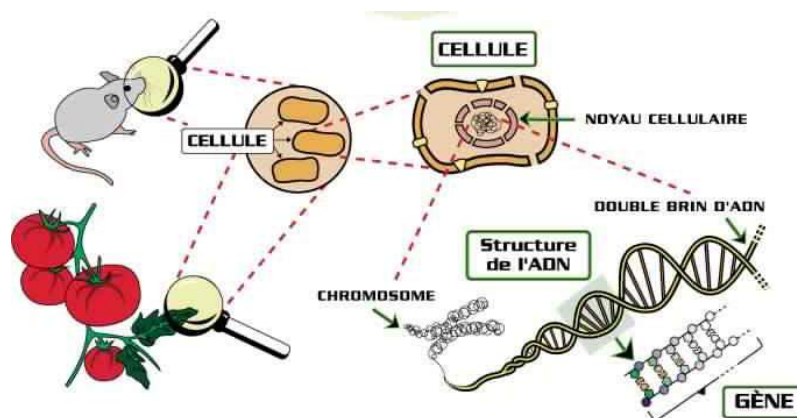


MINISTRERE DE L'ENSEIGNEMENT SUPERIEUR
UNIVERSITE DE LA MANOUBA

INSTITUT SUPERIEUR DE BIOTECHNOLOGIE DE SIDI THABET (ISBST)



SUPPORT DE COURS de
GENIE GENETIQUE



UE : Biotechnologie génique

Master de Recherche en Sciences et Technologies, Mention Sciences du
vivant, Parcours : Biologie moléculaire et santé
Niveau M1

Elaboré par :
Dr. KHOUAJA Fattouma

Année Universitaire 2019/2020

Annexe 2 de la Fiche descriptive de l'UE
Unité d'Enseignement : Biotechnologie génique
Code UE :
ECUE n° 2 Génie génétique

Plan du cours

Objectifs de l'ECUE

Assimiler l'utilité des outils et des techniques du Génie génétique dans le diagnostic et le traitement (thérapie génique) des maladies infectieuses et génétiques et dans le génotypage.

Introduction

Définition du génie génétique : la technologie de l'ADN recombinant

Rappels : les outils et les principales techniques du génie génétique (quelques exemples)

Chapitre I :

- 1- La mutagenèse in vitro
- 2- La cartographie à l'aide des RFLP
- 3- la génétique inverse
- 4- Expression des gènes eucaryotes dans les bactéries par transgénèse
- 5- La technologie de l'ADN recombinant chez les eucaryotes
- 6- La thérapie génique
- 7- Utilisation de l'ADN recombinant pour détecter directement les allèles responsables de maladies

Chapitre II : Les applications de la technologie de l'ADN recombinant

- 1- La transgénèse
- 2- Dépistage et diagnostic génique
- 3- Les applications en biomédecine et en industrie pharmaceutique

Chapitre III : Génotypage et empreinte génétique

- 1- Généralités
- 2- Applications
 - 2-1 Détection des OGM
 - 2-2 Médecine légale (Analyse de l'ADN mitochondrial, Analyse du Chromosome Y)
 - 2-3 Test de paternité par analyse des marqueurs microsatellites
 - 2-4 Applications médicales
 - 2-5 Utilisation des puces à ADN

TP/TD : Application de quelques techniques de base de génie génétique (PCR, RFLP, séquençage) au diagnostic des maladies génétiques

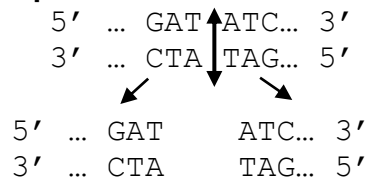
RAPPELS : LES OUTILS DU GENIE GENETIQUE

LES ENZYMES

Modes de coupure de l'ADN par les enzymes de restriction

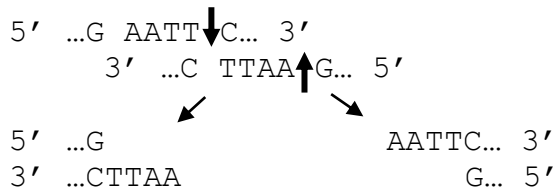
- **Coupure à bouts francs** : coupure au milieu du site de restriction → 2 fragments à extrémités franches.

Exemple : site de restriction de **EcoRV**



- **Coupure à bouts cohésifs (ou collants)** : coupure de part et d'autre du centre de symétrie du site de restriction → 2 fragments à extrémités cohésives (elles vont permettre à toutes les extrémités complémentaires créées par la même enzyme de restriction de venir s'apparier par des liaisons hydrogènes grâce à une enzyme dite ligase).

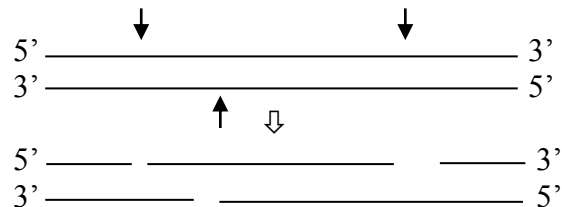
Exemple : site de restriction de **EcoRI**



Les différents types d'enzymes

- **DNase I** = endonucléase (extraite du pancréas de bovins)

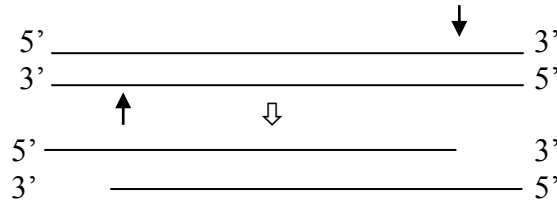
Enzymes de digestion de l'ADN double brin au hasard engendrant des oligonucléotides ayant une extrémité 5'P et une extrémité 3'OH libres. Sur l'ADN natif des chromosomes, la DNase I détruit préférentiellement des gènes actifs c'est-à-dire des gènes en cours de transcription (dépourvus d'histones qui les protègent). Cette enzyme est très utilisée dans les préparations d'ARN ou de protéines (pour éliminer les contaminations par l'ADN) et dans les marquages de sondes avec des radioisotopes.



- **Nucléase S1** : (extraite d'un champignon) n'attaque que l'ADN simple brin

- **Les Exonucléases**

Coupent les nucléotides un par un à partir d'une chaîne polynucléotidique avec une spécificité soit de l'extrémité 5' (coupure 5' → 3'), soit de l'extrémité 3' (coupure 3' → 5'). **L'exonucléase III** par exemple catalyse l'hydrolyse séquentielle des nucléotides de l'ADN à partir de l'extrémité 3'OH libre → obtention de l'ADN simple brin. Elle possède en plus une activité 3' phosphatase.



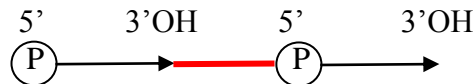
▪ La RNase

Exemples : **RNase A** très spécifique et thermostable qui hydrolyse l'ARN simple brins après les pyrimidines et la **RNase H** qui hydrolyse l'ARN dans les hybrides ARN/ADN

▪ Enzymes de ligature de l'ADN = ligases (Extraites de bactéries)

Assurent la soudure entre 2 fragments d'ADN par établissement de **liaisons phosphodiester** entre une extrémité 3'OH libre et une extrémité 5'P libre en présence d'ATP. Elles peuvent effectuer des ligatures sur des fragments d'ADN à bouts francs ou à bouts cohésifs → la ligase utilisée dépend donc des extrémités de l'ADN. Ainsi :

- la **ligase E. coli** : ne peut lier que 2 fragments d'ADN à extrémité cohésives, alors que
- la **ligase T4 = T4 DNA ligase** : est capable d'assurer la ligature des 2 types d'extrémités.



▪ Enzymes de modification des extrémités de l'ADN

- **Les phosphatases = enzymes de déphosphorylation** (extraites des bactéries ou des animaux)

Utilisées pour préparer de l'ADN recombinant. **Exemple :** la **phosphatase alcaline** (active à pH alcalin) qui retire le groupement phosphate en 5' sur l'ADN et l'ARN. Cette enzyme est surtout utilisée pour déphosphoryler un vecteur qu'on vient d'ouvrir pour éviter sa refermeture.

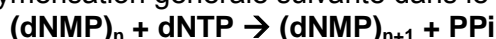
- **Les kinases = enzymes ajoutant des groupements phosphates** (extraites des bactéries)

Permettent de fixer un groupement phosphate en présence d'ATP tel que la **T4 polynucléotide kinase = T4 PK** qui permet le transfert d'un groupement phosphate en position gamma γ à partir d'une molécule d'ATP sur l'hydroxyle 5' d'un ADN préalablement déphosphorylé.

▪ Enzymes de synthèse

Recopient une chaîne d'ADN ou d'ARN dans le sens **5' → 3'** de manière **complémentaire** et **antiparallèle** en présence de désoxynucléosides triphosphates (dNTPs) ou de nucléosides triphosphates (NTPs), respectivement.

① **Les ADN polymérases ADN dépendantes** : synthétisent le nouveau brin d'ADN en présence d'une amorce d'acide nucléique avec une extrémité 3' OH libre. Ces ADN pol catalysent la réaction de polymérisation générale suivante dans le sens 5' → 3' :



Avec N = A, C, T ou G. PPi = groupe pyrophosphate

Exemples :

► **ADN pol I E.coli** (extraite d'*E. coli*), intervient dans la réparation. Donc à part son activité polymérisique 5' → 3', elle possède les 2 activités exonucléasiques (3' → 5' et 5' → 3'). A partir de l'ADN pol I, on a préparé une enzyme appelée **fragment** ou **enzyme de klenow** qui possède uniquement l'activité polymérisique (5' → 3') et l'activité exonucléasique 3' → 5'. Cette dernière permet à l'enzyme, au cours d'une synthèse d'un fragment d'ADN, de

contrôler si l'appariement de la base qui vient d'être ajoutée est conforme aux règles de complémentarité.

► **Séquenase** : (d'origine bactérienne) ne possède pas l'activité exonucléasique $3' \rightarrow 5'$ mais uniquement celle de $5' \rightarrow 3'$. Elle est utilisée dans le séquençage de l'ADN.

► **Taq polymérase** : c'est une ADN polymérase extraite des bactéries vivant dans des sources d'eau chaudes. Elle agit à des températures plus élevées (70°C) que celles usuelles (ambiante ou 37°C) \rightarrow c'est donc une enzyme thermostable utilisée dans les réactions d'amplification *in vitro* de l'ADN par PCR (polymerase chain reaction = réaction de polymérisation en chaîne) et dans les réactions de séquençage de l'ADN. Cette enzyme est en principe dépourvue de l'activité exonucléasique $3' \rightarrow 5'$.

► **Terminal-transférase** : c'est une enzyme extraite du thymus de veau. Elle catalyse l'addition du dNTP à l'extrémité 3'OH de l'ADN \rightarrow elle permet donc de rajouter sur une molécule d'ADN double brin une **queue simple brin = tailing**. Cette enzyme est surtout utilisée pour créer des extrémités collantes.

② **Les ARN polymérases-ADN dépendantes** : (extraites des bactéries), ce sont des enzymes classiques qui réalisent des transcriptions de l'un des 2 brins d'ADN en un brin d'ARN. Elles n'ont pas besoin d'amorce ne peuvent démarrer la transcription que si l'ADN possède le promoteur spécifique de chacune d'elles. Elles nécessitent des NTPs et des ions Mg^{2+} . Elles sont dépourvues d'activité exonucléasique $3' \rightarrow 5'$. Elles sont d'une grande utilité pour la synthèse de l'ARN en grande quantité qui sera utilisé comme sonde (ribosondes) et pour l'étude de la structure de l'ARN et son interaction avec l'ADN et les protéines.

Exemples : Les 3 ARN pol les plus utilisées en génie génétique et en biologie moléculaire : **Sp6 ARN pol** (*Salmonella Thyphimurium*), **T7 ARN pol** (extraite des bactéries *E. coli* infectées par le phage T7) et **T3 ARN pol** (extraite des bactéries *E. coli* infectées par le phage T3).

③ **Les ADN polymérase ARN dépendante**

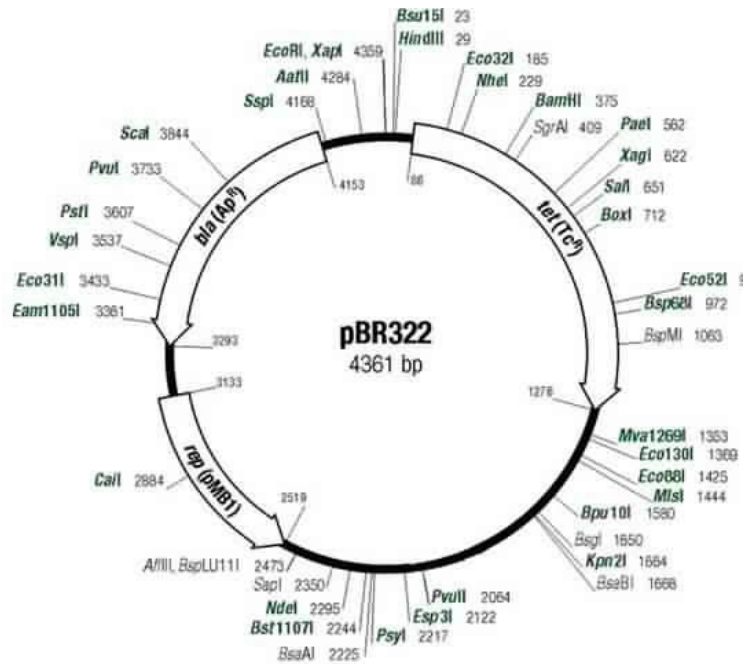
Exemple : La Transcriptase réverse ou rétrotranscriptase : présente surtout dans les rétrovirus (virus à ARN) (codée par le **gène pol**) dont la multiplication nécessite un passage par l'ADN. Elle représente la base des techniques de biologie moléculaire puisqu'elle permet la transcription reverse de l'ARNm en ADN complémentaire. Elle a besoin d'une extrémité 3' OH et d'une amorce pour fixer les dNTP. Elle est dépourvue d'activité exonucléasique $3' \rightarrow 5'$ et pourvue d'une activité RNase.

LES VECTEURS

Propriétés d'un bon vecteur

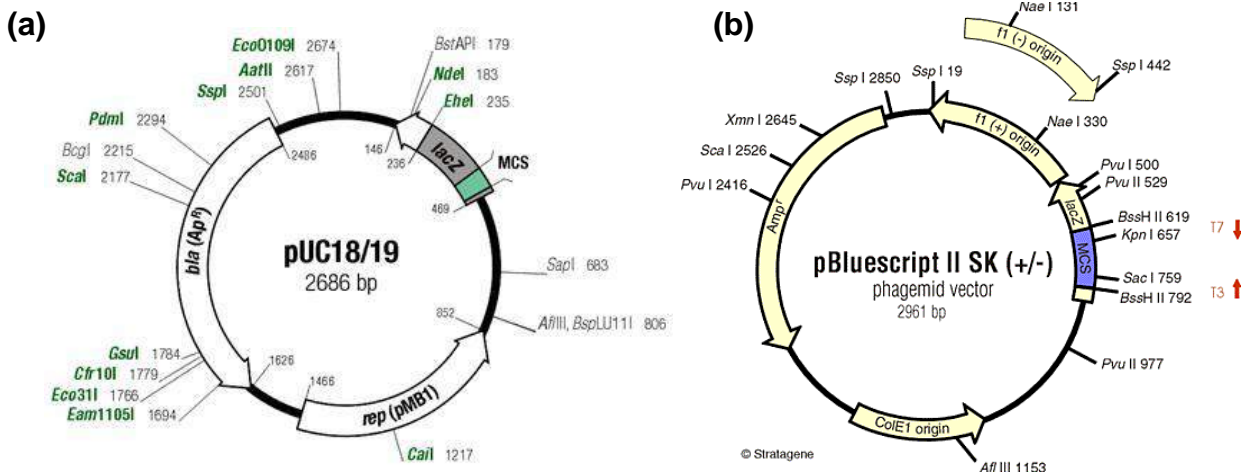
- Réplication active et autonome dans la cellule hôte
- Présence de marqueurs de sélection
- Présence d'un grand nombre de sites polylinkers localisés dans les gènes de sélection
- Sa présence ne doit pas perturber la physiologie de cellule hôte et vice versa
- Facile à isoler sous forme purifiée

Les plasmides



Carte de restriction du plasmide pBR322 (1^{er} génération)

Le plasmide pBR 322 de 4361 pb de taille, est l'un des 1^{ers} plasmides artificiels utilisés comme vecteur de clonage. Il possède 2 gènes de résistance aux antibiotiques: tétracycline et Ampicilline ; une origine de réplication et plusieurs sites de restriction ; des sites uniques de restriction tels que les sites EcoRI, EcoRV, BamH1, Pst1, PvuI, Hind III parmi lesquels BamH I et Pst I sont respectivement localisés au niveau des gènes de résistance à la tétracycline et à l'ampicilline. L'introduction d'un ADN étranger dans l'un de ces sites se traduit donc par la perte de la résistance à l'antibiotique correspondant.



Carte de restriction de plasmides de 2^{ème} génération

Ce sont des plasmides puissants bien construits et permettant de simplifier considérablement le travail. Ils sont constitués des 2 familles : **pUC** (plasmid of University of California) (a) et **bluescript** (b). Ces plasmides présentent un autre gène « **lacZ** » (fragment de l'opéron lactose) qui code pour l'extrémité NH2 de la β-galactosidase. Au sein de ce gène est placée une région renfermant des sites de clonage multiple (**MCS**) ou **site polylinker** site d'insertion de l'ADN étranger). L'interruption de ce gène (par insertion de l'ADN étranger) entraîne une perte de l'activité de la protéine.

Les phages

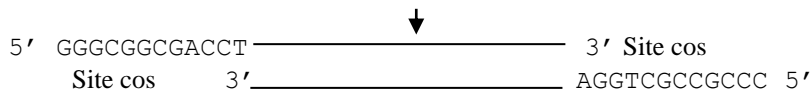
Les phages sont capables d'insérer jusqu'à 25 Kb d'insert (soit le 1/3 de l'ADN du phage peut être remplacé par un ADN étranger). Leur multiplication est rapide et le nombre de copies / cellule est élevé. Le rendement d'une infection phagique est supérieur à ce qui est obtenu lors de la transformation par un plasmide. Les vecteurs phagiques les plus connus et les plus utilisés pour le clonage sont dérivés du **phage lambda (λ)**. Celui-ci peut se multiplier selon 2 modes : selon le **cycle lytique** et le **cycle lysogénique**. Lorsque le phage λ est utilisé comme vecteur de clonage, c'est le cycle lytique qui nous intéresse puisqu'il entraîne la production d'un grand nombre de nouveaux phages qui vont à leur tour infecter d'autres bactéries et ainsi de suite.

L'ADN du phage λ est linéaire double brin de 48,5 Kb de taille. Les gènes de structure du phage sont regroupés aux extrémités de son génome. Les séquences médianes non indispensables pour la réplication du phage peuvent être éliminées et remplacées par des sites de restriction facilitant le clonage. Les extrémités 5' des 2 brins sont plus longues (donc simple brins) d'environ 12 nucléotides et sont complémentaires l'une avec l'autre \rightarrow ce sont donc des extrémités cohésives dites **cos** (L et R) nécessaires pour l'intégration de l'ADN du phage dans le génome de son hôte (donc nécessaires à la réplication et à l'encapsidation).

L'utilisation de phage λ non modifié présente des problèmes parmi lesquels l'existence de plusieurs sites de restriction pour des enzymes identiques, d'où la nécessité de construire des phages λ modifiés possédant un seul site de restriction pour que l'insertion de l'ADN étranger se fasse à un endroit précis du génome du vecteur. **Exemple** de phage λ modifié **λ gt11** qui est le plus utilisé puisqu'il est considéré comme un bon vecteur de clonage. Il possède un génome de 43,7 Kb dans lequel a été placé un marqueur de sélection qui est une partie de l'opéron lactose (codant la β -galactosidase) et possédant un seul site de restriction EcoRI situé dans l'opéron lactose.

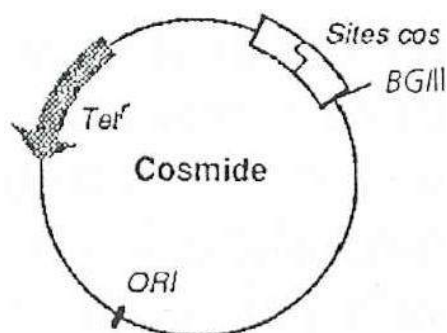
Extrémités cohésives du génome du phage λ gt11

Site EcoRI (site d'insertion)



Les cosmides

Vecteurs artificiels constitués de molécules d'ADN circulaires qui combinent à la fois les avantages des plasmides et des phages. Ce sont donc des hybrides constitués d'un plasmide classique auquel on a ajouté les extrémités cos du phage λ . Ils possèdent une origine de réplication plasmidique, un ou plusieurs marqueurs de sélection, une région MCS et un site cos (pour l'encapsidation *in vitro* de grands fragments d'ADN). Ils permettent l'insertion de grands fragments d'ADN allant jusqu'à 45 Kb.

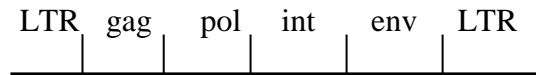


Les vecteurs d'eucaryotes supérieurs (animaux et végétaux)

Les vecteurs de transfert d'ADN dans les cellules eucaryotes sont d'utilisation plus complexe et chaque vecteur est spécifique d'une espèce eucaryote.

Exemples : virus de la vaccine, le virus SV40, adénovirus, baculovirus, rétrovirus et le **plasmide Ti = pTi** (= plasmid tumor inducing) de la bactérie *Agrobacterium tumefaciens*.

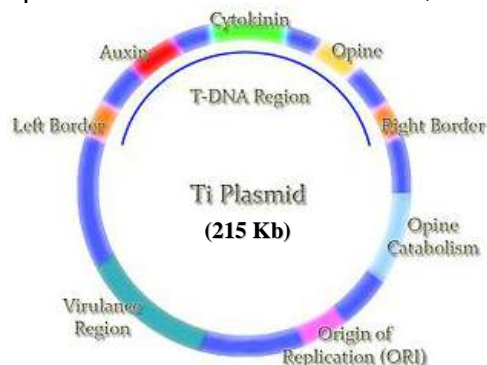
► **Les rétrovirus** : virus à ARN, pénètrent dans la cellule eucaryote et convertissent leur génome en ADN (rétrotranscription) pour pouvoir intégrer leur génome dans le génome hôte (grâce à une **intégrase** virale). Les gènes viraux « **gag** », « **pol** » et « **env** » peuvent être enlevés pour être remplacés par l'ADN à cloner qui sera donc intégré dans la cellule eucaryote hôte.



Structure typique du génome d'un rétrovirus.

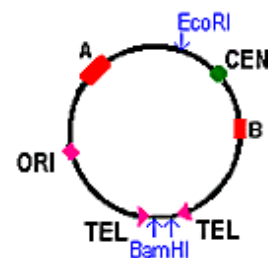
L'ARN du rétrovirus comporte à chaque extrémité de longues séquences répétées terminales **LTR** (long terminal repeat). La partie centrale contient une séquence **gag** qui code pour une **glycoprotéine de structure**, une séquence **pol** codant la **transcriptase inverse**, une séquence **int** responsable de la biosynthèse de l'**intégrase** qui intervient dans l'insertion de l'ADN viral dans l'ADN cellulaire et une séquence **env** qui code une **protéine de l'enveloppe virale**.

