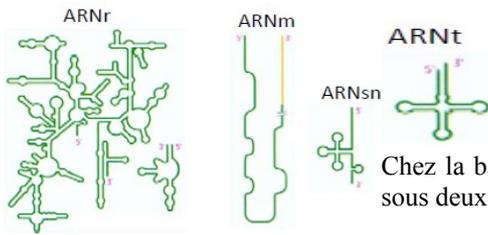


Génie Génétique

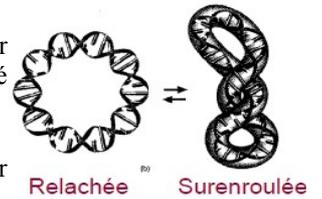
I] Introduction

L'ADN est le support quasi universel de l'information génétique. C'est un polymère de nucléotide bicaténaire, en double hélice.



Chez les virus, l'information génétique peut se trouver sur un ADN ou ARN, circulaire, linéaire, segmenté ou bi-caténaire.

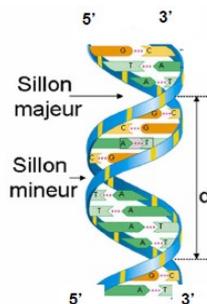
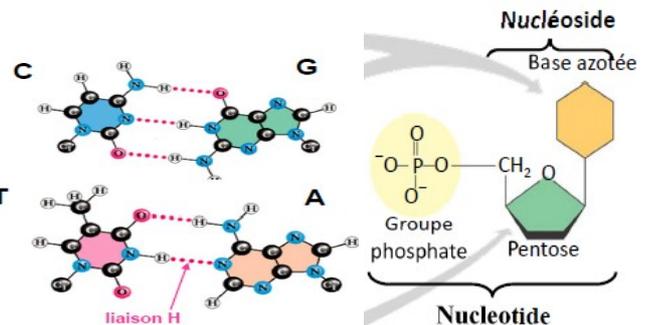
Chez la bactérie, l'ADN est circulaire et nu, et peut se trouver sous deux formes : relâchée ou super enroulé.



Chez les eucaryotes, on trouve des ADN linéaires séparé du reste de la cellule par un noyau. L'ADN se trouve sous forme de chromosomes compactés. Dans la mitochondrie, on retrouve également de l'ADN circulaire.

Structure primaire :

L'ADN est composé de bases azotées (puriques et pyrimidiques) et de désoxyribose. Dans le cas de l'ARN, on aura un ribose. Le groupement phosphate reliera le tout. Base azoté + pentose = nucléoside, et nucléoside + phosphate = nucléotide. Les nucléotides sont reliés entre eux par des liaisons 3'5' phosphodiester.



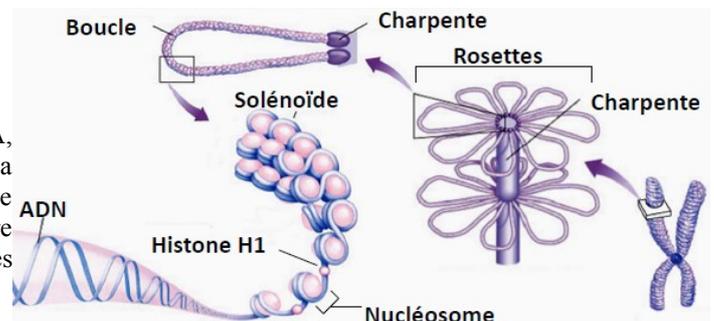
Structure secondaire :

La molécule d'ADN est composée de deux brins complémentaires associés de façon anti-parallèles. Les bases interagissent par des liaisons hydrogènes (2 pour TA, 3 pour CG).

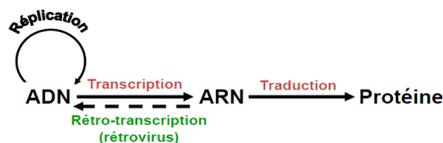
Pour l'ARN, la complémentarité est intramoléculaire. L'ARN va s'organiser en « épingle à cheveux ».

Structure tertiaire :

Enfin, chez les eucaryotes, des protéines, les histones (H2A, H2B, H3, H4), vont s'associer en octamères, autour duquel va s'enrouler l'ADN pour former le nucléosome. L'histone H1 se retrouve entre deux nucléosomes, et vont se rapprocher entre eux pour former le solénoïde. Des protéines non-histones formeront ensuite des rosettes à partir des solénoïdes.



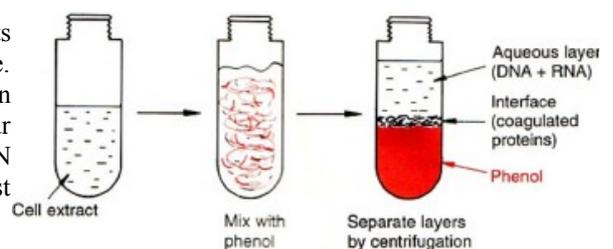
Dogme central :



II] Préparation d'ADN

1) Génomique

Il faut dans un premier temps lyser les cellules avec des détergents (SDS, Tween 20, NP40), puis lyser les protéines à l'aide de protéinase. On purifie ensuite l'ADN en l'extrayant par du phénol/chloroforme (on récupère la phase supérieure aqueuse). On concentre alors l'ADN par précipitation alcoolique (par l'éthanol ou isopropanol). L'ADN génomique obtenu est sous forme de longue molécule, qu'il est nécessaire de fragmenter pour pouvoir pénétrer dans le gel.

