

Génie Génétique

Cours pour 3^{eme} année Biochimie

*Présenté par : Dr. BERBER Naima
Maitre de conférences B
Département de biologie
Université Dr Moulay Taher Saida*

Sommaire

I. Introduction.....	4
II. Propriétés et caractérisation de la molécule d'ADN.....	5
II.1. Structure primaire.....	5
II.2. Structure secondaire.....	6
II.3. Structure tertiaire.....	7
III. Méthodes de préparation d'ADN.....	8
III.1. Préparation d'ADN Génomique.....	9
III.2. Préparation d'ADN complémentaire.....	10
IV. Fragmentation spécifique d'ADN par des enzymes.....	11
IV.1. Les nucléases	11
IV.1.1. Les exonucléases.....	11
IV.1.2. Les endonucléases.....	12
IV.1.2.1. Les DNase non spécifique de l'ADN	13
IV.1.2.2. Les DNases spécifiques de l'ADN : enzyme de restriction.....	14
IV.2. Les polymérase.....	15
IV.2.1. ADN polymérase ADN dépendante ou ADN polymérase.....	16
IV.2.2. ADN polymérase ARN dépendante ou reverse transcriptase.....	17
IV.3. Les ligases.....	19
V. Vecteurs de clonage et analyse des banques	20
V.1. Nécessité et principe du clonage.....	21
V.2. Creation d'un vecteur.....	22
V.3. Différents types de vecteurs	23
V.3.1. Les vecteurs plasmidiques.....	24
V.3.2. Les vecteurs du phage λ.....	24
V.3.3. Vecteurs cosmides.....	25
V.3.4. Chromosomes artificiels bactériens BAC	26
V.3.5. Chromosomes artificiels de levure YAC.....	26
VI. Amplification in vitro de séquences spécifiques par PCR.....	27
VII. Hybridation d'ADN sur support : technique de Southern blot.....	29

VIII. Technique de séquençage.....	30
IX. Applications du génie génétique.....	31
IX.1. Production de protéines thérapeutiques.....	32
IX.2. Production d'animaux transgéniques.....	33
IX.3. Transgénèse chez les plantes.....	34
IX.4. Thérapie génique.....	34
IX.5. Diagnostic génétique	35
X. Travaux dirigés.....	36
TD N° 1: Enzymes de Restriction.....	37
TD N° 2: Clonage d'ADN.....	38
TD N° 3: Technique d'hybridation moléculaire.....	39
TD N° 4: Technique de séquençage.....	40
TD N° 5: Exposés sous formes de poster des différentes applications de génie génétiques.....	41
XI. Références bibliographiques.....	42

I. Introduction

Au début du XX^e, la redécouverte des travaux de Mendel (1822-1888) et les travaux de Morgan (1866-1945) sur des mouches permettent de comprendre que l'hérédité est due à la transmission de particules appelés gènes, disposés de manière linéaire sur les chromosomes.

Dans les années 1950, est mise en évidence la nature chimique des gènes, ainsi que la structure moléculaire de l'ADN. En 1965, découverte des enzymes de restriction confirmée en 1973 par Paul Berg et ses collaborateurs. Ces protéines capables de découper et recoller précisément l'ADN, donnent aux chercheurs les outils qui leur manquaient pour établir une cartographie du génome. Elle ouvre aussi la voie à la transgénèse, qui permet d'intervenir *in vitro* sur des portions d'ADN et donc des gènes. La technologie de l'ADN recombinant permet l'insertion d'une portion d'ADN (un ou plusieurs gènes) dans un autre ADN.

Ces nouvelles technologies d'ingénierie génomique, avec la génomique synthétique (conception de génomes artificiels), figurent actuellement parmi les technologies les plus prometteuses en termes de recherche biologique appliquée et d'innovation industrielle.

Le **génie génétique** (ou ingénierie génétique) est un ensemble de techniques, faisant partie de la biologie moléculaire et ayant pour objet l'utilisation des connaissances acquises en génétique pour utiliser, reproduire, ou modifier le génome des êtres vivants. Il a souvent pour but la modification des génotypes, et donc des phénotypes.

Le génie génétique est un champ très actif de la recherche car les applications possibles sont multiples, notamment en santé humaine (correction d'un gène porteur d'une mutation délétère, production de protéines thérapeutiques, élimination de séquences virales persistantes, etc.), en agriculture biotechnologique (mise au point de nouvelles générations de plantes génétiquement modifiées, etc.) ou encore pour la mise au point d'outils destinés à la recherche (par exemple pour explorer la fonction d'un gène).

II. Propriétés et caractérisation de l'ADN

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, segmente ou bicaténaire.

Chez la bactérie, l'ADN est circulaire et nu, et peut se trouver sous deux formes : relâchée ou super enroule. Chez les eucaryotes, on trouve des ADN linéaires séparés du reste de la cellule par un noyau. L'ADN se trouve sous forme de chromosomes compacts. Dans la mitochondrie, on retrouve également de l'ADN circulaire.

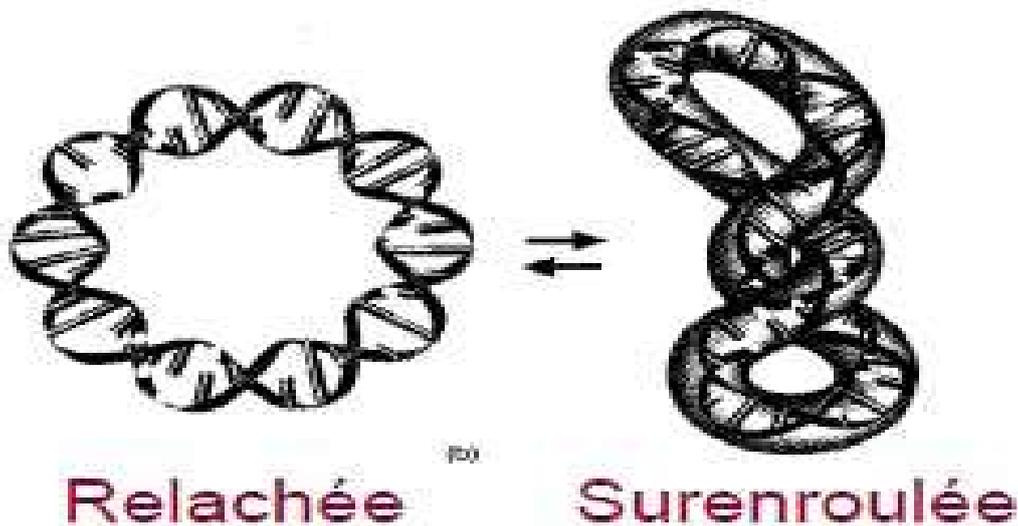


Figure 1. Schéma montrant les formes de la molécule ADN

II.1. 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 azotée + pentose = nucléoside, et nucléoside + phosphate = nucléotide. Les nucléotides sont reliés entre eux par des liaisons 3'5' phosphodiester.

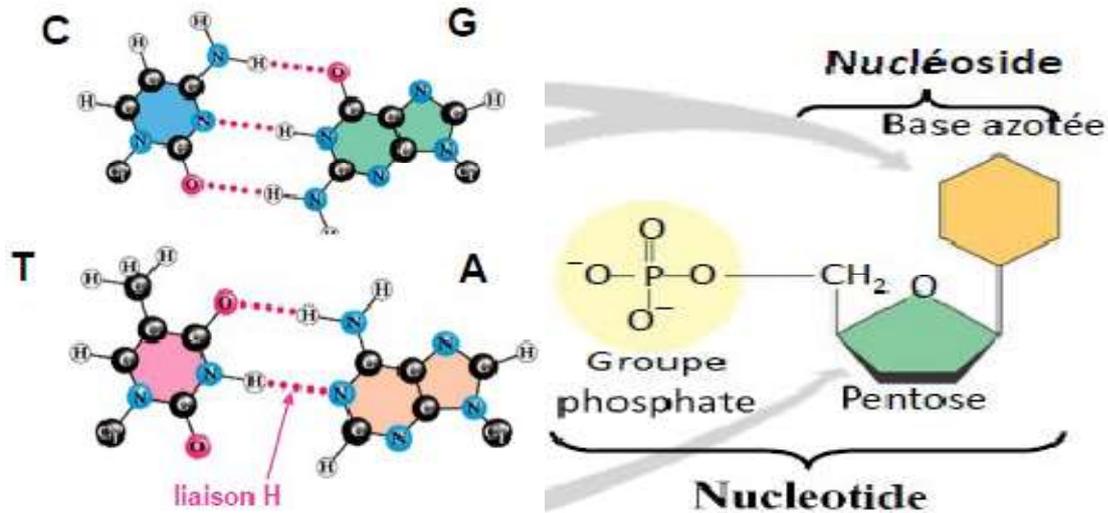


Figure 2. La structure primaire de la molécule ADN

II.2. 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 ».

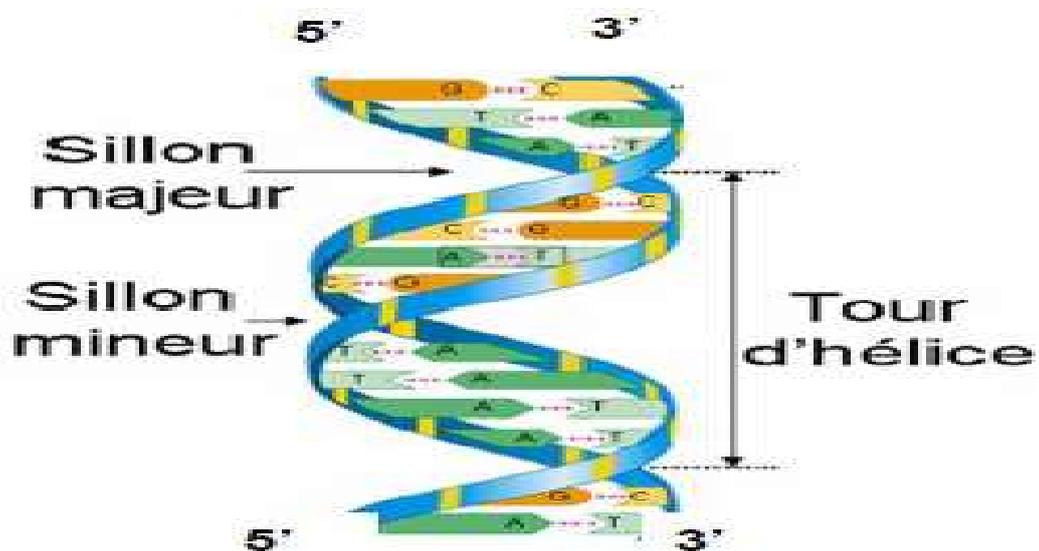


Figure 3. La structure secondaire de la molécule ADN

II.3. 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.

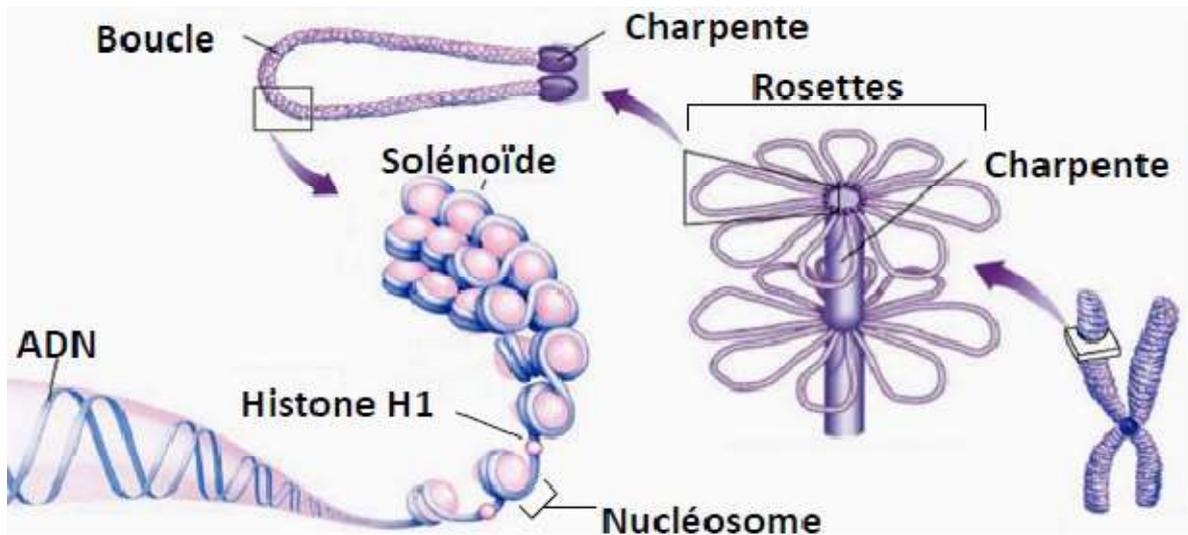


Figure 4. La structure tertiaire de la molécule ADN

III. Méthodes de préparation d'ADN

III.1. Préparation d'ADN 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.