► **Le pTi** : gros plasmide (2150 pb) d'*Agrobacterium tumefaciens* qui est une bactérie tellurique responsable de l'apparition, chez de nombreuses espèces végétales, de la **galle du collet** (Crow gall). Cette maladie se manifeste par l'apparition de tumeurs au niveau du collet dues à une synthèse exagérée d'hormones de croissance codées par une région du plasmide appelée **T-DNA**. L'ADN-T peut être naturellement transféré à la plante (blessure), la bactérie devient donc capable de transférer un fragment d'ADN étranger dans le noyau des cellules de la plante hôte et de l'intégrer dans l'ADN chromosomique de l'hôte. Le pTi ne comporte pas de sites uniques de restriction permettant le clonage direct de l'ADN à transférer à la place de leur région-T. Donc le remplacement du T-DNA par un gène étranger d'intérêt agronomique permet l'intégration de ce gène dans le génome de la plante d'où l'obtention de plantes transgéniques résistantes aux herbicides, aux virus, aux insectes...



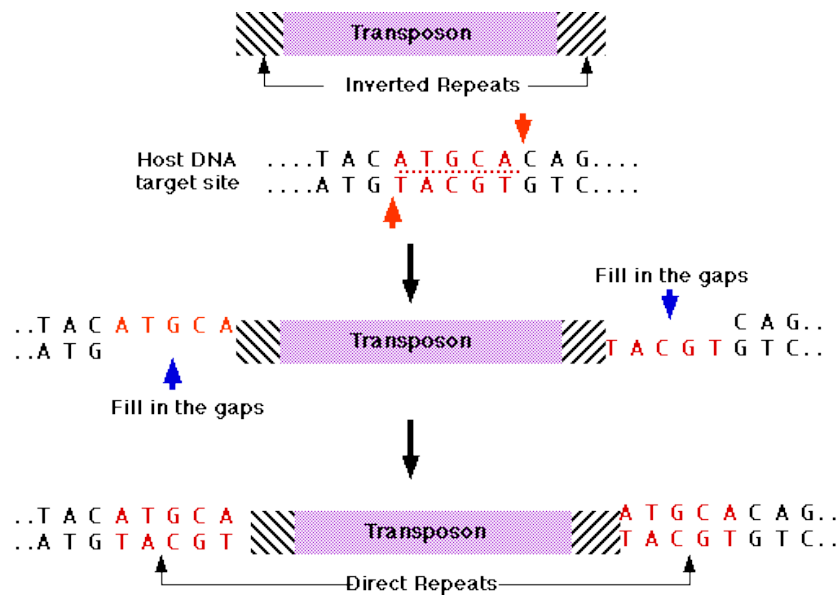
Les chromosomes artificiels

Ce sont des minichromosomes qui copient les chromosomes de la cellule hôte eucaryote et se répartissent régulièrement lors des divisions, comme les chromosomes naturels. Le plus utilisé est le **pYAC** (yeast artificial chromosome) construit à partir des levures et capable d'insérer de grands fragments d'ADN de 140 à 1000 Kb (exemple YAC2 qui fait 11,2 Kb de taille). Il possède une origine de réplication (ORI), une région centromérique (CEN) qui assure la migration correcte du minichromosome, un site unique de clonage (EcoR I) et un ou plusieurs marqueurs de sélection (A et B)



Les transposons

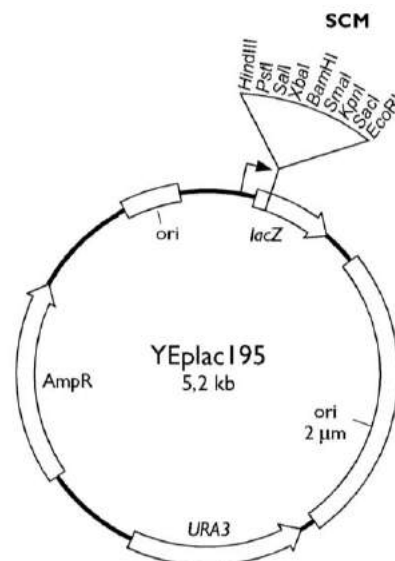
Ce sont des séquences d'ADN mobiles présentes dans les génomes en des sites quelconques



Vecteurs navettes

Ce sont des vecteurs artificiels qui peuvent se répliquer chez les procaryotes et les eucaryotes ; ils peuvent donc se déplacer d'un hôte à l'autre même s'ils sont différents. Ils sont utilisés lorsqu'un gène eucaryote est cloné dans un vecteur procaryote. Il est donc obligatoire de leur ajouter une séquence eucaryote spécifique qui servira d'origine de réplication à ces vecteurs.

Exemple : Vecteur navette bactérie-levure : possède deux origines de réplication : une reconnue par la machinerie de réplication de *E. coli* (ori) et l'autre reconnue par la machinerie de la levure *S. cerevisiae* (ori 2 μ m). Il possède aussi deux gènes de sélection : l'un est utilisé chez la bactérie (AmpR) et l'autre chez la levure (le marqueur de nutritionnel URA3; le milieu de sélection correspondant sera un milieu sans uracile). Un seul site de clonage multiple est nécessaire.



LES CELLULES HOTES

Pour se répliquer, un vecteur doit être incorporé dans une cellule hôte spécifique de chaque type de vecteur. **Exemple** : si le vecteur est un plasmide, la cellule hôte est une bactérie. La cellule hôte ne doit pas modifier ou détruire l'ADN recombiné qui y a été introduit. On distingue 2 catégories d'hôtes :

▪ **Les hôtes bactériens**

Ce sont les hôtes les plus utilisés car ils se multiplient rapidement et sont faciles à manipuler. Les bactéries utilisées ne sont pas pathogènes et sont modifiées afin de diminuer les risques de dissémination accidentelle. La bactérie de choix est *E. coli* qui doit être :

- une souche « **res-** » : ne produit pas des enzymes de restriction pour ne pas détruire l'ADN étranger inséré.
- « **recA-** » : pour éviter la recombinaison de l'ADN bactérien avec l'ADN inséré
- Réceptrice donc « **F-** » : incapables de transmettre le plasmide qu'elles ont reçu ; par opposition aux souches donatrices F+
- Certaines expériences nécessitent des souches « **lac-** »

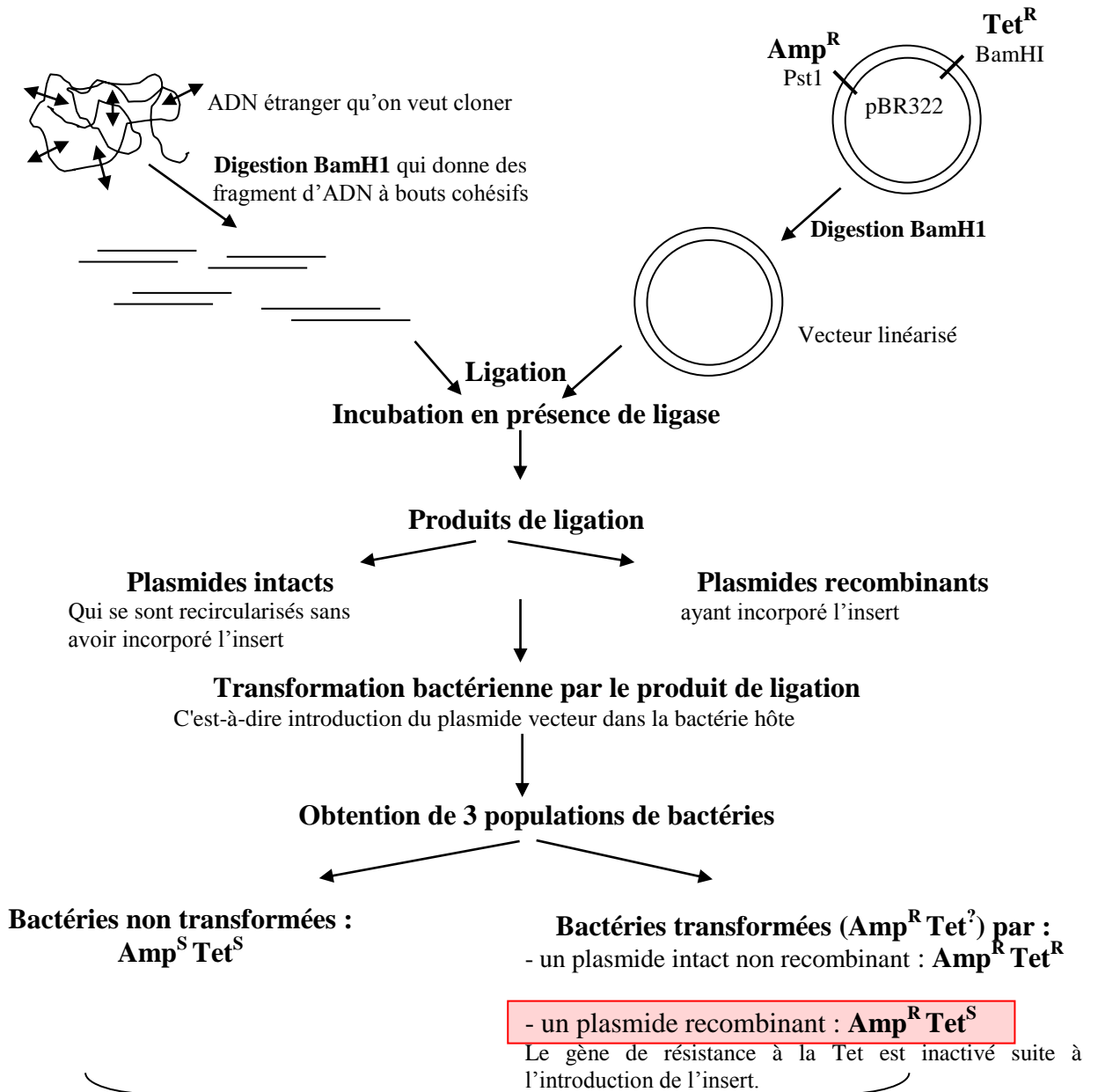
▪ **Les hôtes eucaryotes**

Peuvent être des cellules animales en culture, des levures ou des plantes. Leur manipulation est complexe et coûteuse. Ils ne sont utilisés que :

- lorsqu'on travaille avec des vecteurs ayant une origine de répllication eucaryote ou
- lorsqu'on veut étudier la régulation *in vivo* d'un gène eucaryote ou
- lorsqu'on veut faire exprimer une protéine qui ne peut être fonctionnelle qu'à la suite de certaines modifications post-traductionnelles (telle que la glycosylation) que la bactérie ne peut pas réaliser.

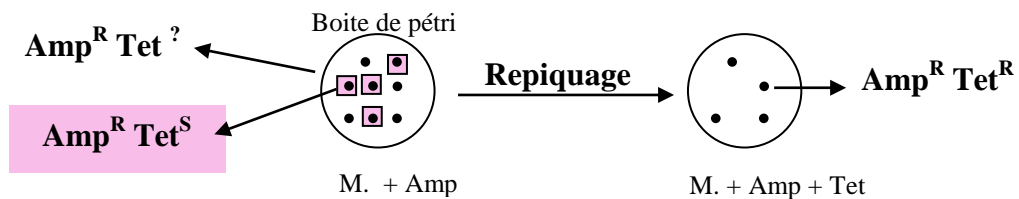
RAPPELS : LA TECHNOLOGIE DE L'ADN RECOMBINANT

Exemple ① : Clonage d'un gène dans le plasmide pBR322



Sélection : par utilisation de la résistance à l'Amp et Tet

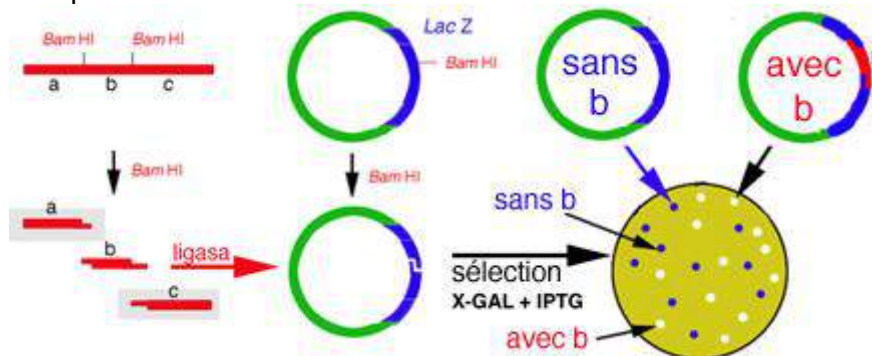
par étalement des bactéries sur un milieu de culture + Amp → seules les bactéries transformées vont se développer sur ce 1^{er} milieu (Amp^R Tet^S). Ensuite, les colonies apparues sur ce milieu seront repiquées sur un milieu contenant Amp + Tet → seules les bactéries transformées par un plasmide non recombinant vont se développer sur ce 2^{ème} milieu (Amp^R Tet^R). les colonies qui ont poussé sur le 1^{er} et non sur le 2^{ème} sont celles transformées par un plasmide recombinant ((Amp^R Tet^S))



Exemple (2) : Clonage d'un gène dans le vecteur pUC19

Ce plasmide comporte : **un gène de résistance à l'Amp**, le **gène lacZ** (fragment de l'opéron lactose et qui code pour la β -galactosidase) contenant des **sites polylinker**. La particularité de ce vecteur par rapport au précédent c'est qu'il possède « un système de sélection enzymatique » appartenant à l'opéron lactose au lieu d'utiliser un second antibiotique. Si ce vecteur intègre un fragment ADN d'intérêt au niveau de la région MCS, il y aura interruption et donc l'inactivation de ce gène lacZ. De ce fait, la β -galactosidase ne sera donc plus fonctionnelle. Donc les cellules hôtes ayant intégré un plasmide recombinant perdent l'activité de la β -galactosidase. Par contre, quand les cellules hôtes intègrent un plasmide intact non recombinant, la β -galactosidase sera active. La sélection des bactéries ayant intégré un plasmide recombinant est réalisée par « un test colorimétrique » ou « test bleu/blanc » : cultiver les bactéries dans un milieu contenant un inducteur l'IPTG (isopropylthio-b-D-galactoside) nécessaire pour métaboliser le substrat X-Galactose = X-gal qui est un β -galactoside (le lactose est un galactoside) dont la couleur passe de l'incolore au bleu quand il est clivé (hydrolysé) par la β -galactosidase mais au lieu de libérer du galactose et du glucose comme dans le cas de l'hydrolyse du lactose, il libère du galactose et une substance X (5-bromo-4-chloro-3-hydroxyindole) qui est colorée en bleu.

Donc, en culture et en présence d'IPTG + X-gal, les bactéries Amp^R qui se développent sur l'agar et qui sont transformées par les plasmides recombinants, se présentent sous forme de colonies blanchâtres car elles ont perdu la capacité de clivage du X-gal par la β -galactosidase. Par contre, les bactéries Amp^R et transformées par des plasmides non recombinants se présentent sous forme de colonies bleues.



LES BANQUES D'ADN

C'est l'ensemble de clones bactériens ayant incorporé différents plasmides recombinants. Une banque d'ADNc est fabriquée à partir des régions transcrites du génome et une banque génomique est fabriquée à partir de séquences géniques et intergéniques.

CRIBLAGE DE LA BANQUE

Il consiste à la recherche du clone ayant inséré le fragment d'ADN qu'on désire étudier. Il nécessite la **réplique de la banque**.

Exemple : Criblage de la banque par Hybridation moléculaire avec une sonde marquée pour la recherche d'un gène particulier tel que le gène qui code pour la synthèse de l'hormone de croissance (HG).

1- Les sondes

Il existe 2 types de sondes :

- **Les sondes qui reconnaissent les protéines = anticorps marqués**

On a recours à cette technique lorsqu'on n'a aucune connaissance sur la séquence des acides aminés de la protéine. Dans ce cas, on purifie la protéine à partir de l'organe pour obtenir des anticorps dirigés contre le produit protéique du gène recherché. Les clones positifs seront révélés grâce au marquage (radioactif ou non radioactif) des anticorps. Donc

en identifiant la protéine, l'anticorps identifie le clone contenant le gène qui a dû synthétiser cette protéine.

- **Les sondes qui reconnaissent les acides nucléiques (ADN et ARN) = sondes nucléotidiques**

Définition : C'est tout fragment d'acide nucléique simple brin (ADNc, fragment PCR, oligonucléotide, ARN) de taille très variable et qui possède une homologie de séquences avec l'ADNc ou le gène recherché.

Obtention :

- par synthèse chimique si on connaît la séquence de l'ADN à étudier. Si cette séquence est inconnue, on peut étudier la protéine correspondante et retrouver la séquence d'ADN grâce au code génétique (Cf. exemple plus loin).
- une sonde peut être un ADNc dont seulement une partie sera utilisée (après digestion et clonage des fragments obtenus)
- une sonde peut être de l'ARNm

2- Hybridation moléculaire de la sonde

La sonde permet de rechercher de manière spécifique un fragment d'acide nucléique que l'on désire étudier. Cette réaction **sonde – fragment** correspond à une réaction « **d'hybridation moléculaire** ». Celle-ci nécessite des conditions physico-chimiques parfaites (tampon, pH, t°...) appelées **strigence** et dépend de la longueur de la sonde et de la complémentarité sonde – fragment avec possibilité de mésappariement. Elle doit être complémentaire et anti// au fragment recherché, qui lui aussi doit être dénaturé (simple brin). Pour repérer cette réaction d'hybridation sonde-ADN recherché, il faut réaliser un marquage de la sonde.

3- Marquage

Dans un mélange complexe où s'effectue l'hybridation moléculaire, la sonde doit être facilement repérable grâce à un **marquage**.

Le marquage d'une sonde est réalisé soit avec un atome radioactif = **radio-isotope** (= isotope radioactif) : c'est le « **marquage radioactif** à chaud » (par une sonde radiomarquée) et le clone positif (contenant le gène recherché) est révélé par une tache noire par autoradiographie; soit avec un « **colorant fluorescent** » : c'est le « **marquage enzymatique** ou **fluorescent** à froid » non radioactif. Dans ce cas, le clone positif est révélé par une tâche fluorescente brillante.

Globalement, le marquage se fait soit à l'extrémité soit à l'intérieur du fragment d'ADN :

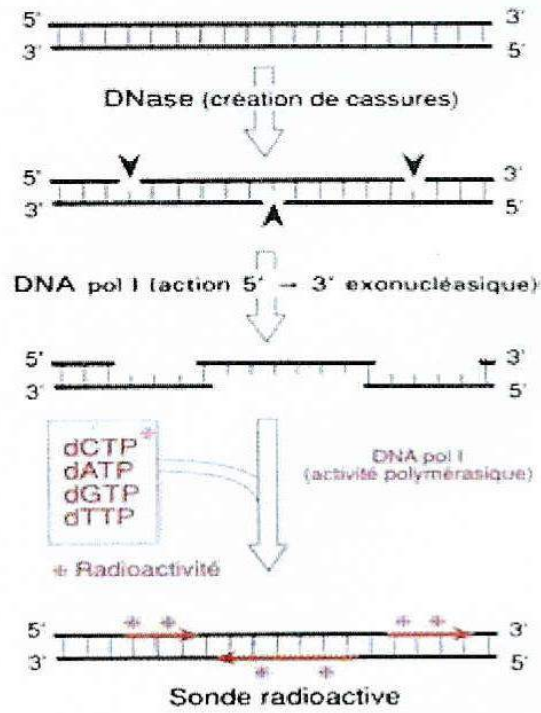
- **marquage aux extrémités** : voir exemple plus loin

- **marquage interne** :

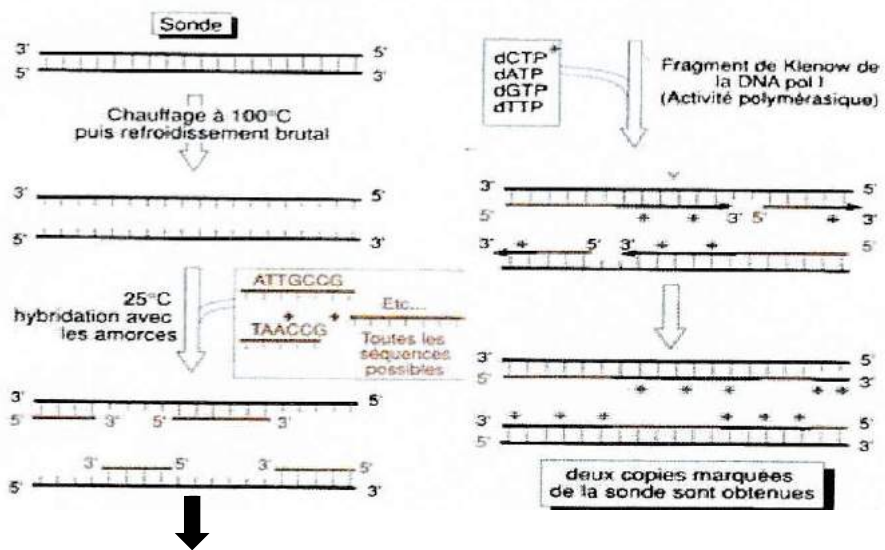
- **Technique « nick translation »** = coupure-réparation : Clivage au hasard de l'ADNdb par la Dnase I, puis réparation des coupures par l'ADN pol I en présence de « **dNTP $\alpha^{32}\text{P}$** » (c'est-à-dire dNTP radiomarquée en position α par l'isotope ^{32}P)

- **Technique « random priming »** (amorçage au hasard) : les 2 brins d'ADN de la sonde sont préalablement séparés par chauffage/refroidissement brutal. Puis on ajoute un mélange d'hexanucléotides de synthèse correspondant à toutes les combinaisons possibles ($4^6 = 4096$). Ces oligo vont s'hybrider avec la sonde par une partie d'entre eux.

Donc on ne peut synthétiser une sonde nucléotidique que lorsqu'on connaît le produit protéique du gène recherché et sa séquence d'acides aminés.



Marquage par transfert de coupure (Nick translation)



Marquage par amorçage aléatoire (random priming)

Principe de la technique d'hybridation moléculaire avec une sonde ADN radiomarquée aux extrémités 5'P

Elle consiste à identifier dans une banque d'ADN, les clones bactériens comportant une séquence d'ADN par hybridation avec une sonde radiomarquée. Pour cela, il est nécessaire de connaître une partie de la séquence recherchée qu'on peut déduire à partir des acides aminés de la protéine correspondante.

① Choix d'un segment d'acides aminés (de la protéine correspondante au gène recherché) avec le minimum de dégénérescence possible : **Exemple** d'une courte séquence de protéine utilisée pour fabriquer un ensemble d'oligonucléotides dégénérés utilisables comme une sonde pour retrouver le gène codant la protéine.

6	2	2	1	2	1	4		
NH2	Ser	– Tyr	– His	– Met	– Glu	– Trp	– Pro	COOH
UCN	UAU	CAU	AUG	GAA	UGG	CCN	} Des sondes ADN synthétiques sont fabriquées à partir des données du code génétique. Cette séquence d'AA est traduite en sens inverse pour obtenir la séquence d'ADN qui l'a codée. Cependant, en raison de la dégénérescence du code génétique, plusieurs séquences d'ADN pourraient avoir codé la protéine en question.	
AGU	UAC	CAC	GAG					
AGC								

② Synthèse chimique (synthétiseur d'oligo) des oligonucléotides de toutes les combinaisons possibles correspondant à cette région (encadrée).