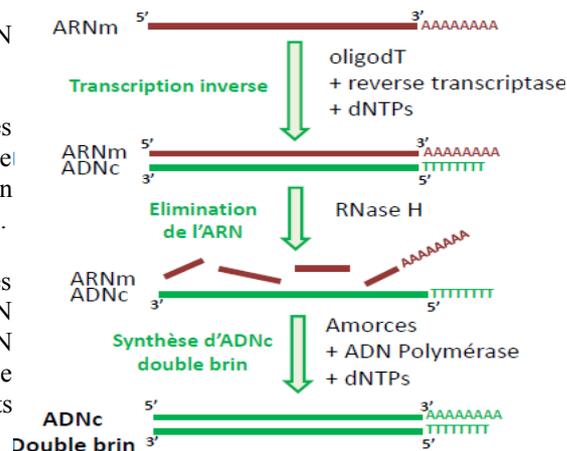


2) Complémentaire

On prépare un ARNm, puis par transcription inverse, on repasse à l'ADN double brin.

Chez les procaryotes, il y a très peu d'ADN non codant. Chez les eucaryotes, on a des exons (codants) mais également des introns (séquence non-codante). Il faut donc supprimer ces parties non-codantes par un mécanisme d'épissage, et maturation (coiffe en 5', et polyadénilation en 3').

Pour réaliser la transcription inverse, on utilise des transcriptases inverses (ADN polymérase dépendante de l'ARN). On obtient une copie de l'ARN en ADN. On obtient un hybride ARN/ADN, et on se débarrasse de l'ARN avec traitement par une RNase H, et on resynthétise une ADNc à l'aide d'ADN polymérase 1 (nécessite des amorces faites par les fragments d'ARN digéré) et nucléotides.

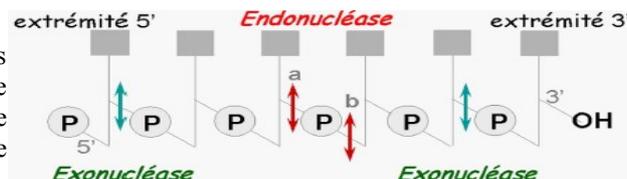


III] Comment manipuler de l'ADN

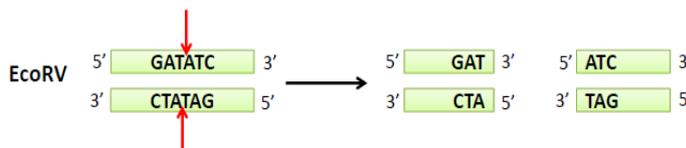
- **Couper de l'ADN** : Nucléases (endonucléases ou exonucléases)
- **Coller de l'ADN** : Ligases
- **Copier de l'ADN en ADN** : ADN Polymérase – fragment klenow
- **De l'ARN en ADN** : Reverse transcriptase
- **De l'ADN en ARN** : Transcriptase
- **Modifier** : élimier des phosphates (Phosphatases), ou ajouter des phosphates (Polynucléotides kinases)
- **Rechercher** : on utilise l'hybridation

1) Nucléases et ligases

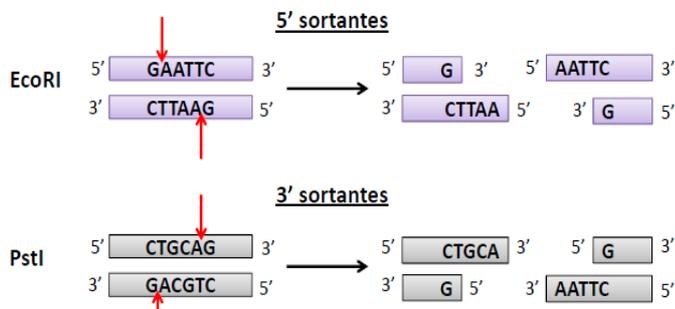
Ce sont des enzymes capables de couper les liaisons phosphodiester. Les exonucléases coupent aux extrémités alors que les endonucléases coupent des liaisons internes. Elles peuvent être larges (Nucléase S1, Dnase 1, Rnase A, H), ou restreintes à une séquence particulière.



- Extrémités franches



- Extrémités cohésives



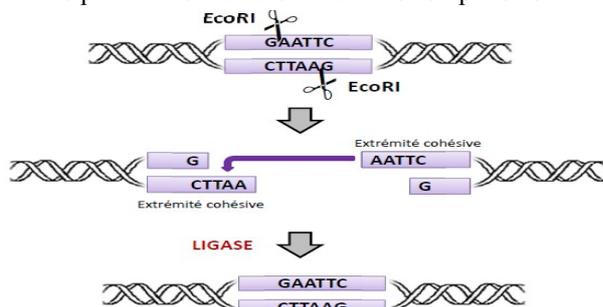
Deux extrémités cohésives peuvent s'apparier, et une **ligase** viendra rattacher les deux brins par une liaison phosphodiester.

Rappel : en 3' on trouve un OH, et en 5' un phosphate. L'association de ces extrémités nécessite une ligase et de l'ATP.

Les enzymes des restrictions hydrolysent des séquences particulières. Elles ont initialement été trouvée dans des bactéries (moyen de défense contre les phages). Le nom de chaque enzyme de restriction a été donné en fonction de la cellule bactérienne dans laquelle elles ont été trouvées.

On appelle **site de restriction** (qui sont des palindromes : lecture identique dans un sens sur un brin, et dans l'autre sur l'autre brin) la zone d'ADN reconnue et coupée par des enzymes de restriction.