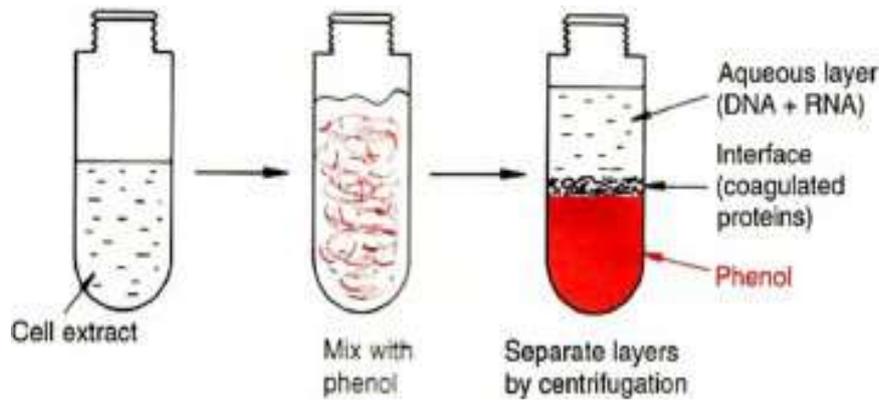


Figure 5. Schéma montrant les étapes d'extraction d'ADN

III.1. Préparation d'ADN Génomique

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 polyadénylation 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.

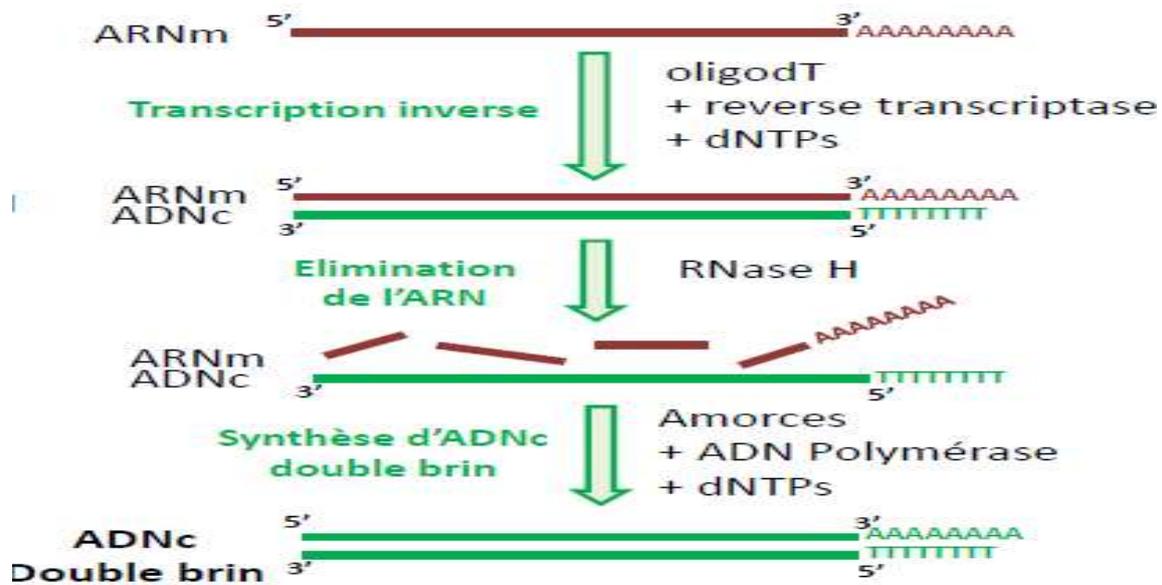


Figure 6. Les étapes de synthèse de la molécule ADNc

IV. Fragmentation spécifique d'ADN par des enzymes

Les enzymes sont les principaux outils de la biologie moléculaire. Ils permettent de travailler sur les acides nucléiques, de les modifier, de les couper, de les associer etc...

Actuellement, il existe trois types des enzymes

1. Les nucléases: Leur fonction consiste à couper l'ADN au niveau d'un phosphate, libérant deux fragments: un se terminant par un 3'OH et l'autre par un 5' phosphate.

2. Les ligases: sont des enzymes qui collent des bouts d'ADN ensemble.

3. Les ADN polymérase: sont des enzymes qui copient l'ADN.

IV.1. Les nucléases

Leur fonction consiste à couper l'ADN au niveau d'un phosphate, libérant deux fragments: un se terminant par un 3'OH et l'autre par un 5' phosphate. On distingue:

1- Les exonucléases qui digèrent l'ADN en retirant les nucléotides à partir d'une extrémité,

2- Les endonucléases qui coupent un des deux brins au milieu du polymère.

IV.1.1. Les exonucléases

Les exonucléases : digèrent l'ADN en retirant les nucléotides à partir d'une extrémité 3'5 ou bien de l'extrémité 5'3'. On cite quelques exemples :

Les exonucléase III est une 3'5' exonucléase, elle retire les nucléotides d'une manière séquentielle à partir de l'extrémité 3' (à la condition qu'elle ne soit pas sortante).

La Bal 31 est aussi une exonucléase dont l'activité principale est une 3' exonucléase mais qui a aussi une activité 5'3' exonucléase. Elle a de plus une activité secondaire d'endonucléase sur le simple brin généré par l'exonucléase

La T7Gene exonucléase est une 5'3' exonucléase qui a comme substrat de l'ADN double brin avec un 5' phosphate ou un 5'OH.

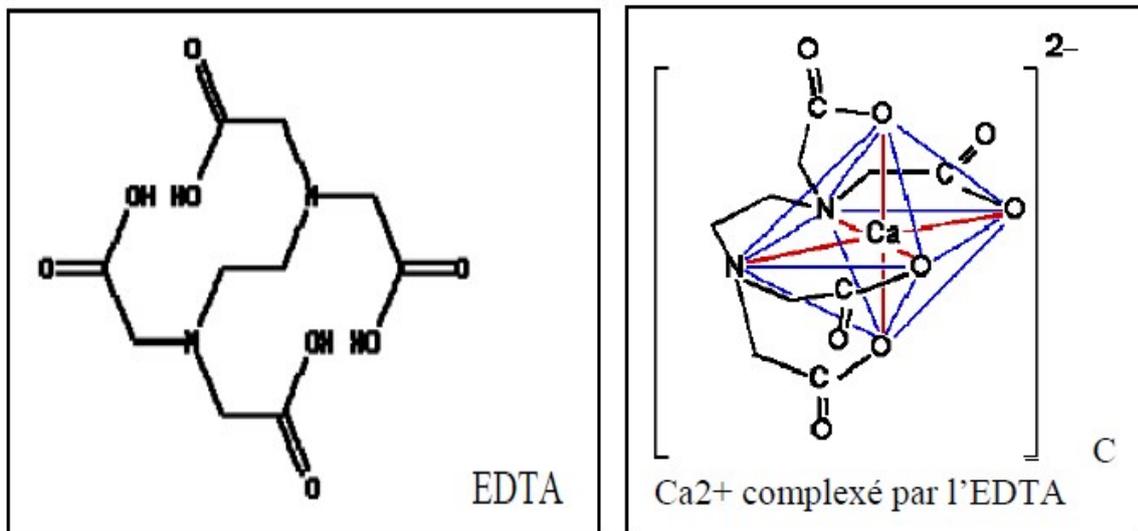


Figure 7. Schéma d'une exonucléase avec la fonction lytique Ca^{2+} .

IV.1.2. Les endonucléases

Les endonucléases ont diverses spécificités de substrat, certaines coupent l'ADN double brin, d'autre l'ADN simple brin, enfin certaines reconnaissent et coupent au niveau de séquences caractéristiques.

IV.1.2.1. Les DNase non spécifique de l'ADN

DNase I : endonucléase qui coupe l'ADN double ou simple brin. On l'utilise par exemple lorsque l'on veut éliminer l'ADN d'une solution ou pour faire des «nicks» sur l'ADN comme dans la technique de marquage par déplacement de coupure (nick translation). Dans ce dernier cas on effectue une digestion ménagée.

En présence de Mg^{2+} , la Dnase I coupe chaque brin de manière indépendante. En présence d'ions Mn^{2+} (10 mM), elle produit préférentiellement des coupures doubles brin. On l'utilise alors pour générer des fragments au hasard.

Nucléase S1 : dégrade principalement l'ADN simple brin et l'ARN. On l'utilise par exemple lorsqu'on veut réduire la taille d'un fragment d'ADN en l'attaquant par ses extrémités. Dans ce cas on fait agir l'exonucléase III qui grignote l'extrémité 3', on obtient une extrémité 5' simple brin sortante puis on fait agir la nucléase S1 pour digérer cette extrémité.

Un deuxième exemple d'application est la cartographie de la structure d'un gène, de sa composition en intron et exon (voir figure ci-dessous). Pour cela, on utilise deux clones, un provenant du génome (clone génomique) et un provenant de l'ARN (clone cDNA) que l'on place dans un tube, après dénaturation et hybridation on obtient trois sortes de molécules double brins, deux homoduplex correspondant aux réassociations des clones génomiques et cDNA et un hétéroduplex correspondant à l'hybride.

Dans l'hétéroduplex, les introns présents dans le clone génomique se retrouvent simple brin et sont donc sensibles à l'attaque par la nucléase S1. Une analyse sur gel d'électrophorèse permettra de connaître le nombre d'introns, le nombre et la taille des exons.

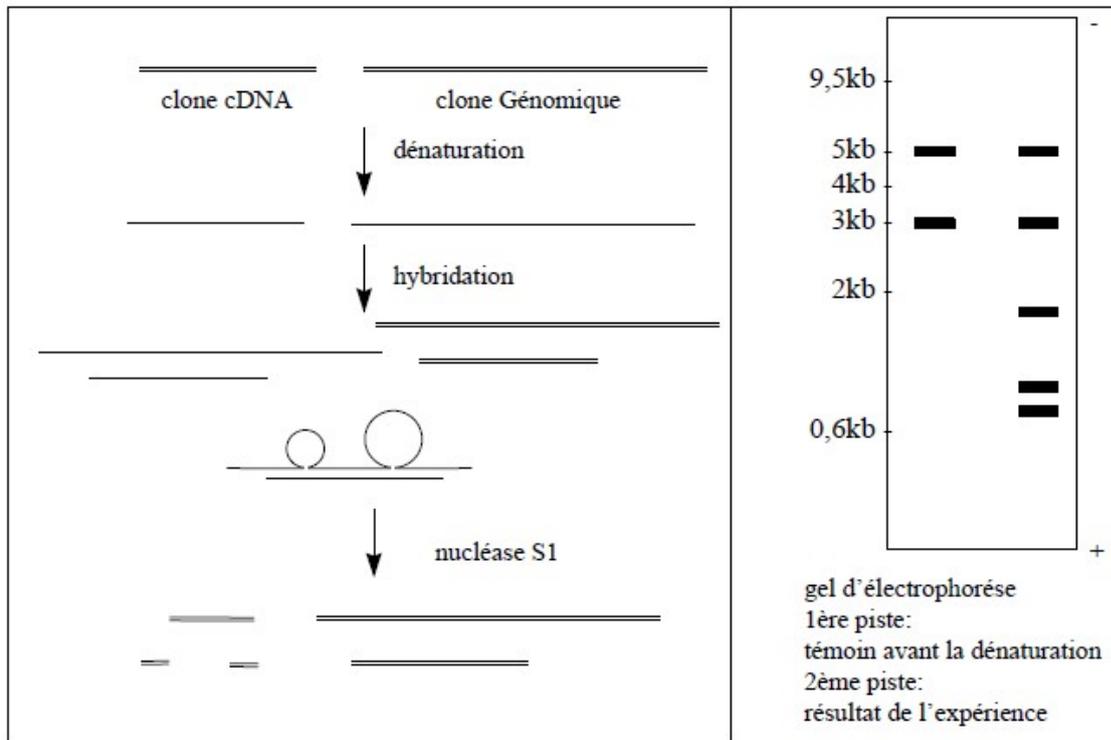


Figure 8. Schéma d'une application de S1 en cartographie.

La **mungbean nucléase** a la même action que la nucléase S1: elle digère le simple brin. Toutefois, s'il y a « nick » sur un fragment double brin, elle aura du mal à digérer la liaison phosphodiester se trouvant en face du « nick » sur l'autre brin. On va l'employer par exemple pour rendre franches les extrémités de fragments d'ADN obtenus par sonication. En effet, dans ce cas on veut rendre franc le fragment sans recouper les fragments au niveau des « nicks ».

