

Chapitre I : Biotechnologie animale

I. La transgénèse

I.1. Définition de la transgénèse

La transgénèse est définie de façon assez large comme l'intégration d'un (ou plusieurs) gène(s) dans le génome d'une plante ou d'un animal par des procédés artificiels c'est-à-dire en dehors de la voie sexuée.

La nouveauté avec la transgénèse est que l'on va agir de façon dirigée sur le logiciel génétique qui pilote le fonctionnement d'un organisme. En général ; les gènes transférés sont des constructions génétiques élaborées pour remplir des fonctions bien précises et s'exprimer à un moment donné dans un organe donné.

I.2. Influence de la région transcrite des transgènes

- Un intron au moins doit accompagner tout ADNc dans un transgène. Les introns, et notamment ceux proches du site d'initiation de la transcription, contiennent souvent des sites de fixation pour des facteurs de transcription participant à l'ouverture de la chromatine, à l'initiation de la transcription et à la transcription.

- Les régions 5' non codantes peuvent contenir des internal ribosomal entry site (IRES), éléments régulateurs de la traduction. L'addition d'IRES favorise la traduction dans certains cas.

- L'addition d'un peptide signal est parfois nécessaire pour l'adressage spécifique de la protéine.

- Les séquences riches en CpG, notamment dans les régions promotrices des transgènes, peuvent dans certains cas inactiver le transgène (par exemple, par méthylation) ; l'élimination de ces séquences est toujours souhaitable.

- Les séquences 3' non codantes sont importantes pour la stabilité des ARNm. Il peut donc être avantageux de les retirer ou au contraire d'en ajouter dans la construction du gène.

I.3. Construction d'un transgène

Chez les mammifères, on considère qu'il existe une centaine de milliers de gènes correspondant à près de $2 \cdot 10^6$ kilobases (kb). Le plus grand nombre de ces gènes est morcelé par des introns. La partie régulatrice du gène est en général elle aussi fractionnée en plusieurs régions activatrices capables d'agir à distance sur le promoteur.

En général la taille du gène est de plusieurs dizaines de kb, d'où l'intérêt de réduire leur taille pour mieux les utiliser en éliminant en général la taille des introns et en ne gardant que les parties essentielles : la partie codante correspondant à l'ADNc de l'ARNm, le promoteur et les éléments de contrôle.

Avant d'être utilisé en transgénèse le minigène artificiel ainsi obtenu est en général d'abord testé en présence d'un gène rapporteur capable d'apprécier son activité dans les cellules transfectées cultivées *in vitro*.

Le minigène est en général fabriqué dans un plasmide d'une souche de colibacille, ce qui permet son amplification et sa purification. L'ADN est précédé en 5' d'une unité régulatrice contenant un promoteur et suivi en 3' d'une séquence de polyadénylation, indispensable à la fin de la transcription. Ces trois éléments sont assemblés à l'aide d'une ADN ligase *in vitro* dans une séquence artificielle du plasmide constituée d'une série de sites reconnus par les enzymes de restriction. Le plasmide comporte aussi une origine de répllication et le gène de sélection d'un antibiotique (par exemple Amp^r pour résistance à l'ampiciline). Le nouveau plasmide est transfecté dans un colibacille, puis amplifié par culture bactérienne. Le fragment d'ADN contenant le minigène, est ensuite purifié par l'action des enzymes de restriction puis par électrophorèse. Certains sites de restriction sont absents dans le minigène, permettant de reconnaître ce dernier des séquences du plasmide bactérien. Le minigène purifié peut ensuite être micro-injecté dans les œufs ou les cellules embryonnaires.

I.4. Techniques de transfert du gène

L'obtention de lignées d'organismes transgéniques suppose qu'un gène étranger ait été transféré de manière telle qu'il soit présent dans les gamètes pour pouvoir être transmis à la descendance. Ce but peut être atteint de plusieurs manières qui dépendent de l'organisme concerné et de la maîtrise que l'on a des différentes techniques de la reproduction.

I.4.1. Le transfert de gène dans les gamètes

La microinjection de gène dans les ovocytes des animaux ne conduit pas fréquemment à l'intégration de l'ADN étranger. L'introduction d'une particule rétrovirale recombinante recouverte de l'enveloppe du VSV (vesicular somatitis virus) entre la zone pellucide et la membrane de l'ovocyte conduit au transfert et à l'intégration des gènes du vecteur (**Figure 1**). L'enveloppe du virus VSV permet une haute efficacité d'infection, et l'intégration est facilitée par l'absence de membrane nucléaire de l'ovocyte au moment choisi pour réaliser l'infection.

La mise en contact direct des spermatozoïdes avec des solutions d'ADN, suivie d'une fécondation *in vitro* ou *in vivo*, n'a conduit qu'à l'obtention d'un très petit nombre d'animaux transgéniques. Plusieurs invertébrés marins, des poissons, des poulets, une vache et un porc transgéniques ont pu être obtenus de cette manière. Dans beaucoup de cas, les gènes intégrés étaient très profondément réarrangés et inexploitable. Il se pourrait qu'une activité DNAsique localisée en périphérie des spermatozoïdes les protège contre une invasion intempestive par des gènes étrangers. Une inhibition de cette enzyme permettrait peut-être à ce procédé d'être utilisable. Une méthode appliquée à la souris, consiste à perméabiliser préalablement la membrane du spermatozoïde avant de l'incuber en présence d'ADN et de procéder à une fécondation par ICSI (intracytoplasmic sperm injection).

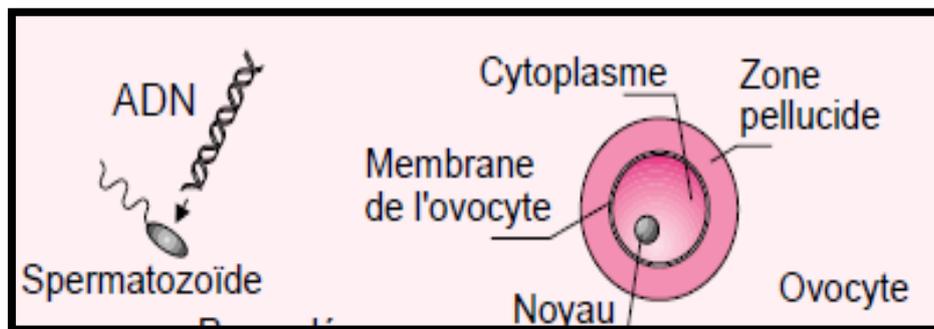


Figure 1

I.4.2. Le transfert de gène dans les embryons au stade une cellule

La technique consiste à injecter dans l'un des deux pronuclei, à l'aide d'un micro-manipulateur un à deux picolitres d'une solution d'ADN (quelques centaines de copies du transgène). La technique nécessite de disposer d'œufs d'âge connu à 2-3h près. Les ovocytes obtenus soit après superovulation, soit par MIV (ovocytes maturés *in vitro*) sont fécondés *in vitro* (FIV).

Il est recommandé de pratiquer cette microinjection quand les pronoyaux sont encore séparés (**Figure 2**). Après microinjection, les embryons sont cultivés jusqu'au stade blastocyste, ce qui permet une élimination spontanée des embryons non viables et un tri des transgéniques, tout en introduisant un gène marqueur avec le gène d'intérêt.

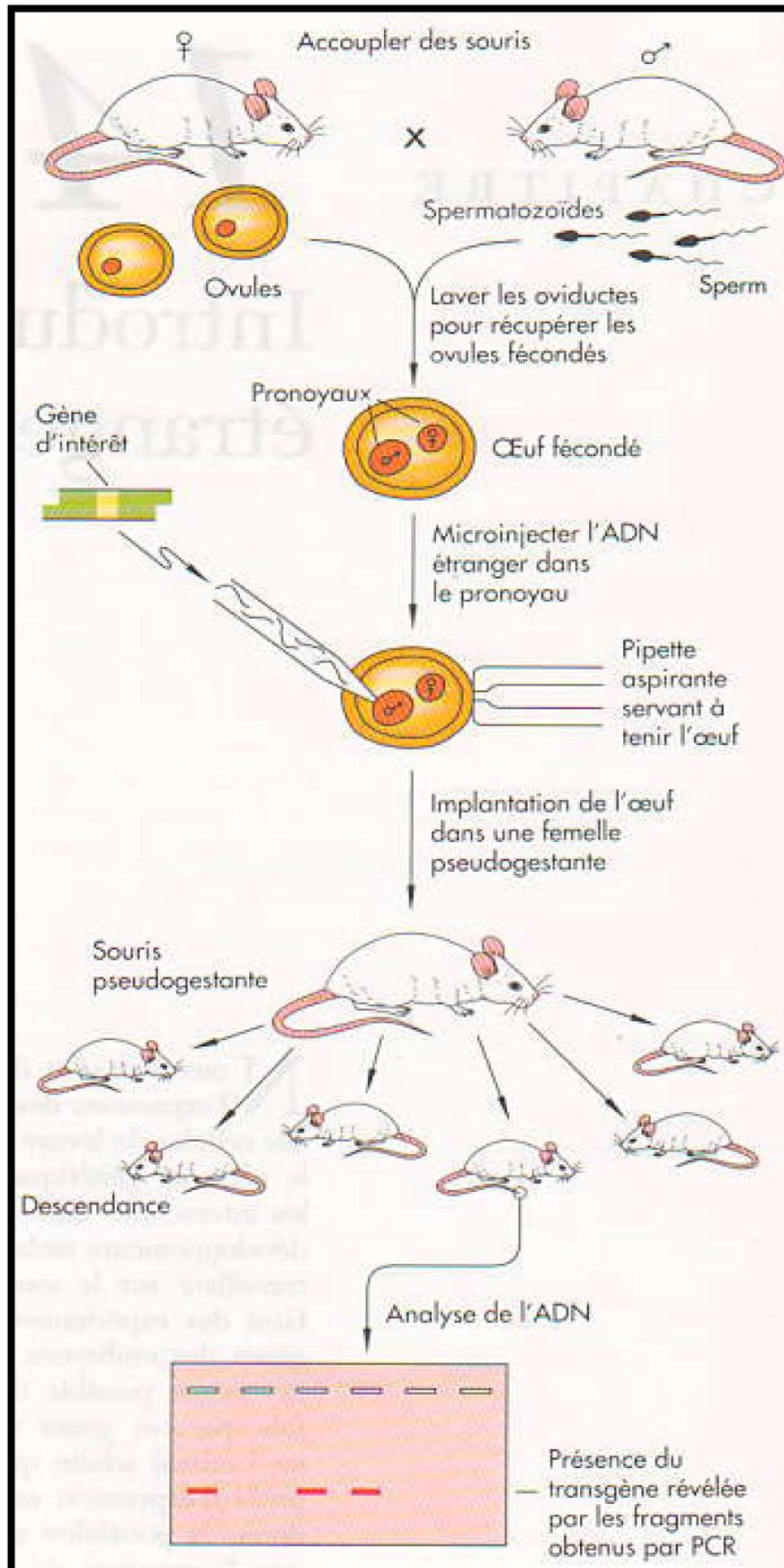


Figure 2

I.4.3. Le transfert de gène par l'intermédiaire des cellules multipotentes

Les cellules multipotentes ont la capacité de pouvoir participer au développement d'un embryon. Elles peuvent être obtenues à partir de la masse cellulaire d'un blastocyste (cellules ES) de souris (**Figure 3**). Elles se cultivent de la même façon que bien d'autres types cellulaires, et pour éviter qu'elles ne se différencient, on les cultive sur une couche nourricière de fibroblastes ou en présence d'un facteur inhibiteur de leucémie dans le milieu de culture. Dans ces conditions, les cellules ES peuvent être maintenues en culture pendant plusieurs semaines en conservant leur remarquable capacité de différenciation.

L'ADN peut être introduit dans ces cellules et on peut contrôler *in vitro* les cellules ayant introduit le transgène à l'aide d'un gène marqueur pour les transférer dans le blastocœle des embryons qui sont ensuite transplantés in utéro. Les souris porteuses du transgène présentent un mosaïsme facilement mis en évidence lorsqu'on utilise un marqueur de pigmentation du pelage.