Elles peuvent couper en faisant des **extrémités franches** (même taille d'ADN sur chaque brin), ou des **extrémités cohésives** (5' sortantes, ou 3' sortantes). Des mêmes sites de restrictions peuvent être clivés par des enzymes différentes et former des extrémités incompatibles, et deux sites différents peuvent former des extrémités compatibles.



b) La transformation

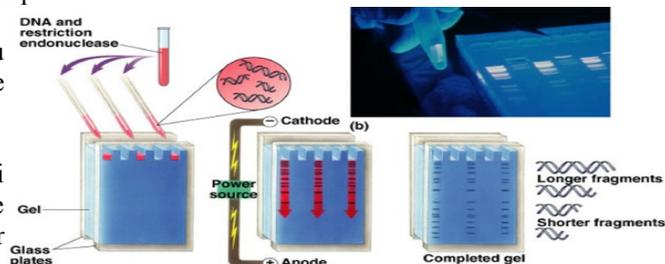
Il faut que la bactérie soit compétente, c'est à dire capable d'incorporer le plasmide nu. Une fois le plasmide intégré (ou non), on multiplie les bactéries.

Pour rendre des bactéries compétentes, on peut utiliser deux techniques :

- **Choc thermique** (bactéries thermocompétentes) : on fragilise la membrane par du CaCl₂, puis on les congèle. On met ensuite les bactéries en présence de l'ADN exogène et on procède au choc thermique.
- **Électroporation** (bactéries électrocompétentes): on centrifuge pour concentrer les bactéries, puis on les resuspend en présence de l'ADN exogène. On fait alors un choc électrique.

On sélectionne à la fin les bactéries en les mettant sur un milieu antibiotique. Les bactéries résistantes auront donc intégré le plasmide.

On peut vérifier que le vecteur recombinant correspond à celui que l'on a voulu intégrer. On réalise une analyse par cartographie de restriction, que l'on analyse par électrophorèse (SDS pour charger l'ADN négativement, qui pourra migrer vers l'anode).



c) Isolation des clones recombinants

Deux types de clonages possibles :

- **Fonctionnel** : il faut connaître la fonction du gène d'intérêt
- **Hybridation ou PCR** : il faut posséder un gène homologue, un fragment du gène ou de l'ARN associé.

α) *Clonage fonctionnel*

Il faut que le fragment s'exprime (donc différents selon si on est chez des eucaryote ou procaryote) : clonage complet du génome (procaryote) ou l'ensemble des ADNc correspondant aux ARNm s'exprimant dans une cellule.

Si on clone l'ensemble de l'ADN génomique, on obtient une « banque d'ADN génomique ». Si on digère partiellement l'ADN, on a plus de chances d'obtenir gène et promoteur (suffisant chez les procaryote). Par reverse-transcription, on obtient une banque d'ADNc (eucaryote), qui possèdent le promoteur qu'à faible probabilité. Si le vecteur contient un promoteur, on parle de « banque d'expression ».

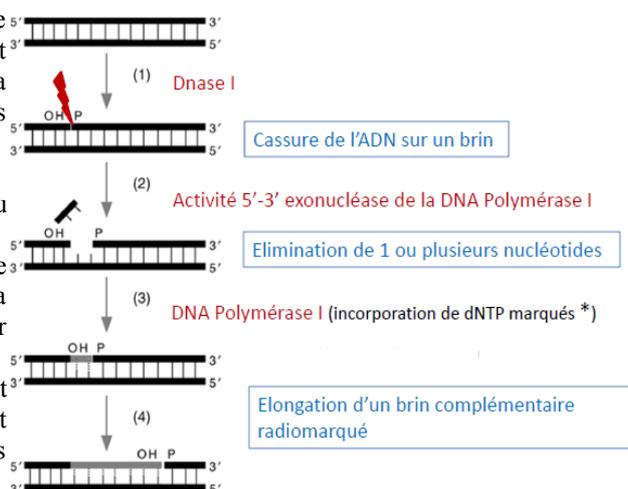
On transforme alors les bactéries (ou transfecter des cellules) avec la totalité de la banque, et on effectue des tests de sélection. On parle de sélection positive si on sélectionne les survivants, et de négative si on sélectionne celles qui meurent.

β) *Clonage par hybridation*

On fabrique une banque. On n'a pas besoin de connaître la fonction du gène. On sélectionne ensuite une enzyme de choix pour cliver l'ADN (Southern Blot).

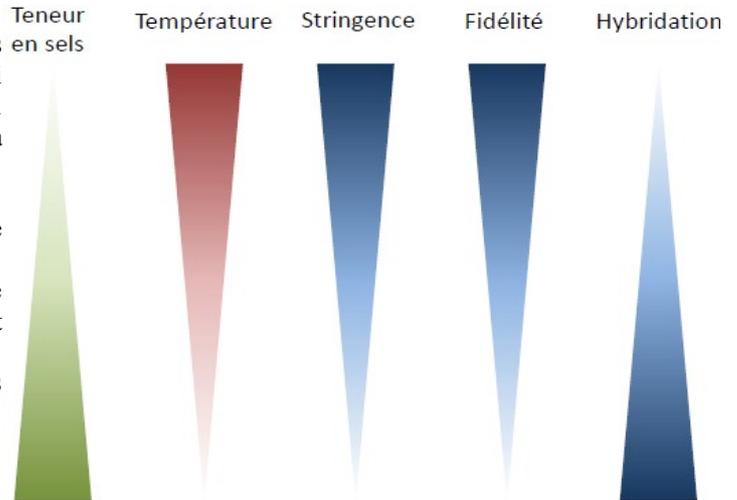
Chaque brin est capable de s'apparier avec son complémentaire inverse. Si on fabrique des brin visualisables, et que l'on reproduit les conditions adéquates pour l'appariement (hybridation), on pourra visualiser les fragments d'intérêts. Ces ADN visualisables sont des sondes : courte séquence nucléotidique radioactive ou fluorescente.