IV.1.2.2. Les DNases spécifiques de l'ADN : enzyme de restriction

Les **DNase de restriction** (ou enzyme de restriction, restrictase) endonucléases qui coupent au niveau d'une séquence particulière ou à proximité.

Il existe trois grandes classes (type I, II et III) seules les enzymes du type II sont utilisées en biologie moléculaire. Elles reconnaissent une séquence et coupent au niveau d'un site fixe appartenant à cette séquence ou à sa proximité.

Les types I et III ne sont pas utilisés en biologie moléculaire, en effet ils portent aussi une activité de méthylation et l'activité de coupure est ATP

dépendante. De plus les enzymes de type I ne coupent pas au niveau du site de reconnaissance, elles digèrent l'ADN au hasard.

Les enzymes de restriction de type II reconnaissent des séquences le plus souvent palindromiques et libèrent des extrémités 5' phosphate et 3' OH et donnent soit des extrémités cohésives soit des extrémités franches. Par exemple :

- Eco RI reconnaît et coupe la séquence G/AATTC (5' sortante)
- Pst I reconnaît et coupe la séquence CTGCA/G (3' sortante)
- Sma I reconnaît et coupe la séquence CCC/GGG ce qui donne des extrémités franches

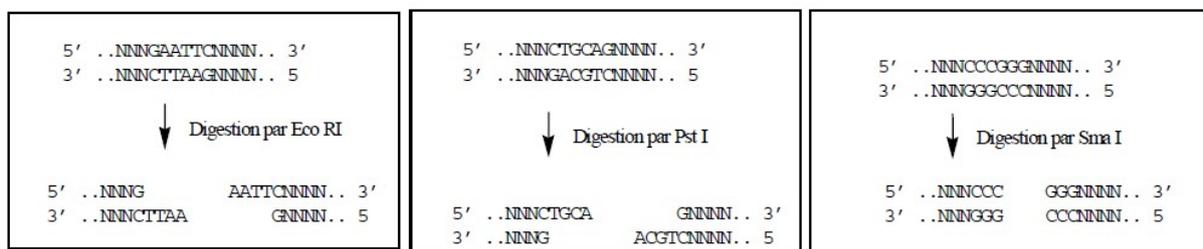


Figure 9. Quelques exemples des enzymes de restriction.

La reconnaissance n'est pas toujours strictement définie [exemple : Nci I CC/(G ou C) GG] et les palindromes peuvent être interrompus [exemple : XmnI reconnaît la séquence suivante : GAANN/NNTTC]. Fragment de restriction : fragment d'ADN généré après coupure d'une molécule d'ADN par des endonucléases de restriction.

Le nom des enzymes de restriction provient de la bactérie dont elles ont été isolées. Par exemple EcoRI provient d'*Escherichia coli* souche R, le I signifie que c'est la première enzyme qui a été découverte. Pour la bactérie, cette enzyme sert à digérer l'ADN des phages qui pourraient l'infecter.

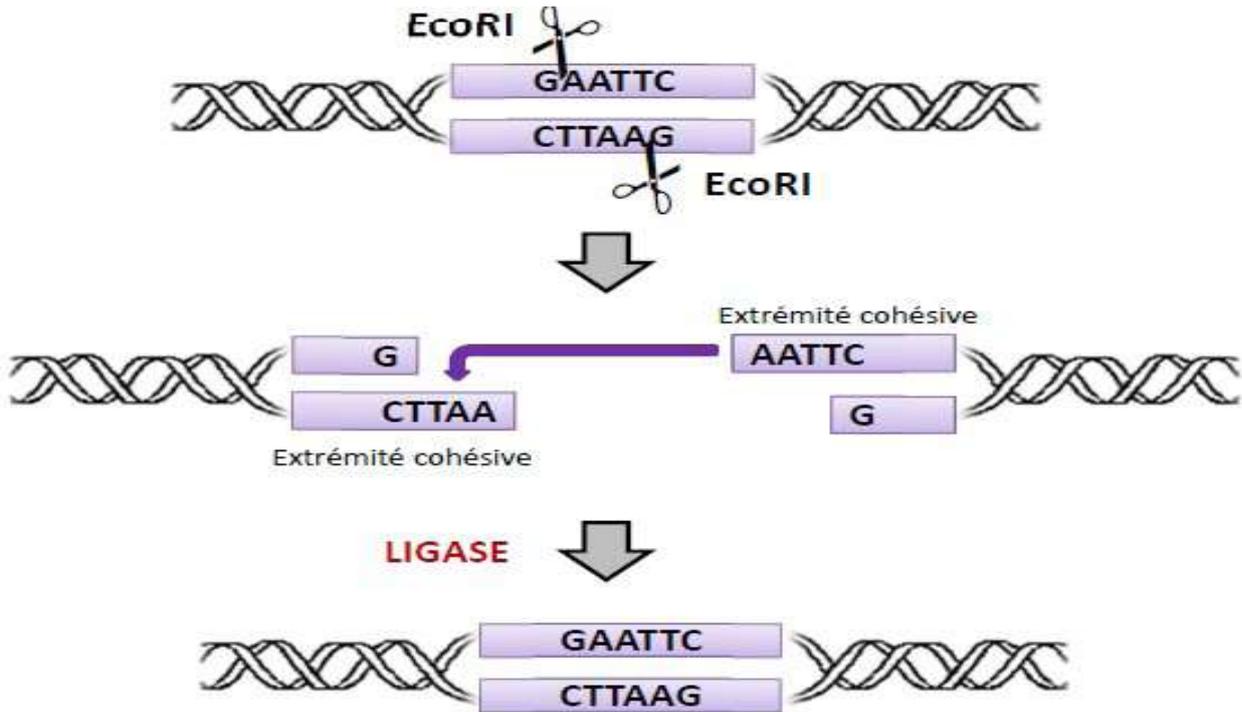


Figure 10. Coupure d'ADN par l'enzyme de restriction EcoRI.

Les Enzymes de restriction ER coupent au niveau de sites spécifiques appelés « sites de restrictions », ces sites sont de nature palindromique, c'est-à-dire que la lecture des deux brins complémentaires dans des sens opposés donne la même séquence.

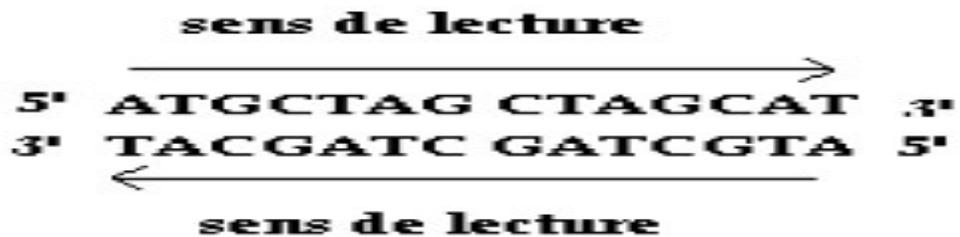


Figure 11. Séquence palindromique

Les enzymes de restriction reconnaissent 4, 6 ou 8 bases. La fréquence de reconnaissance sera d'autant plus faible que le nombre de bases reconnues est élevé. Une enzyme reconnaissant 4 bases coupera en moyenne tous les 44 bases soit tous les 256 nucléotides. De même, une enzyme reconnaissant 6 bases coupera tous les 4096 nucléotides et une enzyme reconnaissant 8 bases tous les 65536 nucléotides.

Les enzymes de restriction reconnaissent souvent une structure de l'ADN et non uniquement la séquence. De ce fait, certains enzymes comme Not I ne reconnaissent pas l'ADN surenroulé. Si un plasmide surenroulé est incubé en présence de Not I, on n'obtient pas de coupure, par contre si le plasmide est linéarisé on observe la coupure.

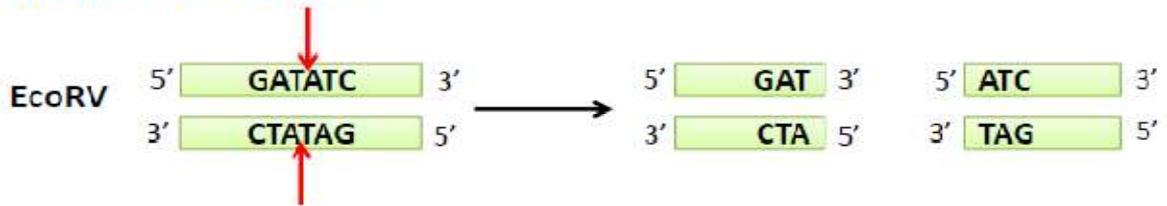
Tableau 1 : Exemple de quelques enzymes de restriction et leurs origines

Enzyme	microorganisme	SITE DE RESTRICTION
<i>EcoRI</i>	<i>Echerichia coli</i>	↓ GAATTC CTTAAG ↑
<i>EcoRII</i>	<i>Echerichia coli</i>	↓ GCCTGGC CGGACCG ↑
<i>PstI</i>	<i>Providentia stuartii</i>	↓ CTGCAG GACGTC ↑
<i>HindIII</i>	<i>Haemophilus influenzae</i>	↓ AAGCTT TTCGAA ↑

A la suite de la coupure d'un ADN génomique, on obtient une population de molécules dont la taille moyenne est de 256 pb pour une reconnaissance de 4 bases. Après les coupures les enzymes de restriction (ER) donnent des extrémités franches ou bien des extrémités cohésives.

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 clives par des enzymes différentes et former des extrémités incompatibles, et deux sites différents peuvent former des extrémités compatibles.

- Extrémités franches



- Extrémités cohésives

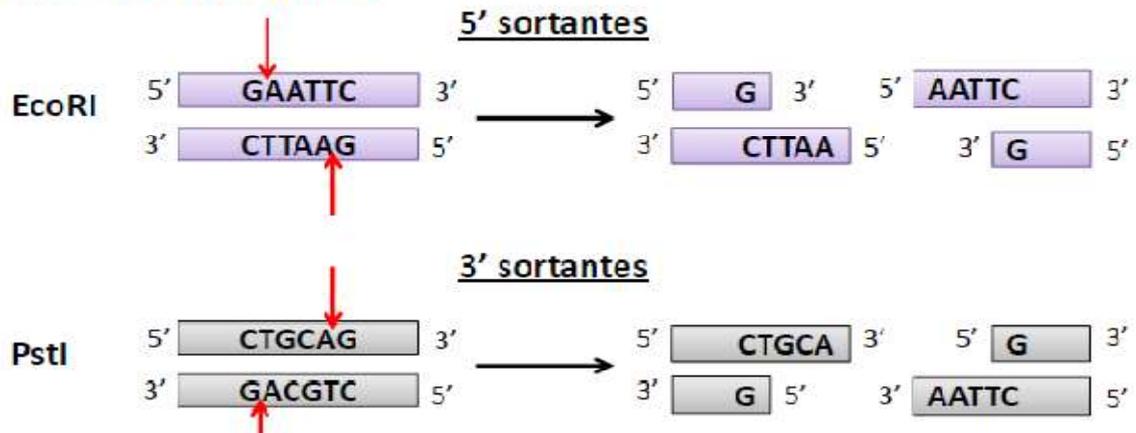


Figure 12. Démonstration de deux types d'extrémités générées par les ER

Il y a deux grands domaines d'utilisation des enzymes de restriction. Le premier est la cartographie d'un fragment d'ADN. On peut en effet identifier un fragment d'ADN grâce à sa carte de restriction, c'est à dire grâce à la position relative des sites de restriction. Le deuxième domaine d'utilisation est la construction de molécule d'ADN recombinante. Un fragment de restriction peut être recollé (ligué) à un autre fragment d'ADN si les deux extrémités sont compatibles, c'est à dire présentent les mêmes extrémités cohésives ou des extrémités franches.

IV.2. Les polymérases

Toutes les polymérases synthétisent les acides nucléiques de 5' vers 3' en utilisant des nucléotides triphosphates. L'énergie est fournie par les nucléotides triphosphates entrants.

Pourquoi la polymérisation ne se fait-elle pas dans l'autre sens ? Dans le sens 5'3', en cas d'erreur il y a excision du dernier nucléotide incorporé libérant un 3' OH. La synthèse peut continuer avec l'ajout d'un nouveau nucléotide. Si la synthèse se faisait dans l'autre sens de 3' vers 5', l'énergie de la réaction serait fournie par

l'extrémité 5' triphosphate de la chaîne en cours de polymérisation. En cas d'erreur, il y aurait excision du dernier nucléotide et on obtiendrait une nouvelle extrémité 5' monophosphate. On aurait donc besoin d'énergie pour repartir dans la synthèse.