Les séquences possibles de la sonde de 17 bases : 5' TAT CAT ATG GAA TGG CC3'

C	C	G
2x	2x	2x

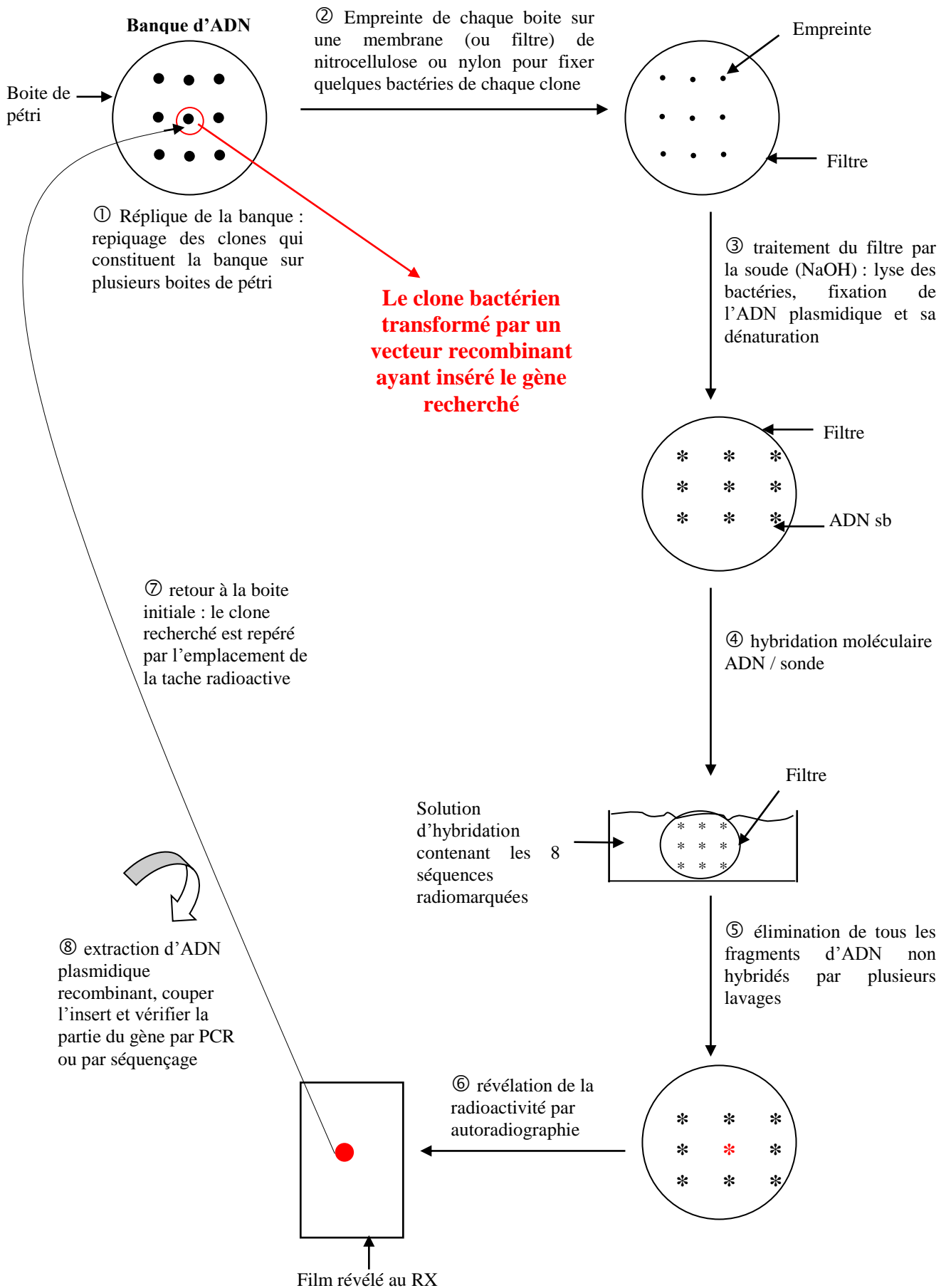
→ Le coefficient de dégénérescence = nombre de combinaisons des séquences de la sonde = $2^3 = 8$.

L'un des oligonucléotides de la sonde correspondra exactement au gène recherché.

③ Marquage des 8 séquences. Le plus simple est le marquage de l'extrémité 5' par la **T₄ PK** (=T₄ **polynucléotide kinase** qui est une kinase qui ajoute des groupements phosphates). Dans notre cas, la **T₄ PK** transfère le groupement phosphate (en position gamma γ = la position la plus externe) à partir d'une molécule d'« **ATP γ ³²P** » (ATP radiomarquée en position γ par l'isotope ³²P) sur l'hydroxyle 5' de l'ADN de la sonde préalablement déphosphorylé.

Au lieu de ce marquage radioactif, on peut réaliser le marquage fluorescent à froid à l'aide du dUTP marqué à froid à l'isotope ³⁵S. Ce marquage est moins sensible mais moins dangereux que le marquage radioactif.

Criblage d'une banque ADN par hybridation moléculaire avec une sonde ADN radiomarquée



4- Southern blot

C'est une technique d'hybridation moléculaire qui consiste à détecter spécifiquement des fragments d'ADN transférés sur filtre par leur hybridation à des séquences complémentaires (sonde) marquées par un radio-isotope.

Exemple : Recherche d'un gène (cloné et provenant d'une espèce de plante) à partir de l'ADN génomique d'une plante en utilisant la technique d'hybridation moléculaire par Southern blot.

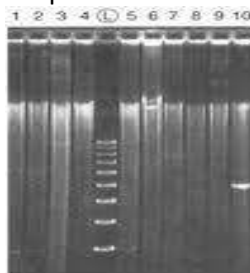
Extraction de l'ADN génomique d'une plante



Digestion par différentes enzymes de restriction



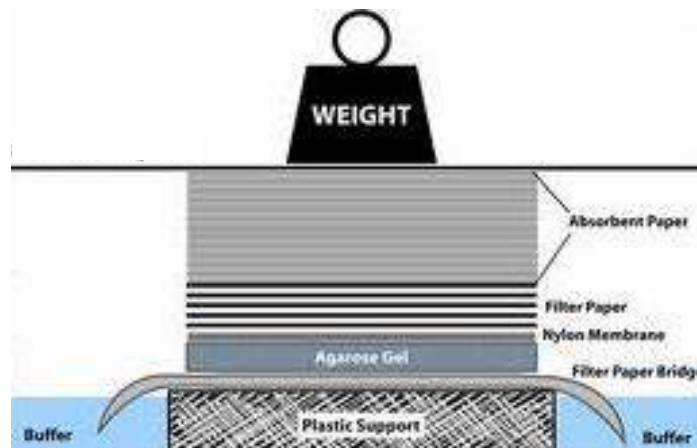
Séparation des fragments d'ADN totaux par électrophorèse sur gel d'agarose (séparation d'une traînée continue de bandes d'ADN en fonction de leur taille et dont l'une d'elles correspond au gène recherché qui sera identifiée par la sonde marquée après Southern blot)



Dénaturation des fragments d'ADN digérés par **traitement alcalin** (acide) du gel d'électrophorèse



Southern blot



Transfert par capillarité des fragments d'ADNs séparés sur une membrane poreuse souple (feuille de nylon).



Fixation de l'ADNs sur la membrane (par cuisson de la membrane ou par exposition sous UV)



Hybridation moléculaire de l'ADN fixé à la membrane **avec une sonde** radio-marquée dénaturée.



Lavages



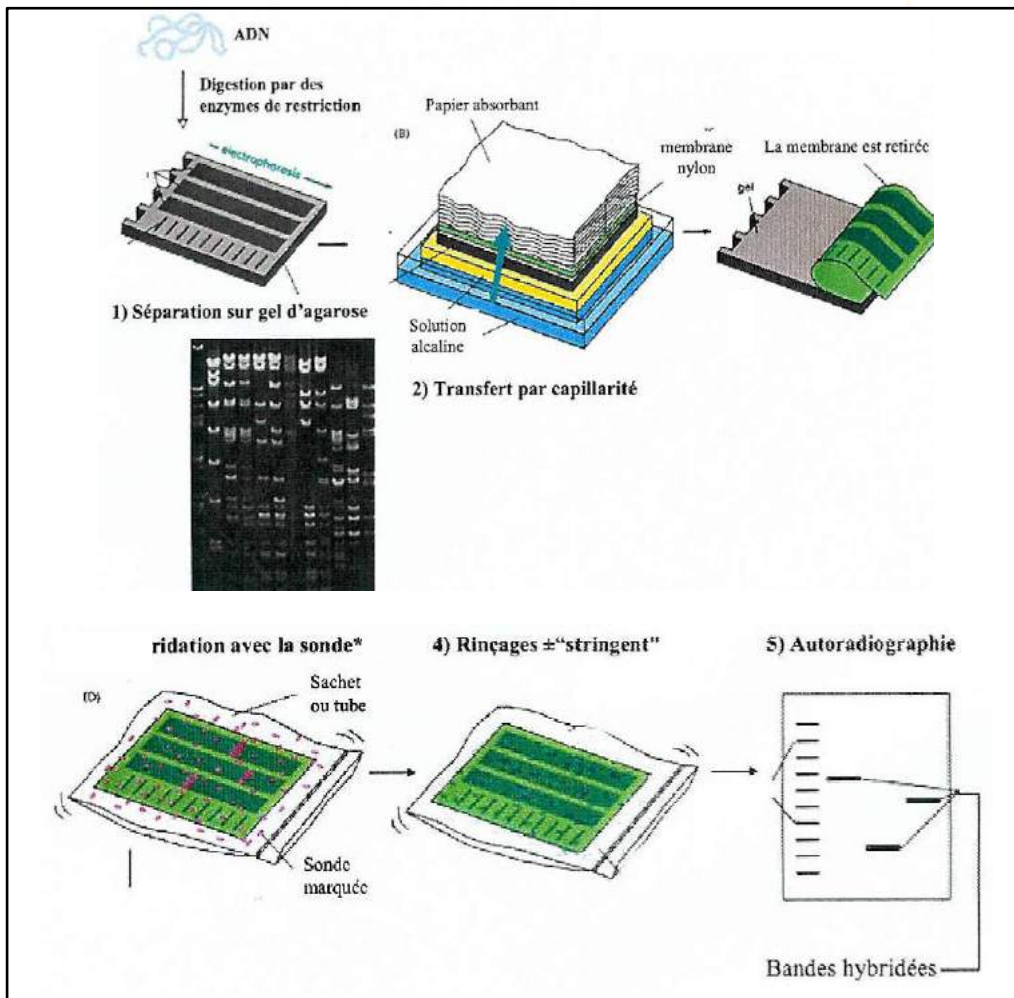
Autoradiographie (La membrane est mise en contact avec un film photographique)

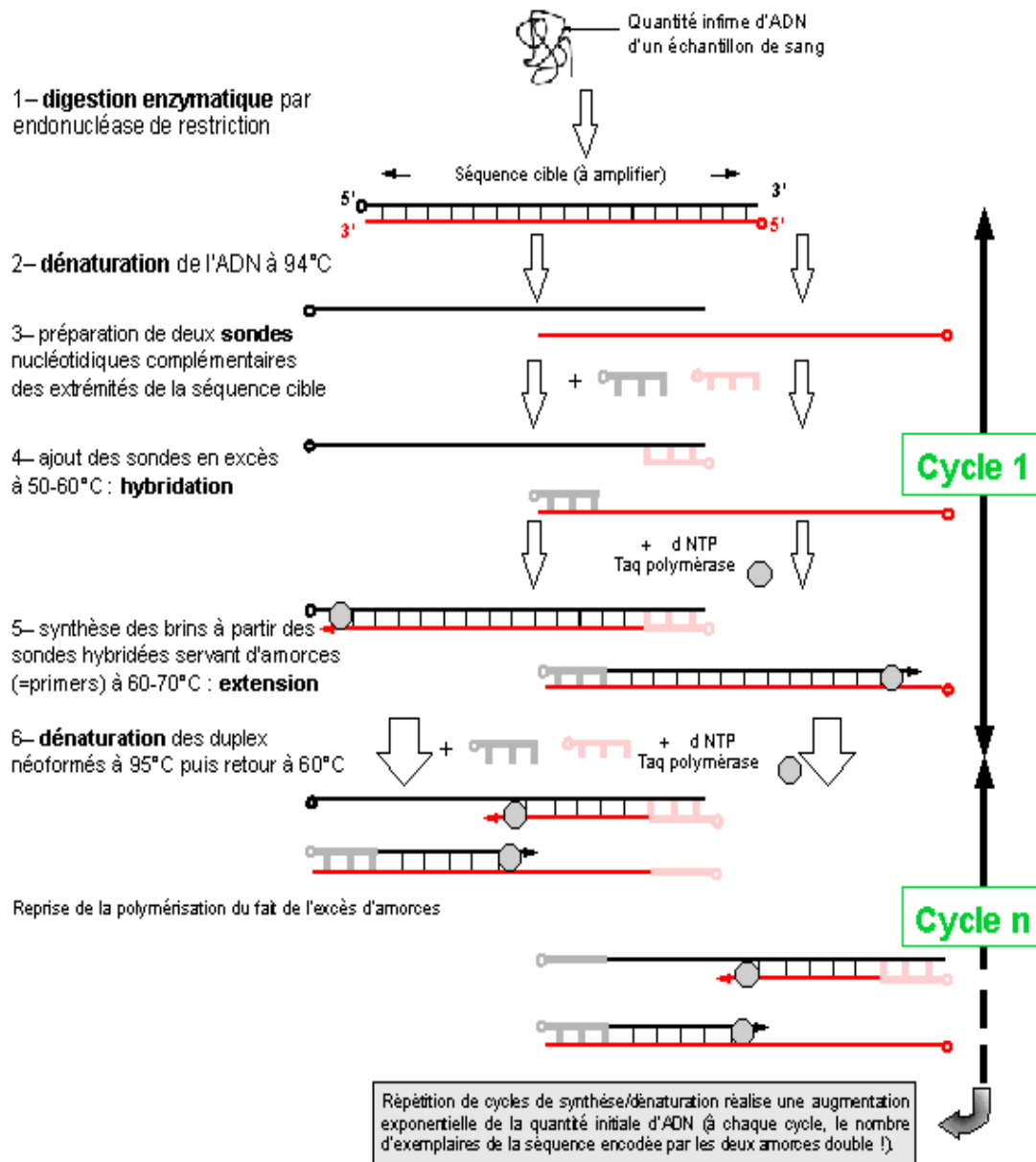


Révélation du film pour révéler les bandes d'ADNs hybridées dont la séquence est homologe à celle de la sonde (ces bandes sont visibles sous forme de bandes noires sur un fond blanc). La position de ces bandes par rapport à des témoins de poids moléculaire permet de déterminer la taille de ces fragments)



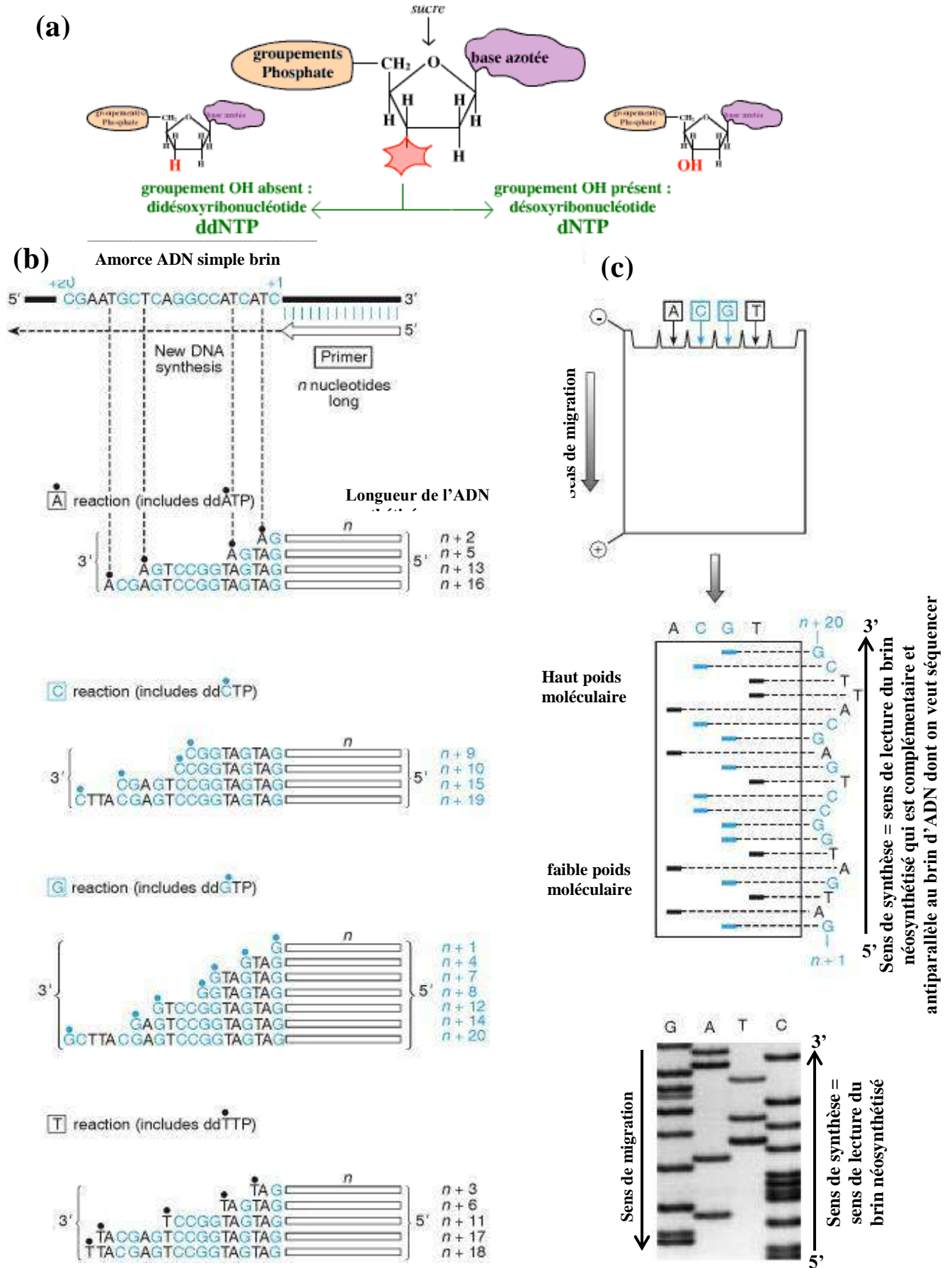
Film photographique révélé



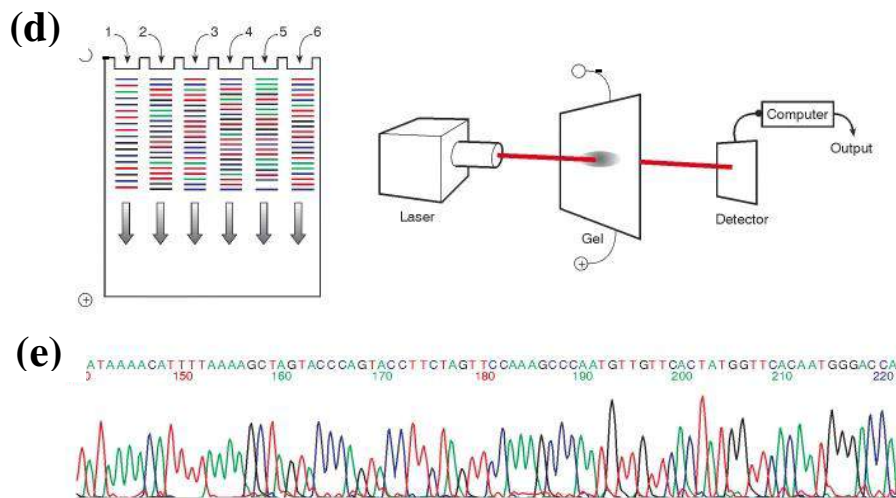


Principe de l'amplification enzymatique de l'ADN par PCR

- **Dénaturation** de l'ADN à 90-95°C pour séparer les 2 brins
- **Hybridation** de chaque brin avec une amorce spécifique (50 à 60°C qui dépend du % en GC). La température d'hybridation T_a doit être < de 5°C de la température de demi-dénaturation T_m pour un oligonucléotide < 30 nt. La T_m est estimée en utilisant la relation suivante : $T_m = 2(A + T) + 4(G + C)$
Les 2 oligonucléotides font ~ 15-25nt et doivent avoir des T_a voisines.
- **Extension** ou **élongation** de ces amorces par l'ADN pol (70°C)



Principe de la technique de séquençage enzymatique de Sanger utilisant le ddNTP (suite page suivante)



Principe de la technique de séquençage enzymatique de Sanger utilisant le ddNTP (suite).

(a) Structure d'un ddNTP.

Cette technique consiste à la synthèse *in vitro* du brin d'ADN complémentaire à l'un des 2 brins de l'ADN que l'on désire séquencer. Elle repose sur l'utilisation dans le milieu réactionnel (contenant les dNTPs, une amorce marquée et une ADN polymérase) de nucléotides de synthèse modifiés et analogues structuraux de nucléotides: les **didésoxyribonucléosides triphosphates (ddNTPs)** dépourvus du groupement OH (fonction alcool) en 3' nécessaire pour former une liaison avec le nucléotide triphosphate suivant. Une fois incorporé dans la chaîne en cours d'élongation, le ddNTP provoque l'arrêt de la synthèse de l'ADN à partir de leur incorporation → c'est un inhibiteur de la réplication.

(b) Etapes du séquençage selon Sanger

Quatre réactions sont préparées à partir d'une matrice d'ADN simple brin (la séquence recherchée) et une fraction d'amorce marquée. Chaque tube reçoit une petite quantité d'un ddNTP différent (ddATP, ddTTP, ddCTP, ddGTP), ainsi que les 4 dNTPs normaux. Un ddNTP sera incorporé au hasard à des sites distincts selon la synthèse ayant lieu dans le tube à essai. Donc tous les tubes contiendront un mélange d'ADN de tailles différentes (correspondant chacune au site d'incorporation du ddNTP spécifique du tube, et donc de l'arrêt de croissance de la chaîne d'ADN) en fonction du point d'arrêt de la polymérisation et qui seront terminés par l'une des 4 bases possibles de l'ADN.

(c) Autoradiogramme d'un gel de séquençage

Les fragments tronqués sont ensuite séparés par électrophorèse sur gel de polyacrylamide (pouvant séparer 2 fragments d'ADN dont la longueur ne diffère que d'un nucléotide) et visualisés par autoradiographie (marquage radioactif de la sonde par fluorescence). L'autoradiogramme contient donc 4 colonnes pour les résultats obtenus dans les 4 tubes (les plus petits fragments migrent toujours le plus vite). La lecture de l'autoradiogramme se fait du bas vers le haut dans le sens 5' → 3' pour le brin complémentaire (radioactif ou fluorescent). Il suffit d'en déduire la séquence du brin matriciel par complémentarité dans le sens 3' → 5' puis de l'écrire dans le sens conventionnel 5' → 3'.

(d) Principe du Séquençage automatisé de l'ADN à l'aide des 4 ddNTPs marqués à des fluorochromes différents

Actuellement, le séquençage est devenu automatisé : « séquençage automatique » qui présente le même principe sauf qu'au lieu de marquer les fragments d'ADN par des amorces radioactives, les machines de séquençage les plus récentes marquent les 4 types de ddNTP par des fluorochromes de couleur différentes spécifiques pour chaque type de nucléotide (ddTTP-ROX (rouge), ddATP-JOE (vert), ddGTP-TAMRA (noir) et ddCTP-5-FAM (bleu)). Les produits de la réaction sont déposés sur un gel d'électrophorèse : gel de séquence et les données de séquence sont enregistrées au cours de l'électrophorèse grâce à un rayon laser qui capte la fluorescence émise par chacun des 4 fluorochromes au cours de la migration des fragments d'ADN. L'information est ainsi enregistrée et interprétée dans la banque de données de l'ordinateur. Les réactions peuvent ainsi se dérouler dans le même tube

(e) Enregistrement obtenu à partir d'un séquenceur automatique : Electrophérogramme de séquençage

Les séquenceurs permettent de lire plusieurs centaines de nucléotides avec une très bonne qualité, jusqu'à 1000 pb pour les appareils les plus performants

CHAPITRE I

1- LA MUTAGENESE IN VITRO

L'une des techniques les plus utilisées est la « **mutagenèse dirigée** » ou « **mutagenèse ciblée** » c'est-à-dire **cibler** la mutagenèse :

- en choisissant le gène à muter à condition qu'il ait déjà été identifié et que sa séquence sauvage soit clonée
- en déterminant le site et la nature de mutation dans le gène.