- On peut utiliser le « **End-labelling** » : remplacement du phosphate 5' par un phosphate radioactif
- La « **Nick-translation** » (ci-contre) : translation de coupure : une DNase coupe l'ADN sur un brin, et la DNAPolymérase complète les trous et remplace le brin par des nucléotides radiomarqués.
- Le « **Random priming** » : amorces aléatoire vont s'hybrider sur le fragment, et les polymérase vont synthétiser le brin complémentaire à partir de nucléotides radiomarqués.

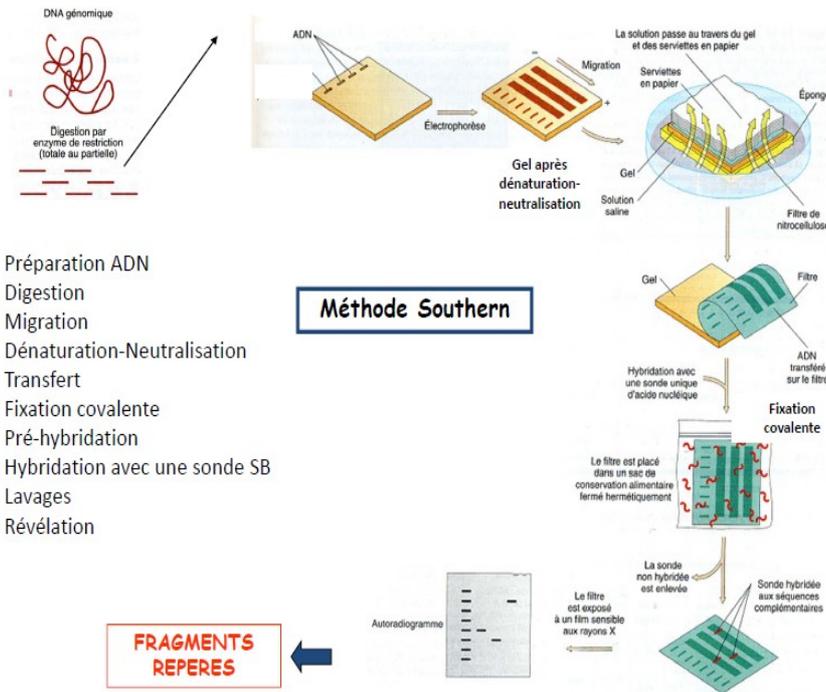


L'hybridation est influencée par :

- La température : si moins élevée, l'ADN est plus stable. T_m est la température de fusion, qui correspond à 50/50 ADN double brin/simple brin. Pour appairer deux brins, il faut diminuer la température lentement.
- La concentration de l'ADN
- La richesse en G-C de la séquence (force de liaison)
- La salinité du milieu : si elle est faible, elle change la T_m , et empêche le mésappariement et donc défavorise l'hybridation.
- Le formamide peut être utilisé pour stabiliser les simples brins



On introduit la notion de stringence : plus elle est élevée, plus la fidélité de l'hybridation sera importante.



Southern Blot : on fait migrer les fragments d'ADN par électrophorèse.

On applique une membrane sur le gel, faisant monter l'ADN sur la membrane par capillarité (avec un tampon).

On hybride alors la membrane avec les sondes qui vont se fixer spécifiquement sur les fragments contenant le gène, et on lave le surplus de sonde.

On révèle ensuite les fragments radiomarqués par autoradiographie.

Le choix de l'enzyme est important, pour restreindre les fragments comprenant le gène d'intérêt dans son intégralité.

Le choix de la sonde est donc déterminant, puisqu'il permet de localiser spécifiquement le gène d'intérêt.

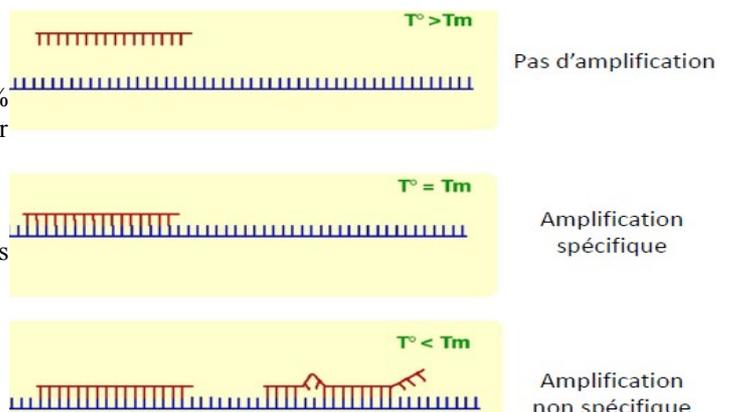
Ainsi, on peut sélectionner seulement les bactéries ayant intégré le plasmide, puisqu'elles seront les seules à fixer la sonde.

4) La PCR

Polymérase chain reaction, permet d'amplifier par polymérisation en chaîne, en utilisant une polymérase résistante à la chaleur (taq polymérase), permettant polymérisation, dépolymérisation en chaîne. On pourra donc obtenir une grande quantité de l'ADN voulu à partir d'une seule copie. On peut également cloner un ARNm après une reverse-transcription (amorce spécifique, et RT grâce à une amorce poly-T).

Le choix des amorces est capital pour la PCR :

- 18 à 25 nucléotides,
- 40 à 60% de contenu en GC,
- Leur T_m (température à laquelle on va avoir 50% double brin et 50% simple brin) ne doit pas différer de plus de 5°
- Ne pas pouvoir se replier sur elles-mêmes



L'étape critique est la température d'hybridation des amorces, qui est de l'ordre de $T_m = 2(A+T) + 4(G+C)$.