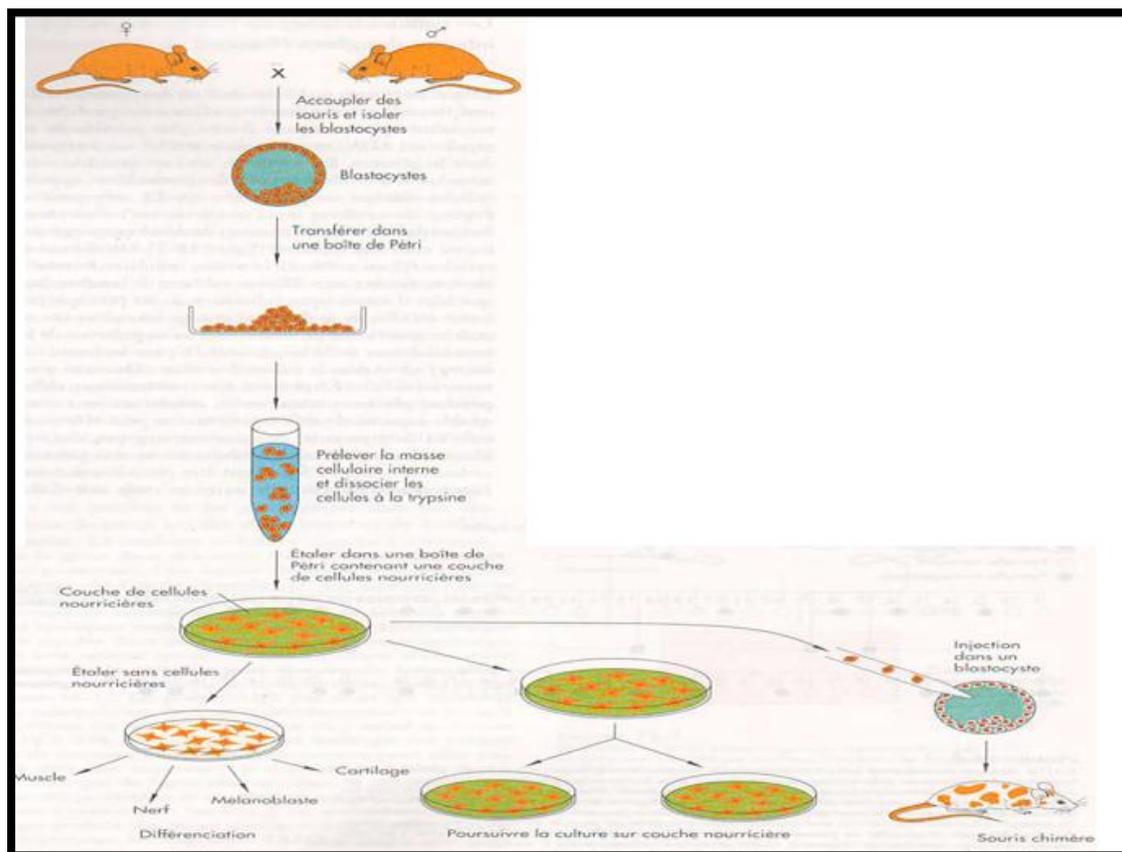


Figure 3

I.5. Caractéristiques des organismes modèles en génétique

I.5.1. La levure comme organisme modèle en génétique

La levure *Saccharomyces cerevisiae*, également connu sous le nom de levure bourgeonnante ou levure de boulanger est l'un des organismes modèles les plus prisés pour la recherche en génétique. Les généticiens font souvent référence au pouvoir impressionnant de la génétique de la levure en raison de la facilité avec laquelle les gènes peuvent être caractérisés chez cet organisme.

Au cours de leur cycle biologique ; les cellules de levure connaissent une phase haploïde (n) et une phase diploïde ($2n$). Durant ces phases, la division cellulaire se produit par bourgeonnement, un processus mitotique au cours duquel une cellule fille plus petite mais génétiquement identique bourgeonne à la surface de la cellule mère, croît, puis se sépare finalement de la cellule mère. Les cellules haploïdes se présentent sous deux types sexuels : a et α . Les cellules a comme α se divisent par mitose en bourgeonnant jusqu'à ce qu'un signal chimique appelé phéromone induise leur croisement. La fusion d'une cellule a avec une cellule α est suivie de la fusion de leurs noyaux puis de la formation d'une cellule de levure diploïde. Selon la disponibilité des nutriments, soit la cellule diploïde continue à bourgeonner soit elle entre en méiose et sporule. Le processus de la méiose conduit à un brassage de l'information génétique et à la formation de quatre spores haploïdes (**Figure 4**).

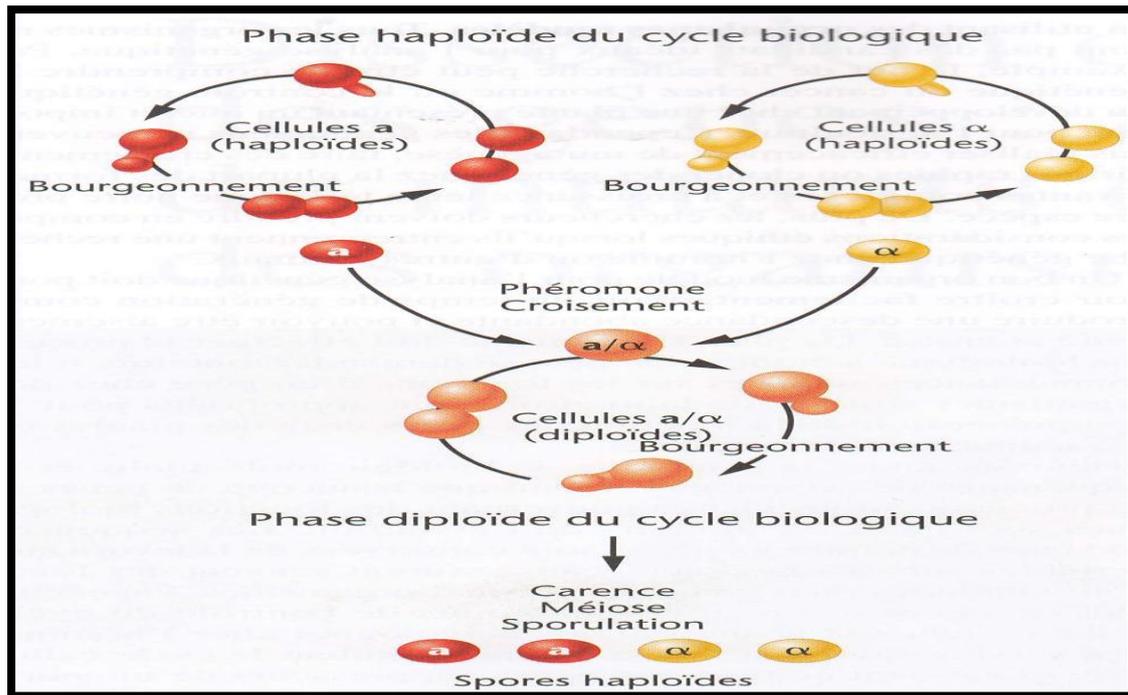


Figure 4

L'alternance de cycles haploïdes et diploïdes dans le cycle biologique de la levure est particulièrement utile pour l'analyse génétique. Il est facile de détecter des mutations ayant un effet dans les cellules haploïdes, puisque l'allèle récessif n'est pas masqué par un allèle de type sauvage. La phase diploïde rend d'autres études possibles, telles que l'analyse par complémentation fonctionnelle. De plus, des mutations ayant un effet létal récessif peuvent être maintenues dans une souche de levure diploïde qui porte la mutation sur un chromosome et l'allèle sauvage sur le chromosome homologue. Un autre avantage majeur de la levure en tant qu'outil génétique est que l'on dispose d'informations sur sa séquence d'ADN ainsi que de collections de mutants et de souches portant des délétions.

I.5.2. La drosophile comme organisme modèle en génétique

La drosophile est un organisme adapté pour la génétique, en partie en raison de la facilité avec laquelle elle peut être élevée et de la taille de son génome environ 13 600 gènes sur quatre chromosomes. La drosophile a un temps de génération court d'environ 10 jours de l'œuf fécondé à l'adulte. Chaque mouche femelle produit approximativement 3 000 descendants au cours de sa vie.

L'extérieur de la mouche révèle aux scientifiques de nombreuses caractéristiques, telles que la couleur des yeux, la forme des ailes, des soies et l'organisation des segments qui peuvent être facilement repérées en utilisant une loupe binoculaire. Des changements de ces caractéristiques reflètent des mutations dans des gènes contrôlant des processus de différenciation et de développement.

Les généticiens tirent également parti des transposons *P* en tant qu'outil chez la drosophile. Les éléments *P* sont des éléments transposables mobiles qui peuvent s'insérer ou s'exciser du génome de la drosophile. Chaque transposition se produit en présence de l'enzyme transposase de l'élément *P* qui reconnaît et agit sur les répétitions inversées de 31 pb à chaque extrémité de l'ADN de l'élément *P*. Les généticiens de la drosophile ont exploité les éléments *P* en tant que vecteurs pour induire des gènes clonés dans le génome de la drosophile. Ils insèrent d'abord un gène cloné d'intérêt au milieu d'un élément *P* qui contient aussi un gène spécifiant un caractère visible tel que la couleur de l'oeil. Ensuite, ils injectent l'ADN de l'élément *P* recombinant dans des œufs, en même temps qu'un plasmide « helper » qui porte le gène codant la transposase. Le gène de la transposase est transcrit et traduit dans les cellules germinales de l'embryon précoce, ce qui permet à l'élément *P* recombinant portant le gène d'intérêt de s'intégrer dans l'ADN des cellules germinales de l'embryon. Puisque le plasmide « helper » ne s'intègre pas dans le génome et ne persiste pas durant le développement, aucune transposition ultérieure ne se produit (**Figure 5**). De plus la transformation assurée par l'élément *P* est l'une des méthodes les plus efficaces pour introduire des gènes clonés chez les eucaryotes supérieurs.

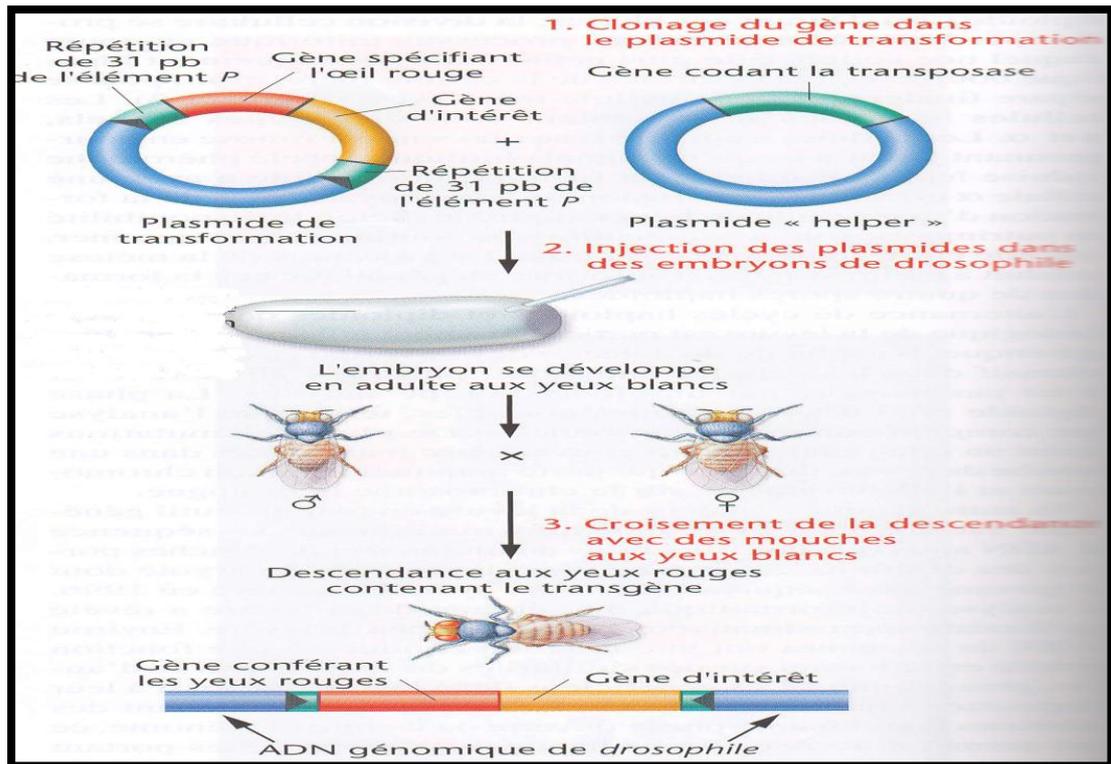


Figure 5

Les généticiens de la drosophile utilisent les éléments *P* comme méthode de mutagenèse et comme outil de clonage de gènes. Pour créer des mutations par insertion d'un élément *P*, ils injectent dans des embryons deux plasmides l'un contient : un élément *P* qui code l'enzyme transposase mais qui n'a pas les extrémités de l'élément *P* pour la transposition. L'autre plasmide contient le gène d'intérêt, le gène spécifiant la couleur rouge des yeux et les répétition de 31pb de l'élément *P*.