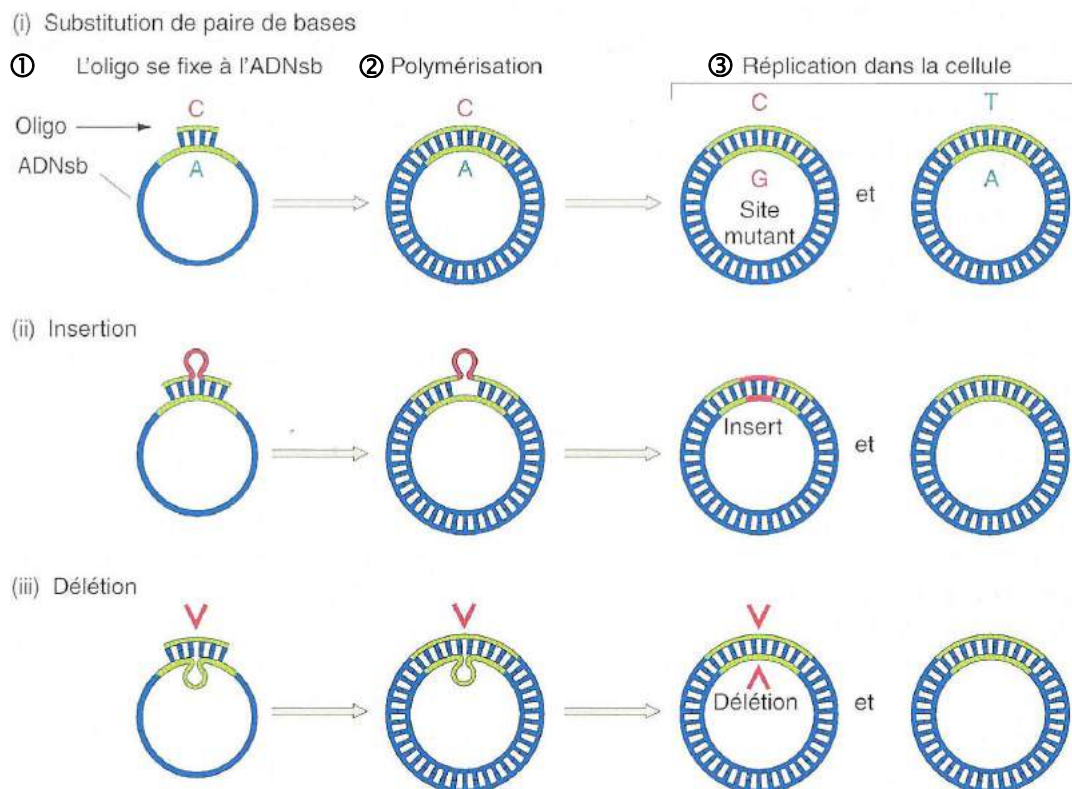
Cette technique consiste donc à l'introduction de différents types de changements, à des sites spécifiques, dans un gène. Elle permet donc de créer des mutations au niveau de n'importe quel site spécifique d'un gène dont on connaît la séquence de type sauvage.

Principe

Le gène est inséré dans un vecteur de clonage qui sera utilisé pour transformer des cellules hôtes de telle sorte que le gène manipulé *in vitro* « **le transgène** » vienne remplacer le gène résidant de type sauvage. La technique de mutagenèse dirigée peut être appliquée à des gènes clonés dans des vecteurs simples brins ou doubles brins.

Exemple : Mutagenèse dirigée *in vitro* utilisant un oligonucléotide

Insertion du gène au niveau duquel on veut créer des mutations, dans le vecteur phagique M13 à ADN circulaire simple brin.



① La séquence connue de type sauvage est utilisée pour synthétiser chimiquement de courts fragments d'ADN appelés « **oligonucléotides** ». Ces oligo synthétiques comportant la mutation désirée : substitution (i), insertion (ii) ou délétion (iii) entraîneront l'apparition de mutations comparables dans le gène cloné visé.

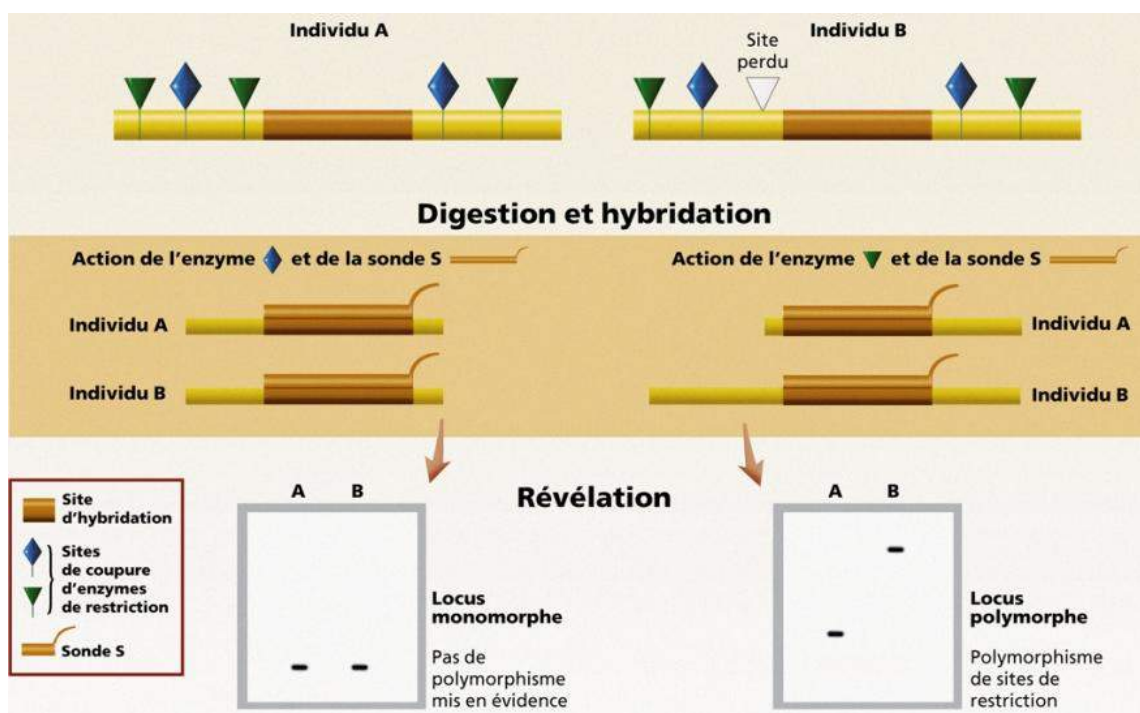
② Grâce à l'hybridation des simple-brins, ces oligo permettent de cibler n'importe quel site à l'intérieur du gène. On laisse cet oligo s'hybrider avec le site mutant. Cet oligo sert donc d'amorce à la synthèse *in vitro* du brin complémentaire du vecteur M13. Bien que

l'hybridation de l'oligo synthétique avec la séquence complémentaire dans le vecteur M13 entraîne l'apparition d'un mauvais appariement, quelques bases mal appariées peuvent être tolérées si cette opération se déroule à basse température et en présence d'une concentration saline élevée.

③ Après la synthèse d'ADN *in vitro* par l'ADN pol, on laisse l'ADN du M13 se répliquer *in vivo* dans des bactéries *E. coli* → obtention d'un grand nombre de phages correspondant au mutant désiré (CG).

L'oligo synthétique peut être utilisé comme une sonde marquée pour distinguer les phages de type sauvage des phages mutants puisqu'à température élevée, l'amorce ne va s'hybrider qu'avec le phage mutant. Alors qu'à faible température, la base mal appariée n'empêche pas l'amorce de s'hybrider avec les 2 types de phages.

2- LA CARTHOGRAPHIE A L'AIDE DES RFLP



Analyse du polymorphisme par RFLP

Se base sur l'utilisation du couple enzyme/sonde pour détecter une mutation qui supprime ou crée un site de restriction

Les étapes de la technique RFLP

- ① Extraction de l'ADN des différents génotypes à analyser
- ② Digestion de l'ADN par des enzymes de restriction à des sites de restriction (ou PCR dans le cas d'une PCR/RFLP)
- ③ Séparation des fragments engendrés selon leur taille par électrophorèse sur gel d'agarose pour séparer les fragments de restriction. La révélation au BET ne permet pas de distinguer individuellement les fragments car ils sont trop nombreux (traînée de coloration le long de la piste de migration), en plus il est impossible de repérer les fragments homologues. Pour cela, il est nécessaire d'utiliser des sondes moléculaires. D'où le
- ④ Transfert selon Southern
- ⑤ Hybridation moléculaire par des sondes ADN marquées (radioactives ou fluorescentes)
- ⑥ Révélation par Autoradiographie.

3- LA GENETIQUE INVERSE

C'est une nouvelle approche qui est née grâce à la technologie de l'ADN recombinant et qui procède en sens inverse par rapport à la génétique classique. Elle part du gène pour descendre vers la protéine dont on ne dispose d'aucune information génétique, alors que la génétique classique part de la protéine pour arriver au gène (exemples : les hémoglobinopathies, les hémophilies).

→ **la génétique inverse permet de découvrir le rôle normal de séquences inconnues d'ADN ou de protéines, en mutant leur séquence *in vitro* et en recherchant les changements induits au niveau phénotypique.**

Soit un gène sauvage cloné détecté lors du séquençage sous la forme d'un cadre de lecture ouvert (ORF) de fonction inconnue. Ce gène peut être soumis à une mutagenèse *in vitro*, puis réinséré dans l'organisme de départ pour en observer les conséquences sur le phénotype.

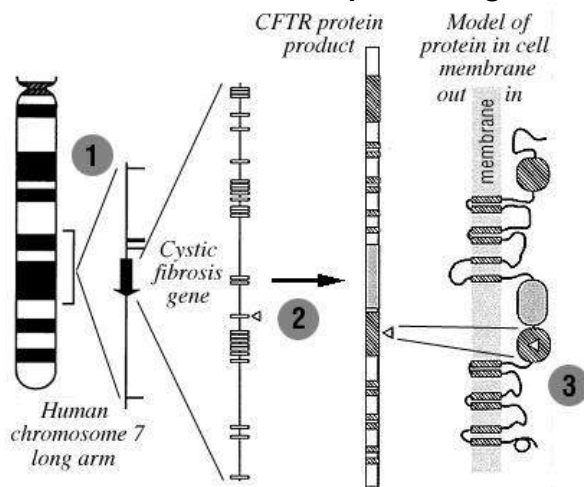
Si on prend le cas d'une protéine de fonction inconnue :

- la séquence d'acides aminés peut être traduite en sens inverse en une séquence d'ADN
- synthèse *in vitro* de cette séquence d'ADN.
- utilisation de cette séquence d'ADN comme sonde pour trouver le gène correspondant
- ce gène est alors cloné puis soumis à une mutagenèse *in vitro* pour déterminer les conséquences des mutations sur le phénotype de l'organisme qui l'abrite.
- Le séquençage du gène muté révèle la présence de nombreux ORF inconnus qui n'ont aucune ressemblance avec des gènes connus.

→ On peut ainsi déterminer la fonction de ces gènes grâce à la génétique inverse.

Les principaux outils de la génétique inverse sont :

- la mutagenèse *in vitro* et
- l'inactivation complète des gènes = « knock-out » (KO) des gènes



- ① le gène de la mucoviscidose se trouve sur le bras long du chromosome humain 7 (150 millions de pb)
- ② Grâce à l'utilisation de techniques microbiologiques, on a d'abord localisé la région (230,000 pb) du gène de la mucoviscidose (Tsui en 1989). Cette région code la protéine cftr (Cystic Fibrosis Transmembrane conductance Regulator). Le petit triangle montre l'endroit où se trouve la mutation de la délétion des trois paires de base.
- ③ Modèle de protéine cftr dans la membrane de la cellule : un gène normal fabrique du cftr qui régule le passage des ions chlorure et donc la sécrétion du mucus dans les cellules épithéliales (de surface) qui recouvrent les intestins, les poumons etc... Un acide aminé manquant à cet endroit cause la majorité des cas de mucoviscidose. Le reste des cas sont causés par plus de 1200 mutations différentes sur le gène cftr, chacune de ces mutations entraînant un petit nombre de cas.

Exemple : la mucoviscidose est une maladie génétique (récessive autosomique) qui se traduit par la sécrétion d'un mucus trop visqueux au niveau du pancréas et de l'épithélium bronchitique ce qui gêne la respiration du malade. Elle a été découverte grâce à la génétique inverse puisque la séquence du gène, donc la séquence en acides aminés de la protéine a été connue avant même que la protéine ne soit isolée. La stratégie de la génétique inverse utilisée consiste à : (a) localiser le locus morbide (responsable du phénotype pathologique) sur un chromosome, puis sur une région chromosomique (b) à le cerner ; (c) à l'isoler à partir d'un ADN normal ; (d) à le cloner et le séquencer ; (e) en fin de compte à déchiffrer la protéine correspondant au message génétique. La première étape a été particulièrement longue à franchir. En effet, en l'absence de tout indice permettant de suspecter un chromosome plutôt qu'un autre, il a fallu adopter la démarche aléatoire de la stratégie d'exclusion qui consiste à rechercher dans les familles de mucoviscidosiques la co-ségrégation d'un RFLP avec le locus CF (Cystic Fibrosis gene). Cette étude permet l'évaluation quantitative de la distance génétique séparant le site polymorphe du locus morbide.

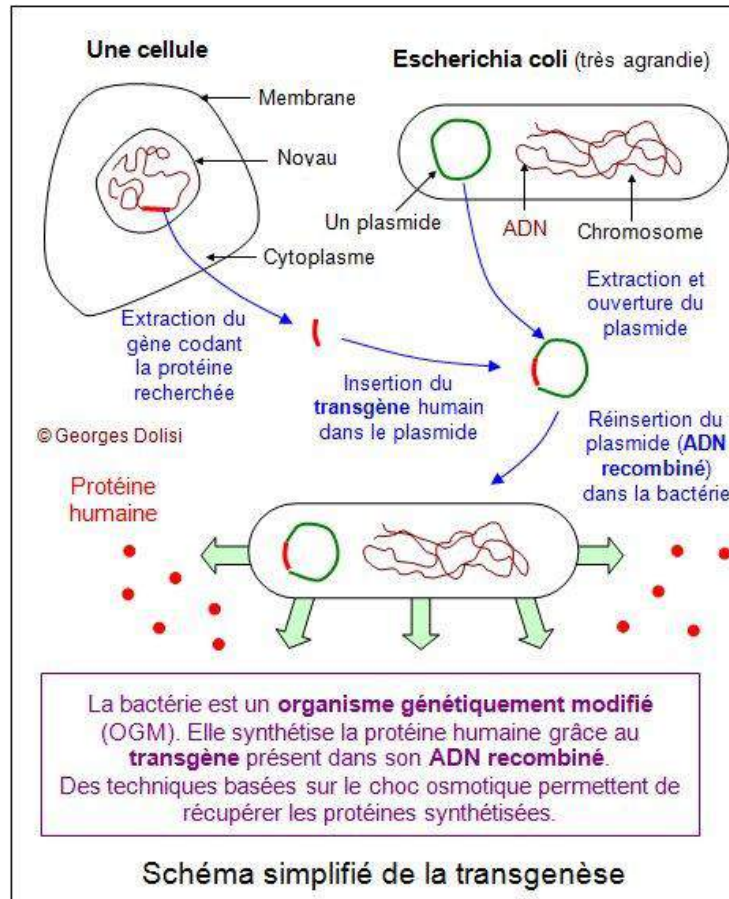
4- ETUDE DE L'EXPRESSION DES GENES EUCARYOTES DANS LES BACTERIES PAR TRANSGENESE

Les organismes vivants sont utilisés comme des usines de production qui permettent de valoriser une ressource génétique particulière.

L'expression des gènes clonés dans des systèmes procaryotes ou eucaryotes nécessite l'utilisation des vecteurs d'expression.

Quand on dispose d'une banque d'ADN d'intérêt (ADNg ou ADNc) que l'on veut faire exprimer, il faut transformer à nouveau des cellules avec le vecteur de clonage purifié. Cependant, pour que le gène s'exprime dans la cellule, il doit être entouré d'un ensemble de signaux que la cellule transformée devra reconnaître :

- **Promoteurs de transcription** : les vecteurs d'expression possèdent un ou plusieurs promoteurs qui doivent être puissants et aisément régulés. En effet, chaque cellule possède des promoteurs faibles de transcription de gènes codant pour des molécules requises en faible quantité par la cellule et des promoteurs forts pour des substances devant être produites en quantité importantes. De plus, l'expression d'un gène peut être induite ou réprimée par la présence d'un composé spécifique. Un des promoteurs le plus couramment utilisé est le promoteur Lac (séquence contrôlant la transcription du gène Lac Z codant pour la β -galactosidase bactérienne). Ce promoteur est induit par l'IPTG. Il est aussi possible d'utiliser d'autres promoteurs de transcription comme celui du gène de la tryptophane synthétase (Tryp E). Il est possible d'augmenter la puissance du promoteur par des séries de mutations ou de délétions dans la séquence promotrice. On arrive alors dans certains cas à obtenir une efficacité de transcription multipliée par un facteur 10. Le gène de la protéine à exprimer doit donc être cloné de telle sorte que sa transcription soit sous la dépendance de ce promoteur et introduit dans la bonne phase de lecture. Ensuite, l'expression peut être améliorée sous les conditions environnementales qui normalement activent le promoteur : c'est le cas de l'addition d'IPTG dans le cas du promoteur Lac.
- **Terminateurs de transcription** à la fin des gènes clonés : permettent la synthèse de longs transcrits non nécessaires exigeant de l'énergie et des structures secondaires indésirables pouvant se former dans le transcrit ce qui peut diminuer l'efficacité de la traduction du côté 5'.



Exemple 1 : Synthèse de l'insuline humaine par *E. coli*

L'insuline est une hormone peptidique constituée de 2 chaînes (A) et (B) reliées par des ponts disulfures et formées à partir de la proinsuline par clivage d'une partie de la chaîne peptidique. Utilisée pour le traitement des diabètes insulino-dépendants, elle fut tout d'abord extraite du bœuf (3 acides aminés différents avec l'insuline humaine) puis du porc (un seul acide aminé différent). Cependant, du fait de ces différences, ces insulines exogènes étaient immunogènes. **Le génie génétique** a permis de produire de l'insuline humaine à partir de bactéries génétiquement modifiées. La stratégie expérimentale suivante a été utilisée :

① **Isolement du gène de l'insuline** : L'ARNm codant pour la proinsuline est isolé de cellules du pancréas. L'ADNc est obtenu par la reverse transcriptase, l'ADNc bicaténaire est obtenu par la polymérase.

② **Vecteur de clonage**: Le plasmide pBR322. Le gène de résistance à la tétracycline contient un site de 6 bases reconnu par l'enzyme de restriction Bam HI. Il contient aussi un promoteur puissant.

③ **Insertion dans le vecteur** : Le plasmide est ouvert au site Bam HI. L'ADNc codant pour la proinsuline est flanqué de 2 linkers (séquence d'ADN bicaténaire artificiellement ajoutée) pour avoir la séquence de nucléotides appropriée à la ligation (compatibilité). La ligation des 2 molécules d'ADN (vecteurs et ADNc) se fait par une ligase.

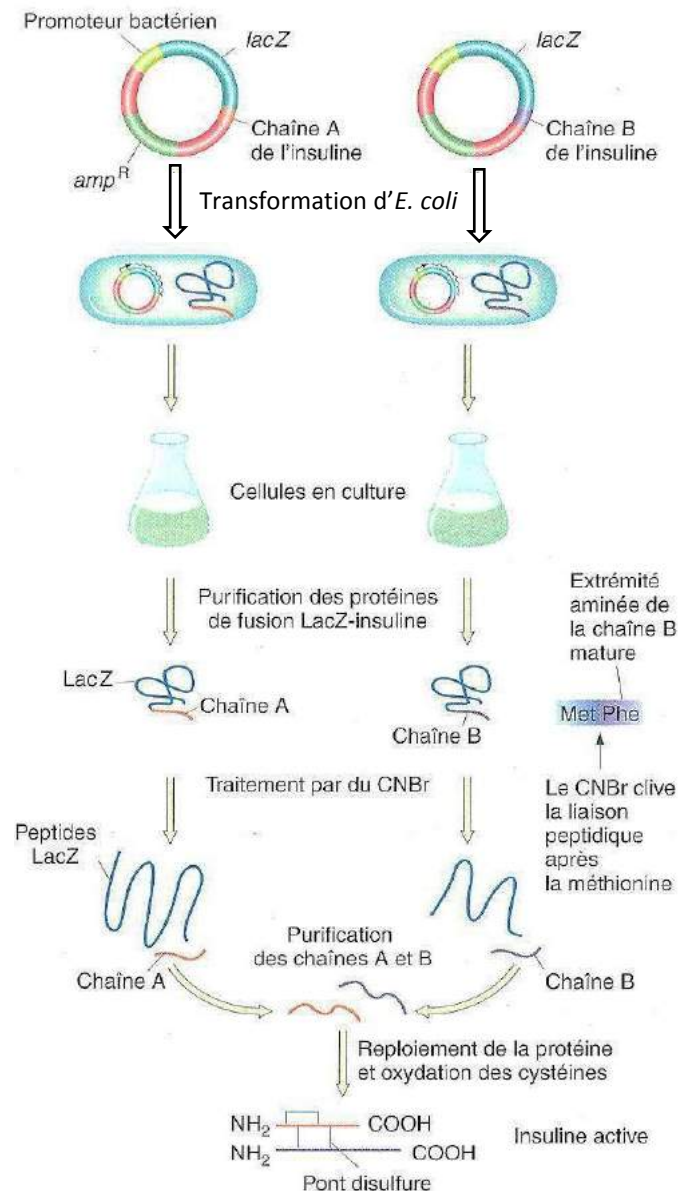
④ **Hôte** : cellule transformée : Le plasmide ainsi modifié est introduit dans une bactérie. La souche choisie est *E. coli* K12 initialement sensible à la tétracycline et à l'ampicilline.

⑤ **Sélection des clones** : Le plasmide pBR322 apporte initialement les 2 résistances à l'ampicilline et à la tétracycline. Cependant, l'insertion du gène de la proinsuline au site BamHI du plasmide ne lui permet pas d'apporter cette seconde résistance à la tétracycline. Ainsi, les colonies bactériennes capables de se développer sur un milieu contenant de l'ampicilline sont des bactéries ayant intégré le plasmide. Par réplique sur un milieu contenant de la tétracycline, on repère alors les colonies d'*E. coli* qui sont à la fois résistante

à l'ampicilline et sensible à la tétracycline ; Ces dernières sont alors celles qui ont intégré le plasmide transformé avec le gène de la proinsuline.

⑥ **Vérification de la production de proinsuline** : Il faut ensuite vérifier que les colonies sélectionnées produisent bien de la proinsuline. Pour cela, on applique sur les colonies choisies une membrane sur laquelle ont été fixés des anticorps capables de reconnaître la proinsuline. A l'aide d'un second anticorps marqué (méthode indirecte), on détecte alors les clones bactériens capables de produire de la proinsuline.