- $T > T_m$, pas d'amplification,
- $T = T_m$ amplification spécifique
- $T < T_m$ amplification non spécifique

Il existe un ordre dans l'organisation du code :

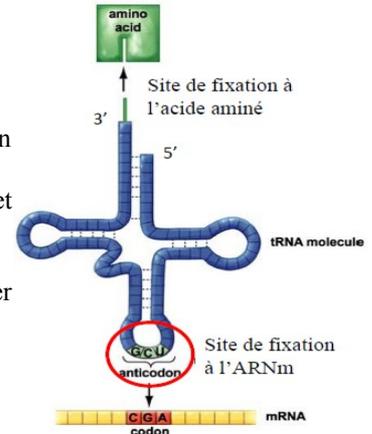
- La troisième base, pour beaucoup de codons, n'influe pas sur le choix de l'aa (GCU et GCC codent pour l'alanine).
- La deuxième base code pour un aa hydrophobe si c'est une pyrimidine, et polaire si c'est une purine.
- La première base, si elle est changée code souvent pour un autre acide aminé

Le code a été « décodé » en mettant en présence ribosomes, ARNt, cytosol et des bases sélectionnées : UUU → que phénylalanine.

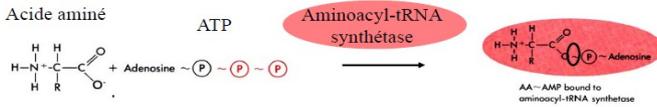
2) Mécanisme de traduction

Processus séquentiel : les ARNm sont lus et les aa attachés à la suite. La traduction nécessite :

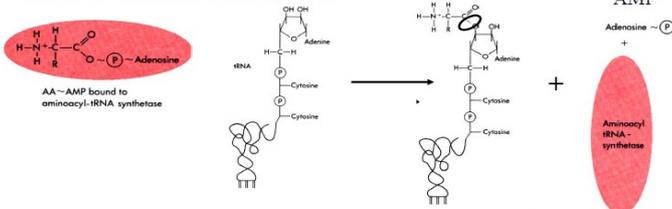
- **ARNt chargés** : ces ARNt interagissent avec des aa auxquels ils sont spécifiques, et possèdent un anticodon pour interagir avec le brin d'ARNm.
- **ARNm**
- **Enzymes** : amino acid ARNt synthétase (autant que d'acide aminé) permet de lier ARNt et aa, avec trois sites de reconnaissance (ARNt, ATP, aa).
- **ATP ou GTP**



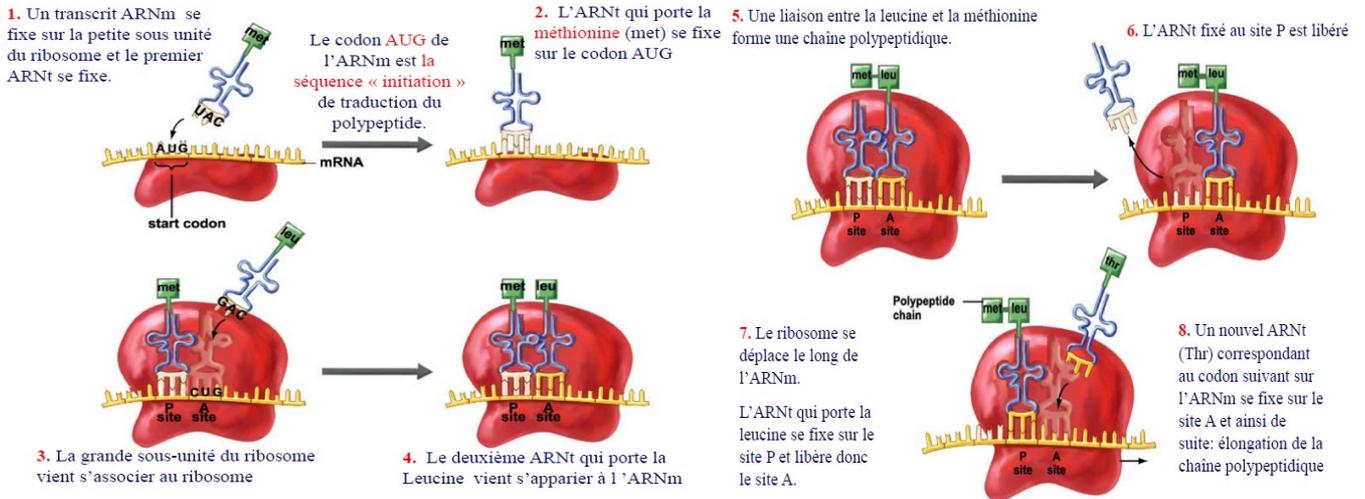
Étape 1 : activation de l'aa



Étape 2 : transfert de l'aa sur l'ARNt



La charge de l'acide aminé se fait en deux étapes : activation de l'aa par un AMP, puis transfert de l'aa sur l'ARNt.



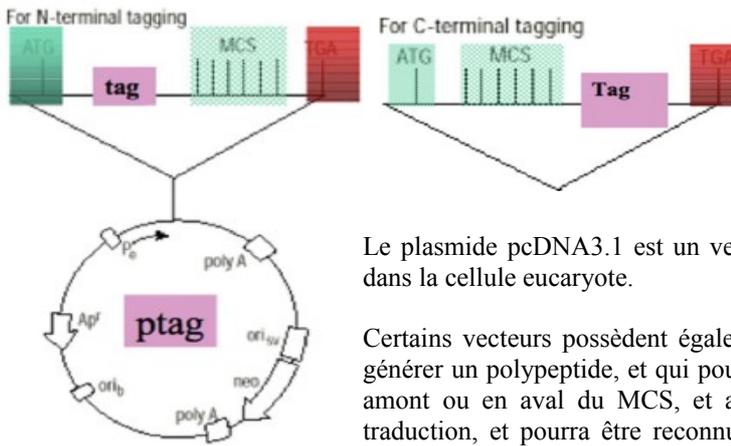
En PCR, on peut utiliser des oligonucléotides dégénérés déduits à partir de la séquence protéique. On peut donc faire des PCR sans connaître vraiment la séquence.