Ce second élément *P* contient aussi un gène conférant la résistance à un antibiotique et des séquences ORI venant de l'ADN d'un plasmide bactérien. Les cellules germinales synthétisent la transposase et portent l'élément *P* qui va s'exciser et s'insérer à des positions au hasard dans tout le génome. Certaines de ces insertions se feront à l'intérieur ou autour de gènes importants inactivant ainsi leur fonction. Les chercheurs criblent ensuite parmi les descendants issus du croisement des mouches aux yeux blancs ceux qui présentent le marqueur visible (par exemple, la couleur de l'œil rouge) et ceux qui ont des phénotypes dus à l'insertion d'un élément *P* à une nouvelle position dans le génome.

L'étape suivante consiste à cloner le gène dans lequel l'élément *P* s'est inséré. L'ADN génomique de la nouvelle souche mutante de drosophile est alors purifié, avant d'être digéré par une enzyme de restriction qui coupe une fois à l'intérieur de l'élément *P* et à diverses positions en dehors de l'élément *P* dans ou à proximité du gène inconnu inactivé. Cette digestion libère un fragment d'ADN contenant une partie de l'élément *P* et une partie du gène inconnu. Cet ADN digéré est traité par une ligase d'ADN préalablement digérés sur eux-mêmes, pour former des cercles fermés. Les molécules d'ADN circulaires résultantes sont ensuite introduites dans les bactéries. Seules les bactéries qui ont reçu les cercles fermés portant le gène de résistance à un antibiotique (en même temps que le gène inconnu associé) vont former des colonies sur un milieu contenant l'antibiotique correspondant. Après l'isolement du plasmide à partir des bactéries qui ont survécu, le gène inconnu est séquencé et identifié (**Figure 6**).

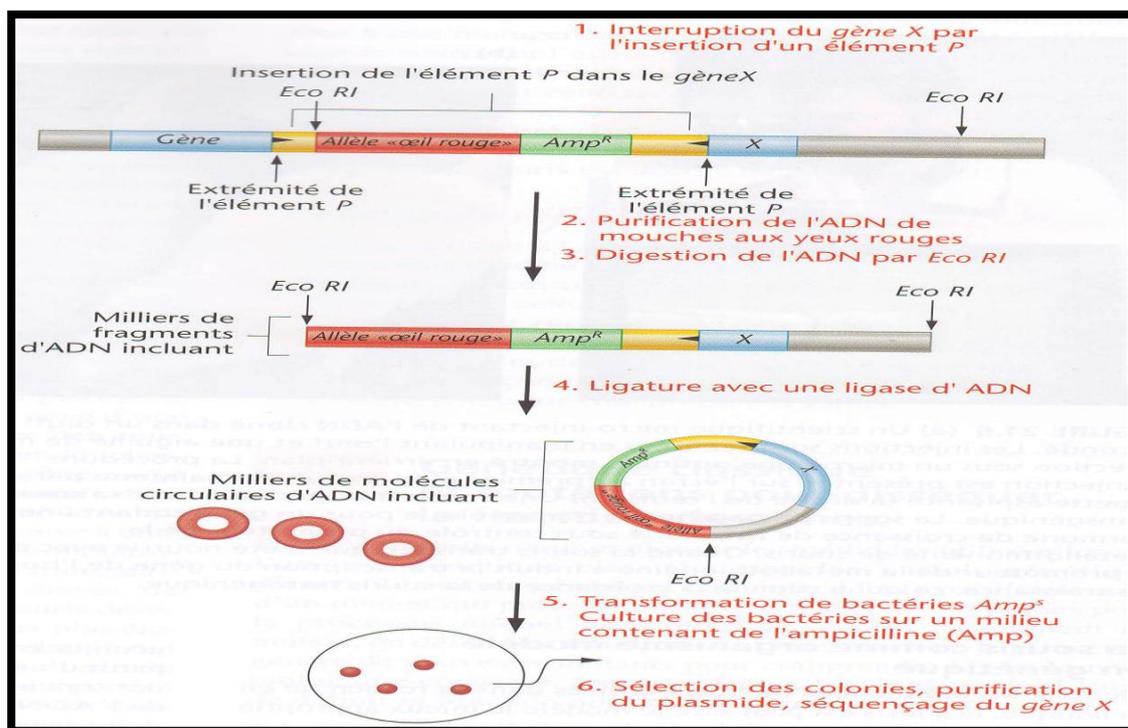


Figure 6

Le « Berkeley Drosophila Genome Project » a collecté des souches de drosophiles portant des insertions d'éléments *P* dans environ 25% des 13000 gènes de la drosophile. Le but de ce projet est de générer des mutations nulles dans chaque gène essentiel, produisant ainsi un « kit de gènes inactivés ».

I.5.3. La souris comme organisme modèle en génétique

De tous les organismes modèles utilisés dans la recherche en génétique, la souris est peut-être le modèle le mieux approprié et le plus accessible pour l'étude de maladies humaines. Les souris ont des temps de génération relativement courts, de huit à neuf semaines. Elles peuvent être facilement maintenues dans des conditions de laboratoire et ont au moins huit descendants par croisement. De plus, les souris et les humains ont non seulement des plans d'organisation similaires mais effectuent aussi des étapes de développement embryonnaire similaires.

Les génomes murin et humain ont approximativement la même taille (environ 3 milliards de paires de bases d'ADN) avec des nombres similaires de chromosomes (20 paires chez la souris et 23 paires chez l'homme) et la plupart des gènes humains ont des orthologues chez la souris. Ainsi, il est particulièrement significatif de noter que les gènes qui sont liés sur le même chromosome chez l'homme sont aussi souvent liés chez la souris. Les généticiens utilisent cette similarité dans l'organisation des gènes pour identifier et localiser des gènes dans une espèce, une fois que les gènes ont été localisés dans l'autre espèce.

Les gènes qui présentent des similarités de séquence entre l'homme et la souris contrôlent souvent, mais pas toujours, les mêmes processus biologiques. Puisque les séquences des exons sont généralement très bien conservées entre la souris et l'homme, les sondes préparées à partir des gènes clonés dans une espèce peuvent souvent être utilisées pour détecter et cloner les gènes correspondants dans l'autre espèce.

En dépit des nombreux avantages que présentent les souris pour l'expérimentation, elles demeurent plus difficiles à élever, à croiser et à muter que des eucaryotes moins complexes tels que la levure ou la drosophile. On ne peut pas facilement réaliser des criblages génétiques à grande échelle chez la souris pour identifier des gènes d'intérêt. Cependant, chez la souris, des gènes spécifiques peuvent être insérés, délétés ou soumis à une invalidation ciblée. Globalement, ceci fait de la souris un système utile pour déterminer la fonction et la régulation de gènes spécifiques.

II. Mutagenèse et quelques exemples d'étude de la fonction des protéines

II.1. Création des mutants

II.1.1. Radiations ionisantes :

Cassures chromosomiques, délétions, translocations, autres réarrangements majeurs (mutations de perte de fonction)

II.1.2. UV + agents chimiques (EMS) :

Mutations ponctuelles+ délétions (mutants conditionnels)

II.1.3. Transposons : (élément P)

Mutations de perte de fonction

II.2. Différentes classifications des mutations

II.2.1. Classification en fonction de la localisation

- **Maladies de transmission autosomique dominante**

Les maladies génétiques humaines qui résultent d'une mutation dans un gène spécifique présentent plusieurs modes de transmission selon la nature et la position chromosomique des allèles responsables. L'un des modes de transmission caractéristique est présenté par un allèle dominant présent dans un autosome. Un allèle autosomique dominant est exprimé chez un hétérozygote ; c'est pourquoi en général au moins l'un des parents d'un individu affecté est lui aussi atteint de la maladie.

La chorée de Huntington (HD pour Huntington's disease) est un exemple de maladie autosomique dominante. Si l'un ou l'autre parent porte un allèle mutant HD, chacun de ses enfants (indépendamment du sexe) a un risque de 50% de recevoir l'allèle mutant et d'être affecté par la maladie

- **Maladies de transmission autosomique récessive**

Un allèle récessif dans un autosome présente un mode de ségrégation très différent. Dans le cas d'un allèle autosomique récessif, les deux parents doivent être porteurs hétérozygotes de l'allèle pour que leur enfant ait un risque d'être affecté par la maladie. Chaque enfant de parent hétérozygote a une probabilité de 25% de recevoir les deux allèles récessifs et donc d'être affecté. Une probabilité de 50% de recevoir un allèle normal et un allèle mutant et donc d'être porteur et enfin, une probabilité de 25% de

recevoir les deux allèles normaux. La mucoviscidose est un bon exemple de maladie autosomique récessive.

- Maladies liées au chromosome X

Le troisième type de ségrégation courant est celui d'un allèle récessif lié à l'X. Un allèle récessif sur le chromosome X s'exprimera plus souvent chez les garçons qui ont reçu un seul chromosome X de leur mère, que chez les filles qui reçoivent un chromosome X à la fois de leur père et de leur mère. Ceci conduit à un profil de ségrégation lié au sexe, dans lequel la maladie apparaît le plus souvent chez les garçons que chez les filles. Par exemple, la dystrophie musculaire de Duchenne (DMD) qui affecte spécifiquement les garçons (**Figure 7**).

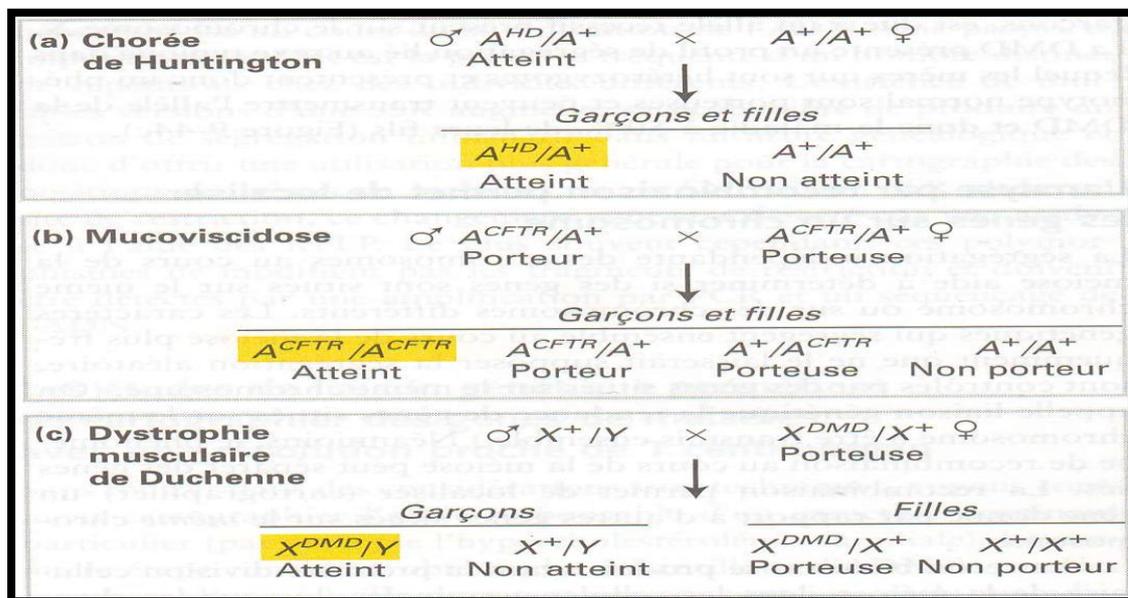


Figure 7

II.2.2. Classification en fonction de la nature du changement moléculaire

Les généticiens classent les mutations en fonction de la nature nucléotidique qui crée la mutation. Une substitution d'une paire de bases en une autre dans la molécule d'ADN est appelée mutation ponctuelle ou substitution nucléotidique.