On distingue les ADN polymérase ADN dépendantes (ou ADN polymérase), des ADN polymérase ARN dépendantes (ou reverse transcriptases) des ADN polymérase n'ayant pas besoin de matrice (les terminal transférases)

IV.2.1. ADN polymérase ADN dépendante ou ADN polymérase

Elles fabriquent de l'ADN en prenant comme matrice de l'ADN. Les ADN polymérase ont besoin d'une base déjà hybridée ou amorce. Le plus souvent l'amorce est un oligonucléotide, simple brin qu'on est capable de synthétiser chimiquement. (Homonymes : primer ou, par extension, oligonucléotide). Toutes les ADN polymérase ont :

- Une activité de polymérisation de 5' vers 3', Les DNA polymérase utilisées en biologie moléculaire sont des enzymes utilisées dans la cellule pour les réparations, elles ont donc une processivité faible : elles synthétisent une vingtaine de nucléotides à la suite avant de se décrocher de l'ADN, une autre molécule reconnaît alors le substrat et continue la synthèse pour une vingtaine de nucléotides (Bambara et al., 1978).
- Une activité terminal-transférase sur double brin avec une préférence pour l'incorporation d'une adénine. Si on incube un fragment d'ADN double brin, présentant une extrémité franche avec une ADN polymérase on obtient une extrémité avec un A sortant en 3'.

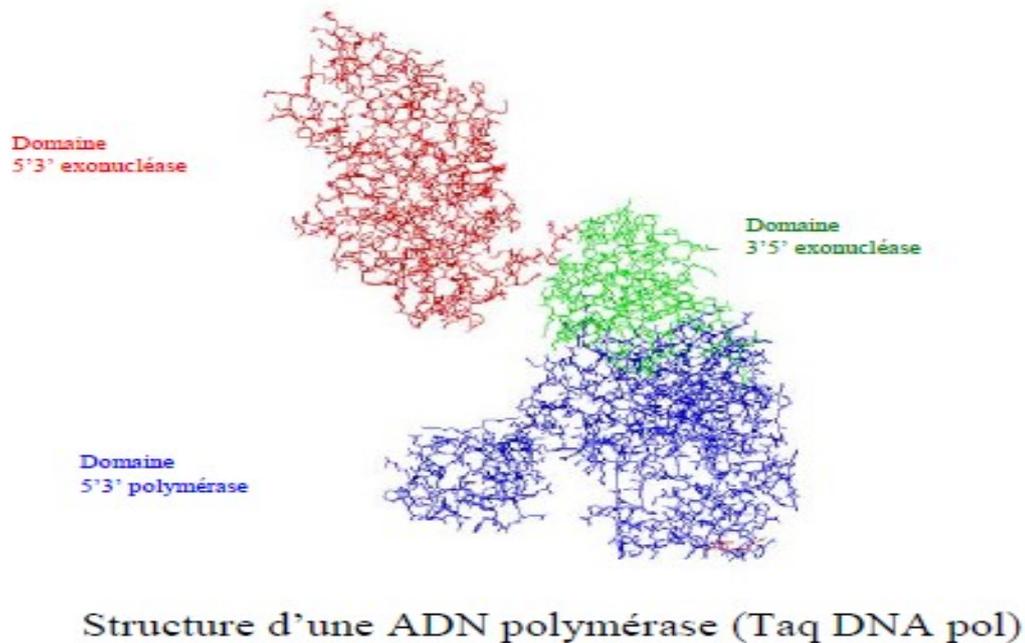


Figure13. Structure d'ADN polymérase utilisées en biologie moléculaire.

Certaines DNA polymérase ont en plus :

- Une activité 3'5' exonucléase, cette activité est souvent appelée activité de correction. En effet si la polymérase fait une erreur, le dernier nucléotide n'est plus hybridé, la polymérisation est bloquée. L'activité 3'5'exonucléase retire le nucléotide non hybridé, la polymérisation peut continuer. Cette activité 3'5' exonucléasique retire aussi le A ajouté en 3' par l'activité terminal-transférase
- Une activité 5'3' exonucléase

La réaction de polymérisation peut fonctionner à l'envers. Si on ajoute une ADN polymérase, un fragment d'ADN et du pyrophosphate, on obtiendra la réaction inverse : digestion de 3' vers 5' du fragment d'ADN et formation de dNTP.

IV.2.2.ADN polymérase ARN dépendante ou reverse transcriptase

Cette enzyme a la même activité que l'ADN polymérase ADN dépendante, mais elle fabrique un ADN à partir d'un ARN. Elle a donc une activité 5'3' ADN polymérase sur un substrat composé d'une amorce hybridée sur un ARN.

L'utilisation principale de cette enzyme est la synthèse d'ADN complémentaire (ADNc) à partir des ARN messagers de cellules eucaryotes. On peut utiliser trois sortes d'amorce :

- soit un oligodT dans ce cas on obtient une population d'ADNc.
- soit une amorce au hasard, on obtient alors une population d'ADNc dont l'extrémité 5' est variable
- soit une amorce spécifique à une séquence présente sur un ARN, on obtient alors des ADNc de séquence unique correspondant à un seul gène à partir d'ARN de cellules eucaryotes.

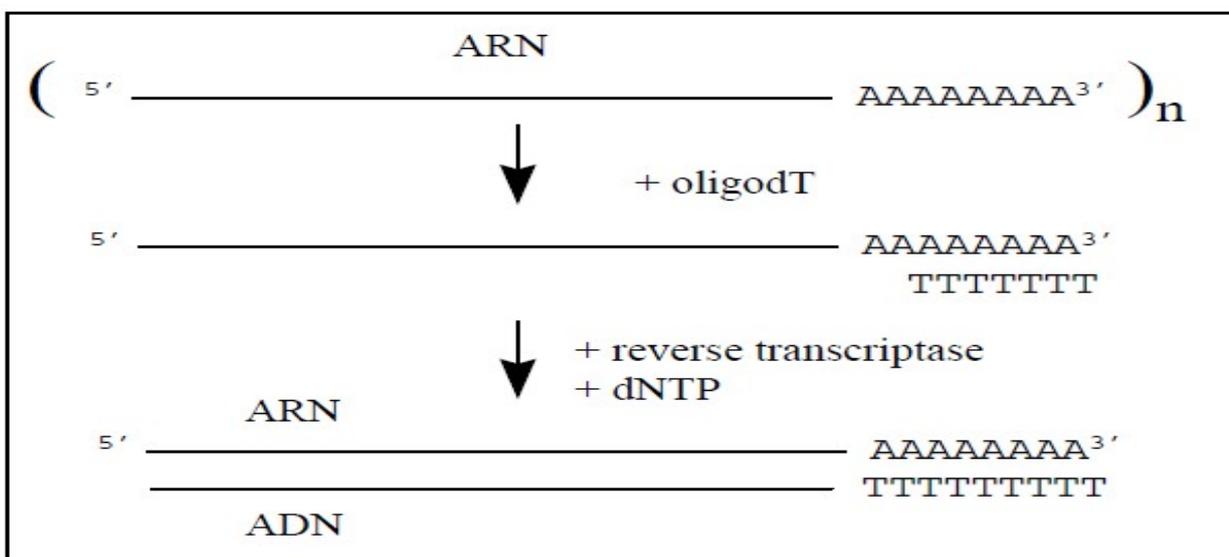


Figure14. L'utilisation principale de l'enzyme Reverse transcriptase

Ces séquences pourront être amplifiées par PCR, l'association successive des deux techniques a pris le nom de RT-PCR. La Reverse transcriptase manque d'activité 3'5' exonucléase si bien qu'il n'y pas de correction d'erreur. On peut augmenter ces erreurs en présence de Mn⁺ et en forte concentration de dNTP, dans ce cas, une base sur 500 est mal incorporée.

Les reverse transcriptases sont des enzymes fragiles. Pour améliorer la synthèse et obtenir de longs cDNA on peut ajouter des stabilisateurs de protéines tels que le tréhalose (170 mM) et le sorbitol (750 mM) (Carninci et al., 2002).

IV.3. Les ligases

Elle catalyse la formation d'un pont phosphodiester entre un 3' OH et un 5' phosphate, elle a besoin d'ATP et d'ions divalents. Elle ligue de l'ADN double brin.

Ces enzymes utilisées pour la ligation d'extrémités cohésives ou d'extrémités franches de fragments de restriction.

Dans le cas d'extrémités franches, la ligation est souvent effectuée à 15-20°C pendant 4 à 16 heures pour favoriser l'interaction entre les extrémités. Si les deux extrémités sont déphosphorylées, la ligation ne peut avoir lieu, par contre si une seule est déphosphorylée, la ligation a lieu sur un des deux brins.

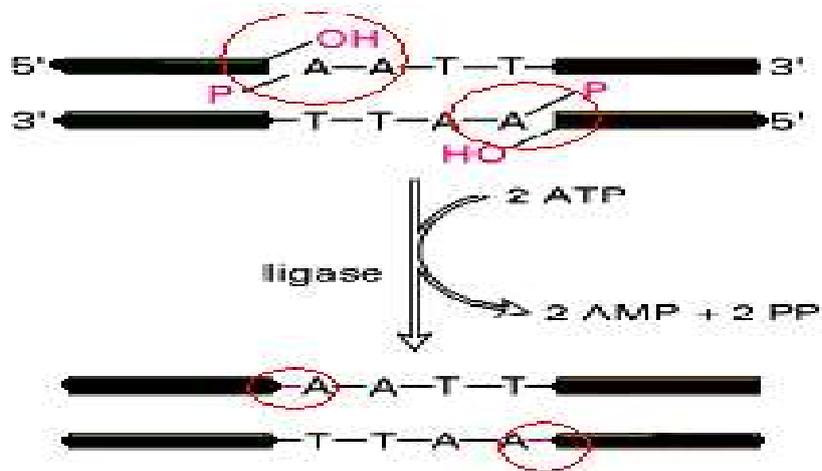


Figure 15. Réalisation de liaisons phosphodiesters par la ligase

V. Vecteurs de clonage et analyse des banques

V.1. Nécessité et principe du clonage

Le clonage moléculaire est une des bases du génie génétique. Il consiste à insérer un fragment d'ADN (dénommé insert) dans un vecteur approprié comme un plasmide par exemple. Le nouveau plasmide ainsi créé sera ensuite introduit dans une cellule hôte, en générale la bactérie *Escherichia coli*. Celle-ci sera alors sélectionnée et multipliée afin d'obtenir une grande quantité du plasmide d'intérêt. Cloner un gène consiste donc à l'insérer dans un plasmide. Un clone sera le transformant bactérien qui contient ce plasmide particulier. Dans ce cas on parle de clone car tous les individus de la colonie bactérienne sont génétiquement identiques.

Le clonage moléculaire est donc différent du clonage reproductif (créer un individu génétiquement identique à un autre mais d'un âge différent) ou du clonage thérapeutique (fabriquer des tissus à partir de cellules souches pour effectuer des greffes compatibles avec le receveur).

V.2. Création d'un vecteur

Les vecteurs sont des molécules d'ADN construites à partir de plasmides ou de virus (cosmides), qui permettent le transfert de gènes entre cellules.

Tous les vecteurs possèdent un polylinkerrefermant plusieurs sites de restriction, où on peut ouvrir le vecteur pour insérer une séquence d'ADN et un gène de résistance aux antibiotiques. Le gène de résistance permet de sélectionner les cellules qui ont reçu le vecteur après transfection. Il existe deux grandes catégories de vecteurs :

A) Vecteur de clonage : renferme une origine de réplication (Ori V), il permet de cloner un segment d'ADN ou un gène qui y est intégré.

B) Vecteur d'expression : Renferme un promoteur, il permet de faire exprimer un gène.