V] L'expression d'un gène cloné

On rappelle qu'un vecteur doit posséder un site multiple de clonage, un gène de sélection, et doit être capable de se répliquer dans une cellule hôte. De plus, selon le type de vecteur, le nombre de paires de bases insérables pourra varier de 10pb pour un plasmide bactérien à 1000 pb pour un chromosome de levure.

Pour exprimer le gène dans la bactérie, on clone l'intégralité d'un gène, promoteur compris chez les procaryotes. Quand on veut exprimer un gène eucaryote, il faut mettre un promoteur en amont du MCS et supprimer les introns.

Ainsi, pour faire une banque d'expression dans la bactérie, il faut choisir le vecteur d'expression, cloner l'ADNc d'intérêt, l'exprimer dans la bactérie et le reconnaître sur gel avec des anticorps.



On peut exprimer des gènes d'intérêt dans des bactéries, des levures, des champignons et même des cellules animales. Ainsi, on pourra accéder à la machinerie nécessaire à la maturation eucaryote post-traductionnelle. Le produit peut également y être plus stable. Cependant, le coût reste plus élevé, et les méthodes de culture plus complexes.

Le plasmide pcDNA3.1 est un vecteur navette, capable de s'exprimer chez la bactérie et dans la cellule eucaryote.

Certains vecteurs possèdent également un TAG : séquence nucléotidique définie, qui va générer un polypeptide, et qui pourra être reconnu par un anticorps. Le TAG est inséré en amont ou en aval du MCS, et ainsi, la protéine d'intérêt sera collé au tag lors de la traduction, et pourra être reconnue par les anticorps. Il faut cependant veiller à ce que l'ORF soit totalement conservée.

VI] Applications du génie génétique : Biotechnologie

Toutes les techniques évoquées ont permis de séquencer le génome et d'apprendre à manipuler de l'ADN et ainsi de créer de domaine des biotechnologies.

On retrouve les biotechnologies aussi bien dans la recherche, que la médecine et l'agriculture

1) Production de protéines humaine

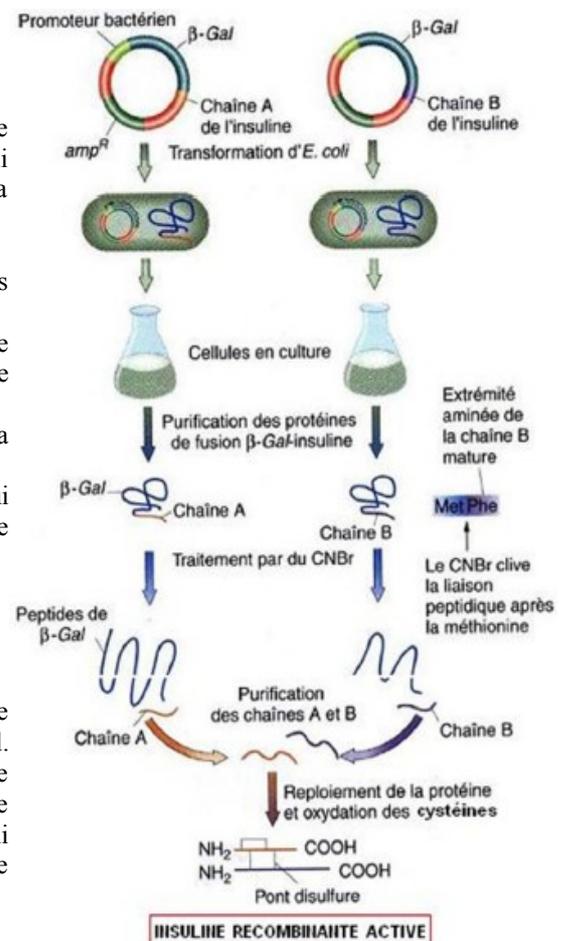
L'insuline est la première protéine qui a été créée par génie génétique (1982), et a été posé comme alternative à l'insuline porcine ou bovine, qui posait à la fois des problèmes immunitaires et éthiques. Cette découverte a eu d'énormes impacts sanitaires et économiques.

Bien qu'elle ait été la première synthétisée, ce n'était pourtant pas la plus simple :

- **Préproinsuline (108 aa)**: Elle est d'abord synthétisée sous forme de préproinsuline avec une séquence signal permettant le passage des membranes
- **Proinsuline (84 aa)** : clivage de la séquence signal, et la proinsuline sera stockée dans des granules de sécrétion
- **Insuline (21+30)** : clivage de la protéine en 3 partie, deux qui s'associeront par deux ponts disulfures, forme qui sera la forme finale de l'hormone

a) Mécanisme de synthèse

E.coli est utilisée, avec des vecteurs recombinants, l'un contenant la chaîne A, et l'autre contenant la chaîne B, chaque gène collés à celui de la β -gal. Les vecteurs sont ensuite intégrés dans les bactéries. Ainsi, chaque bactérie peut produire une β -gal collée à l'une des chaînes. La protéine synthétisée est ensuite clivée par du CNBr. On purifie ensuite les chaînes A et B qui pourront ensuite se replier et s'assembler par des ponts disulfures. L'insuline est alors active.



b) Autres protéines et protocoles

De nombreuses autres protéines peuvent être synthétiser par des bactéries, mais on peut également les faire synthétiser par des animaux transgéniques : vache (lactoferrine), rat (hormone de croissance)...