⑦ **Modification et caractérisation du produit** : Les clones sélectionnés sont mis à cultiver à grande échelle. Le produit obtenu est purifié puis converti en insuline par un mélange de trypsine et de carboxypeptidase B (enzymes protéolytiques). L'insuline est ensuite analysée par HPLC et sa stéréochimie (disposition dans l'espace des atomes qui la constituent) vérifiée. Enfin, des tests de toxicité et d'activité biologique sont effectués sur animal en mêmes temps que des études pharmacologiques et cliniques.

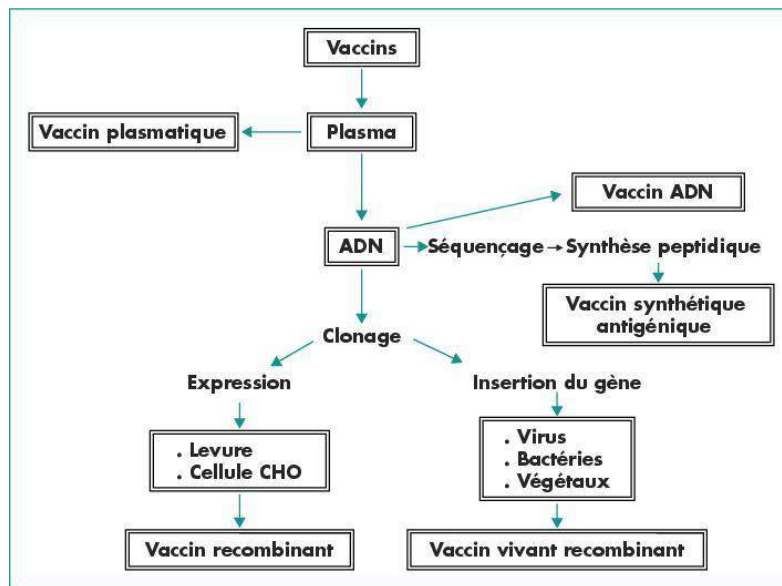


Expression de l'insuline humaine chez *Escherichia coli* : Les deux chaînes d'insuline sont synthétisées séparément sous forme de protéines de fusion avec la β -galactosidase. Elles sont ensuite libérées par un traitement chimique puis mélangées. L'insuline active est obtenue par une réaction d'oxydation chimique des cystéines. Actuellement l'insuline est produite en grande quantité par *E. coli* par une insertion de son gène dans un plasmide.

Autres applications possibles : de nombreuses molécules sont aujourd'hui synthétisées par génie génétique (vaccins, hormones, vitamines...).

Exemple 2 : Production du vaccin contre le virus de l'hépatite B par génie génétique

Le vaccin contre le virus de l'hépatite B (VHB) est obtenu par génie génétique. Avec le schéma actuel comportant deux injections distantes de 1 mois d'intervalle et un rappel à 6 mois, le pourcentage de séroconversion (désigne la phase au cours d'une maladie infectieuse où les anticorps apparaissent suffisamment dans le sang pour qu'on puisse les doser) se situe entre 99 et 99,5 %.



Voies actuelles de fabrication des vaccins

Cellules CHO (Chinese hamster ovary cells) sont des lignées cellulaires originaires de hamster chinois et sont actuellement les plus utilisées pour l'expression de protéines thérapeutiques recombinantes.

Exemple : La production du vaccin contre l'hépatite B est encore plus compliquée. Les virus de l'hépatite B ne se reproduisent pas dans les œufs de poule. Autrefois, les virus devaient être récoltés dans le sang de patients atteints d'hépatite et la production durait presque une année. Ce qui signifie que longtemps, le vaccin n'était disponible qu'en quantité restreinte. Aujourd'hui, c'est possible grâce au génie génétique (figure ci-dessous).

Le mode de fabrication du vaccin VHB présente les avantages suivants : son prix économique et absence de risque de contamination par des virus vivants.

Depuis 2007, les vaccins anti-grippe sont également produits par génie génétique et ce, par culture de cellules au lieu d'œufs de poule. De tels systèmes sont déjà utilisés dans la production de vaccins contre la variole, l'hépatite A et la poliomyélite. Le vaccin HPV (contre le papillomavirus humain) disponible depuis 2007 et efficace contre le cancer du col de l'utérus est lui aussi fabriqué par génie génétique dans des cellules de levure.

Digestion de l'ADN viral de l'hépatite B



Insertion dans un plasmide (de levure) du fragment d'ADN qui code pour l'antigène viral (qui déclenche sur l'individu infecté une réaction immunitaire contre le VHB)



Production de l'antigène par les cellules hôtes de levure



Antigène purifié



Application de l'antigène purifié en tant que vaccin



Production du vaccin anti-hépatite B par génie génétique

5- LA TECHNOLOGIE DE L'ADN RECOMBINANT CHEZ LES EUCARYOTES

Les techniques de manipulation, clonage et expression des gènes ont d'abord été développés chez les bactéries puis utilisés en routine chez de nombreux modèles eucaryotes. Les génomes des eucaryotes sont plus grands et plus complexes que ceux des bactéries d'où la complexité des techniques de manipulation.

Comparaison des différents systèmes d'expression

En étudiant les différents critères, la production de protéines recombinantes par les plantes semble être la plus adéquate. Les plantes auraient l'avantage de ne pas contenir de pathogène transmissible à l'homme, la quantité de protéines produite est largement supérieure à l'organisme unicellulaire pour un coût de production plus faible.

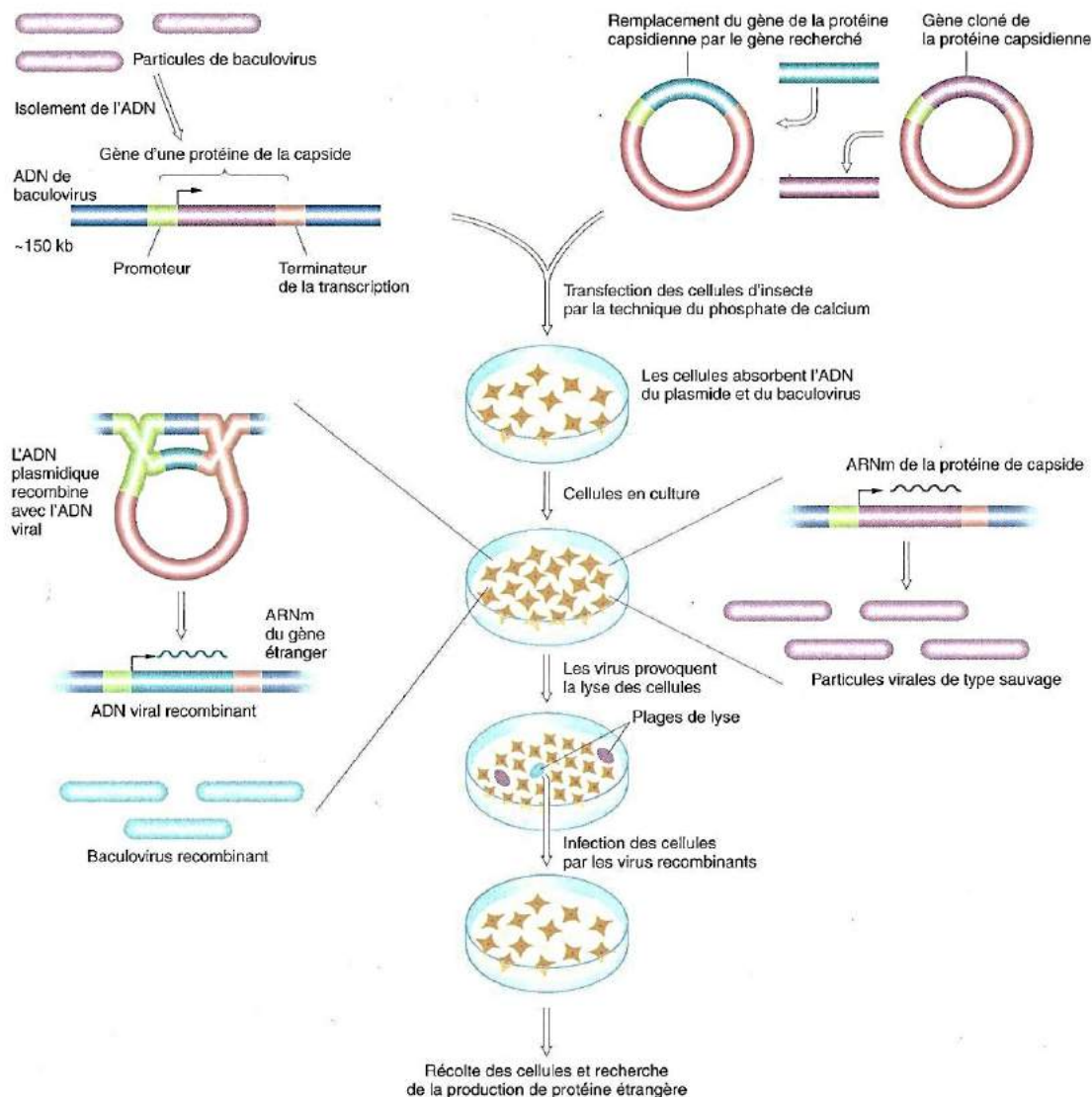
Comparaison des différents systèmes de production de protéines recombinantes

Système expression	Avantages	inconvénients	Applications	Coût/g
Bactéries	Voie de régulation établie; la génétique bien comprise; bon marché et facilité de développement	Les protéines non habituellement sécrétés; Contiennent endotoxines; pas modifications post-traductionnelles	Insuline (E. coli); hormone de croissance; facteur de croissance ; interféron	200 – 3000€
Levure	Reconnu sans risque, Longue histoire d'utilisation ; rapide; peu coûteuse; modifications post-traductionnelles	La sur glycosylation peut modifier la bioactivité et la sécurité; Peut contenir des immunogènes/antigènes	fermentation; vaccins recombinants contre le virus hépatite B ; insuline humaine	50–100€
Cellules d'insecte	Modifications post-traductionnelles; Protéines correctement formées; niveau assez élevé d'expression	Faible voie de régulation; croissance lente; moyen assez coûteux; Infection au bacillo-virus; virus de mammifères peuvent infecter les cellules	production de particules virales	
Cellules Mammifères	Protéines correctement formées; modifications post-traductionnelles correctes; bonnes connaissances des voies de régulation; le seul choix pour les grosses protéines	Moyen coûteux; croissance lente Peut contenir des allergènes/contaminant; purification compliquée	activateur tissu plasminogène; facteur VIII (glycoprotéine); anticorps monoclonal	500- 5,000€
Animaux Transgéniques	Processus utiles pour protéines complexes; Très haut niveau d'expression; coût de production faible	Peu d'expérience dans la régulation ; contaminations virales possibles; isolation dans les fermes	Lipase (mouton, lapin ; hormone croissance (chèvres); facteur VIII (bétail)	20–50€
Plantes Transgénique	Court cycles de développement; Stockage facile; bon niveau d'expression; pas de virus connue chez les plantes qui infecteraient les humains	Potentiel pour de nouveaux contaminants (champignon, bactéries, pesticides); modifications post-traductionnelles; peut contenir des allergènes	Vaccins pour le choléra (tabac); lipase gastrique (maïs); hépatite B (pomme de terre)	10–50€

Exemple ① : vecteur d'expression viral : les baculovirus (virus à ADN 150Kb) des insectes est largement utilisé pour la production des protéines eucaryotes. Des gènes sont insérés dans ces vecteurs et exprimés à des fréquences élevées dans des cellules d'insectes en culture. Bien que les gènes eucaryotes soient clonés et séquencés dans des hôtes bactériens, il est souvent souhaitable de réintroduire ce type de gène dans l'hôte eucaryote d'origine ou dans un autre eucaryote, d'où l'obtention d'un eucaryote transgénique.

Pour exprimer un gène étranger dans un baculovirus, ce gène est cloné à la place du gène de la capsid virale dans un plasmide contenant une petite partie du génome viral. Le plasmide recombinant est ensuite co-transfecté dans des cellules d'insectes avec de l'ADN de baculovirus de type sauvage. Les ADN plasmidique et viral recombinent à une fréquence faible au niveau des séquences homologues, conduisant à l'insertion du gène étranger dans le génome viral. Des plages de lyse virales se développent et les plages contenant les virus

recombinants ont un aspect différent en raison de l'absence de la protéine capsidienne. On prélève les plages contenant ces virus recombinants et on les laisse s'étendre. La réserve de virus est ensuite utilisée pour infecter une nouvelle culture de cellules d'insectes, aboutissant à une expression élevée de la protéine étrangère.



Actuellement, cette voie d'applications permet la création d'organismes transgéniques dans lesquels on introduit dès les premiers stades embryonnaires des gènes modifiés.

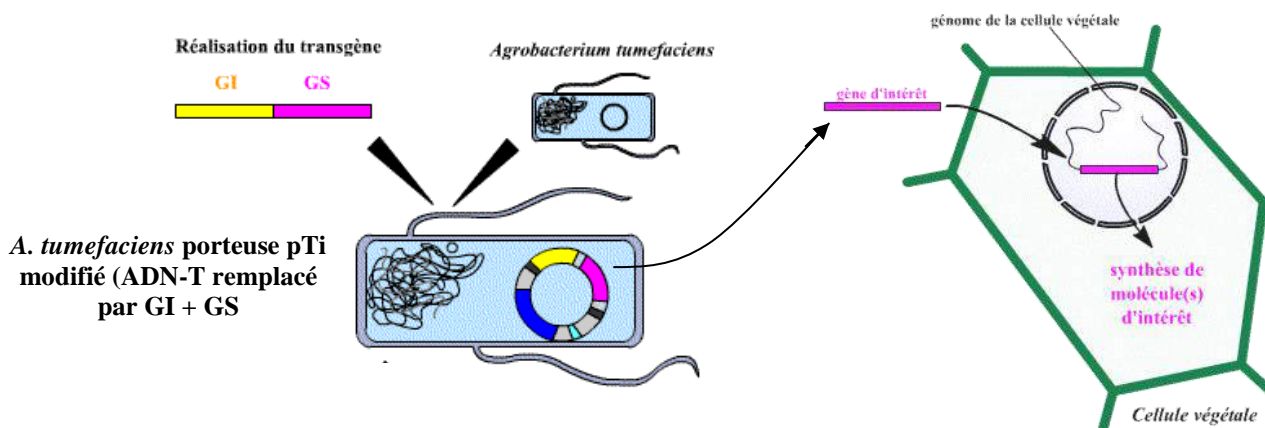
Exemple ② transgénèse végétale : Transformation biologique = Transfert d'un transgène dans une cellule végétale hôte par le plasmide Ti (pTi) modifié d'*Agrobacterium tumefaciens*

La bactérie *Agrobacterium* sp. est capable de réaliser naturellement la transformation génétique des plantes dicotylédones afin de les parasiter tel que l'espèce *A. tumefaciens* responsable de la **galle du collet (crown gall)** par le plasmide **pTi**. La bactérie se fixe à la cellule végétale et y transfère un petit fragment d'ADN (**ADN-T**) du pTi. Cet ADN-T s'intègre à l'ADN de la cellule végétale. L'information génétique portée par l'ADN-T s'exprime dans la cellule végétale transformée en y induisant une tumeur qui est due à l'expression de 2 hormones végétales (**auxine** et **cytokinine**) codées par des gènes de l'ADN-T. D'autres gènes de l'ADN-T codent pour la synthèse des **opines**, utilisés par les agrobactéries comme substrats de croissance azotée et carbonée.

D'autres gènes : **gènes vir** non transférés à la plante sont indispensables à l'expression du pouvoir pathogène et responsables de la reconnaissance de séquences

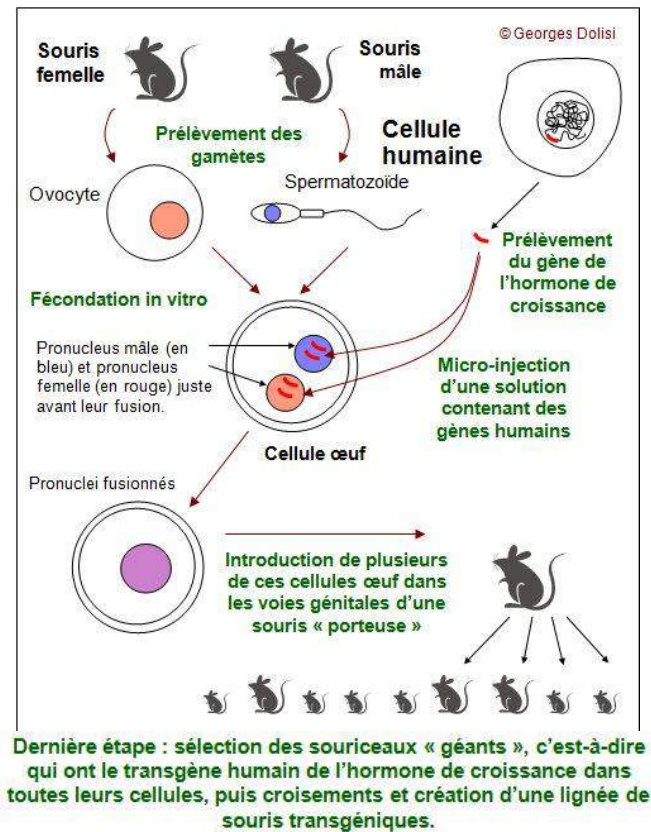
répétées situées aux extrémités de l'ADN-T appelées frontières. Une des protéines codées par les gènes *vir* possède une activité endonucléase qui conduit à l'excision de l'ADN-T sous forme d'ADN simple brin qui va recombinaison avec l'ADN génomique de la plante. L'interaction *Agrobacterium* / plante permet donc le transfert d'ADN de la bactérie vers la plante et l'insertion dans le génome de la plante de n'importe quel fragment d'ADN placé entre les frontières de l'ADN-T → ceci constitue le fondement de tous les protocoles de génie génétique faisant intervenir *Agrobacterium*.

Le pTi est utilisé comme vecteur de la construction génétique désirée. Les gènes oncogènes sont remplacés par le gène d'intérêt agronomique + un ou plusieurs gènes de sélection : c'est le plasmide modifié → une construction génétique introduite dans la bactérie avirulente, sera transférée dans la plante et intégrée à son génome. Cependant, Il existe quelques plantes telle que les céréales, qui sont insensibles aux bactéries de la galle du collet et ne peuvent donc pas être ainsi modifiées. D'où l'utilisation d'autres techniques de transfert de gènes d'intérêt : Biolistique (canon à particules), Electroporation (electrotransfert) et Microinjection.



Les systèmes utilisant des plantes transgéniques permettent de produire des protéines en quantités très importantes. Cependant, cultivés en milieu « ouvert », des risques de contaminations physico-chimiques (pesticides...) ou biologiques (mycotoxines, transmission de virus animaux par les déjections) provenant de l'environnement, ou encore de virus végétaux, la possibilité de transmission et de pathogénicité vis-à-vis de l'homme ne sont pas à exclure. Ils peuvent également poser des problèmes d'allergies ou de néo-antigénicité en raison des différences de glycosylations des protéines végétales et des protéines animales. Enfin, le risque de dissémination dans l'environnement d'une plante transgénique codant pour une protéine-médicament ne semble pas majeur, même s'il ne doit pas être écarté.

Exemple ③ Transgénèse animale : Micro-injection du gène de l'hormone de croissance d'un rat ou d'un humain dans les œufs de souris. Les pourcentages de réussite (animaux qui possèdent le transgène dans toutes leurs cellules) est de l'ordre de 20 à 30% pour les souris. Il est bien moindre pour d'autres espèces sur lesquelles on a essayé la transgénèse.



6- LA THÉRAPIE GÉNÉTIQUE

Une autre voie d'application est celle de la thérapie génique permettant d'introduire des gènes sains dans un organisme contenant des gènes défectueux provoquant des maladies génétiques. Le principe de la thérapie génique **somatique** (Seule la transgénèse des cellules somatiques est autorisée en France) consiste à introduire un allèle fonctionnel dans les cellules cibles impliquées dans la pathologie, grâce à des virus vecteurs, modifiés et inoffensifs. Le transgène introduit dans les cellules somatiques n'est pas transmis à la descendance. La thérapie génique **germinale** (introduction d'un gène médicament dans une cellule œuf ou un embryon malade) n'est pas autorisée en France (Problèmes d'éthique).

La thérapie génique est aujourd'hui envisagée pour le traitement d'une grande variété de maladies acquises et héréditaires. Différents systèmes viraux ont été développés pour un usage *ex vivo* (tests biologiques mis en place en dehors de l'organisme) et *in vivo*, comme les vecteurs rétroviraux, adénoviraux, adénoviraux-associés, vecteurs dérivés du virus de l'herpès simplex, des poxvirus et des alphavirus.

Elle est basée sur la construction de vecteurs de transfert de gènes permettant l'introduction de gènes thérapeutiques dans des cellules cibles du patient. Elle permet de conférer la capacité aux cellules cibles d'agir comme des réservoirs de protéines thérapeutiques manquante (cellule déficiente), de modifier les fonctions cellulaires ou les programmes de développement, d'empêcher l'apoptose (mort cellulaire programmée), de sensibiliser des cellules malades à des agents thérapeutiques, ou d'induire la mort cellulaire (cellule infectée ou cancéreuse).

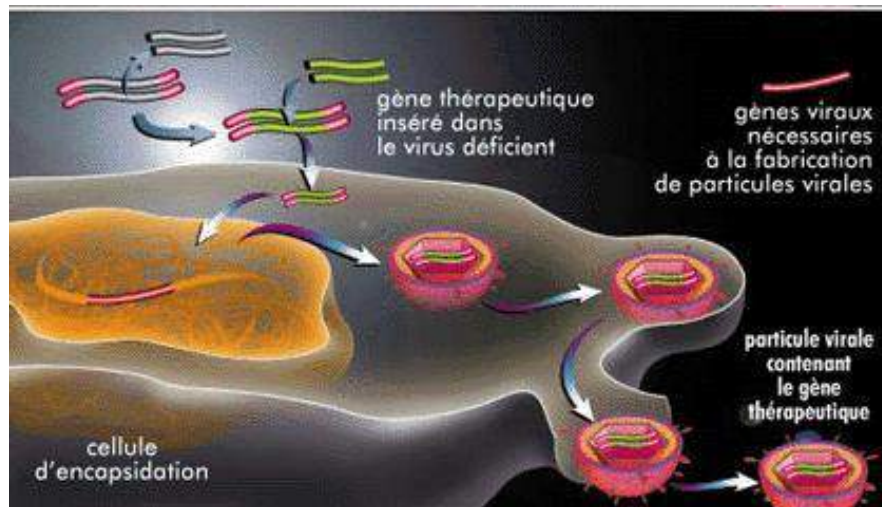
Étapes de la thérapie génique

1. Isoler et cloner le gène d'intérêt thérapeutique.

2. construction d'un vecteur chargé d'amener le transgène dans le noyau cellulaire.