Le changement d'un nucléotide dans la partie du gène codant la protéine peut conduire à la création d'un nouveau codon codant un autre acide aminé dans la protéine. Il s'agit alors **d'une mutation faux sens**.

Le codon peut aussi être changé en codon stop, ce qui conduira à un arrêt de la traduction de la protéine. Il s'agit alors **d'une mutation non-sens.**

Si la mutation ponctuelle affecte un codon mais ne change pas la nature de l'acide aminé (cela est possible à cause de la dégénérescence du code génétique) on dira qu'il s'agit **d'une mutation silencieuse.**

Deux autres termes sont souvent employés pour décrire les substitutions nucléotidiques. Si une base pyrimidique est remplacée par une autre base pyrimidique ou si une base purique est remplacée par une autre purique, il s'agit d'une **transition.** si une base purique et une base pyrimidique sont interchangées il s'agit d'une **transversion.**

Un autre type de mutation est la délétion ou l'insertion d'un ou plusieurs nucléotides à un endroit quelconque du gène. On les appelle **des mutations de décalage du cadre de lecture** car le cadre de lecture des bases trois par trois est décalé, ce qui altère la traduction. Une mutation de décalage du cadre de lecture surviendra lorsqu'il y aura insertion ou délétion d'un nombre de nucléotides non multiples de trois, puisqu'un multiple de trois nucléotides restaurera le cadre de lecture initial. Ces insertions et délétions peuvent changer tous les codons suivants du gène. Il est possible qu'un des codons altérés soit l'un des trois codons stop UAA, UAG ou UGA ; ce qui conduit à l'arrêt de la synthèse peptidique.

II.2.3. Classification en fonction des effets phénotypiques

Selon leur localisation et leur nature, les mutations peuvent avoir des effets phénotypiques très variables, depuis l'absence d'effet (mutations silencieuses) jusqu'à un effet létal dominant.

Une mutation de perte de fonction supprime la fonction du produit du gène. Toute mutation, depuis la mutation ponctuelle jusqu'à la délétion totale du gène, peut conduire à une perte de fonction. Ces mutations aussi connues sous les noms d'allèles nuls ou d'allèles KO (knock-out).

Une mutation par gain de fonction est due à une fonction nouvellement acquise du produit du gène. Il peut s'agir d'un changement d'acides aminés conduisant à une nouvelle activité de la protéine ou bien d'une mutation de la séquence régulatrice du

gène responsable d'une expression anormale pouvant affecter le lieu, la quantité ou le moment de l'expression.

Une autre catégorie de mutations est constituée par **les mutations ayant un effet nutritionnel ou biochimique**. Une mutation biochimique typique retrouvée chez la levure ou chez les bactéries est l'incapacité à synthétiser un acide aminé ou vitamine. Chez l'homme, la drépanocytose ou l'hémophilie sont des exemples de ces mutations.

Une mutation peut aussi interrompre un processus vital pour l'organisme. On parle alors de **mutation létale**. À titre d'exemple une bactérie mutante ayant perdu la capacité à synthétiser un acide aminé essentiel sera incapable de se développer et donc mourra sur un milieu de culture ne possédant pas cet acide aminé. De nombreuses maladies métaboliques héréditaires humaines sont dues à des mutations létales. La maladie de Tay-Sachs ou la maladie de Huntington sont toutes létales bien que la létalité survienne à des moments différents de la vie.

Des mutations particulièrement intéressantes sont celles dont l'expression dépend de l'environnement de l'organisme. De telles mutations sont appelées **mutations conditionnelles** parce qu'elles sont présentes dans l'organisme mais ne peuvent être détectées que dans certaines conditions. L'un des meilleurs exemples de mutants conditionnels est celui des mutants thermosensibles : à une certaine température, dite « permissive », le produit du gène est fonctionnel, tandis qu'à une température dite « restrictive », il devient non fonctionnel et l'impact de la mutation devient visible

II.3. Criblage des mutants

II.3.1. Les mutations conditionnelles

Les procédures utilisées pour identifier et isoler des mutants, que l'on appelle le criblage génétique, varient suivant l'haploïdie ou la diploïdie d'un organisme et, dans ce dernier cas, suivant la récessivité ou la dominance de la mutation. Les gènes qui codent les protéines essentielles à la vie sont parmi les plus intéressants et les plus importants à étudier. Puisque l'expression phénotypique des mutations dans des gènes essentiels conduit à la mort de l'individu, il faut des criblages génétiques astucieux pour isoler et conserver des organismes porteurs d'une mutation létale.

Dans les cellules haploïdes de levure, on peut étudier les gènes essentiels grâce à l'utilisation de mutations conditionnelles. Les mutations thermosensibles sont parmi les

mutations conditionnelles les plus courantes les chercheurs ont réalisé des expériences sur la levure en suivant le protocole décrit à la figure 5. Une température à laquelle on observe le phénotype mutant est dite non permissive. Une température permissive est une température à laquelle on n'observe pas le phénotype mutant même si l'allèle mutant est présent. On peut donc conserver des souches mutantes à une température permissive puis en réaliser une sous-culture et la placer à une température non permissive afin d'analyser le phénotype mutant.

Pour poursuivre le criblage les chercheurs décidèrent de rechercher des gènes importants pour la régulation du cycle cellulaire au cours duquel une cellule synthétise ses protéines, réplique son ADN puis subit une division cellulaire mitotique, lors de laquelle chaque cellule fille reçoit une copie de chacun des chromosomes.

La croissance exponentielle d'une seule cellule de levure pendant 20 à 30 divisions cellulaires forme une colonie de levure visible sur un milieu solide d'agar. Puisque des mutants présentant un blocage complet de leur cycle cellulaire sont incapables de former une colonie. Pour rechercher ce type de mutants, les chercheurs sélectionnèrent d'abord des cellules de levure ayant subi une mutagenèse, qui pouvaient croître normalement à 23°C mais ne pouvaient former de colonie à 36°C.

Une fois les mutants thermosensibles isolés, d'autres analyses révélèrent que leur division cellulaire était effectivement déficiente. Chez *S. cerevisiae*, la division cellulaire se déroule par le biais d'un processus de bourgeonnement et la taille du bourgeon, que l'on voit bien sous microscope photonique, indique l'étape du cycle cellulaire en cours. Chacun des mutants incapables de croître à 36°C était examiné au microscope après plusieurs heures passées à la température non permissive. L'examen de nombreux mutants thermosensibles différents révéla qu'ils présentaient un blocage effectif du cycle cellulaire. On les appela donc mutants cdc (cycle de division cellulaire).

Les cellules de levure possédant une mutation thermosensible dans un gène contrôlant une étape du cycle cellulaire de se diviser avec un même phénotype, caractéristique d'une étape du cycle cellulaire (**Figure 8**). Par exemple à 36°C, toutes les cellules bloquées à la transition G2/M présentaient un gros bourgeon et un noyau à la jonction entre la cellule mère et la cellule fille.

En utilisant ces méthodes les chercheurs ont identifié environ 150 mutants thermosensibles présentant des étapes du cycle cellulaire (cdc).

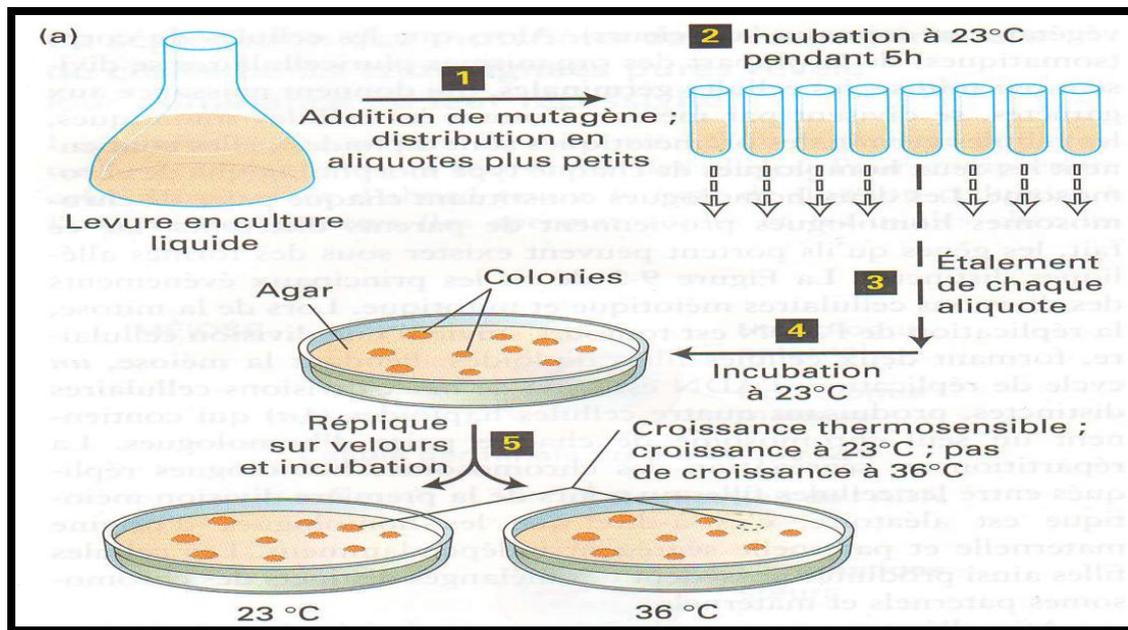


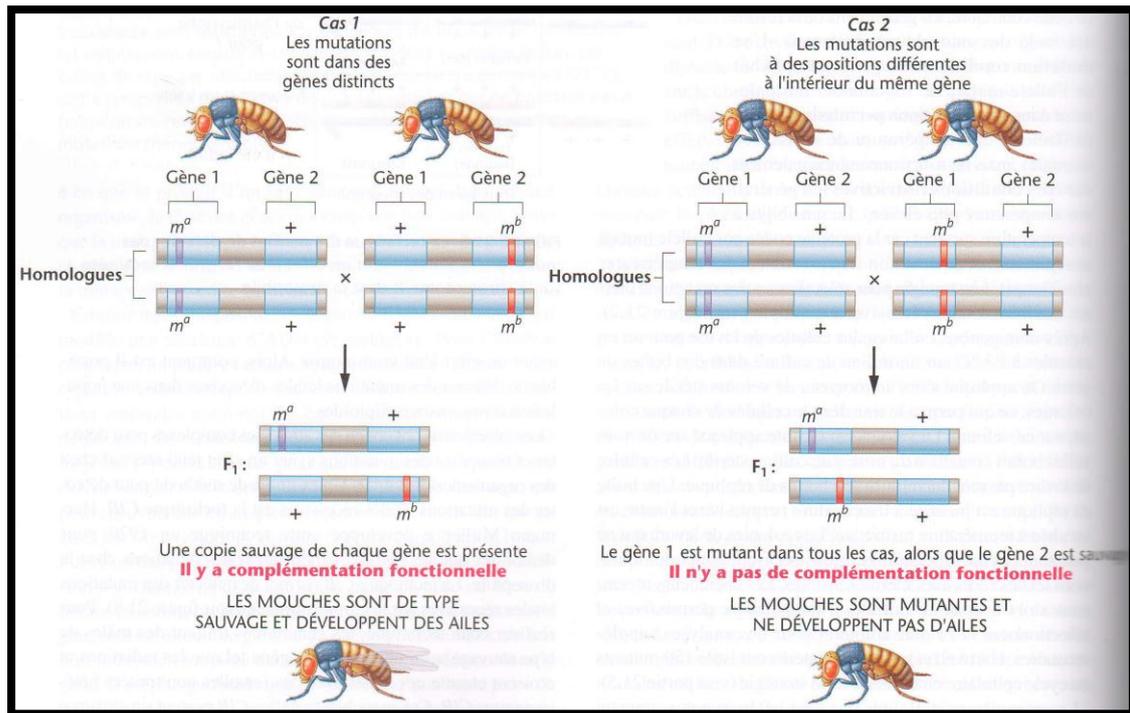
Figure 8

II.3.3. Définir les gènes par le test de complémentation

Un test de complémentation permet aux chercheurs de déterminer si deux mutations touchent le même gène - c'est-à-dire de définir s'il s'agit de deux allèles du même gène ou bien si elles correspondent à des mutations dans des gènes distincts.