2) Production d'animaux transgéniques

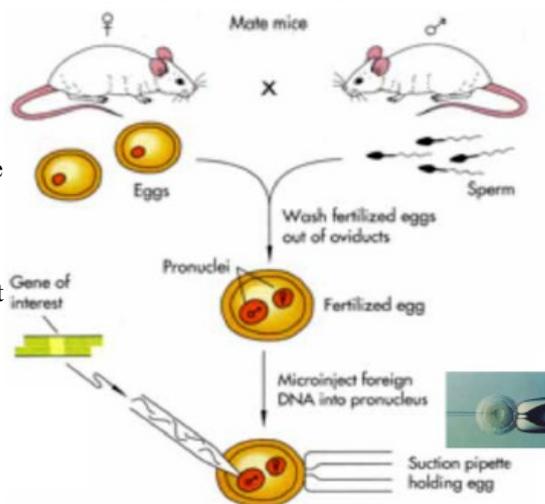
Ils vont exprimer un gène qui ne fait pas partie de leur patrimoine génétique. Pour cela, il faut que l'organisme croissent rapidement, avec un temps court de génération, ait un grand nombre de descendants et des mécanismes similaires à celui de l'homme.

La souris est un bon récepteur de gène, puisqu'elle répond à toutes les conditions. La première souris transgénique a été faite en 1982, et la souris knockout (inactivation totale d'un gène) quelques années plus tard.

On distingue ainsi deux types de transgénèse : transgénèse classique aléatoire ou ciblée (qui remplace un gène par un autre).

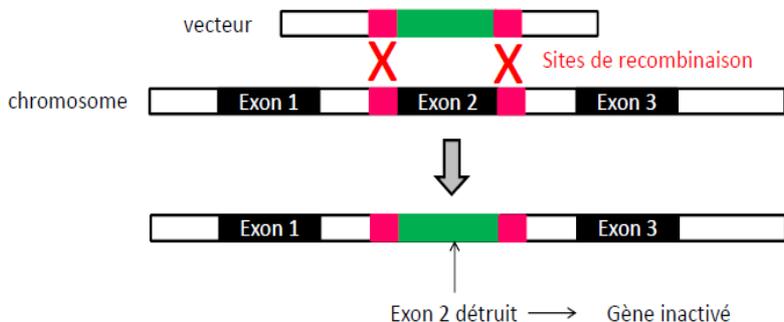
a) Transgénèse classique

- Construction du transgène
- Introduction de milliers de transgène dans le pronucléus mâle d'un zygote fécondé par microinjection)
- Le transgène s'introduit aléatoirement dans le génome du zygote
- On réimplante le zygote dans une mère porteuse
- On vérifie la présence du transgène par PCR et southern blot, et son expression par PCR et northern blot
- On purifie et on amplifie la lignée transgénique
- Analyse du phénotype des souris



b) Transgénèse ciblée

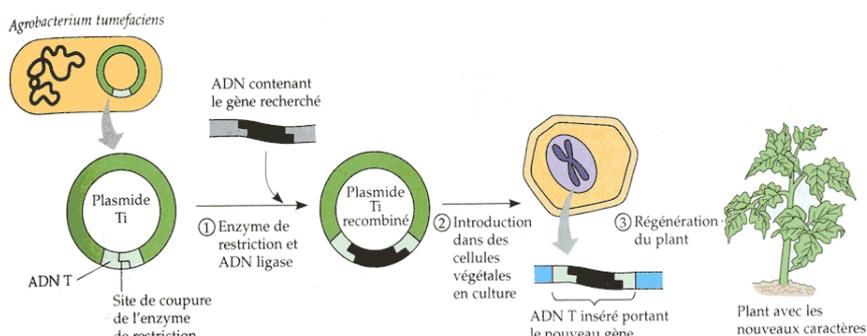
Elle implique un événement de recombinaison homologue (échange de fragment d'ADN situé entre deux molécules différente au niveau de l'appariement de deux séquences nucléotidiques identiques). Ainsi, on met le gène d'intérêt entre deux séquences qui sont homologues à un bout d'un exon.



- Construction du transgène
- Intégration dans des cellules souches embryonnaires en culture
- Sélection des clones de cellules contenant l'évènement de recombinaison désiré par PCR, nombre taille et localisation par southern blot, et vérification de la séquence par séquençage.
- Injection des cellules souches dans des blastocystes et dépôts chez les femelles porteuses
- Amplification et purification de la lignée
- Vérification de la présence du transgène par PCR et de la délétion du gène

Le but est donc de reproduire des maladie pour pouvoir ensuite les étudier et tester des traitements, étudier le rôle physiologique d'un gène particulier ou d'une mutation, marquer des cellules pour ensuite pouvoir les étudier.

3) Transgénèse chez les plantes



Chez les plantes, on peut ainsi créer des plantes résistantes aux herbicides, aux conditions climatiques, des plantes aux capacités nutritionnelles améliorées...

On peut utiliser des bactéries, des projections d'ADN avec des canons à particules (particule enrobés d'ADN d'intérêt) ou par perméabilisation des membranes.