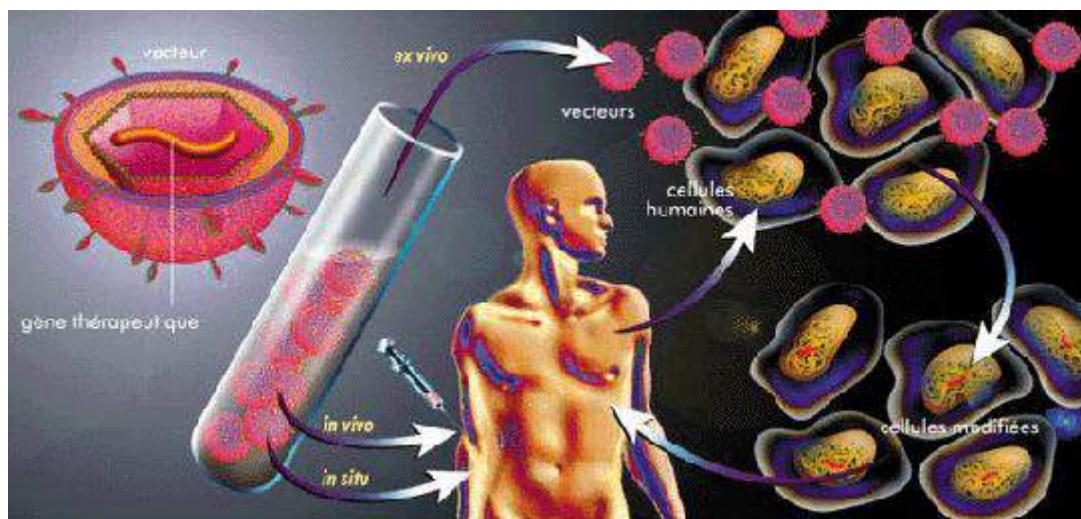
Le vecteur est généralement un virus. Pour produire des vecteurs viraux, on utilise des cellules modifiées, dites d'encapsidation. Normalement, ces cellules expriment de façon

stable les protéines virales formant la capsid. En absence de génome viral, elles ne produisent que des particules virales vides. L'introduction dans ces cellules d'une construction génétique contenant le génome viral porteur du gène thérapeutique conduit à la formation de particules virales complètes contenant le vecteur.



3. Administrer le vecteur selon un protocole :

- la thérapie génique **ex vivo** : consiste à prélever sur le patient les cellules cibles, à les modifier génétiquement avec le vecteur viral porteur du gène d'intérêt thérapeutique, puis à les réintroduire chez le patient. Cette méthode est utilisée en particulier pour les cellules sanguines qui sont faciles à prélever et à réintroduire.
- la thérapie génique **in situ** : le vecteur de transfert est directement injecté au sein du tissu cible
- la thérapie génique **in vivo** : consiste à injecter le vecteur portant le gène d'intérêt thérapeutique directement dans les organes cibles donc dans la circulation sanguine, celui-ci devant atteindre spécifiquement les cellules cibles.



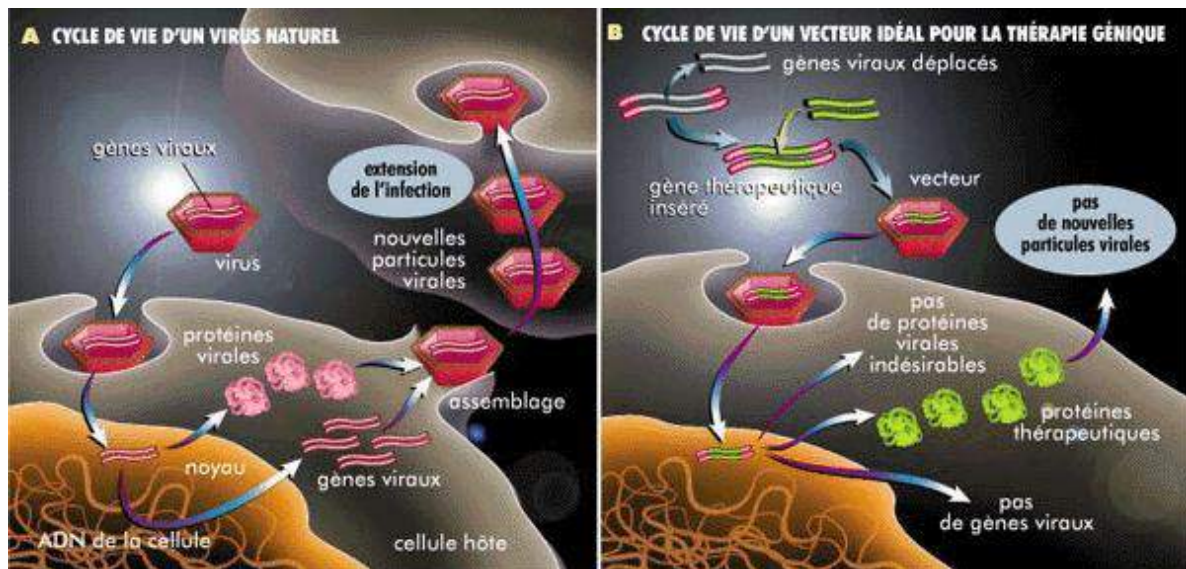
4. Vérifier l'intensité et la durée de l'expression du gène thérapeutique ainsi que les éventuels effets secondaires.

Comment un virus ou un vecteur viral pénètre dans les cellules :

A) Pour infecter une cellule, la particule virale se fixe d'abord à la membrane cellulaire. Les gènes viraux sont libérés dans le noyau et, qu'ils soient intégrés ou non au génome

cellulaire, ils utilisent la machinerie de réplication de la cellule pour produire de nouvelles particules virales. Ces dernières pourront aller infecter d'autres cellules.

B) Lorsque le virus est modifié pour être utilisé comme vecteur de transfert, les gènes codant pour les protéines virales sont remplacés par le gène d'intérêt thérapeutique. Ce vecteur pénètre dans la cellule de la même façon que le virus naturel. Il permet la synthèse de la protéine d'intérêt thérapeutique, sans production de particules virales.



Utilisation des vecteurs d'expression pour la thérapie génique des cancers

La description des premiers résultats, obtenus lors d'essais cliniques de thérapie génique du cancer, montrent la nécessité d'améliorer les vecteurs de transfert de gènes actuels. De nombreuses études visent à augmenter les titres viraux obtenus, à administrer le plus efficacement possible les gènes thérapeutiques aux cellules cibles, à améliorer l'expression à long terme du transgène, à cibler la transduction sur un type cellulaire donné et à réduire l'immunogénicité des vecteurs.

Systemes viraux de transfert de gènes utilisés en thérapie génique

Systeme viral	Expression <i>in vivo</i>	Propriétés intéressantes
Rétrovirus (autre que lentivirus)	Durable	Intégration du génome viral dans le génome des cellules en division
Lentivirus	Durable	Intégration dans le génome de certaines cellules quiescentes et en division ; système de production encore mal caractérisé
Adénovirus	Transitoire	Haute efficacité <i>in vivo</i> , production à hauts titres, tropisme modifiable, induit une réaction inflammatoire et est immunogène, peut être produit en tant que vecteur répliquatif
Poxvirus	Transitoire	Haute efficacité <i>in vivo</i> , effets cliniques du virus parental étudiés extensivement, grande capacité d'insertion, induit une réaction inflammatoire et est immunogène, production à hauts titres
Virus associés à l'adénovirus (AAV)	Durable	Non pathogène, faible capacité d'insertion, production à hauts titres
Virus de l'herpès (HSV 1)	Transitoire	Haute efficacité <i>in vivo</i> , grande capacité d'insertion
Vecteurs chimères	Durable	Combinaison des propriétés de deux vecteurs (adénovirus + rétrovirus)

La capacité naturelle des virus à délivrer leur matériel génétique aux cellules qu'ils infectent a été exploitée pour mettre au point les premiers systèmes de transfert de gènes. Environ 68% des essais cliniques de thérapie génique du cancer, et notamment ceux utilisant des stratégies immunes, sont réalisés au moyen de vecteurs viraux. Cependant, les vecteurs viraux utilisés chez l'homme présentent des inconvénients : toxicité à fortes doses, immunogénicité et possibilité de générer des virus sauvages durant le processus de production. De nombreux travaux visent donc à améliorer ces vecteurs afin de les cibler sur

un type cellulaire particulier, d'augmenter conditionnellement leur potentiel infectieux, de diminuer leur immunogénicité.

Exemple : Oncorétrovirus

Les oncorétrovirus sont les 1^{ers} vecteurs de transfert de gènes vu leur capacité à s'intégrer dans le génome de l'hôte pour traiter les maladies génétiques monogéniques. Ce sont des dérivés des oncorétrovirus murins (MoMLV pour Moloney murine leukemia virus) constitués de particules enveloppées contenant deux molécules identiques d'ARN monocaténaire linéaire de polarité positive, de la transcriptase inverse, de l'intégrase et de la protéase. Le génome comprend à ses extrémités deux régions répétées : **LTR** (long terminal repeat) qui contiennent les séquences **U3**, **R**, **U5** nécessaires à l'intégration chromosomique et le **site d'initiation de la transcription** avec une **séquence « activatrice »** contenue dans la région U3, le **gène gag** (code les protéines de capsid, produites sous forme d'un unique précurseur clivé ensuite en 3 à 5 protéines : matrice, nucléocapside, capsid), le **gène pol** (codant les protéines à activité enzymatique : la transcriptase inverse, une intégrase permettant l'intégration du provirus dans l'ADN cellulaire), le **gène pro** (codant une protéase responsable du clivage des polyprotéines gag et pol) et le **gène env** (codant la glycoprotéine d'enveloppe qui sera clivée pour générer une protéine de surface SU et une protéine transmembranaire TM).

Après pénétration d'un virion dans la cellule cible, l'ARN viral est rétro-transcrit en ADN par l'action de la transcriptase inverse. Cette molécule d'ADN double brin est alors intégrée, de façon aléatoire, dans le génome de la cellule hôte et les gènes peuvent être transcrits par l'utilisation de la machinerie cellulaire.

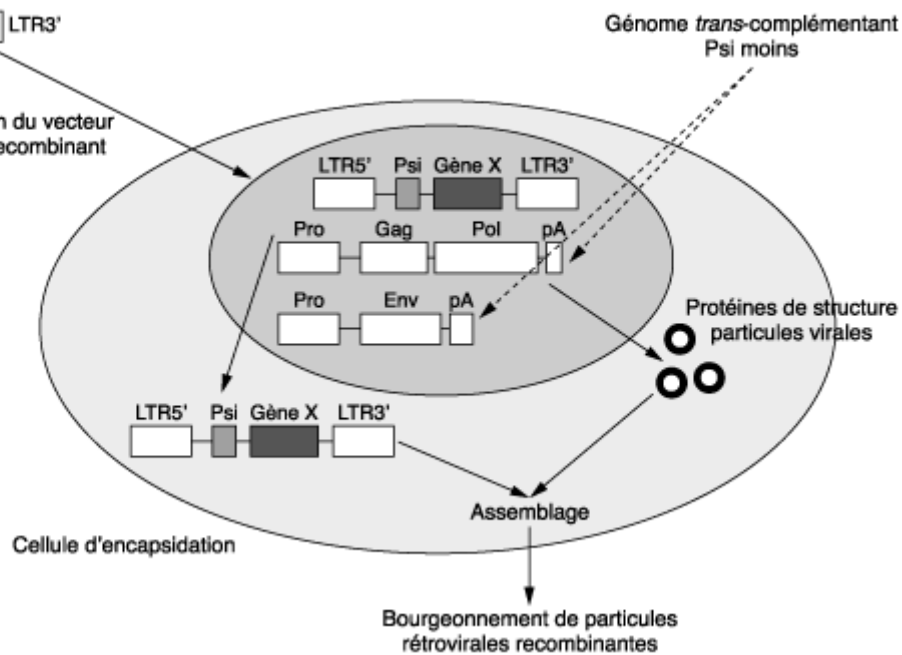
Dans les oncorétrovirus recombinants, les séquences virales LTR nécessaires en cis sont conservées alors que les régions codant les protéines virales gag, pol et env sont délétées pour permettre l'insertion de séquences exogènes allant jusqu'à 8 kb. Le génome recombinant porte ainsi la cassette d'expression du transgène ou gène d'intérêt et, souvent, une cassette de sélection (gène de résistance à un antibiotique, marqueur de surface). L'utilisation d'une séquence IRES (internal ribosome entry site = séquences qui permettent dans les cellules eucaryotes, le démarrage de la traduction d'un ARNm de manière interne) permet de placer plusieurs gènes sous le contrôle des mêmes séquences régulatrices. Les protéines virales nécessaires à la réplication sont fournies en trans par les lignées d'encapsulation, conçues pour diminuer les risques de recombinaison avec le génome recombinant. Par exemple, la séparation des gènes viraux gag-pol et env sur deux plasmides d'expression différents augmente le nombre d'événements de recombinaison nécessaire à la génération d'un génome compétent et diminue ainsi le risque d'apparition d'oncorétrovirus compétents pour la réplication (RCR). L'utilisation de cellules humaines pour générer les lignées d'encapsulation de vecteurs murins permet de diminuer le risque de RCR et la fréquence d'encapsulation de séquences rétrovirales endogènes. L'origine humaine de la lignée productrice permet de préserver les particules rétrovirales de l'inactivation par le complément après injection in vivo chez l'homme. Finalement, les dernières avancées dans ce domaine de recherche ont permis, très récemment, de produire des vecteurs rétroviraux dépourvus de toute séquence codante virale, empêchant ainsi théoriquement toute génération de RCR.

Séquence

d'encapsidation Psi Gène thérapeutique



Transfection du vecteur rétroviral recombinant

**Production de particules oncorétrovirales recombinantes dans une lignée de complémentation**

Problèmes :

- méthode de transfert peu efficace
- expression du transgène limitée dans le temps
- insertion du transgène non ciblée (se fait n'importe où dans le génome)

Bien qu'elle ne soit encore qu'à ses balbutiements, la thérapie génique s'est, depuis quelques années, notamment axée sur une possibilité de traitement de la mucoviscidose ou de certains cancers.

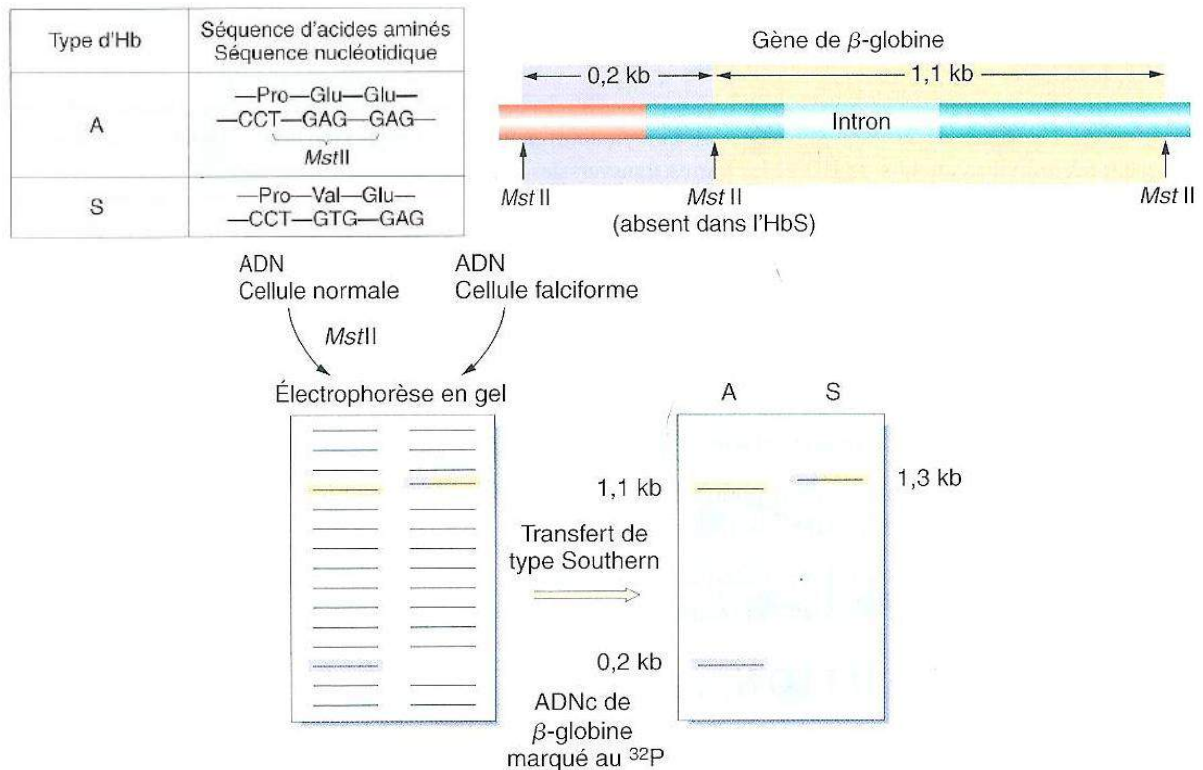
7- UTILISATION DE L'ADN RECOMBINANT POUR DETECTER DIRECTEMENT LES ALLELES RESPONSABLES DE MALADIES

Le dépistage des maladies génétiques (protéines modifiées ou absentes) peut être réalisé par **amniocentèse** (prélèvement, séparation et culture de cellules fœtales du liquide amniotique pour analyser les chromosomes, les protéines, les réactions enzymatiques et autres propriétés biochimiques) ou par **analyse des villosités chorioniques** (CVS : chorionic villus sampling : aspiration des cellules du placenta). Cependant, l'étude des cellules en culture se limite aux maladies qui affectent les caractères ou les protéines exprimées dans les cellules en culture. D'où la technologie de l'ADN recombinant qui élargit la recherche des maladies génétiques *in utero*, puisque l'ADN est directement analysé.

Exemple : Etude des modifications des sites de restriction par mutation dans le cas de la drépanocytose

La drépanocytose ou anémie à cellules falciformes (0,25% des Noirs américains) est une maladie génétique qui se traduit par une hémoglobine modifiée due à la substitution d'un acide glutamique (Glu : GAG) par une valine (Val : GTG) en position 6 de la chaîne de β -globuline. Ce changement supprime un site de coupure pour l'enzyme de restriction MstII qui coupe la séquence CCTNAGG. Le changement CCTG**A**GG en CCTG**T**GG est détecté par

analyse par transfert selon Southern en utilisant comme sonde de l'ADNc marqué de β -globuline.



La détection du gène d'hémoglobine responsable de l'anémie à cellules falciformes, par hybridation sur un transfert de type Southern. La substitution A>T à l'origine de la maladie détruit un site cible de MstII présent dans le gène normal de la β -globuline. Ainsi, l'ADN provenant de personnes atteintes a un fragment d'ADN de moins que celui des personnes saines et présente un grand fragment (non clivé) invisible dans l'ADN normal.

Résumé

- Les gènes clonés peuvent être modifiés par l'introduction de changements mutationnels spécifiques au niveau de sites précis, à l'aide de la technique de mutagenèse dirigée.
- Les RFLP sont très utiles comme marqueurs, notamment pour la cartographie des chromosomes. La liaison génétique d'un RFLP à un gène de maladie humaine est un outil utile pour conseiller de futurs parents et diagnostiquer des maladies in utero.
- Dans la génétique inverse, on part d'un gène ou d'une protéine de fonction inconnue, on induit un changement spécifique (tel que l'inactivation complète d'un gène) puis on suit l'évolution du phénotype pour essayer de comprendre la fonction du gène.
- La biotechnologie est l'utilisation des techniques de l'ADN recombinant pour modifier des animaux, végétaux ou micro-organismes importants dans le commerce. Les organismes transgéniques, modifiés par l'insertion d'ADN exogène spécifique, occupent une place centrale en biotechnologie, car ils permettent des modifications hautement spécifiques du génome.
- La thérapie génique humaine est une application particulière de la technologie transgénique. La thérapie génique germinale a pour but d'incorporer des cellules transgéniques dans la lignée germinale, pour qu'elles soient transmises aux descendants. La thérapie somatique consiste à introduire des cellules génétiquement modifiées dans le corps.
- L'amniocentèse et l'analyse des villosités chorioniques utilisent les techniques de l'ADN recombinant pour diagnostiquer directement la présence d'un allèle de maladie.

CHAPITRE II

LES APPLICATIONS DE LA TECHNOLOGIE DE L'ADN RECOMBINANT

1- Les applications de la transgénèse

Les applications de la transgénèse peuvent être regroupées dans quatre grands domaines :

- améliorations agronomiques et environnementales
- qualités alimentaires
- production de molécules à intérêt industriel
- production de molécules destinées à la santé humaine



① L'agriculture et l'environnement

De nombreux travaux de transgénèse concernent l'introduction de gènes de résistance aux herbicides ou aux insectes, et dans une moindre mesure, à certains virus et maladies. Associées à un usage raisonné d'herbicides et de pesticides, ces plantes transgéniques vont améliorer l'efficacité de l'agriculture, tout en respectant encore mieux l'environnement.

- La résistance des plantes à des insectes ravageurs

La lutte contre les ravageurs, notamment les insectes, est réalisée essentiellement par l'utilisation d'insecticides chimiques. Suivant les cultures, les zones géographiques et les années, la fréquence et la sévérité des attaques d'insectes sont très variables. Selon les cas, l'application d'insecticide est systématique ou décidée sur la base de comptages des insectes (ou de larves) "nuisibles" lors de contrôles dans les champs.

La transgénèse offre aujourd'hui un outil supplémentaire aux agriculteurs pour limiter les traitements chimiques et protéger leurs récoltes contre les insectes et les maladies et ainsi réduire les pertes.

Pour rendre une plante résistante à un insecte, un gène codant une protéine toxique pour cet insecte a été introduit dans le génome de la plante. Jusqu'à présent, pour toutes les variétés mises sur le marché, les gènes introduits proviennent de la bactérie *Bacillus thuringiensis* (d'où l'abréviation Bt donnée aux plantes de ce type), bien connue depuis

longtemps pour ses propriétés insecticides et largement utilisée en agriculture biologique, ainsi que par les exploitants forestiers et les jardiniers. Cette bactérie constitue donc un véritable réservoir de gènes de résistance aux insectes. En effet, les différentes souches de cette bactérie du sol recèlent plusieurs protéines insecticides ayant différents modes d'action, et affectant uniquement certains insectes. Chacune de ces protéines est codée par un seul gène, c'est donc un caractère facilement transférable par génie génétique. Plusieurs équipes ont obtenu des tabacs, des pommes de terre, des cotons, des tomates, des maïs résistants à des insectes grâce à cette source de gènes.

Dans le cas du maïs, la résistance à la pyrale est conférée par le gène Cry A, appelé communément Bt. Ce gène permet dans les cellules du maïs, la production d'une protéine qui se transforme en toxine dans le tube digestif de la pyrale. Chez les autres animaux et chez l'homme, cette protéine est simplement digérée sans aucun effet toxique.

Une étude sur l'impact du coton bt (résistant aux insectes) montre qu'en 1999, les agriculteurs chinois ayant adopté des variétés bt ont consommé en moyenne 10 kg/ha d'insecticides contre 58 kg/ha pour les agriculteurs ayant cultivé des variétés non transgéniques.

- La résistance à des maladies

Les virus, les champignons et les bactéries sont responsables de pertes importantes en production végétale. Or, il n'existe aucune méthode de traitement des maladies dues à des virus chez les plantes cultivées. Par transgénèse, il est possible d'obtenir des plantes résistantes aux virus. Ces plantes transgéniques synthétisent des protéines qui bloquent la multiplication et le développement des virus. Ainsi, il a été possible d'obtenir des courgettes et des melons résistant au virus de la mosaïque du concombre. L'obtention de plantes résistant aux champignons et aux bactéries est en cours de développement.

- La résistance à des herbicides

Le glufosinate (Basta ou Liberty) et le glyphosate (Roundup) sont des herbicides totaux qui détruisent aussi bien les mauvaises herbes que les plantes cultivées. Les gènes de résistance à l'herbicide introduits dans une plante empêchent la matière active d'agir sur celle-ci, transformant l'herbicide total en herbicide sélectif sur cette plante. Ainsi l'herbicide détruit toutes les mauvaises herbes présentes tout en respectant totalement la plante cultivée. De plus, ces désherbants totaux ont la propriété de ne pas être rémanents. De nombreuses plantes transgéniques ont été développées pour obtenir une tolérance à ces herbicides. Il s'agit de variétés de betterave, colza, coton, maïs, pomme de terre et de soja.