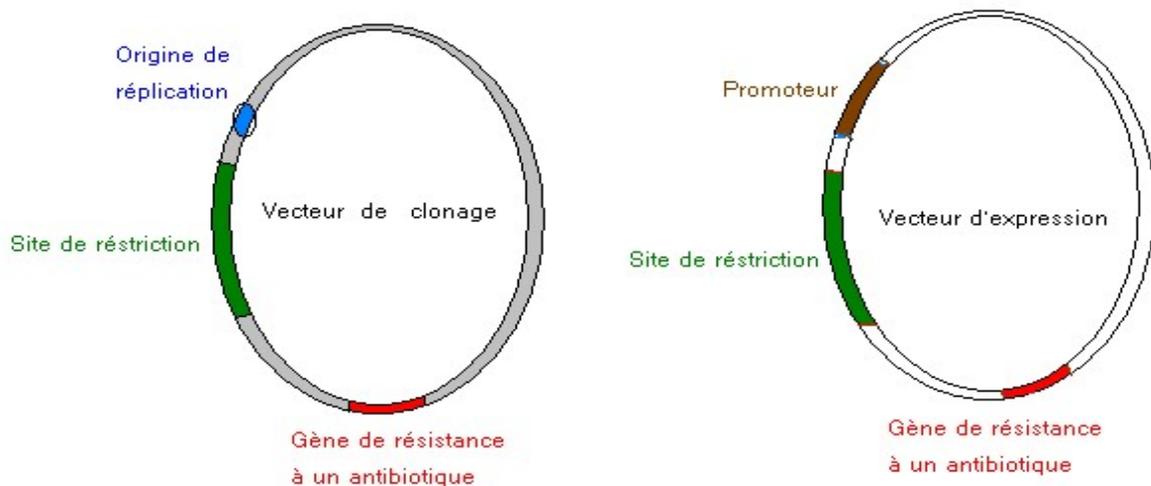


Figure 16. Schéma simplifié des deux types de vecteurs

On a obtenu une très grande quantité de petits fragments d'ADN. Il faudra alors pouvoir isoler les fragments d'intérêt, et les avoir en suffisamment grande

quantité pour pouvoir les utiliser. Pour cela, on introduit les fragments dans des vecteurs grâce à des ligases (ligation). Chaque vecteur recombinant est introduit dans une bactérie (transformation). Les bactéries sont ensuite isolées et reproduites pour identifier et sélectionner les gènes d'intérêt (amplification). Ces trois étapes

On coupe le vecteur par une enzyme de restriction. Le vecteur linéarisé est mis en présence de phosphatase pour empêcher la ligation de l'ADN avant que l'on ait pu le modifier. On introduit alors les fragments d'ADN d'intérêt (coupe par une même enzyme de restriction). L'ADN introduit possédant un groupe phosphate, la ligation pourra se faire.

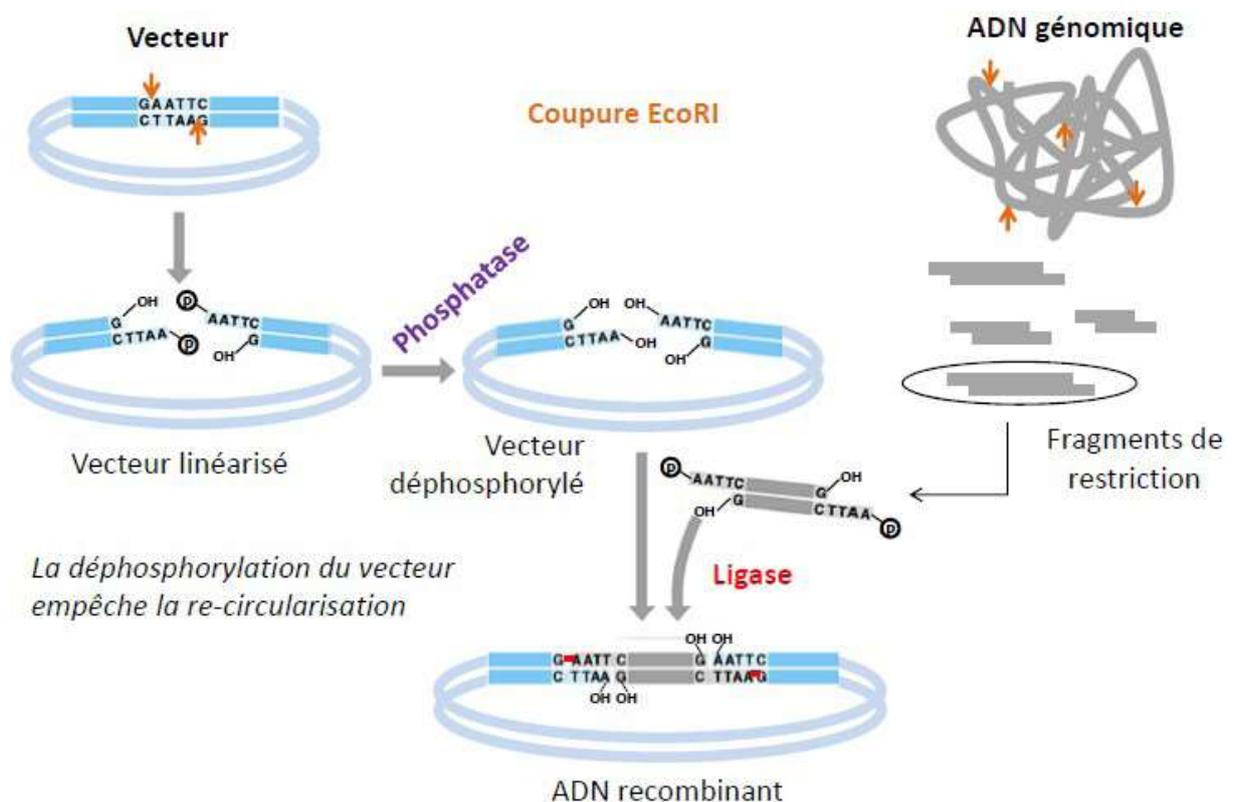


Figure 17. Les étapes de création d'un vecteur pour le clonage

V.3. Différents types de vecteurs

V.3.1. Vecteurs plasmidiques

Les plasmides sont de petits morceaux d'ADN circulaires, double brin que l'on trouve dans les bactéries en dehors du chromosome (épisomes). Ils se répliquent grâce aux enzymes présents dans la bactérie.

On est capable de les introduire dans une bactérie par des méthodes chimiques (CaCl₂) ou des méthodes physiques (électroporation). Les parties essentielles :

- Une origine de réplication, importante pour l'initiation de la réplication.
- Un site de restriction pour pouvoir insérer un fragment d'ADN (ce fragment d'ADN est souvent appelé insert).
- Un gène de résistance aux antibiotiques : permet de sélectionner les bactéries ayant incorporés le plasmide de celles qui ne l'ont pas incorporé. (La transformation est un événement rare, il n'y a qu'un faible nombre de bactéries qui sont transformées)

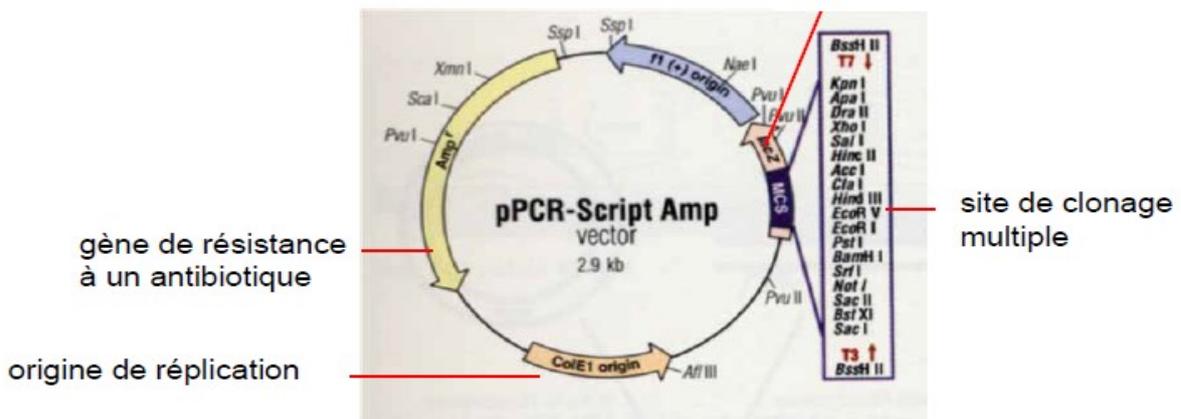


Figure 18. Schéma d'un vecteur plasmidique

V.3.2. Vecteurs dérivés du phage λ

Le phage λ c'est un virus à ADN double brin, linéaire de 50 kb avec à ses deux extrémités 12 nucléotides simple brin complémentaires (extrémité cohésives ou cos). Il se fixe sur le récepteur de la bactérie pour faire entrer de l'ADN dans la bactérie.

Les connaissances accumulées sur ce phage ont permis de l'utiliser comme un vecteur. Pour l'utiliser en tant que vecteur on retire ces 20 kb et on insère à la place le fragment d'ADN d'intérêt.

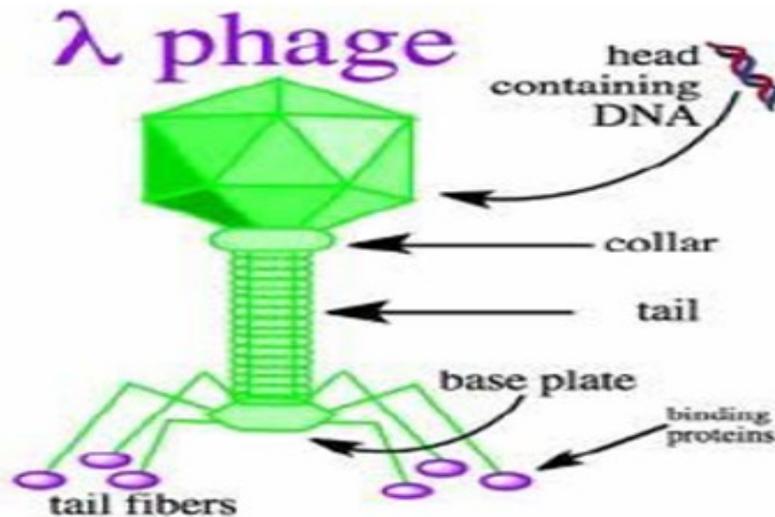


Figure 19. Schéma du phage λ

Phage λ principalement pour faire des banques à ADN, on profite ainsi des plages de lyse qu'il est plus facile à cribler que des colonies. On l'utilise pour des banques génomiques dans ce cas les inserts font de 15 à 20 kb, ou pour des banques d'ADNc. Dans ce dernier cas, les inserts sont plus petits (<10 kb) et n'influencent pas l'empaquetage in vitro.

V.3.3. Vecteur cosmides

Un cosmide est un plasmide (molécule d'ADN capable de se cloner) permettant de cloner des fragments d'ADN de grande taille. Ce type de structure ne se trouve que dans les bactéries.

La particularité d'un cosmide est de contenir une séquence d'ADN cos (d'où son nom) contenant environ 200 paires de bases d'ADN issues du phage Lambda (particule virale qui infecte spécifiquement la bactérie *Escherichia Coli*). Par ailleurs, les cosmides permettent d'incorporer davantage d'information que les plasmides avec en moyenne 45 milliers de bases (kilobases) d'ADN.

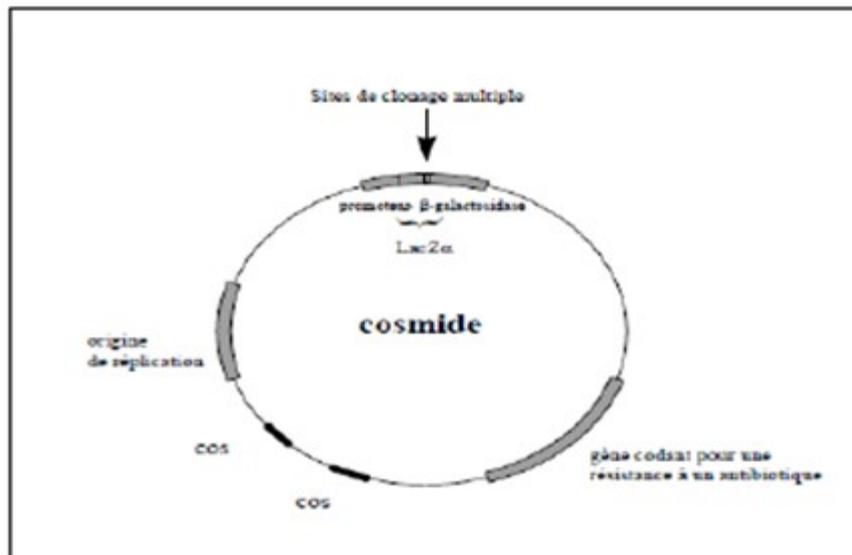


Figure 20. Schéma du vecteurs cosmide

V.3.4. Chromosomes artificiels bactériens (BAC)

Les BAC (Bacterial artificial chromosome). Leur taille est d'environ 100 kb, et ils se maintiennent chez *E. coli* comme simple copie. Composés de plusieurs parties : *parA*, *parB* and *repE* sont des gènes requis pour la stabilité du BAC. *OriS* est une origine de réplication et le CM : gène de résistance au chloramphenicol. Les avantages de ce vecteur sont:

- Grande capacité d'insertion
- Facilité de manipulation (comme plasmide)
- Hôte bactérien
- Copie unique

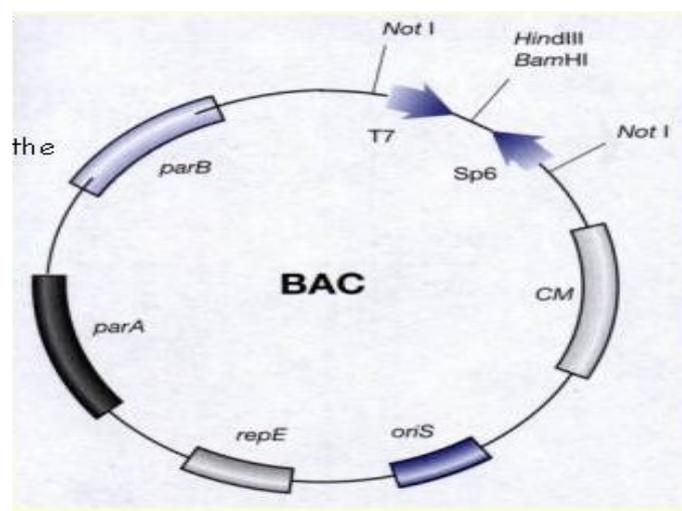


Figure 21. Schéma du vecteurs BAC

V.3.5. Chromosomes artificiels de levure (YAC)

Les YAC (Yeast Artificial Chromosome) sont dérivés des chromosomes de *Saccharomyces cerevisiae*. Ils contiennent une origine de réplication, un centromère et deux télomères. Ces vecteurs ont été principalement utilisés pour réaliser des clonages positionnels mais cette application est aujourd'hui abandonnée car elle est difficile à réaliser et qu'il y avait trop de recombinaisons.