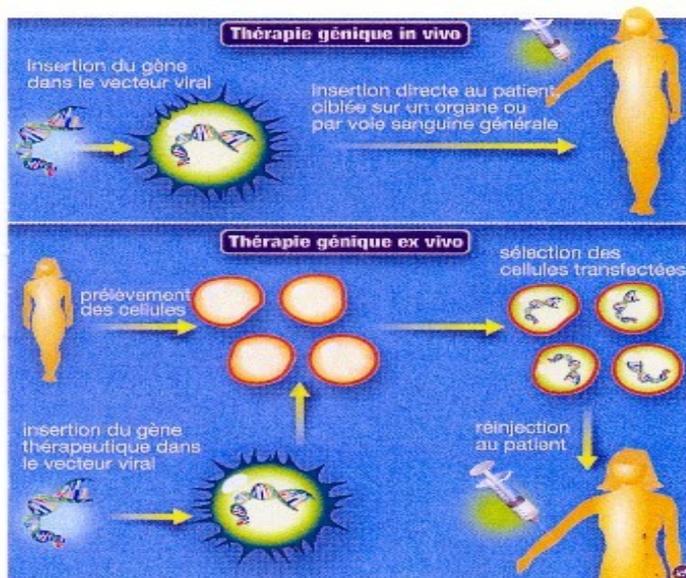
4) Thérapie génique

Le but est d'intégrer des gènes dans des cellules malades, pour corriger des défauts physiologiques majeurs, en corrigeant l'absence ou la déficience d'un gène, en inactivant un gène par la technique du Knock out.

- Il faut un transgène
- Un vecteur (virus inactivés, liposomes ou de l'ADN nu)
- Cibler les cellules

Ces thérapies peuvent se faire *in vivo* (simple), ou *ex vivo* (en extrayant les cellules à traiter, puis les réintégrant une fois modifiées).

Les maladies génétiques sont les premières à pouvoir être ciblées, ainsi que des maladies neurodégénératives, des maladies cardiovasculaires et infectieuses acquises (SIDA). Cependant, le cancer représente 60 à 70% des test effectués.



Les vecteurs viraux sont les plus utilisés, mais les problèmes proviennent essentiellement de ces vecteurs viraux : mauvaise intégration de l'ADN, réaction immunitaire provoquée, insertion aléatoire (possibilité d'inactivation d'un gène indispensable), faible capacité de stockage du virus, et risque de reconstituer un virus actif après recombinaison. On s'intéresse actuellement au VIH qui s'intègre bien et au AAV (associées aux adénovirus) qui provoquent très peu de réactions immunitaires..

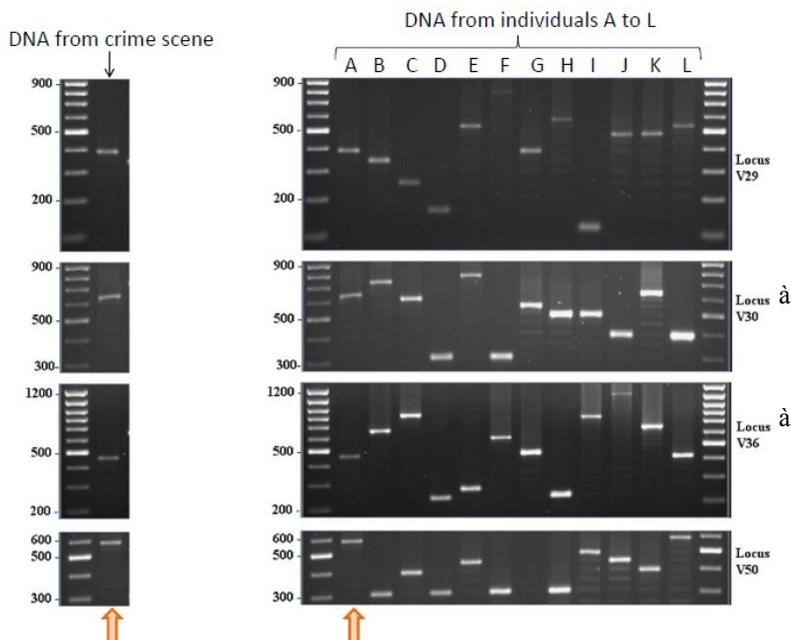
En plus des difficultés scientifiques et technique, la recherche se heurte à des problèmes éthiques et financiers.

5) Diagnostic génétique et empreinte génétique

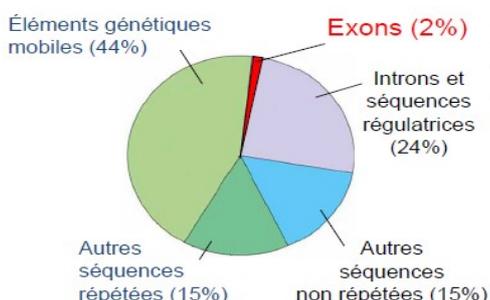
Analyse d'une partie ou de l'intégralité du génome (pré-implantatoire, pré-natal, post-natal). La drépanocytose est la première maladie ayant été diagnostiqué par des outils biotechnologiques (par restriction de gènes donnant des fragments de taille différentes : direct par clivage du gène, ou indirect par clivage à proximité du gène).

Les empreintes génétiques sont de plus utilisées en médecine légale et pour les test de paternité. Seuls 4 5% de l'ADN code pour des protéines. On trouve dans le reste des région non-codantes comme les mini-satellites VSTR (15 à 40 pb) et micro-satellites STR (4 pb en moyenne). Ces séquences sont propres chaque personne, et ainsi, on clive l'ADN, trie des fragments par taille, puis hybridation. L'enfant possèdera l'empreinte d'au moins un parent.

Actuellement, on utilise plus la PCR qui se base sur les STR, et plus précisément leur nombre et emplacement dans le génome.



6) Séquençage



Les techniques actuelles permettent de séquencer des génomes et les cartographier. Seuls les modèles (levures, bactéries, arabidopsis...) et l'homme ont été séquencés.