Pour réaliser un test de complémentation, les chercheurs croisent simplement deux homozygotes mutantes entre elles et examinent la descendance F1. Il y a deux résultats et interprétations différentes pour un tel croisement : si toute la descendance est de type sauvage. Dans ce cas, il y a eu complémentation ; par conséquent, les deux mutations doivent toucher des gènes distincts. La complémentation a lieu quand la descendance F1 est hétérozygote pour chaque locus et que l'allèle sauvage à chaque locus dirige la synthèse d'une protéine sauvage qui est nécessaire pour obtenir le phénotype sauvage.

L'autre résultat possible est que toute la descendance F1 soit mutante. Dans ce cas, il n'y a pas eu complémentation ; par conséquent, les deux mutations se trouvent dans le même gène. Il n'y a pas eu de complémentation quand les deux allèles du gène sont mutants et qu'aucun produit sauvage du gène ne peut être synthétisé (**Figure 9**).

**Figure 9**

L'analyse par complémentation d'un groupe de mutants de même phénotype permet de distinguer les différents gènes dans un ensemble de gènes qui doivent tous être fonctionnels pour produire un caractère phénotypique donné.

Par exemple, le criblage de mutations *cdc* chez *Saccharomyces* décrit ci-dessus a produit de nombreux mutants thermosensibles récessifs qui apparaissaient avec une interruption au même stade du cycle cellulaire (**Figure 10**). Pour déterminer le nombre de gènes affectés par ces mutations, les chercheurs réalisèrent des tests de complémentation sur toutes les combinaisons deux à deux de mutants *cdc* à l'aide du protocole général décrit dans la figure 8. Ces tests permirent d'identifier plus de 20 gènes *CDC* différents.

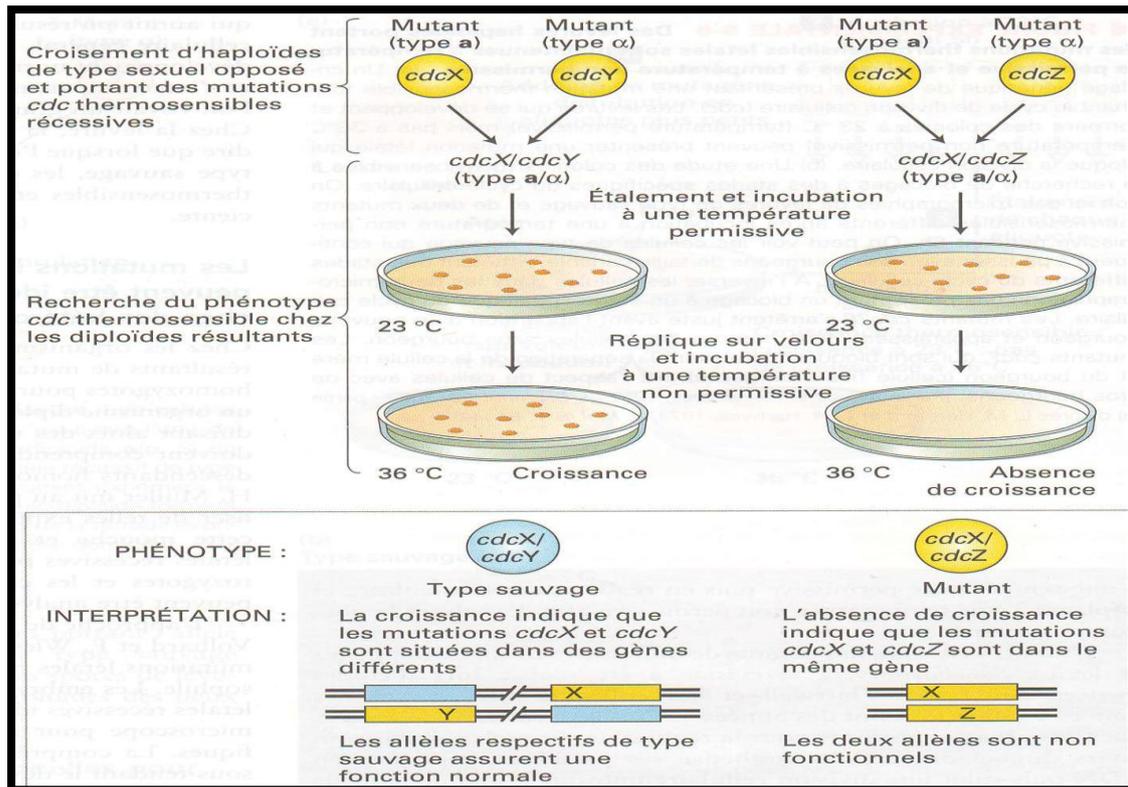


Figure 10

III. Invalidation des gènes dans différents organismes

Trois approches élémentaires sous-tendent ces techniques d'inactivation (ou invalidation) de gènes :

- (1) Le remplacement d'un gène normal par d'autres séquences ;
- (2) L'introduction d'un allèle dont la protéine correspondante inhibe le fonctionnement de la protéine exprimée normalement
- (3) La destruction de l'ARN exprimé à partir d'un gène.

Le gène endogène normal est modifié dans les techniques basées sur la première approche mais reste inchangé dans les autres approches.

III.1. Le remplacement par recombinaison génétique des gènes normaux de levure par des allèles mutants

Modifier le génome de la levure *Saccharomyces* est particulièrement facile pour deux raisons :

- Les cellules de levure absorbent facilement l'ADN exogène dans certaines conditions,

- L'ADN introduit est échangé efficacement contre le site chromosomique homologue dans la cellule receveuse.

Cette recombinaison ciblée spécifique de segments identiques d'ADN permet de remplacer par un allèle mutant, n'importe quel gène dans les chromosomes de levure (la recombinaison entre chromosomes homologues se produit naturellement au cours de la méiose).

On explique l'une des méthodes courantes utilisées pour modifier les gènes de levure : la PCR est utilisée pour fabriquer une construction inactivante contenant le marqueur de sélection (comme le gène *kanMX*, qui confère comme *neo^r* la résistance au G-418) encadré par environ 20 paires de bases complémentaires des extrémités du gène cible de levure (**Figure 11**).

Les cellules diploïdes transformées de levure, dans lesquelles l'une des deux copies du gène endogène cible a été remplacée par la construction inactivante, sont identifiées par leur résistance au G-418. Ces cellules diploïdes hétérozygotes de levure poussent normalement indépendamment de la fonction du gène cible, mais la moitié des spores haploïdes issues de ces cellules porteront uniquement l'allèle inactivé. Si un gène est essentiel à la viabilité, alors les spores portant un allèle inactivé ne survivront pas.

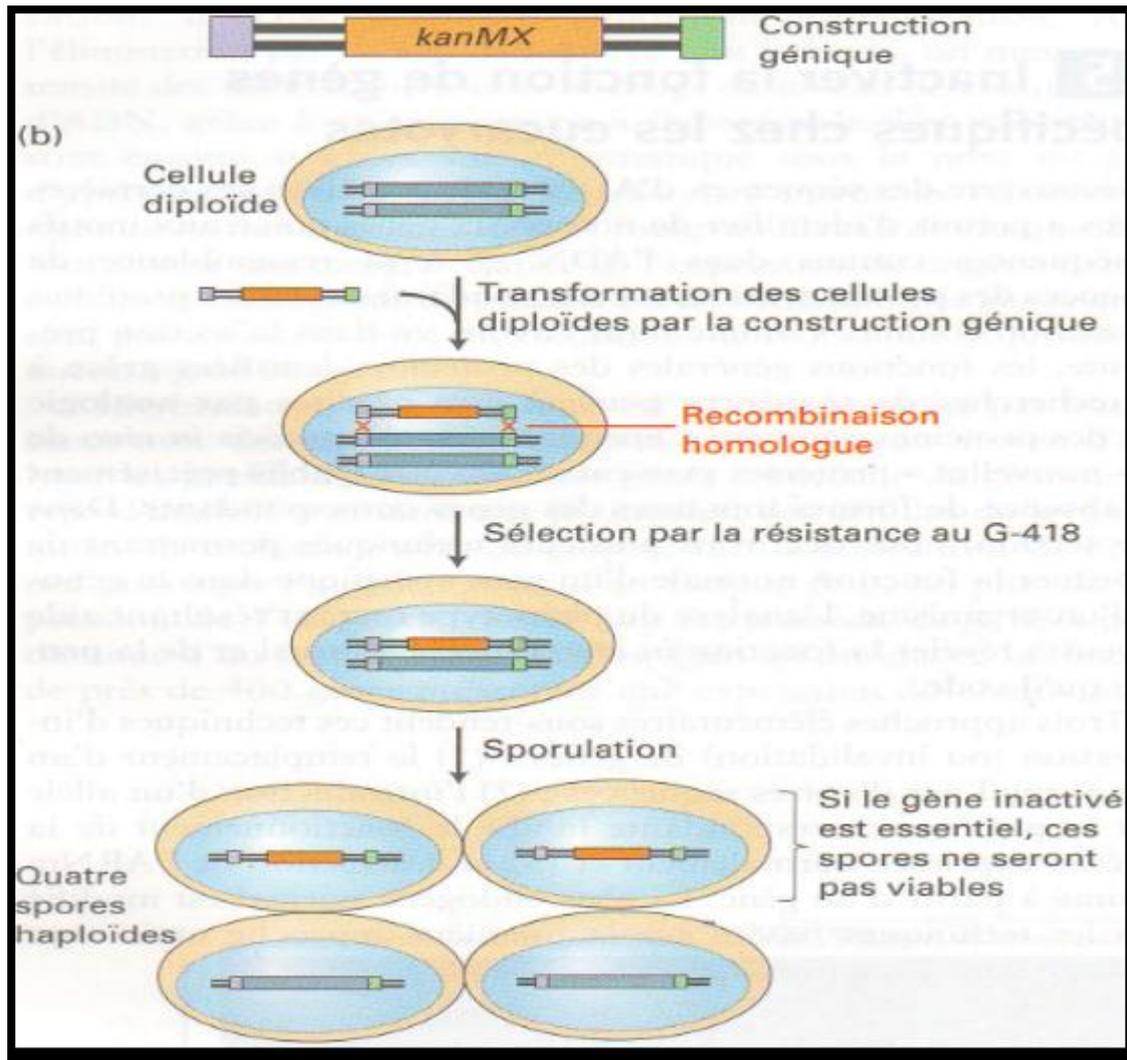


Figure 11

III.2. L'inactivation des gènes spécifiques dans la lignée germinale d'une souris

Un grand nombre des techniques d'inactivation des gènes chez la levure peuvent être appliquées aux gènes d'eucaryotes supérieurs.