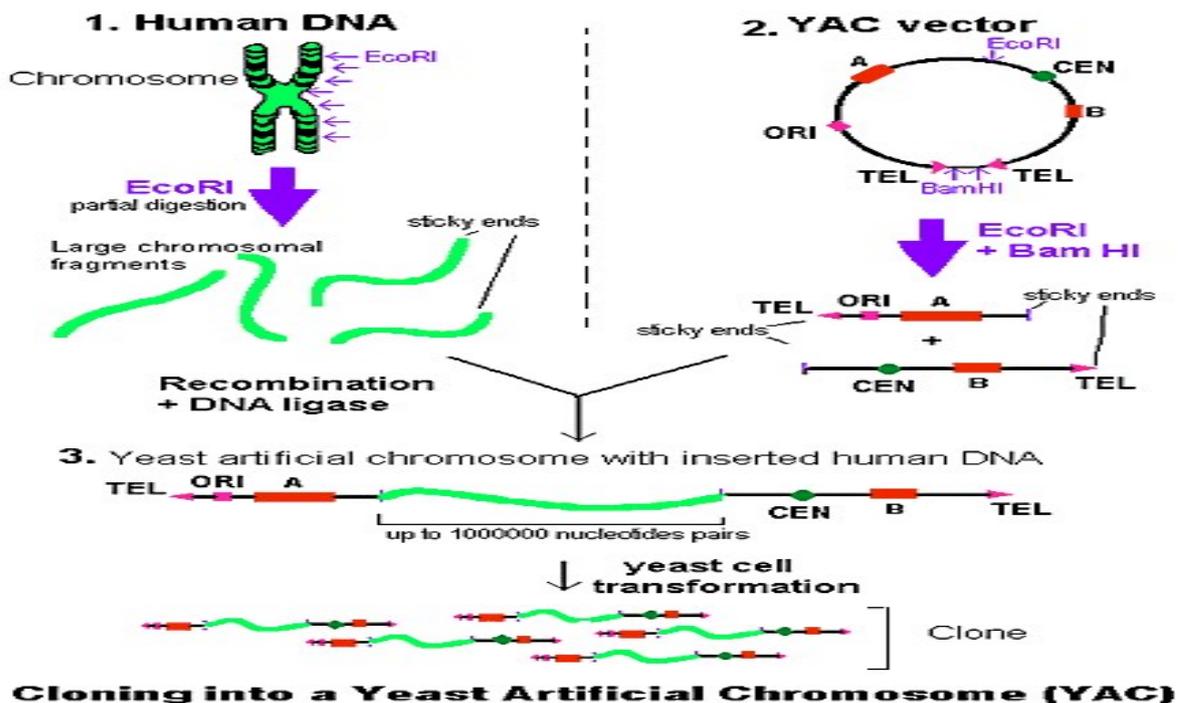


Figure 22. Schéma du vecteurs YAC

VI. L'amplification *in vitro* de séquences spécifiques par PCR

La technique PCR permet d'amplifier de courtes séquences d'ADN (quelques kilobases) *in vitro*, à partir d'une très petite quantité d'ADN même issue d'une seule cellule, par une série de cycles se déroulant en trois étapes : dénaturation, hybridation de l'amorce, réplication. Le nombre de molécule obtenu après n cycles est 2^n .

- Dénaturation : l'ADN à amplifier est dénaturé à la chaleur, les deux brins se séparent et servent de matrice.
- Hybridation des amorces : on baisse la température pour permettre à l'amorce de s'hybrider spécifiquement aux extrémités de la séquence cible à amplifier.

- Elongation ou polymérisation à 70°C : à cette température optimale de l'activité de l'enzyme Taq polymérase, l'amorce est allongée par complémentarité au brin matrice, ce qui produit deux nouvelles molécules d'ADN.

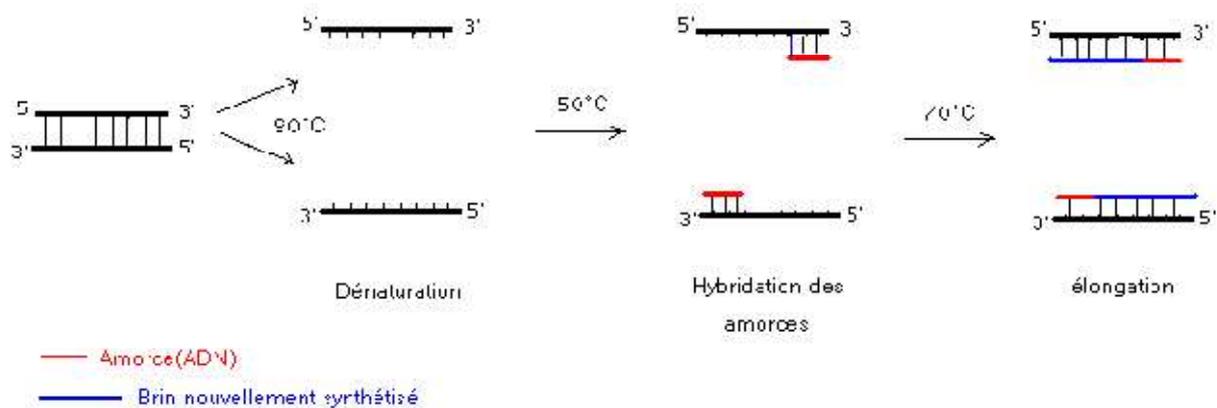


Figure 23 : Représentation schématique des 3 étapes d'un cycle PCR

La PCR permet également d'analyser un vecteur recombinant, pour vérifier la présence d'un insert, et vérifier l'orientation. La PCR permet donc de :

- Récupérer un fragment d'intérêt si l'on connaît partiellement sa séquence (criblage fonctionnel et par hybridation).
- Lorsqu'on connaît le fragment, de l'obtenir ce fragment en grande quantité, sans passer par une banque, et être cloné par une cible cellulaire.

VII. Hybridation d'ADN sur support: technique de SouthernBlot

Le but de la technique est le transfert de l'ADN du gel sur une feuille de nitrocellulose ou de nylon afin de visualiser des séquences précises par hybridation moléculaire avec une sonde marquée.

1. Fragmentation de l'ADN par une endonucléase
2. Séparation des fragments de restrictions en fonction du poids moléculaire (électrophorèse sur gel d'agarose)
3. Transfert des fragments du gel sur une feuille de nitrocellulose (ou membrane de nylon)
4. Dénaturation de l'ADN fixé sur la membrane.

5. Mise en contact de l'ADN et des sondes marquées en milieu liquide et sous agitation.
6. Lavage de la feuille pour éliminer les hybridations non spécifiques.
7. Visualisation des sondes fixées par autoradiographie.

Dans la technique 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 contenant 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.

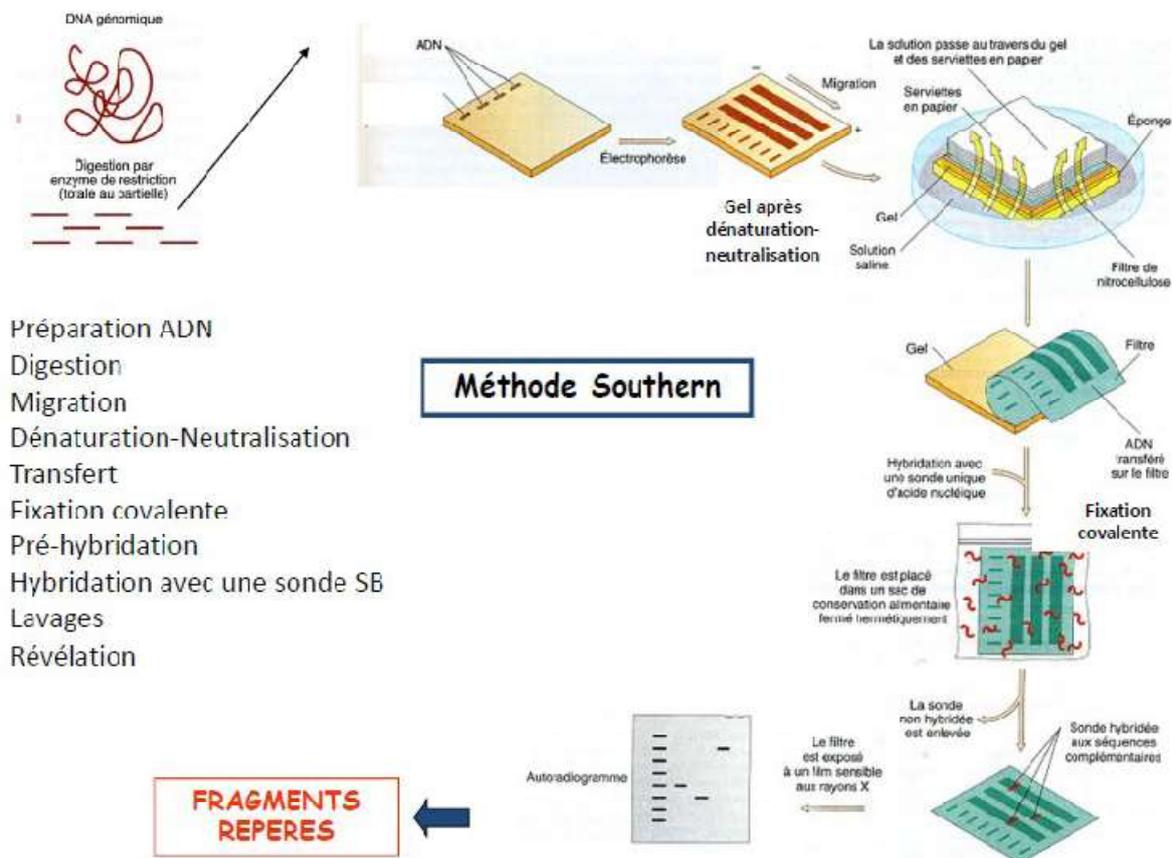
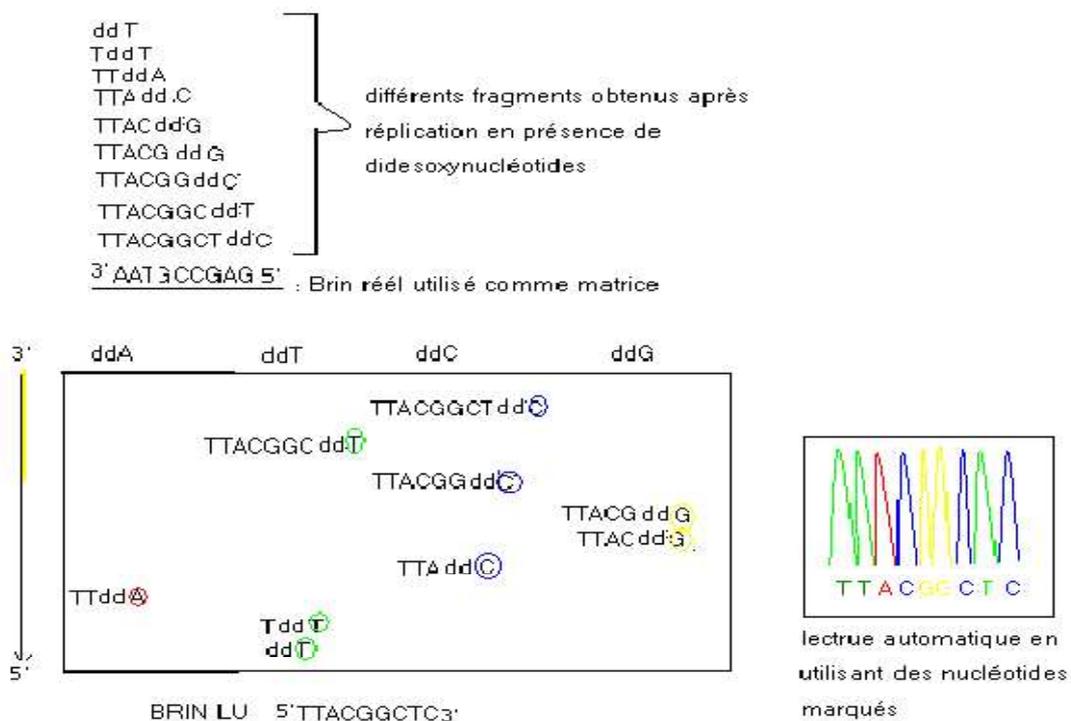


Figure 24 : Stratégie de la technique Southern Blot

VIII. Technique de séquençage

Permet de déterminer la séquence d'ADN en réalisant une série de réplication en présence de didésoxynucléotides qui bloquent la réplication de l'ADN lorsqu'ils sont incorporés dans le nouveau brin. Ainsi on utilise différents didésoxynucléotides (ddA, ddT, ddG, ddC) on obtient plusieurs fragments se terminant chacun par l'un des didésoxynucléotides.

