non.