Ces gènes peuvent être introduits dans la lignée germinale par recombinaison homologue, afin de produire des animaux avec une inactivation de gènes (knockout) ou animaux KO. Les souris KO chez lesquelles un gène spécifique est inactivé constituent un système expérimental performant pour étudier le développement des mammifères, leur comportement et leur physiologie. Elles sont également précieuses pour étudier l'origine moléculaire de certaines maladies génétiques humaines.

Les souris KO pour un gène déterminé (inactivation ciblée) sont créées en deux étapes. Tout d'abord une construction d'ADN contenant un allèle inactivé d'un gène cible particulier est introduite dans des cellules de la lignée embryonnaire (Embryonic

Stem) ou cellules ES. Ces cellules, issues du blastocyste, peuvent être mises en culture pendant de nombreuses générations. Dans une petite fraction des cellules transfectées, l'ADN introduit subit une recombinaison homologue avec le gène cible, même si les recombinaisons au niveau de sites chromosomiques non homologues sont les plus fréquentes. Pour sélectionner des cellules ayant subi une insertion homologue ciblée, la construction d'ADN recombinant introduite dans les cellules ES doit comprendre deux gènes marqueurs sélectionnables.

L'un de ces gènes (*neo^r*) qui confère la résistance au G-418, est inséré dans le gène cible (X), ce qui l'inactive. L'autre gène sélectionnable, le gène de la thymidine kinase provenant du virus herpes simplex (*tk^{HSV}*) qui phosphoryle sélectivement des analogues de nucléosides tels que le ganciclovir. Il est inséré dans la construction hors de la séquence du gène cible ; et si une cellule exprimant le gène *tk* est mise en culture en présence de ganciclovir, l'analogue de nucléoside phosphorylé devient un inhibiteur toxique de la réplication de l'ADN cellulaire. Seules les cellules ES ayant subi une insertion homologue peuvent survivre en présence de G-418 et de ganciclovir. Dans ces cellules, un allèle du gène X est inactivé (**Figure 12**).

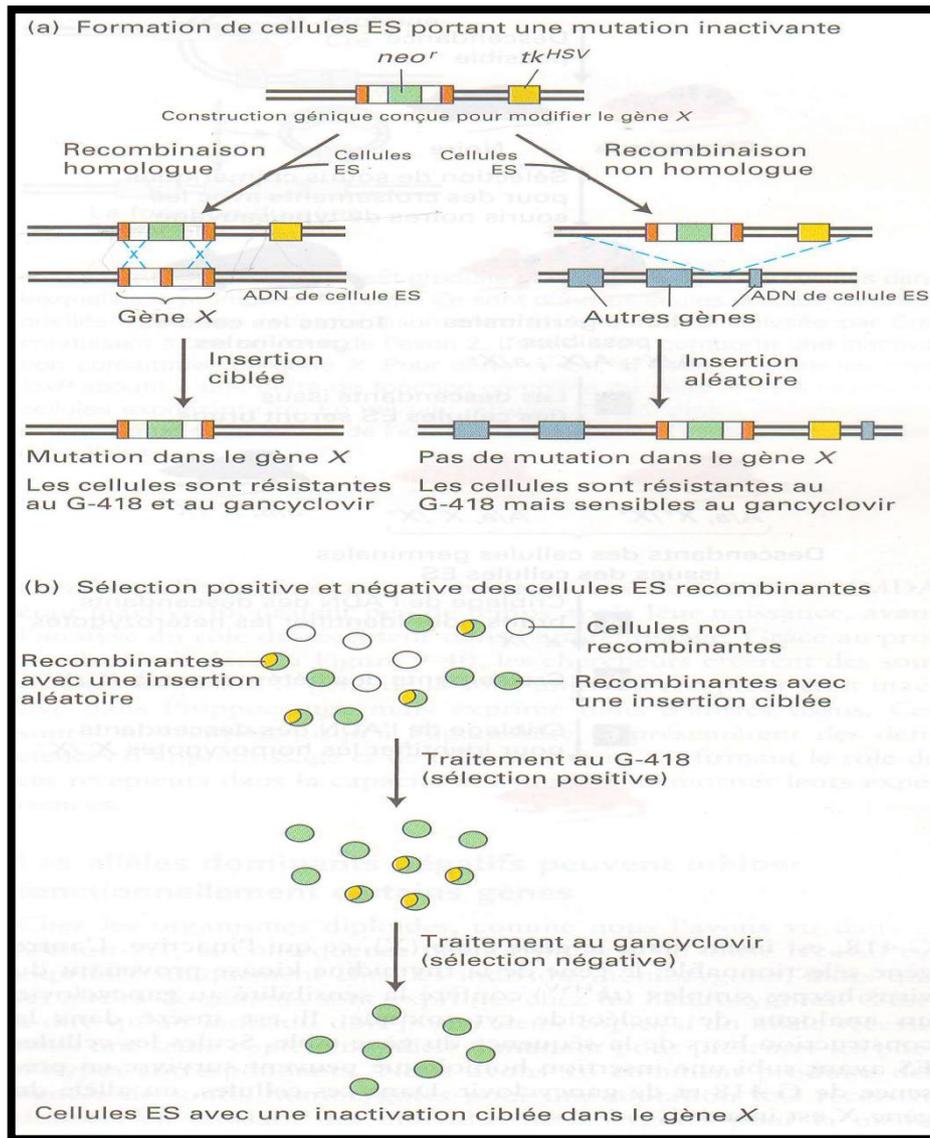


Figure 12

Lors de la deuxième étape de la création de souris KO, les cellules ES hétérozygotes pour une mutation inactivante dans le gène X sont injectées dans un blastocyste receveur d'une souris de type sauvage qui est ensuite transféré dans une souris femelle porteuse. Les descendants résultants sont des chimères (ou mosaïques). Ils contiennent des tissus dérivés des cellules ES transplantées et des cellules de l'hôte. Si les cellules ES sont également homozygotes pour un caractère marqueur visible (comme la couleur du pelage), alors les descendants chimériques chez lesquels les cellules ES ont survécu et proliféré sont facilement identifiables. Les souris chimériques sont ensuite croisées avec des souris homozygotes pour un autre allèle du

caractère marqueur afin de déterminer si la mutation inactivante est incorporée dans la lignée germinale. Enfin, le croisement de souris hétérozygotes pour l'allèle inactivé, produit des descendants homozygotes pour la mutation inactivante (**Figure 13**).

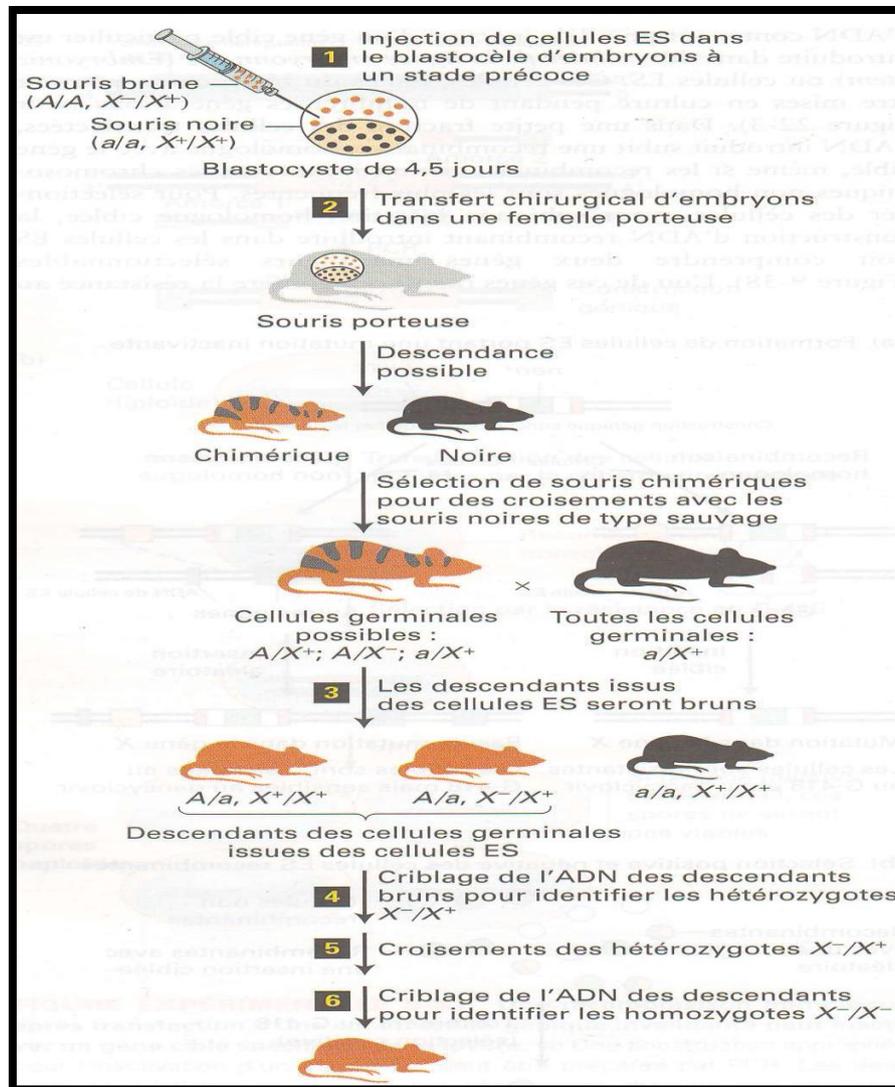


Figure 13

III.3. L'inhibition fonctionnelle de certains gènes par des allèles dominants négatifs

Pour certains gènes, on peut contourner les difficultés pour produire des mutants homozygotes KO en utilisant un allèle portant une mutation dominante négative. Ces allèles sont génétiquement dominants, c'est-à-dire qu'ils produisent un phénotype mutant même dans des cellules portant une copie sauvage du gène.

Mais, au contraire des autres types d'allèles dominants, les allèles dominants négatifs produisent un phénotype équivalent à celui d'une mutation perte de fonction.

Des allèles dominants négatifs utiles ont été identifiés pour différents gènes et peuvent être introduits dans des cellules en culture par transfection ou dans la lignée germinale des souris ou d'autres organismes. Dans les deux cas, le gène introduit est intégré dans le génome par recombinaison non homologue.

Les transgènes portant un allèle dominant négatif sont généralement modifiés génétiquement de sorte que l'allèle est contrôlé par un promoteur régulé, ce qui permet l'expression d'une protéine mutante dans différents tissus et à des moments variés. Comme nous l'avons vu plus haut, l'intégration aléatoire d'ADN exogène par recombinaison non homologue a lieu à une fréquence nettement supérieure à l'insertion par recombinaison homologue. En raison de ce phénomène, la production de souris transgéniques est un processus efficace et direct (**Figure 14**).