Après séparation des brins néosynthétisés et leur migration sur gel d'électrophorèse, chaque fragment indique la position d'une base au niveau du brin néosynthétisé (brin lu) dont l'ordre est donné par sa position sur le gèle d'électrophorèse



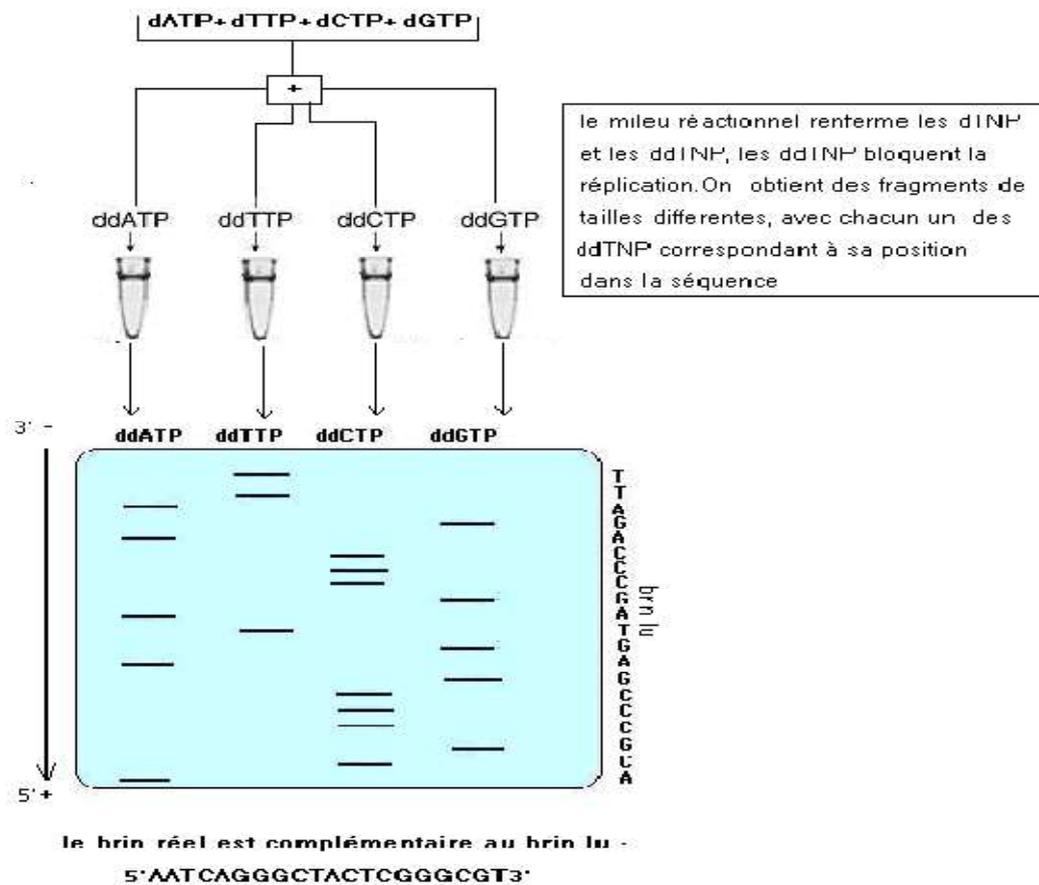


Figure 25: Technique de séquençage classique (Sanger)

IX.Applications du génie génétique

Grace aux outils du génie génétique en particulier et à la biologie moléculaire en général, l'humanité a connu une avancée importante dans le domaine de la recherche, que la médecine et l'agriculture. 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.

IX.1.Production de protéines thérapeutiques

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 synthétisé, ce n'était pourtant pas la plus simple :

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.

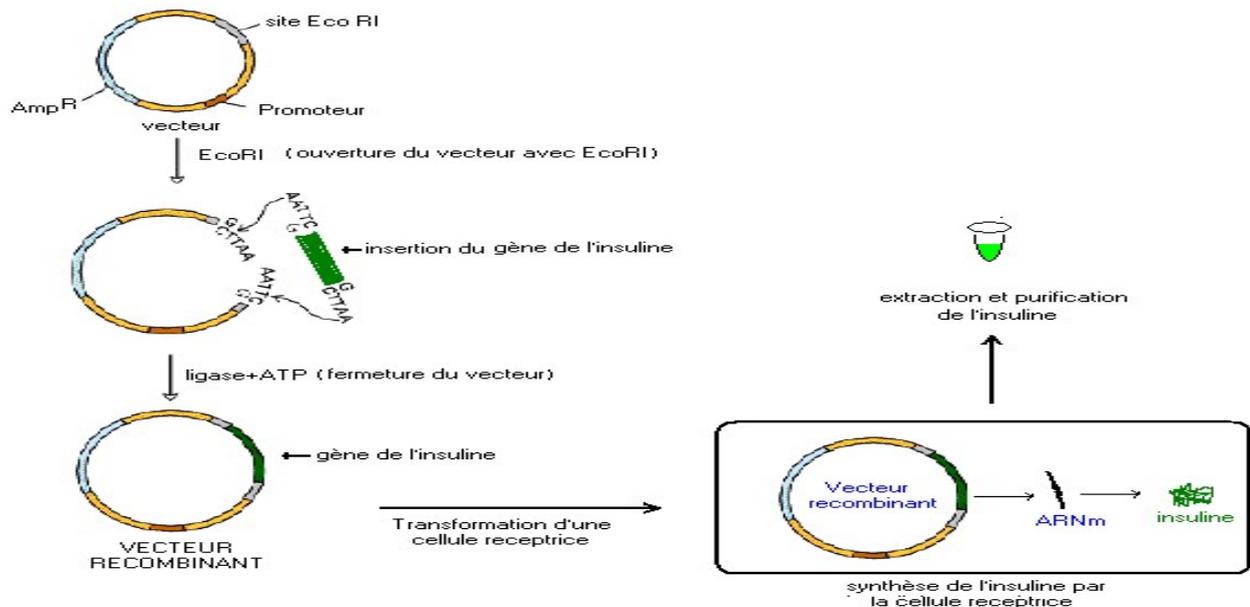


Figure 26 : Stratégie de la synthèse de l'insuline par génie génétique

IX.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 croisse 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.

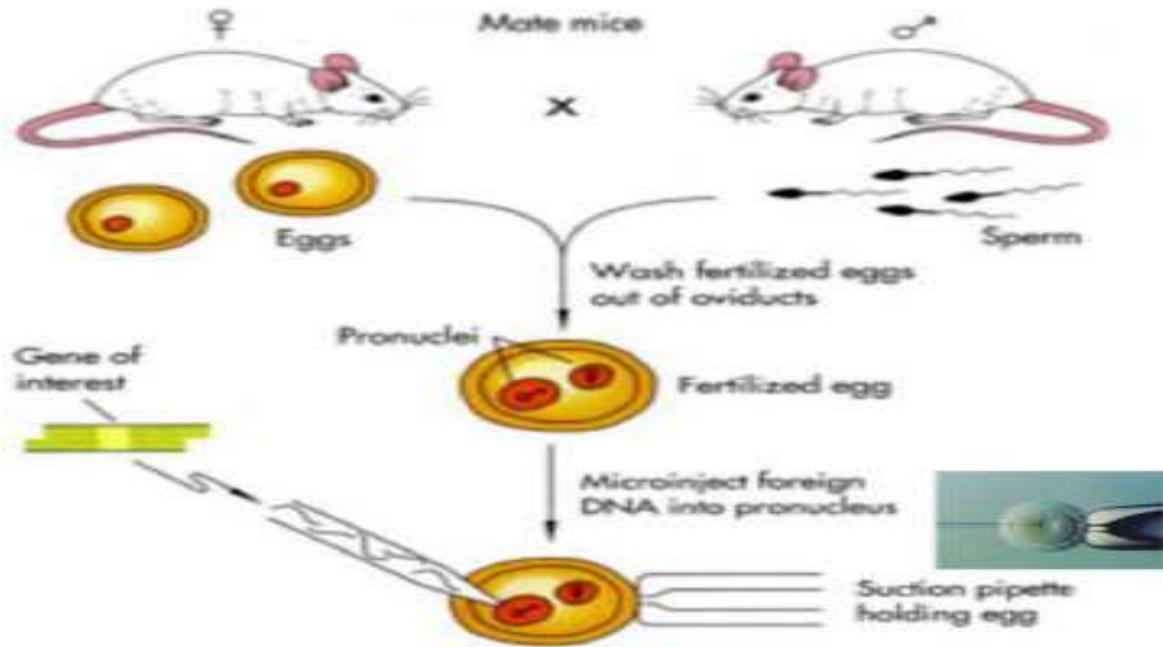


Figure 27 : Stratégie de production d'un animale transgénique

IX.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.

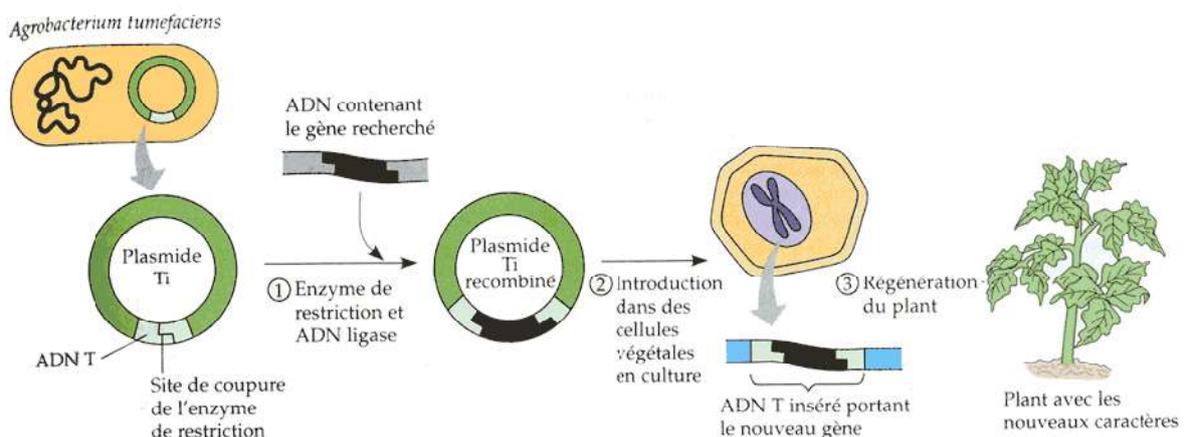


Figure 28 : Stratégie de production d'une plante transgénique

IX.4. Thérapie génique

Le principe de la thérapie génique est simple, le gène malade est pris à partir des cellules de l'individu malade, la mutation est corrigée puis le gène normal est

réintroduit dans la cellule par l'intermédiaire d'un vecteur. La cellule ainsi réparée est réimplantée chez le patient. Pour réaliser cette technique il faut maîtriser certaines contraintes :

- Connaître la séquence du gène
- Trouver le bon vecteur
- Trouver le mode d'administration le plus approprié
- Prolonger la durée de l'effet en utilisant des cellules- souches

Le premier traitement par thérapie génique c'est produit en 2000 ; il s'agit de deux enfants atteints d'une immunodéficience sévère 'enfants bulle' due à la nondifférenciation des lymphocytes. Cette anomalie est provoquée par une mutation des récepteurs gamma C cytokines, qui normalement captent les signaux de différenciation.