La création de plantes tolérantes aux herbicides permet l'utilisation de matières actives au profil écotoxicologique favorable, c'est à dire à faible durée de vie, à biodégradabilité rapide, respectant l'environnement et à large efficacité. Ces cultures peuvent supporter ce traitement grâce à l'introduction d'un gène de tolérance spécifique. En 1996, un nouveau système de désherbage a été lancé en Amérique du Nord sur des cultures comme le soja, le colza et le maïs.

Une étude sur 5 ans (1996 à 2001) auprès de 450 cultivateurs américains de soja montre que pour 63 % d'entre eux, le développement des techniques culturales sans labour, qui permet une réduction de l'érosion des sols de l'ordre de 90 %, est rendu possible en premier lieu par l'introduction de variétés de soja transgéniques tolérant à un herbicide.

- L'enrichissement du patrimoine végétal

La sélection classique a déjà fait la preuve de sa capacité à enrichir les espèces et variétés. Par la création de variétés nouvelles qui constitue l'objectif premier de son activité, la sélection classique a ainsi doté le "patrimoine végétal" de spécimens nouveaux. Dès les prémices de l'agriculture, cette activité de sélection empirique a été à l'origine de nombreuses variétés aujourd'hui partie intégrante de ce patrimoine.

- Le maïs transgénique utilisant mieux l'eau disponible

L'amélioration des rendements du maïs est due à un grand nombre de facteurs, dont l'un des plus importants est la tolérance aux stress environnementaux. Cette amélioration est surtout attribuée à une augmentation de la performance des lignées parentales, qui sont en partie tolérantes à un déficit d'eau. En effet, pour le maïs, la période pendant la pollinisation et le début du remplissage des grains est la plus sensible à un stress hydrique. Pourtant, l'eau est une ressource limitée dont l'agriculture est la première utilisatrice, devant l'industrie et la consommation humaine. Pour la culture du maïs, une façon de diminuer l'utilisation d'eau consiste à créer des variétés qui tolèrent une disponibilité réduite en eau, sans que leurs capacités de production n'en soient affectées. Grâce à des études génétiques, plusieurs variétés de maïs transgéniques ont été créées par l'introduction de gènes impliqués dans la réponse à un déficit hydrique. Par exemple, un maïs plus tolérant à la sécheresse a été mis au point grâce à l'introduction par transgénèse d'un gène de sorgho, céréale africaine particulièrement tolérante à la sécheresse. Ce gène code pour une protéine impliquée dans la photosynthèse, la PEPc (Phosphoenolpyruvate carboxylase). Les plantes transgéniques obtenues surexpriment cette protéine. Des analyses du comportement photosynthétique de ces plantes en situation de contrainte hydrique en serre ont permis de montrer que l'efficacité d'utilisation de l'eau est significativement augmentée (+ 25 %).

Conclusion : Le recours à des variétés transgéniques permet ainsi une moindre utilisation d'insecticides et d'herbicides et ouvre le champ de la recherche sur les pratiques culturales simplifiées, d'où le respect de l'environnement

② L'alimentation

Il s'agit de modifier la composition d'une plante afin de lui apporter des avantages nutritionnels et gustatifs ou de lui conférer de nouvelles caractéristiques qui permettent de diversifier les débouchés.

- Les qualités nutritionnelles

En alimentation animale, les recherches vont dans le sens d'un développement de plantes permettant un meilleur rendement nutritionnel et évitant l'apport de compléments nutritifs. Ainsi, il est possible d'obtenir des plantes de maïs, colza, soja à teneurs élevées en acides aminés, notamment en méthionine et lysine, et des maïs enrichis en huile. Concernant l'alimentation humaine, des travaux sont menés pour diminuer les propriétés allergènes du riz et du soja. Pour obtenir ce résultat, on cherche à introduire dans la plante un transgène qui inhibe la synthèse de la protéine allergisante.

- La maturation des fruits

Ce sont les résultats les plus avancés concernant la qualité alimentaire. Sur le melon et la tomate, on a pu obtenir des variétés transgéniques à maturation retardée. Ces fruits peuvent être récoltés à un stade de maturation plus avancé, donc être plus savoureux. D'autre part, il en résulte une meilleure conservation et une aptitude au transport améliorée, réduisant les pertes. Le melon est le premier fruit génétiquement modifié obtenu par un laboratoire de recherche français. Un gène capable de bloquer la synthèse de l'éthylène a été introduit, ce qui ralentit la maturation. Le détachement du fruit est retardé et le melon maintenu sur pied continue d'accumuler des sucres.

- La transformation agro-alimentaire

Dans ce domaine, les champs d'application potentiels sont très variés : il peut s'agir de la production des protéines impliquées dans des procédés agro-alimentaires, ou de la modification des caractéristiques des végétaux pour optimiser leur utilisation. Ainsi, des travaux ont permis de modifier la teneur en amidon chez la pomme de terre, afin

d'augmenter la teneur en matière sèche, et de disposer ainsi de pommes de terre mieux adaptées à la fabrication de féculé, de purée ou de chips. Des gènes ont également été transférés chez le colza pour modifier la teneur en acides gras ou pour obtenir des huiles contenant des nouveaux acides gras recherchés en alimentation humaine.

③ L'industrie

Les biotechnologies ouvrent de nombreuses perspectives dans les domaines de l'industrie, en produisant des molécules nouvelles (Molecular Farming) et en améliorant les procédés industriels et la qualité des produits.

- Les pâtes à papier

Les lignines sont l'un des constituants majeurs du bois, mais elles gênent l'industrie papetière qui ne peut les valoriser et doit les éliminer par des méthodes coûteuses et polluantes. Des travaux conduits par la recherche publique française ont permis de connaître les gènes impliqués dans la synthèse des lignines et de développer des variétés de peupliers transgéniques, chez lesquels le taux de lignine est fortement réduit. Ceci facilite le blanchissement de la pâte à papier et donc réduit l'impact sur l'environnement. Le même type de travail a été réalisé sur l'eucalyptus.

- Les huiles industrielles

Elles sont synthétisées à partir de matières premières fossiles, dont les ressources sont limitées. Il est donc nécessaire de s'orienter vers d'autres ressources renouvelables. Parmi les nombreux programmes de recherche, on peut citer celui destiné à l'obtention d'un colza transgénique à haute teneur en acide gras érucique ou ricinoléique pour la production de lubrifiants, de matières plastiques, etc. Cette stratégie devrait favoriser le développement de lubrifiants et de plastiques biodégradables.

- Les colorants

Un exemple original est l'obtention de cotons transgéniques de couleur grâce à l'introduction d'un gène bactérien ou végétal codant pour un pigment. Ceci évitera l'utilisation de teintures chimiques difficilement recyclables.

④ La santé

Génétiquement modifiées, des plantes de tabac, de maïs, ou de pomme de terre peuvent produire des molécules thérapeutiques ou des vaccins. Le grand avantage de la production de ces molécules est l'absence de risques de contamination par des virus pathogènes pour l'homme.

- Les produits sanguins

Des recherches menées en France ont déjà permis de faire produire des protéines plasmatiques à des plants de tabac transgéniques, permettant l'obtention d'hémoglobine humaine recombinée. Des travaux montrent qu'il est possible de synthétiser de l'albumine humaine, employée lors du traitement des traumatismes, à partir de tabac ou de pomme de terre. Cette albumine devrait être moins chère que celle issue du plasma sanguin. Cette nouvelle source permettrait de répondre à l'augmentation des besoins.

- Les vaccins

Des chercheurs américains travaillent à la mise au point d'une banane vaccin pour l'homme, prévenant les cas de gastro-entérites provoquées par la bactérie *E. coli*. Il serait alors envisageable de vacciner à faible coût les populations de pays en voie de développement, les plus touchées par ces diarrhées d'origine bactérienne.

- Les protéines humaines

Des travaux sont actuellement en cours pour faire produire des protéines ou des glycoprotéines à usage thérapeutique à partir de soja, de tabac, de pomme de terre, de riz ou de colza.

2- Dépistage et diagnostic génique

Les biotechnologies chez l'Homme servent à diagnostiquer une maladie.

On appelle maladie génique, une maladie transmise des parents à l'enfant par l'intermédiaire des gènes portés par les gamètes. L'examen d'un arbre généalogique permet parfois **d'évaluer le risque statistique** de maladie génique chez un individu ou un fœtus. Le diagnostic génique se fait à partir de l'utilisation de sondes spécifiques et d'enzymes de restriction. Ces maladies géniques sont :

- soit **autosomique récessive** (relatif à un chromosome non sexuel). Les malades sont homozygotes pour l'allèle morbide. Exemple : la Mucoviscidose (troubles respiratoires et digestifs, 1 cas sur 2500 naissances, gène porté par le chromosome 7). Les mutations responsables de cette maladie peuvent faire apparaître ou disparaître certains sites de restriction dans les gènes et modifier ainsi le nombre et la longueur des fragments d'ADN (exo TD).
- soit **autosomique dominante** (les malades sont hétérozygotes). Exemple : la Corée de Huntington (nécrose des cellules nerveuses du cerveau, débute vers 35 40 ans, 1 cas sur 12 à 15000, mouvement brusque et involontaire, démence (oubli).
- soit **gonosomique récessive** (relatif à un chromosome sexuel). Exemple : la Myopathie de Duchène, 1 garçon sur 3500, dégénérescence progressive des muscles

Le diagnostic précoce des maladies géniques pose de nombreux problèmes éthiques car souvent il n'existe pas de traitement.

3- Application de la transgénèse dans l'élevage de l'agriculture et de la biomédecine.

La biomédecine est une branche scientifique médicale qui consiste à l'application des connaissances de la médecine à la biologie.

Exemple : étude réalisée en Allemagne en 2003 :

La micro-injection d'ADN étranger dans le pronucléi d'un ovocyte fécondé a surtout été utilisée pour la production de bétails transgéniques. Cette technologie fonctionne de manière fiable, mais est inefficace et conduit à une intégration aléatoire et modes d'expression variable dans la progéniture transgénique. Néanmoins, des progrès remarquables ont été réalisés avec cette technologie. En ciblant l'expression de la glande mammaire, de nombreuses protéines humaines recombinantes hétérologues ont été produites en grandes quantités qui pourraient être purifiées à partir du lait de chèvres transgéniques, des moutons, des bovins et des lapins. Les produits tels que l'anti-thrombine humaine III, alpha-anti-trypsin et activateur tissulaire du plasminogène (protéine plasmidique inactive participant à la fibrinolyse en se transformant en plasmine sous l'effet d'activateurs) sont actuellement en essais cliniques avancés et devraient être sur le marché dans les prochaines années. Porcs transgéniques qui expriment des protéines de régulation du complément de l'homme ont été testés quant à leur capacité à servir de donneurs dans la transplantation d'organes humains (exemple : la xénotransplantation). Des expériences *in vitro* et *in vivo* montrent de façon convaincante que la réponse de rejet suraigu peut être cliniquement résolue par cette stratégie. Il est prévu que les porcs transgéniques seront disponibles en tant que donneurs

de xénogreffes fonctionnels d'ici quelques années. Les développements récents en matière de transfert nucléaire et sa fusion avec les données génomiques de croissance permettent une production transgénique ciblée et régulée. Les systèmes de recombinaison homologue efficace dans les cellules somatiques sont en cours d'élaboration et l'adaptation des outils moléculaires sophistiqués, déjà explorés chez la souris, pour la production de bétail transgénique est en cours. La disponibilité de ces technologies sont essentielles pour maintenir la «sécurité génétique» et de veiller à l'absence d'effets secondaires indésirables.

4- Les applications en industrie pharmaceutique

La technologie de l'ADN recombinant est largement utilisée par l'industrie pharmaceutique. Aujourd'hui, un médicament sur six est issu du génie génétique. Tout a commencé en 1941, avec la création de thérapies à base de pénicilline, un antibiotique découvert à la fin des années vingt. Dès lors, ces pratiques n'ont jamais cessé et se sont même amplifiées. Aujourd'hui, la transgénèse permet de produire à grande échelle des hormones, des enzymes et des molécules diverses qui sont utiles à la création de médicaments permettant de soulager ou de guérir certains dysfonctionnements de l'organisme (hormone de croissance, insuline). L'industrie pharmaceutique a également fait appel à la transgénèse dans le cadre de la création de nouveaux vaccins, comme le vaccin contre l'hépatite B, il y a quelques années (les recherches sur un éventuel vaccin contre l'hépatite C sont, à ce jour, encourageantes). Les nouvelles techniques mises au point permettent de n'insérer que l'antigène et non pas le virus dans sa totalité. Grâce au génie génétique, il est donc beaucoup plus facile de contrôler les caractéristiques d'un vaccin et de travailler avec plus de précision.

CHAPITRE III

GENOTYPAGE ET EMPREINTE GENETIQUE

Génotypage : c'est l'ensemble des analyses génétiques moléculaires visant à déterminer l'identité d'une variation génétique, à une position spécifique dans le génome, pour un individu ou un groupe d'individus donné appartenant à une espèce animale, végétale, fongique... Il est effectué de manière standardisée et automatisée ou par d'autres techniques tel que le génotypage microsatellite dans des laboratoires.

Empreinte génétique (= profil génétique) : c'est le résultat d'une analyse génétique, rendant possible l'identification d'une personne à partir d'une petite quantité de ses tissus biologiques (bulbe de cheveux, sang, salive, sperme). L'analyse de cette empreinte permet de comparer des séquences spécifiques d'un individu.

Les empreintes génétiques sont utilisées en médecine légale pour identifier ou innocenter des suspects grâce à leur sang, leur salive, leurs poils ou leur sperme. Elles permettent également d'identifier des restes humains, de faire des test de paternité, d'organiser le don d'organe, d'étudier des populations d'animaux et de végétaux et même de détecter des aberrations chromosomiques responsables de maladies génétiques.

Applications

1- Détection des OGM

La détection des OGM se base sur la PCR moyennant des amorces spécifiques du transgène introduit dans l'organisme hôte. Elle cible directement le transgène ou la protéine exprimée. On distingue ainsi 2 types de techniques de détection :

- **Détection basée sur l'ADN** : par la PCR qualitative (identifier le type d'OGM) ou quantitative (quantification exacte de l'ADN amplifié par PCR en temps réel).
- **Détection basée sur la protéine exprimée par l'ADN** : par des tests **immunologiques** (anticorps efficaces correspondant à la protéine recherchée). Cette méthode est facile à réaliser, plus rapide et moins coûteuse que celle basée sur l'ADN. Cependant, la même protéine peut se retrouver dans différents OGM. La modification génétique ne conduit pas toujours à la synthèse d'une protéine nouvelle. Il peut s'agir au contraire de diminuer ou d'augmenter la synthèse d'une protéine initialement présente.

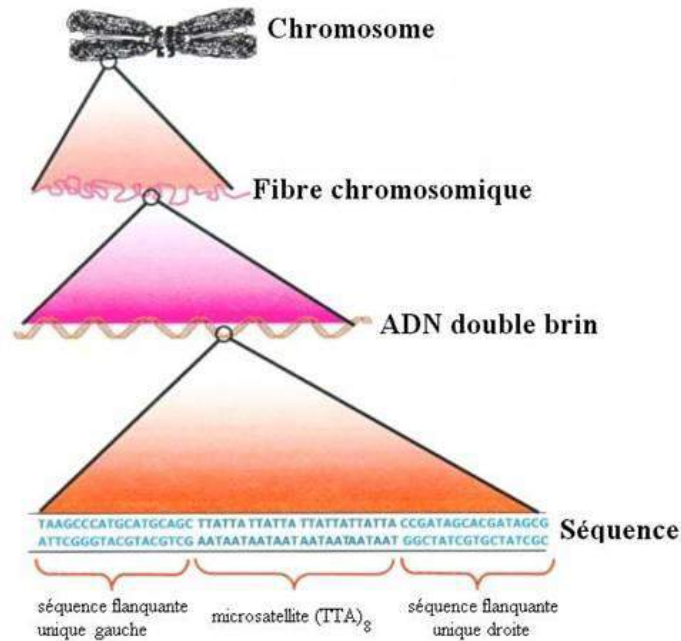
2- Applications légales

L'analyse génotypique à des fins d'identification en médecine légale est effectuée:

- À des fins médicales ou de recherches scientifiques.
- Dans le cadre d'une enquête ou d'une instruction judiciaire (vols, crimes, viols)
- Afin d'identifier un militaire décédé dans le cadre d'opérations armées
- Pour établir ou contester un lien de filiation : test de paternité.

Comparée à la RFLP (difficile à interpréter avec une probabilité de coïncidence de $1/10^{10}$, due au fait que 2 marqueurs dans 2 échantillons génèrent une bande au même endroit), l'analyse des empreintes génétiques par microsatellites permet une meilleure précision de distinction d'ADN avec une probabilité de coïncidence de $1/10^{13}$.

Les microsatellites : sont des séquences d'ADN constituées de répétitions continues de courts motifs en tandem (n fois) composés le plus souvent de 1 à 4 bases.



Les minisatellites sont des séquences d'ADN caractérisées par la répétition en tandem de motifs de 10 à 100 pb sur une longueur de quelques centaines à quelques milliers de pb. Ils peuvent être polymorphes par l'existence d'une variabilité soit dans la séquence soit dans le nombre de répétitions du motif de base répété. Mais généralement, pour un minisatellite donné, les allèles correspondent à des séquences avec un nombre différent de répétitions du motif de base. Elles sont présentes chez toutes les espèces et sont particulièrement étudiées chez l'homme et les bactéries.

```

5'--- CGCCTCGGCCTCCCAGAGTGTGAGATTACAGGCGTGAACCACCATGCCTAGCCGTTAGCTCCCA
      CTTATGAGTGAGAACAGGTGATGTTTTGGTTTTCCATTCCCTGAGTTACTTTACCCAGAATTGTTGT
      CTCCAATCTCATCGAGGTCCTCTGCCAATGCCAGTAATTCATTCCTTTTTATGGCTAAGTAGTATT
      CCATCGTATATATACATATACATATATATGTATACACACATATACATATATATGTATACACACAT
      ATACATATATATGTATACACACATATACATATATATGTATACACACATATACATATATATGTATA
      CACACATATACATATATATGTATACACACATATACATATATATGTATACACACATATACATATAT
      ATGTATACACACATATACATATATATGTATACACACATATACATATATATGTATACACACATATA
      CATATATATGTATACACACATATACATATATATGTATACACACATATACATATATATGTATACAC
      ACATATACATATATATGTATACACACATATACATATATATGTATACACACATATACATATATATG
      TATACACACATATACATATATATGTATACACACATATACATATATAGTATACACACATATACATA
      TATATGTATACACACATATACATATATATGTATACACACATATACATATATATGTATACACACAT
      ATACATATATATGTATACACACATATACATATATATGTATACACACATATACATATATACATATA
      TACACAAACACACATGCACCGCACTTTTTTTTTTTTTTTTTTTGAGATGGGAGTCTCACTCTAT
      CACCAGGCTGGAGTGCAAGTGGTGTGATCTTGGCTCACTGCAACCTCTGTCTCCTGGGTTCAAGCT
      ATTCTCCTGCTTCAGCCTCCTGAGTAGCTGGGATTACAGGTGCTCACCACCATGCCAGCTAATT
      TTTGATTTTTTACCATGTTGGCCAGGATGATCTCCATCTCCTGACCCTCTGATCCTCCTGCCTTG
      GGGTCCCAAAGTGCAGGGATTGCAGGCATGAGCCACTGCATGTGGCCACACCACACTTCCTTTAT
      CCACTTGTGATGGATGGGCATTTGTGTTGAAAAGAAGTCATCTCTGTTTTTCTTCCCTGAACG---3'
    
```

← CCTTTTTATGGCTAAGTAGTATTCCATCGT →
← TTCAGTAGAGACAAAAAGAAAGGGACTTGC
} Amorces

Séquence correspondant à un allèle du minisatellite CEB 310 avec les séquences flanquantes dont la connaissance est indispensable pour le choix des amorces PCR.

Le minisatellite est ici constitué par la répétition (25 fois) du motif de base de séquence consensus de 24 nucléotides : GTATACACACATATACATATATAT avec un très faible % de variation interne (les substitutions sont en bleu foncé). Les séquences (30 nucléotides chacune) utilisées pour le choix des amorces sont en rouge. Il est facile de prévoir quelle sera la taille du fragment amplifié par PCR qui révélera la présence de cet allèle du marqueur CEB 310. On le retrouvera sur le gel, après électrophorèse au niveau du marqueur de taille 1000 pb

▪ **Analyse de l'ADN mitochondrial**

L'analyse de l'ADNmt ne permet pas d'identifier à 100 % un individu, mais permet d'exclure une hypothèse, car cet ADN ne présente pas assez de variabilité dans les populations (~1 risque / 2000 que 2 personnes non apparentées présentent le même ADNmt).

Les médecins légistes étudient les sites HV1 et HV2 de l'ADNmt. Ce sont des séquences hypervariables dont les différentes versions sont nombreuses et avec des fréquences relativement faibles (quelques % de la population générale) → ces marqueurs ne confirment pas si un échantillon provient d'une unique personne. Mais, si on obtient 2 ADNmt différents on peut être sûr qu'il ne s'agit pas de la même personne.

▪ **Analyse du Chromosome Y**

Il existe sur le Chromosome Y des régions hypervariables dites STRY pouvant être utilisées en criminologie. Le chromosome Y est transmis par le père, et l'analyse de ce chromosome permet de faire les tests de paternité, lorsque le père n'est plus vivant. Cependant, ce chromosome n'existe que chez les individus de sexe masculin → Il ne pourra donc pas être effectué sur une femme.

3- Test de paternité

Un test de paternité consiste à analyser l'ADN de 2 personnes dans le but d'établir un lien de parenté génétique (filiation) avec les conséquences juridiques qui peuvent en découler.

A l'exception des jumeaux monozygotes, tous les humains ont un génome sensiblement différent en raison du polymorphisme génétique.

Contrairement aux méthodes simples (gènes récessifs et groupes sanguins) permettant d'apporter des réponses complètes et fiables aux questions de filiation, l'analyse de l'ADN est une méthode qui nécessite une analyse poussée de l'ADN de l'individu qui recherche ses parents ainsi que celui de son père et/ou de sa mère potentiels. Elle permet d'arriver à des réponses tranchantes : soit écarter la filiation, soit au contraire conclure à une très forte probabilité de filiation (> 99 %). Grâce à la PCR, la détermination des « empreintes génétiques » d'un individu est réalisée par analyse des minisatellites et microsatellites.

Principe : Le test de paternité consiste donc à évaluer la taille de certains mini ou microsatellites bien spécifiques et à comparer pour chaque chromosome, ces caractéristiques entre les chromosomes de l'enfant et ceux de ses parents potentiels.

Technique : Extraction d'ADN d'un échantillon prélevé sur l'individu (sang, salive, etc.) puis amplification PCR de l'ADN de chaque paire de chromosome puis digestion de l'ADN amplifié par une enzyme de restriction qui va isoler les minisatellites ou microsatellites en préservant leur longueur et enfin analyse des fragments d'ADN digérés par électrophorèse.

Exemple de test de filiation par analyse des marqueurs microsatellites

Voici les cartes génétiques de 4 personnes (Aurore, Bastien, Christine et Denis) et de l'enfant de 2 d'entre eux, Émilie. Dans chaque tableau est indiquée pour chaque chromosome, la taille du minisatellite ACGCC. Ainsi, Aurore possède sur son premier chromosome 14 la séquence ACGCC ACGCC ACGCC (n=3) et sur son autre chromosome 14, la séquence ACGCC ACGCC ACGCC ACGCC (n=4).

La comparaison des tableaux permet de conclure que Christine et Denis sont les parents d'Émilie car :

- d'après le chromosome 8, Aurore ne peut être la mère de l'enfant, car elle lui aurait transmis soit 13 soit 14 répétitions, or Émilie en possède 5 et 17.
- D'après le chromosome 13, Bastien ne peut pas être le père de l'enfant car il aurait nécessairement transmis une de ses répétitions 3 et 9.