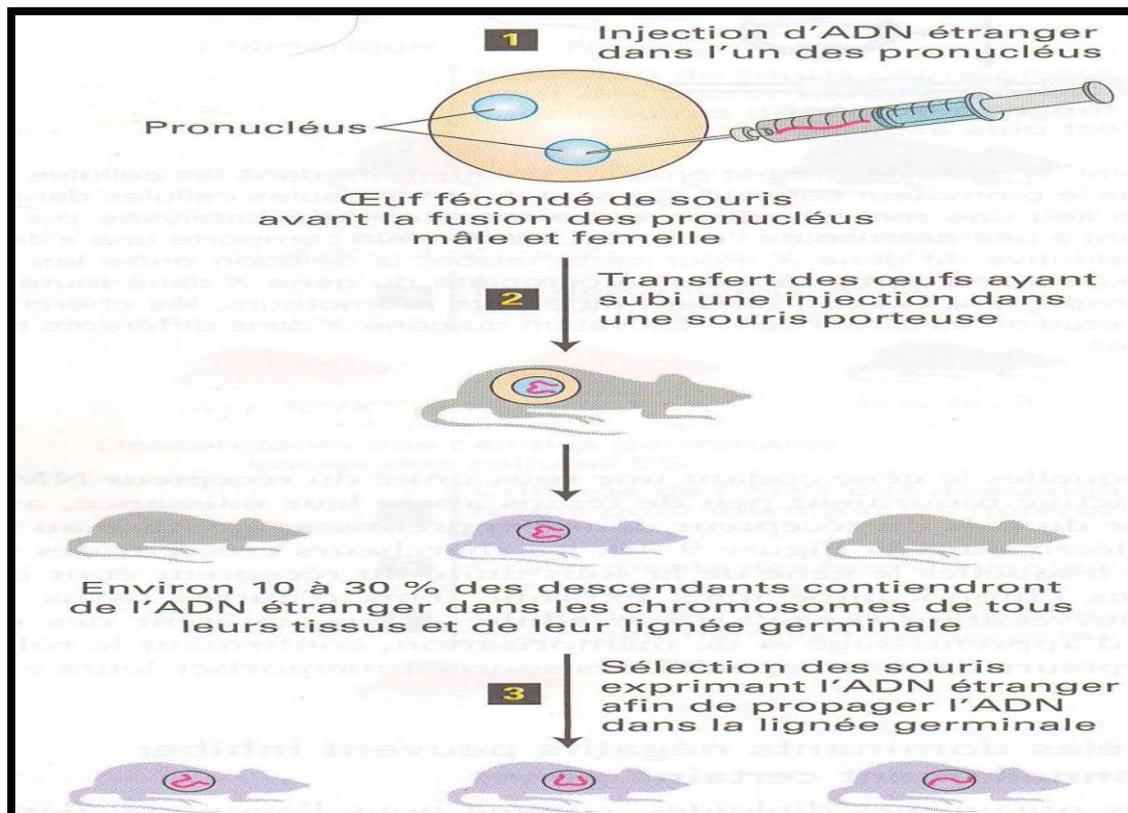


Figure 14

Parmi les gènes susceptibles d'être inactivés fonctionnellement par l'introduction d'un allèle dominant négatif, on trouve ceux qui codent les petites protéines (monomériques)

qui fixent le GTP et appartiennent à la superfamille des GTPases. Ces protéines (comme Ras, Rac et Rab) jouent le rôle de commutateurs intracellulaires. La conversion des petites GTPases de l'état inactif lié à du GDP vers l'état actif lié à du GTP dépend de leur interaction avec une protéine équivalente, le facteur d'échange des nucléotides guanine (GEF). Une petite GTPase mutante qui se fixe définitivement à la protéine GEF bloquera la conversion des petites GTPases endogènes de type sauvage à l'état actif lié au GTP, les empêchant ainsi de jouer le rôle de commutateurs (**Figure 15**).

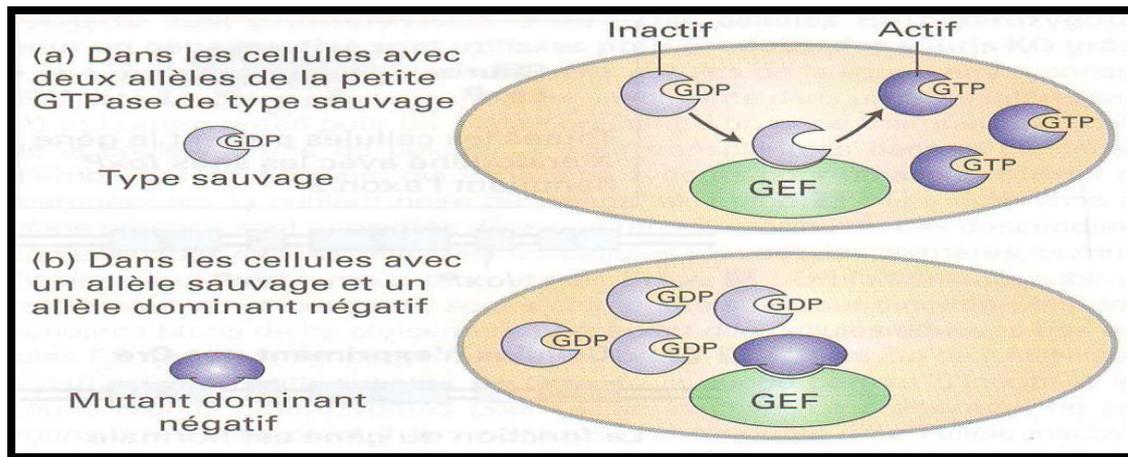


Figure 15

III.4. L'inhibition par les ARNi

Les chercheurs exploitent un phénomène appelé interférence d'ARN (ARNi) afin d'inhiber la fonction de gènes spécifiques. Cette approche est techniquement plus simple que les méthodes décrites ci-dessus pour inactiver les gènes. Cette technique peut également être utile pour la thérapie génique où des gènes responsables de maladies peuvent être inactivés. Avec cette approche, un avantage précieux est que seule une partie de la séquence du gène à inactiver doit être connue.

Dans la nature, les cellules utilisent un mécanisme de régulation génique par extinction de gènes dépendante des ARN pour se défendre contre les virus ou les transposons invasifs : le processus commence avec un double-brin d'ARN (environ 70 nucléotides) qui est dégradé par une protéine (appelée Dicer) qui présente une activité RNase double-brin. Dicer se lie aux molécules d'ARN double brins et les coupe en petites molécules d'environ 21 nucléotides appelées ARNsi (small/ interfering RNA). Les ARNsi se lient à un complexe multiprotéique, le RISC (RNA-induced silencing complex), qui sépare deux brins d'ARN pour former des molécules d'ARN capables de

se fixer sur des ARNm présentant des régions complémentaires, réalisant ainsi un marquage en vue de leur dégradation (**Figure16 au milieu**)

Dans le cytoplasme, il existe d'autres systèmes utilisés pour l'extinction de gènes :

- Un précurseur partiellement double brin est digéré par Dicer pour générer des microARN (ARNmi) qui se lient aux régions complémentaires 3' non traduites des ARNm, inhibant leur traduction (**Figure16 à droite**).

- Les petits ARNsi, dégradés par Dicer jouent un rôle dans la méthylation de l'ADN (RdDM RNA-directed DNA Methylation). Ces ARN interagissent avec des ADN méthyltransférases (DMTases) pour méthyler les résidus cytosines des régions promotrices (cercles violets) et éteindre les gènes (**Figure16 à gauche**).

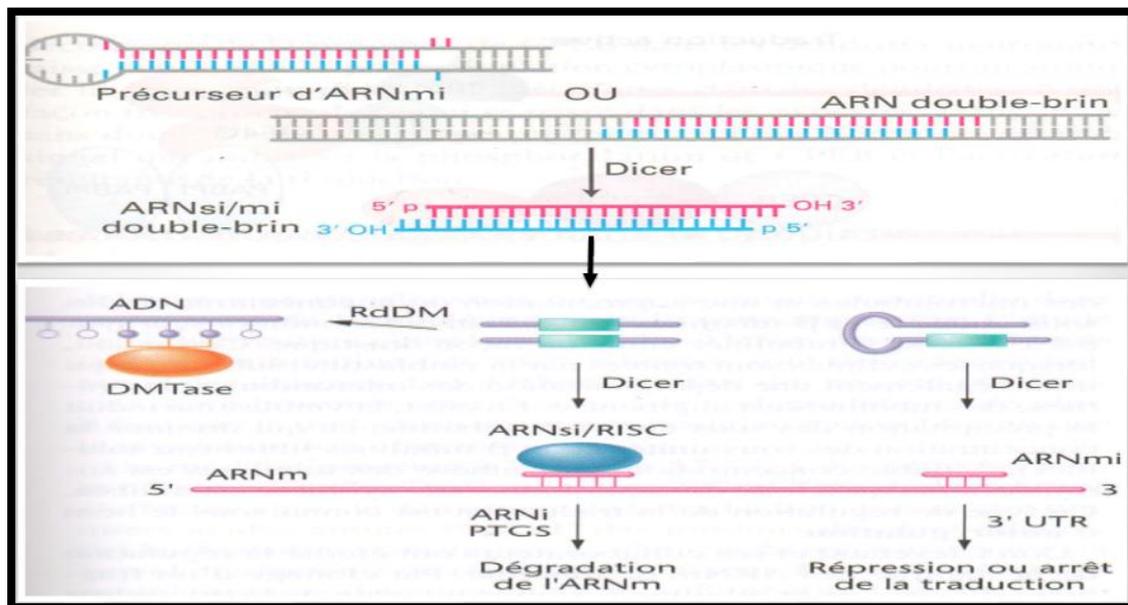


Figure 16

Afin d'utiliser l'ARNi comme outil d'analyse, les chercheurs introduisent dans les cellules de petites molécules d'ARN double-brin synthétiques (**Figure 17**). Ces molécules induisent la même voie de dégradation de l'ARN que celle qui est induite par les ARN double-brin viraux ou de transposons. Le brin antisens de l'ARN double-brin s'hybride à l'ARNm complémentaire dans les cellules, conduisant à la dégradation de cet ARNm.

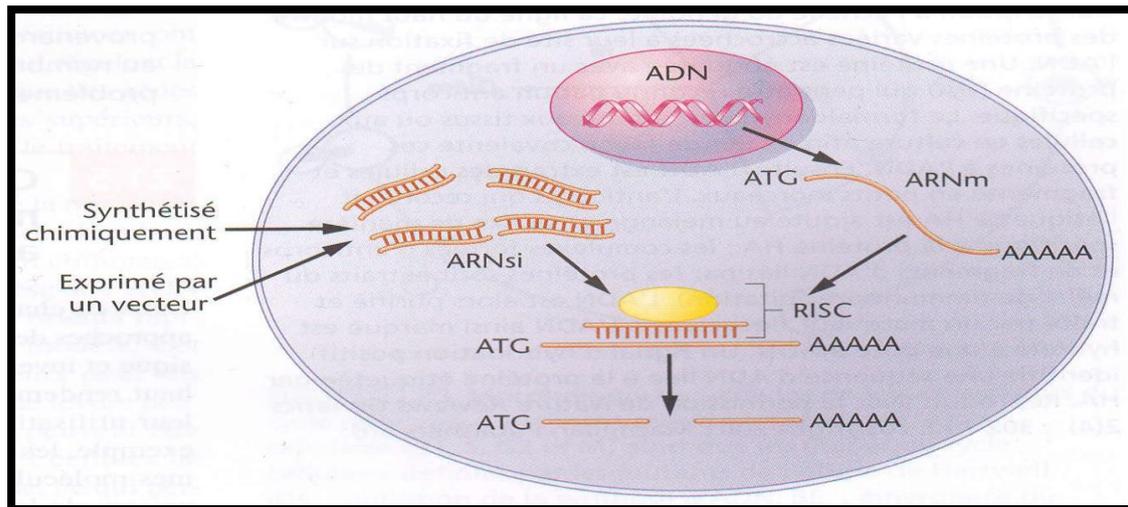


Figure 17

Une autre approche consiste à transformer des cellules par des vecteurs d'ADN qui expriment des ARN en épingle à cheveux, qui sont partiellement double-brin. Si ces vecteurs s'intègrent dans le génome d'une cellule, toute la descendance de cette cellule exprimera les molécules d'ARNi et un gène spécifique sera éteint dans ces cellules. L'inactivation peut être due à la dégradation de l'ARN ou à l'inhibition de la transcription ou de la traduction.