Pour ce traitement, le vecteur utilisé est un rétrovirus modifié renfermant une copie normale du gène codant pour le récepteur. Ce vecteur a été ensuite introduit dans des cellules prélevées chez les enfants. Une fois l'expression du gène confirmée, les cellules sont réimplantées chez les enfants. Un mois plus tard, des cellules lymphocytaires matures sont détectées chez les deux enfants.

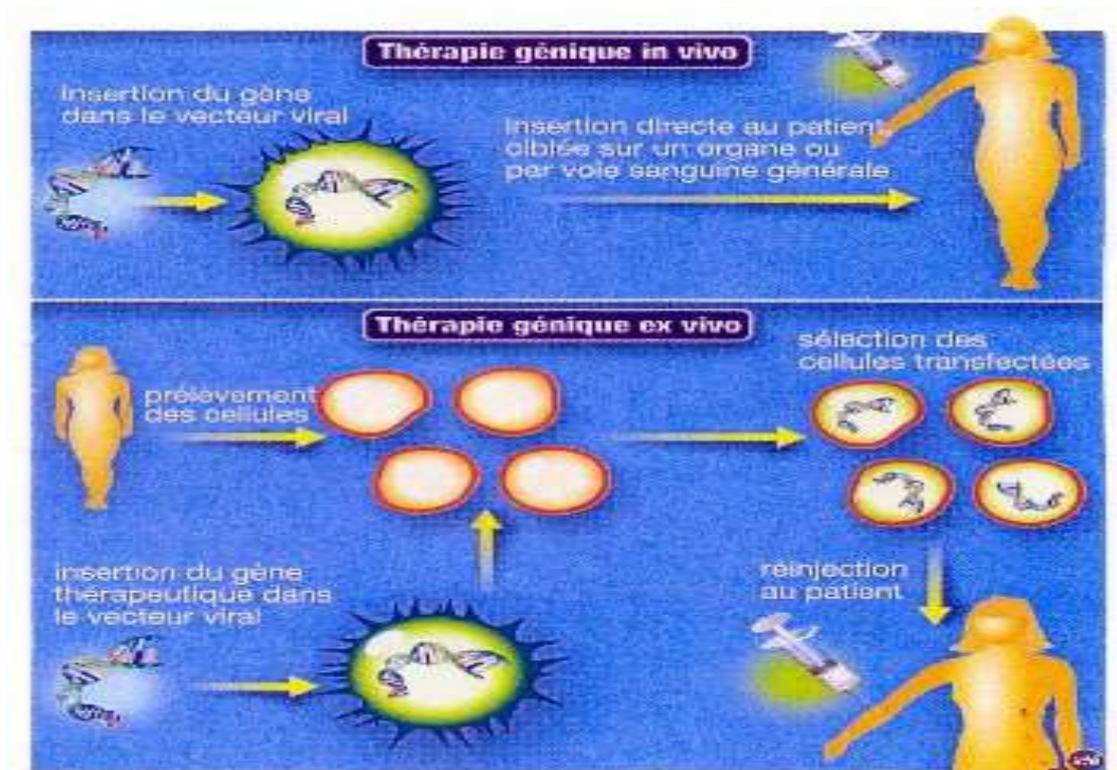


Figure 29 : Stratégie de thérapie génique

IX.5. Diagnostic génétique

Analyse d'une partie ou de l'intégralité du génome (pré-implantatoire, prénatal, post-natal). La drépanocytose est la première maladie ayant été diagnostiquée 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).

L'ADN du patient est extrait à partir du sang périphérique, il est fragmenté par une enzyme de restriction et les fragments ainsi obtenus sont séparés par électrophorèse sur gel d'agarose (voir IV-1). L'ADN est ensuite transféré sur une feuille de nitrocellulose par la technique de Southern blot (voir IV-4). Cet ADN est dénaturé par chauffage pour permettre à la sonde d'agir par complémentarité au gène cible.

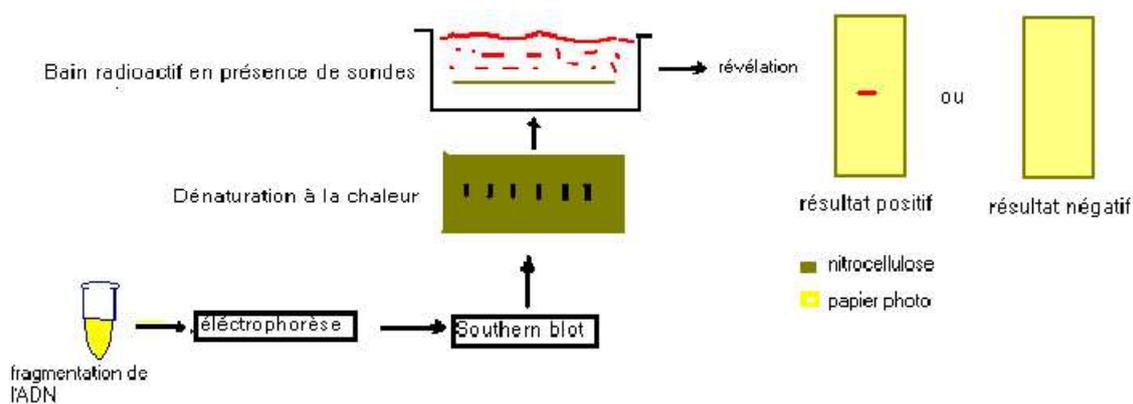


Figure 30 : Diagnostic génétique

Module Génie Génétique

TD N° 1: Enzymes de Restriction

EXERCICE 1 : Représenter les extrémités des fragments de restriction obtenus après digestion par les enzymes de restrictions suivantes en les classant selon la nature de ces extrémités (franches, 5'- ou 3'-sortantes).

BamH1 G/GATCC

HpaII C/CGG

MspI C/CGG

SmaI CCC/GGG

AccI GT/CGAC

KpnI GGTAC/C

EcoRI G/AATTC

EXERCICE 2 : La séquence d'un ADN bicaténaire (double brin), correspondant à un gène, est partiellement reportée ci-dessous.

5' ATACGGGATCCGAGCTCTCGATCGTCTGCAGAAATTCC 3'

- 1) Écrire la séquence et l'orientation du second brin de ce fragment.
- 2) Donner le brin complémentaire d'ARN
- 3) Soient les enzymes de restriction BamH I, Pst I, Xho I et Mbo I dont les sites reconnus sont :

BamH I : 5' G/GATCC 3' ; Pst I : 5' CTGCA/G 3' ; Xho I : 5' C/TCGAG 3' ; Mbo I : 5' /GATC 3'.

Recopier la séquence de l'ADN et encadrer les sites de restriction en indiquant la position des coupures.

- 4) Pour chaque enzyme, écrire les séquences des extrémités des molécules d'ADN digérées et préciser le type d'extrémités obtenu.

TD N° 2: Clonage d'ADN

EXERCICE 1 : Si on utilise une enzyme de restriction possédant un site de restriction de 4 bases, quelles seront la fréquence de coupure d'un ADN génomique et la taille moyenne des fragments de restriction obtenus ?

EXERCICE 2 : En surligné, les sites reconnus respectivement par les enzymes avec en noir, la coupure correspondante

5' ATACGGGATCCGAGCTCTCGATCGTCTGCAGAAATTCC 3'

3' TATGCCCTAGGCTCGAGAGCTAGCAGACGTCTTTAAGG 5'

BamHIMboIPst I

BamHI : 5' G/GATCC 3'

3' CCTAG/G 5'

PstI : 5' CTGCA/G 3'

3' G/ACGTC 5'

XhoI : 5' C/TCGAG 3'

3' GAGCT/C 5'

TD N° 3: Technique d'hybridation moléculaire

EXERCICE 1 :

- 1) On a extrait et purifié un fragment de 15 kbp d'un vecteur λ contenant le gène A. Comment peut-on le fragmenter pour avoir des sous fragments plus petits mais dont certains contiennent encore le gène A entier ?
- 2) Comment intégrer ces sous fragments dans pBR322, repérer les bactéries qui contiennent un plasmide recombiné et ressortir ce sous fragment?
- 3) Trois de ces sous fragments ont été retenus. On en a préparé 3 digestions différentes respectivement par ERI, ER II et ER III, ainsi que des doubles digestions.

digestion par	Fragment A	Fragment B	Fragment C
ERI	3,2+0,4+0,3	2,6+2,2+0,3	2,6+1,8+1,7
ER II	1,8+1,1+1	3,3+1+0,8	3,3+1,7+0,8+0,3
ERI + ER II	1,4+1,1+0,7+0,4+0,3	1,9+1,4+0,8+0,7+0,3	1,9+1,4+1+0,8+0,7+0,3
ER III	2,4+1,5	3+1,9+0,2	2,7+2,2+1,2
ERI + ER III	2+1,2+0,4+0,3	2+1,6+1+0,3+0,2	1,7+1,6+1,2+1+0,6
ER II + ER III	1,8+1+0,6+0,5	2,4+1+0,9+0,6+0,2	2,4+1,3+0,9+0,8+0,4+0,3

Quelle est la carte de restriction des fragments A, B, C. Placez les uns par rapport aux autres. Quelle est la partie qui contient le gène A ?

TD N° 4: Technique de séquençage

EXERCICE 1 :

Soit un ADN monocaténaire dont on décide de déterminer sa séquence nucléotidique primaire. Par hybridation moléculaire, on fixe sur l'extrémité 3' une courte amorce (primer) complémentaire marquée au ^{32}P sur son extrémité 5' et possédant un 3'OH libre. On réalise 4 réactions indépendantes de polymérisation. Les 4 réactions contiennent les éléments suivants :

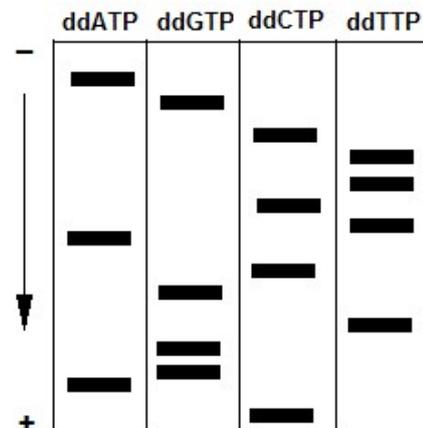
- ADN monocaténaire avec une amorce fixée.
- dATP, dTTP, dCTP et dGTP.
- DNA polymérase I d'*E. coli*, enzyme synthétisant une copie d'ADN complémentaire à l'ADN simple brin
- Pour chacune des 4 réactions, on ajoute un des 4 didésoxyribonucléotides triphosphate (ddXTP) ne possédant pas de OH en position 3' (ddATP, ddTTP, ddCYP et ddGTP).

1- Pourquoi utilise-t-on une amorce dont l'extrémité porte une fonction hydroxyle libre ?

2- Sachant que la polymérase utilisée reconnaît également les ddXTP comme substrat, quelle sera la conséquence de l'incorporation dans la chaîne d'ADN néosynthétisée, de ddXTP donné à la place d'un dXTP correspondant ?

3- Les 4 réactions sont réalisées simultanément. Le milieu réactionnel est chauffé ensuite à 100°C , refroidi rapidement puis soumis à une électrophorèse sur gel de polyacrylamide dénaturant (gel de s"quençage). Après migration, le gel est analysé par autoradiographie.

- Quel est le but de traitement thermique du milieu réactionnel ?
- Après autoradiographie, quelles sont les chaînes nucléotidiques qui seront visualisées.
- A partir de l'autoradiogramme ci dessous, déduisez une partie de la séquence nucléotidique primaire correspondant à l'ADN monocaténaire de départ.



TD N° 5: Exposés sous formes de poster des différentes applications de génie génétiques

Les thèmes exploités :

- Clonage et amplification d'une séquence d'ADN
- Création des banques d'ADN génomiques et d'ADNc
- Introduction d'un gène dans des cellules, des plantes ou des organismes (animaux transgéniques)
- Production de protéines codées par les gènes insérés

XI. Références bibliographiques

1. **Genes VI** (traduction de la 6ème édition anglaise). B. Lewin 1999, De Boeck, ISBN / EAN13 : 9782744500244
2. **Introduction à l'analyse génétique** (4ème édition - traduction de la 8ème américaine)
A.J.F. Griffiths et al. 2006, De Boeck, 9782804150198
3. **Expression des gènes et génie génétique**. M. Crépin. 1997, Hermann, Collection Hermann Sc./Exc, 9782705660611
4. **Molecular biology of the gene** (5th edition). J.D. Watson et al. 2004, Benjamin Cummings, 9780321248640
5. **L'essentiel de la biologie cellulaire : introduction à la biologie moléculaire de la cellule**. Ed. Bruce Alberts. 2005, Flammarion, Collection Médecine et Sciences, 9782257151179
6. **La Cellule - Biologie moléculaire**. J. Darnell, L. Lodish. 2005, De Boeck, 9782804148027