- Les patrimoines de Christine et Denis expliquent complètement les caractéristiques des chromosomes d'Émilie.

Afin d'augmenter encore les probabilités de filiation, il faudrait comparer ces minisatellites sur davantage de chromosomes. Les laboratoires d'analyse parviennent à assurer des tests à l'issue desquels ils peuvent annoncer une probabilité de filiation excédant les 99 %.

ADN d'Aurore		ADN de Christine	
Chromosome analysé	Nb de répétitions	Chromosome analysé	Nb de répétitions
chromosome 1	8 et 15	chromosome 1	2 et 8
chromosome 8	13 et 14	chromosome 8	17 et 20
chromosome 13	4 et 15	chromosome 13	4 et 4
chromosome 14	3 et 4	chromosome 14	2 et 14

ADN de Bastien		ADN de Denis	
Chromosome analysé	Nb de répétitions	Chromosome analysé	Nb de répétitions
chromosome 1	8 et 18	chromosome 1	5 et 13
chromosome 8	5 et 7	chromosome 8	5 et 5
chromosome 13	3 et 9	chromosome 13	4 et 18
chromosome 14	2 et 8	chromosome 14	3 et 8

ADN d'Émilie	
Chromosome analysé	Nb de répétitions
chromosome 1	8 et 13
chromosome 8	5 et 17
chromosome 13	4 et 4
chromosome 14	2 et 3

4- Applications médicales

Les applications médicales consistent à un diagnostic prénatal (DPN) qui est l'ensemble des pratiques médicales ayant pour but de détecter *in utero*, chez l'embryon ou le fœtus, une affection d'une particulière gravité.

Une alternative au DPN est le diagnostic préimplantatoire (DPI) qui succède une fécondation *in vitro*, afin de ne réimplanter *in utero* qu'un embryon non atteint et éviter ainsi le recours à une interruption de grossesse. Il est plus rarement utilisé du fait de ses contraintes (stimulation ovarienne, tri des embryons après prélèvement d'un ou deux blastomères trois jours après une FIV et l'implantation intra utérine des embryons sélectionnés). Il nécessite une analyse génotypique à partir de quelques pg d'ADN reposant sur les performances et la qualité de la PCR.

Techniques de prélèvements ovulaires

Le DPN nécessite habituellement une amniocentèse avec culture et examen des cellules du liquide amniotique ou une **choriocentèse**. Donc l'ADN fœtal est préparé soit à partir de cellules amniotiques, soit à partir des cellules trophoblastiques. La mutation recherchée

ayant déjà été identifiée (chez des parents hétérozygotes par exemple), le diagnostic génotypique est simple et direct.

- Les examens réalisés sur le liquide amniotique consistent au dosage de certaines protéines (α fœto-protéine, α 1antitrypsine), à la recherche de l'acétylcholinestérase, au diagnostic de certaines maladies métaboliques, à l'analyse de caryotype après culture cellulaire, à l'étude du système HLA (Ensemble de groupes d'antigènes tissulaires constituant le facteur majeur de l'histocompatibilité, parfois lié à certaines maladies) et à l'extraction de l'ADN après culture cellulaire.

- La choriocentèse est réservée dans le cas où le % d'aberration chr est \uparrow .

- Le prélèvement du sang fœtal : permet le diagnostic des déficits immunitaires, l'hémophilie, les infections virales et parasitaires et le caryotypage rapide en fin de grossesse. Le dosage des IgM totales ou spécifiques sur le sang fœtal indique une atteinte hépatique fœtale par la toxoplasmose

Techniques d'analyse de l'ADN

Ces analyses sont réalisées par l'utilisation des enzymes de restriction et des sondes moléculaires radiomarquées (ADNc, ADNg, oligo de synthèse, ARNm); par la RFLP qui détecte des mutations dans des régions non codantes donc sans conséquences phénotypiques et qui se transmettent dans une même famille comme un caractère mendélien, par dot blot, southern, northern, PCR. Si le gène est connu, il est cloné et l'analyse est soit directe mettant en évidence la lésion moléculaire par PCR soit semi-directe par une sonde du gène muté mettant en évidence un polymorphisme intragénique ou très proche du gène. Si le gène est inconnu mais localisé, le diagnostic indirect est basé sur les sondes non spécifiques du gène muté.

Exemple d'anomalies géniques : maladies récessives liées à l'X

Quand le produit du gène est inconnu et que le gène est localisé et qu'on possède des marqueurs génétiques qui lui sont liés \rightarrow c'est la méthode de diagnostic indirect. Quand le produit du gène est inconnu mais ne s'exprime pas dans un tissu fœtal accessible, l'ADN nécessaire à cette étude est extrait des villosités choriales ou des cellules du liquide amniotique en culture.

5- Une perspective intéressante : l'utilisation des puces à ADN

Les puces à ADN (1995) constituent une véritable révolution technologique d'avenir en matière de séquençage et d'analyse de l'expression des gènes. Ce sont des combinaisons de techniques issues de la micro-électronique, de l'informatique, de la chimie et de la biologie moléculaire. Cette technologie provient d'une adaptation du Northern Blot (ADN fragmenté fixé à un support puis hybridé avec un ARNc).

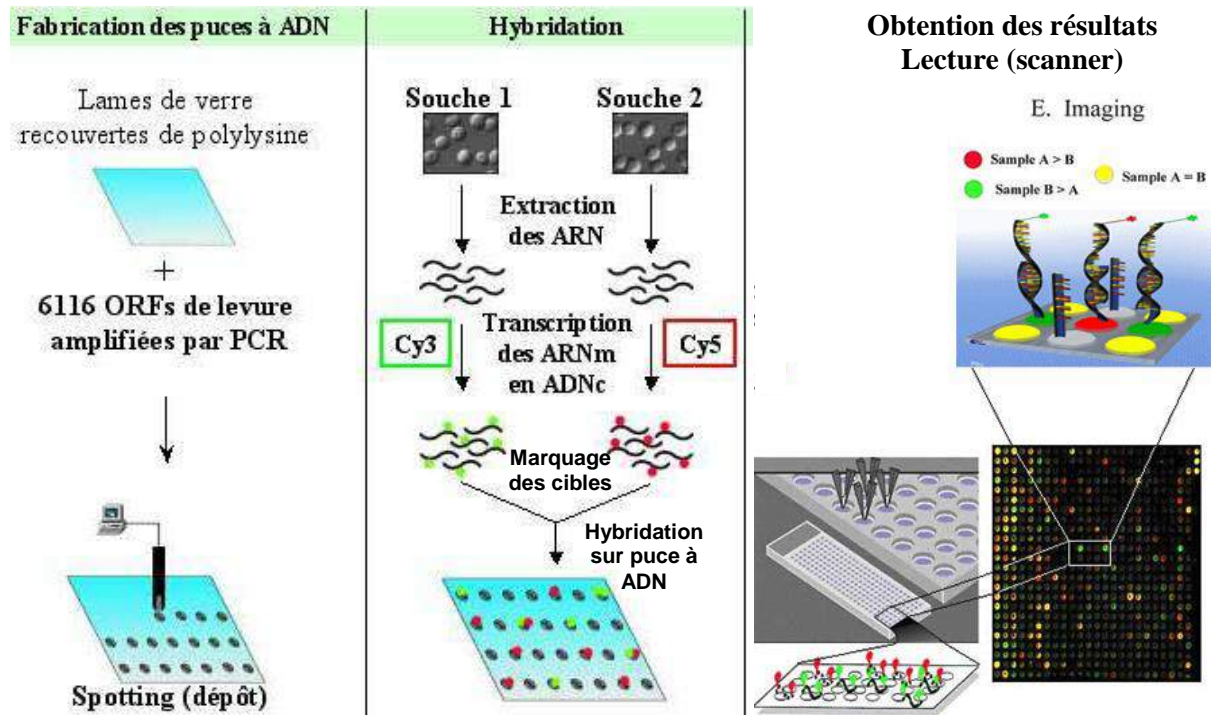
Une puce à ADN (ou micromatrice d'ADN ou puces à gènes, ou biopuces, ou microréseau d'ADN ou micromatrice d'ADN ou DNA chip, DNA-microarray, biochip) est un ensemble de molécules d'ADN caractéristiques de gènes de séquence connue et fixées en rangées ordonnées sur une petite surface (en verre, en silicium ou en plastique pour microscope d'environ 25 x 75 mm) permettant d'analyser, de mesurer et de visualiser très rapidement les différences d'expression entre les gènes (transcrits) dans une cellule, un tissu, un organe, un organisme ou un mélange complexe, à un moment donné et dans un état donné par rapport à un échantillon de référence \rightarrow Le transcriptome (= ensemble des ARNm d'une cellule) varie donc d'un type cellulaire à l'autre, au cours de la vie de la cellule et de son état de santé.

Principe des puces à ADN

Leur principe est basé sur la complémentarité des brins ADN : des séquences d'ADN greffées sur une puce constituent des sondes dont le rôle est de détecter d'autres séquences cibles qui leur sont complémentaires. Le choix des sondes portera sur un échantillon représentatif du génome et les cibles proviendront de deux populations : des cellules témoins et des cellules traitées par différentes molécules ou malades. Les effets de

ces dernières sur l'expression des gènes sont ensuite observés, ce qui révèle leur efficacité et leurs effets secondaires éventuels. Concernant le génome humain, ces sondes sont choisies en fonction de la nature de l'étude parmi les 25 000 gènes humains. Le génome d'animaux, de plantes ou d'organismes unicellulaires utilisent également ce support.

Fabrication des puces à ADN et Exploitation des données



① **Fabrication de la puce :** Les fragments d'ADN (sondes) amplifiés par la technique de PCR sont déposés sur une lame de microscope (ou autre type) préalablement recouverte de polylysine (assurer la fixation de l'ADN déposé via des interactions électrostatiques). Les fragments de PCR sont dans notre cas la partie exprimée (l'ORF) des 6116 gènes de *Saccharomyces cerevisiae* (la levure de boulanger). Juste avant l'hybridation, on dénature l'ADN pour qu'il se trouve sous la forme simple brin sur la puce, ce qui lui permettra de s'accrocher au brin complémentaire contenu dans la cible.

② **Préparation de la cible :** Les ARN totaux sont extraits des deux cultures de levures dont on veut comparer l'expression. Les ARNm sont ensuite transformés en ADNc par transcription inverse. Lors de cette étape, l'ADN de la 1^{ère} culture est marqué avec un fluorochrome vert (Cyanine 3), tandis que l'ADN de la 2^{ème} culture est lui marqué en rouge (Cyanine 5).

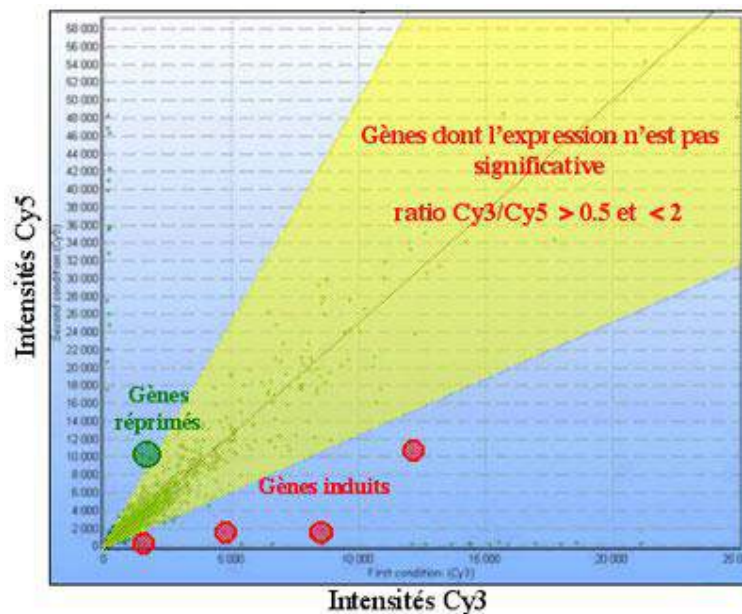
③ **Hybridation :** Les ADN marqués en vert et ceux marqués en rouges sont mélangés (la cible) et placés sur la matrice d'ADN simple brin déposée (la sonde). La puce est alors incubée une nuit à 60°C : les ADN fluorescents vont ainsi s'hybrider sur les gènes déposés.

④ **Lecture :** Chaque point (spot) de la puce excité par un laser va être analysé individuellement par un scanner à très haute résolution, et ce à la longueur d'onde d'excitation de la Cyanine 3 puis de la Cyanine 5. La fluorescence émise par les traceurs est détectée puis photographiée par un photo-multiplicateur (PMT) couplé à un microscope confocal. On obtient alors deux images dont le niveau de gris représente l'intensité de la fluorescence lue. Si on remplace les niveaux de gris par des niveaux de vert pour la première image et des niveaux de rouge pour la seconde (chaque image est représentative d'un transcriptome), on obtient en les superposant une image en fausses couleurs composée de spots allant du vert (seulement de l'ADN de la 1^{ère} condition fixé) au rouge (seulement de

l'ADN de la 2^{ème} condition fixé) en passant par le jaune (ADN des deux conditions fixé en quantité égale).

⑤ **Analyse des données** : À chaque spot est attribuée une valeur d'intensité normalisée par rapport à l'ADN "étalon" : on parle de spike. Chacune des valeurs peut être analysée par des techniques de bio-informatique, ce qui permet d'estimer avec plus ou moins de précision l'intensité d'expression d'un gène.

On a maintenant deux images de la puce à partir desquelles il faut réussir à calculer le nombre de molécule d'ADN au départ. Pour cela, on mesure la quantité de signal dans la longueur d'onde d'émission du fluorochrome vert et celle dans la longueur d'onde d'émission du fluorochrome rouge. Puis on normalise ces quantités en fonction de divers paramètres (quantité de levure de départ dans chaque condition, puissance d'émission de chaque fluorochrome, ...). On suppose alors que la quantité d'ADN fluorescent fixée est proportionnelle à la quantité d'ARNm correspondant dans la cellule de départ et on calcule le ratio fluorescence rouge / fluorescence verte. Si ce ratio est supérieur à 1 (rouge sur l'image en fausses couleurs), le gène est plus exprimé dans la seconde culture, si ce ratio est inférieur à 1 (vert sur l'image en fausses couleurs), le gène est moins exprimé dans la seconde culture. On peut visualiser ces différences d'expression grâce à un logiciel développé dans le laboratoire appelé Arrayplot (cf image ci-dessous). Il permet d'obtenir à partir de la liste des intensités de chaque point le nuage de point intensité dans le rouge = f (intensité dans le vert).



⑥ **Le regroupement en fonction des profils d'expression** : On peut ensuite essayer de regrouper des gènes ayant le même profil d'expression sur plusieurs expériences ayant un rapport biologique. Ce regroupement peut se faire de proche en proche comme pour une phylogénie, ce qui consiste à calculer un critère de similitude entre les réponses et à rassembler les profils les plus similaires. On peut également faire appel à des techniques plus complexes comme l'analyse en composante principale (ACP : est une méthode de la famille de l'analyse des données et plus généralement de la statistique multivariée, qui consiste à transformer des variables statistiquement corrélées en nouvelles variables non corrélées les unes des autres. Ces nouvelles variables sont nommées "composantes principales", ou axes principaux. Elle permet au praticien de réduire le nombre de variables et de rendre l'information moins redondante.) ou les réseaux neuronaux. Au final on représente en général le clustering hiérarchique sous la forme d'une matrice où chaque colonne correspond à une expérience et chaque ligne correspond à un gène. On

représente les ratios grâce à une échelle de couleur du vert (gènes réprimés) au rouge (gènes induits).

Enfin, les résultats seront souvent confirmés gène par gène par des méthodes telles que la PCR quantitative ou le Northern Blot. Les puces apportent principalement des données qualitatives (variation d'expression d'un gène) mais il est difficile de quantifier avec précision l'expression d'un gène avec la technologie des puces à ADN. Généralement, après une étude de puce à ADN, la bio-informatique extrait une liste de gènes intéressants (en fonction de ce que l'on cherche). Pour confirmer ces gènes on fait appel à la technique de la PCR quantitative ou Real-Time PCR (PCR en temps réel).

Exemple : on peut marquer l'ADNc du malade en vert et du traité en rouge, ou bien, du témoin en rouge et du traité en vert. Ce marquage se fait habituellement grâce à une enzyme : la polymérase T7 qui amplifie l'ARNm et incorpore les cyanines pour un marquage optimal. Une fois marqués, ces ADNc sont déposés sur la lame de verre sur laquelle ont été préalablement fixés des fragments de génome humain recouvrant tous les gènes présents dans une cellule.

Domaines d'application

Les avantages de puces sont : la très grande spécificité des sondes et le gain de temps. En effet, des dizaines de milliers de sondes peuvent être fixées sur une même puce. Cela permet de tester différentes cultures cellulaires sur une même lame et de faire des réplicats permettant ainsi, en une seule expérience qui dure environ 2 jours, d'avoir une estimation sur l'expression de plus de 30000 gènes simultanément.

La mesure de l'expression de gènes par puce à ADN s'applique à de nombreux domaines de la biologie et de la médecine comme l'étude de traitements, de maladies ou de stades de développement.

▪ Domaine médical

Diagnostiquer des maladies selon le profil d'expression du patient, étudier la différence d'expression des gènes de cellules cancéreuses par rapport à des cellules saines. Dans le domaine pharmaceutique, biotechnologique et cosmétologique, les puces à ADN servent essentiellement aux criblages moléculaires et cellulaires dans le but de découvrir par exemple de nouveaux médicaments.

L'utilisation des puces à ADN connaît un essor croissant notamment dans le domaine de la cancérologie pour le typage tumoral d'après leur profil génétique. L'utilisation des puces à ADN comme outil de diagnostic présente l'avantage de faire appel à de nombreux marqueurs : plusieurs milliers de gènes peuvent être criblés simultanément pour fournir une signature du type cellulaire étudié. Si l'on considère que chaque type de tumeur présente une signature génétique unique, ce système permet virtuellement de distinguer et classer tous les types de tumeurs.

Les puces à ADN permettent donc de comparer l'expression des gènes de deux types cellulaires différents, de faire de l'étude des gènes exprimés sur un grand nombre de patients pour observer l'effet d'un médicament (anti-cancéreux par exemple), de regarder l'effet d'un traitement sur l'expression des gènes, de comparer tissus sains contre tissus malades, traités contre non-traités etc...

Exemple : étude des effets d'irradiations sur les cellules : Grâce aux biopuces, on a réussi à mettre en évidence de nouveaux effets des irradiations sur l'expression de l'ADN. Pour cela, on a comparé le profil d'expression des cellules avant et après l'exposition au rayonnement en exposant les échantillons à différentes doses d'irradiation. Si les effets à fortes doses sont bien connus, ils le sont moins pour des doses plus petites. La liste des gènes réprimés ou activés, suite à une faible exposition, fournit de précieux renseignements. L'étude se fait aussi en fonction des tissus et des types cellulaires. Après irradiation, les chercheurs ont remarqué une variation de la radiosensibilité. Par exemple dans le sang, les lymphocytes

(globules blancs) sont les premiers à disparaître. Ils sont plus sensibles que les autres cellules sanguines. Les chercheurs veulent maintenant savoir pourquoi afin d'ajuster, par exemple, la radiothérapie à chaque patient atteint de cancer. Jusqu'à présent, les médecins soignaient les tumeurs par rayonnement avant même de savoir si le traitement serait efficace, alors que certaines tumeurs sont résistantes. Une analyse préliminaire de la tumeur par les biopuces permettrait d'éviter au patient une radiothérapie inutile.

Plus encore, ces chercheurs étudient les effets des irradiations par rapport au temps. Ce qui était impossible sans les puces à ADN. Par exemple, une heure après l'exposition, le profil d'expression de nombreux gènes change. Vingt-quatre heures après, il revient à la normale. Cela a permis de découvrir, notamment, un nouveau gène impliqué dans la réparation de l'ADN, abîmé par les rayonnements.

- **Agro-alimentaire**

- **Agronomie**

Analyser des traits agronomiques de sélection.

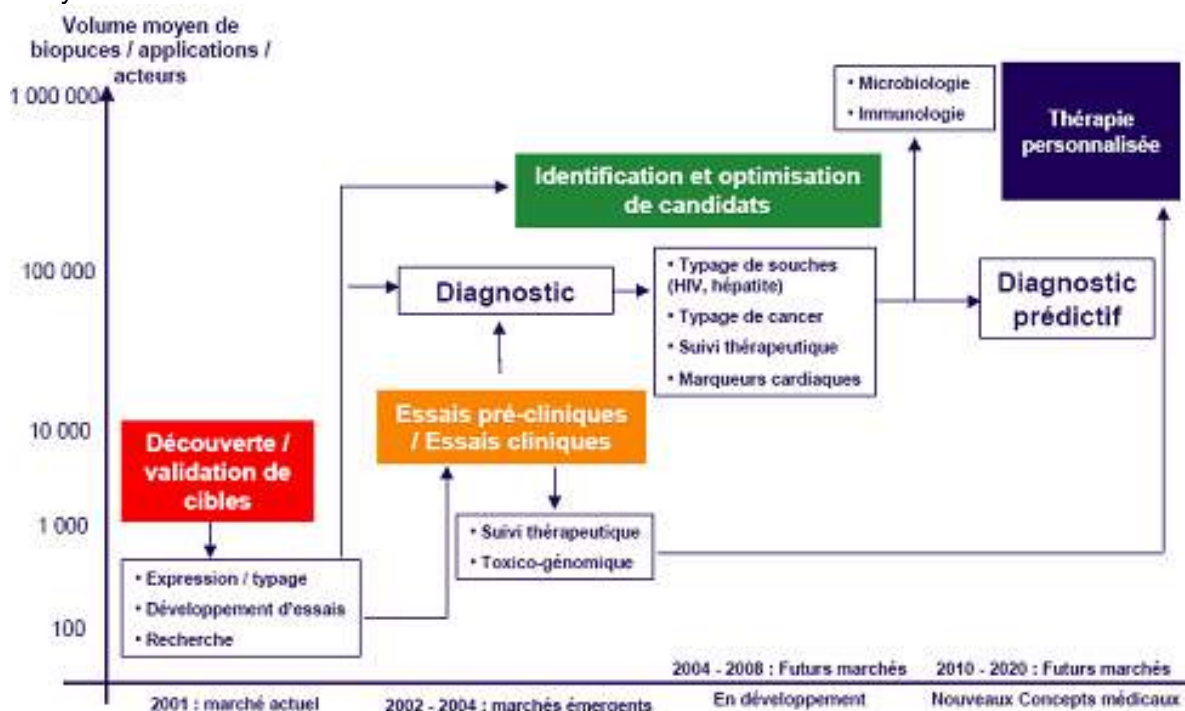
- **Soins vétérinaires**

- **Environnement**

Améliorer la traçabilité et la surveillance d'OGM (les sondes correspondront à différentes séquences introduites dans les plantes OGM), de microorganisme (analyses bactériennes d'eau potable), de pathogènes (détection d'agents infectieux dans l'air ou dans l'eau), de toxines ou de pesticides lors du contrôle qualité des aliments.

- **Toxicogénomique**

C'est l'étude génétique de l'influence des facteurs environnementaux : les cellules sont soumises à un stress par une irradiation, un polluant ou un médicament, dont la nocivité est analysée.



Conclusion

Les techniques de génie génétique permettent :

- le développement d'OGM dont les enjeux sont la production de plantes cultivables améliorées et la production de médicaments
- l'analyse de l'ADN humain pour diagnostiquer et dépister les anomalies géniques
- de soigner certaines maladies grâce à la thérapie génique.