III.5. CRISPR-Cas9 et nucléases guidées par ARN

Le système CRISPR/Cas (CRISPR : clustered regularly interspaced short palindromic repeats) est un mécanisme d'immunité adaptative décrit chez des procaryotes (il existe plus de trois systèmes différents et c'est le CRISPR/CRISPR-associated Cas system Type II qui est utilisé). Ce mécanisme repose sur l'expression du locus CRISPR du génome bactérien. Ce locus est constitué de l'insertion de séquences d'ADN provenant du génome d'agents « infectieux » des bactéries et de motifs d'ADN spécialisés. La transcription du locus CRISPR permet la synthèse des CRISPR-RNA ou crRNA, composés d'une séquence de ciblage du génome de l'agent infectieux et de motifs spécifiques. Le crRNA est associé à une seconde molécule, le TracrRNA qui s'hybride au crRNA. Ces deux molécules associées sont reconnues par l'endonuclease bactérienne Cas9 pour former un complexe qui interagit et clive l'ADN double brin du génome de l'agent lors d'une infection. C'est la complémentarité de la séquence du crRNA avec la séquence de l'agent qui donne la spécificité de coupure de la Cas9 qui contribue ainsi à éliminer l'agent infectieux. Cette interaction requiert des

motifs particuliers de la séquence cible, dont la séquence PAM (Protospacer adjacent Motif) et les éléments d'espacements du crRNA (spacer) pour assurer un fonctionnement optimal du complexe. Le fonctionnement du locus CRISPR est assimilé à un système de protection cible contre certains agents infectieux.

Les technologies d'ingénierie moléculaire utilisant ce système reposent sur la possibilité de contrôler le ciblage et le fonctionnement de ces composants et de les réunir dans la cellule dont on souhaite modifier le génome. Le crRNA est ainsi modifié pour cibler le complexe vers une séquence génétique choisie par l'utilisateur et exprime sous la forme ARN dit guide (sgRNA, small guide RNA) rassemblant les motifs du TracrRNA et du crRNA en une seule molécule. Techniquement, il est nécessaire que tous les éléments décrits ci-dessus soient présents concomitamment dans la cellule et l'introduction de cet ARN guide est effectué en parallèle avec le transfert d'une unité d'expression de la protéine Cas9. Il existe des outils informatiques permettant le dessin d'ARN guides à disposition de la communauté scientifique (**Figure 18**).

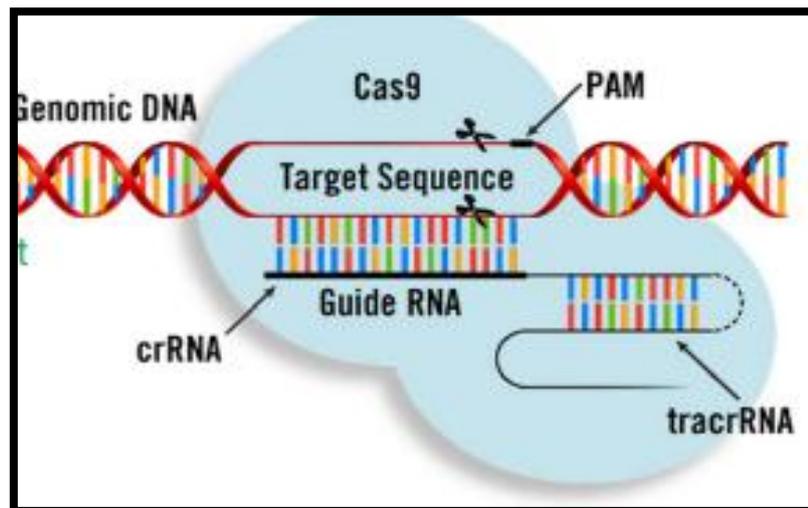


Figure 18

Dans cette stratégie, l'endonucléase Cas9 recrutée vers cette séquence coupe le double brin d'ADN (DSB, Double Strand Break) et active le système de réparation d'ADN de la cellule. Les mécanismes de réparation par « collage » des extrémités d'ADN (NHEJ) permettent l'inactivation de gènes.

En présence d'une séquence matrice, montrant une forte similarité de séquence avec la séquence coupée, une recombinaison homologue (HR) peut entraîner la substitution de la séquence coupée par la séquence matrice. Ceci permet d'introduire des changements

précis dans la séquence (modification dirigée de la séquence d'un gène) ou même d'insérer des fragments d'ADN hétérologues (transgène ciblée) (**Figure 19**).

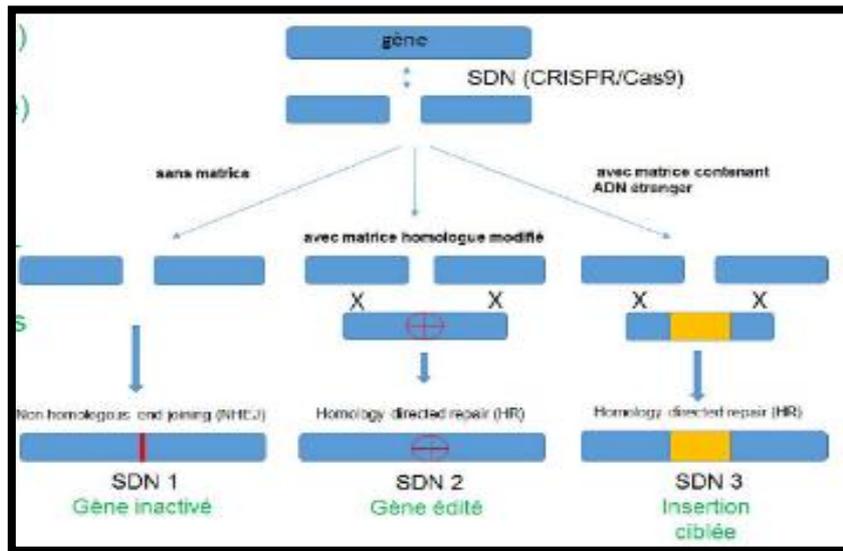


Figure 19

Lors de l'identification du gène cible, il est nécessaire de vérifier l'existence de PAM dans sa séquence. La séquence PAM est absolument nécessaire à la Cas9 pour se lier à l'ADN. Il est donc nécessaire d'identifier toutes les séquences PAM au sein de la zone à cibler. Si aucune séquence PAM n'est présente dans cette zone, on peut se tourner vers une Cas9 d'une espèce différente ou provenant d'un variant de la même espèce. Dès que toutes les séquences PAM potentielles ont été identifiées, c'est le moment de choisir quel site sera à même de permettre le clivage le plus spécifique. Le gARN doit être complémentaire de la séquence cible, mais il est aussi obligatoire de vérifier qu'une séquence identique n'existe pas, quelque part, ailleurs dans le génome.

V. Exemples d'application :

V.1. Modification des composants naturels du lait

V.1.1. Déficit enzymatique et intolérance au lactose

Le lait est un aliment riche en macro et micronutriments. Le lactose est le sucre majeur du lait (5 g/100 ml de lait de vache) et joue le rôle de pompe osmotique dans la glande mammaire pour fluidifier le lait. Pour être digéré, le lactose doit être hydrolysé en glucose et en galactose par la lactase, β -galactosidase digestive présente exclusivement dans l'intestin grêle chez les mammifères.

L'activité lactasique est importante à la naissance, à l'exception des cas rares d'alactasie congénitale. Elle se maintient à un niveau élevé chez 30 % des individus adultes, mais, chez 70 % d'entre eux, chute vers l'âge de 2-5 ans et atteint chez l'adulte une valeur 10 à 20 fois plus faible que celle du nourrisson. Le déclin d'activité lactasique définit le phénotype d'hypolactasie et entraîne une maldigestion du lactose. Chez les individus conservant une activité lactasique élevée à l'âge adulte, une hypolactasie transitoire peut se déclarer lors de maladies digestives.

Le développement de méthodes permettant de réduire la teneur en lactose. Elles sont de deux types :

- D'une part, l'utilisation de préparations orales (liquides ou tablettes) de β -galactosidase de levure consommées en même temps que le lait ;
- D'autre part, le traitement industriel du lait qui élimine le lactose par l'action d'une β -galactosidase ou par ultracentrifugation. Toutefois, si ces produits sont bien tolérés par les individus hypolactasiques, leur distribution reste limitée en raison de leur coût généralement élevé.

Les contraintes liées au traitement industriel du lait ont conduit à rechercher des méthodes alternatives de production in vivo de lait appauvri en lactose, en utilisant la technologie de la transgénèse chez la souris. Une façon d'y parvenir consiste à empêcher la synthèse de lactose dans les cellules mammaires en invalidant le gène de la sous-unité régulatrice de la lactose synthétase : l' α -lactalbumine. Les souris déficientes en alpha-lactalbumine ont été créées en perturbant le gène par recombinaison homologue dans des cellules souches embryonnaires. Les souris mutantes homozygotes sont viables et fertiles, mais les femelles ne peuvent pas nourrir leur progéniture. Ils produisent un lait très visqueux que les chiots semblent être incapables de retirer de la glande mammaire. Ce lait est riche en graisse et en protéines et est dépourvu d'alpha-lactalbumine et de lactose. Le phénotype des souris hétérozygotes a été trouvé intermédiaire, avec une diminution de 40% de l'alpha-lactalbumine, mais seulement une diminution de 10-20% de la teneur en lactose de leur lait par rapport aux animaux de type sauvage.

D'autres recherches ont tenté de limiter, sans empêcher totalement, la production d' α -lactalbumine, ce qui assurerait une synthèse réduite de lactose et normaliserait ainsi l'osmolarité du lait. L'objectif de la seconde approche est de diminuer la teneur en

lactose du lait en le « pré-digérant » dans la glande mammaire. Le principe consiste à construire des animaux transgéniques qui sécrètent dans le lait l'enzyme lactase capable d'hydrolyser le lactose. Celle-ci est synthétisée sous la forme d'un précurseur enzymatiquement actif. La synthèse de lactase est placée sous le contrôle d'un promoteur spécifique de la glande mammaire : le promoteur de l' α -lactalbumine.

Chez les souris transgéniques allaitantes, le précurseur de la lactase est effectivement synthétisé puis exporté dans le lait à la surface des globules lipidiques. L'activité lactasique présente dans le lait provoque une diminution spécifique de la teneur en lactose de 50 % à 85 %, sans altérer sa viscosité ni sa composition en protéines et en graisses. Le glucose et le galactose libérés lors de l'hydrolyse du lactose sont largement réabsorbés par les cellules mammaires.

Le lait modifié conserve ses propriétés nutritionnelles et son absorption n'entraîne aucun effet secondaire.

Ces résultats valident la méthode employée et suggèrent qu'elle pourrait être appliquée aux bovins qui produiraient alors *in vivo* un lait appauvri en lactose. En effet, une réduction de 50% à 70% de la teneur en lactose est suffisante pour éliminer les symptômes cliniques que crée la maldigestion du lactose.

V.1.2. Modification de la composition en caséines et qualité de fromage

Une augmentation de la sécrétion de la caséine κ est susceptible de changer la taille et la stabilité des micelles de caséines et d'augmenter ainsi la masse et la qualité de la pâte à fromage.

Dans le même ordre d'idée, un lait enrichi en caséine- β donne une pâte à fromage de meilleure qualité. Le gène de la caséine- β a été isolé et il s'exprime de manière abondante chez des souris. Il peut donc être transféré des chèvres pour tenter de surproduire cette protéine.

Il est également concevable de modifier quelques acides aminés dans certaines caséines pour augmenter leur sensibilité aux protéases et accélérer ainsi la maturation de la pâte à fromage.

V.2. Addition de composants nouveaux dans le lait destiné à la consommation humaine et animale

Une opération plus volontaire peut consister à ajouter dans le lait des protéines qui ne s'y trouvent pas normalement, ces protéines pouvant ou non être naturellement

présentes dans le lait selon les espèces animales. C'est le cas de la lactoferrine humaine qui peut être secrétée dans le lait de vache grâce au transfert du gène correspondant. Cette opération a pour but essentiel d'ajouter au lait de vache une protéine importante du lait humain. La lactoferrine est en effet une protéine transporteuse de fer. Elle peut donc apporter du fer aux nourrissons (ou aux petits animaux) mais aussi réduire le taux de fer circulant. La lactoferrine (comme la transferrine) est en effet secrétée dans un état où elle n'est pas complètement chargée en fer. Elle capte donc avidement le fer libre et réduit ainsi très notablement le développement des bactéries dont la croissance dépend de cet ion métallique. La lactoferrine pourrait donc assurer une protection du tractus digestif des nourrissons.