

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE Batna 2
FACULTE DES SCIENCES DE LA NATURE ET DE LA VIE
DEPARTEMENT D'Ecologie et environnement

POLYCOPIE DE COURS

MODULE

BIOLOGIE MOLÉCULAIRE
MASTER I (EDP)

Présenté par

Mr. Ghedadba Nabil

Polycopie de cours de Biologie moléculaire

Intitulé du Master : Ecophysiologie et développement des plantes

Semestre : 02

Intitulé de l'UE : UED1.

Intitulé de la matière : Biologie moléculaire

Crédits : 2

Coefficients : 2

Objectifs de l'enseignement : Initiation théorique et pratique aux techniques de la biologie moléculaire.

Connaissances préalables recommandées : Biologie moléculaire.

Contenu de la matière :

- i. Définition
 1. DNA recombinant, clonage,
 2. Expression,
 3. Banques Génomique
- ii. Les outils de la biologie moléculaire
 1. Enzymes de restriction
 2. Les ligases
 3. Phosphatases
 4. Kinases
 5. Les vecteurs de clonage
 - 5.1. Les plasmides
 - 5.2. Les phages
 - 5.3. Phagemides
 - 5.4. Les cosmides
 - 5.5. Autres vecteurs de clonage
 6. les cellules hôtes
 7. Les sondes nucléotidiques
- iii. Techniques de biologie moléculaire
 1. criblage de banques
 2. cDNA
 3. Purification des AN , analyse quantitative séquençage
 4. Technique de Southern blot et Northern blot
 5. PCR
 6. Applications : recherche d'un gène, transfert de gène

Mode d'évaluation : 50% Contrôle continu + 50% examen final

1- Quelques définitions

1-1- L'ADN recombinant (ou recombiné)

L'**ADN recombinant** (ADNr) désigne de façon générale, ce qui est obtenu par génie génétique. C'est un fragment de matériel génétique d'un organisme qui est introduit artificiellement dans l'ADN d'un autre organisme, généralement une **bactérie** ou un **virus** dans lequel il va s'intégrer.

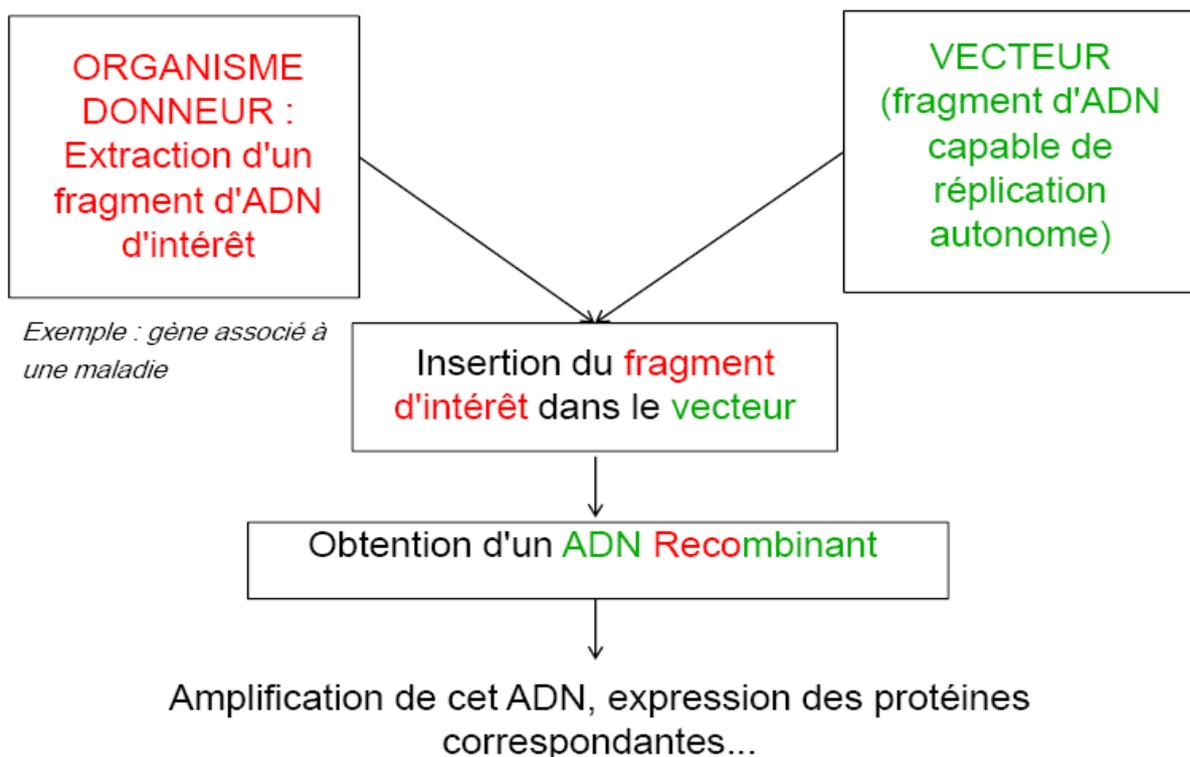
On appelle « ADN recombinant » un ADN hybride obtenu *in vitro* par combinaison de deux molécules d'ADN appartenant à deux espèces différentes. Par exemple un fragment d'ADN humain à étudier, dit ADN d'intérêt, est greffé sur un ADN viral.

L'ADN inséré est appelé « insert » ou « ADN exogène » ou « ADN étranger ».

L'ADN recombinant est préparé dans le but d'effectuer « un clonage », ce clonage peut être ou non accompagné d'une expression de protéines.

ADNr = Vecteur + insert (ADN d'intérêt)

Principe



1-2- Clonage de l'ADN recombinant

Un « clone » est, par définition, un grand nombre de cellules, ou de molécules identiques, provenant d'un seul ancêtre, que cet ancêtre soit une cellule ou une molécule.

L'ADN recombinant a le grand avantage de pouvoir se répliquer dans une bactérie comme l'aurait fait le bactériophage. C'est donc un moyen commode de produire une amplification du fragment d'ADN à étudier.

Une amplification d'ADN peut aussi être obtenue par voie enzymatique *in vitro*, avec la technique de **polymerase chain reaction** ou **PCR**.

Principe:

Le clonage consiste à:

- 1) insérer un fragment d'ADN étranger à cloner (insert) dans un vecteur de manière à obtenir un vecteur recombinant.
- 2) introduire le vecteur recombinant dans une cellule hôte.
- 3) amplifier le vecteur recombinant par division de la cellule hôte, afin d'obtenir un clone recombinant.

ex: si le vecteur est un plasmide:

- 1) création du plasmide recombinant.
- 2) faire pénétrer le plasmide recombinant dans une cellule d'E. coli (CaCl₂ ou Electroporèse).
- 3) division d'E. coli => clone recombinant.

1-3- Expression d'ADN recombinant

L'expression de la séquence d'ADN insert nécessite la présence de signaux de transcription en amont de cette séquence placés donc en général dans le vecteur. Ces signaux peuvent être spécifiques des machineries transcriptionnelles de bactéries, levures ou encore de cellules de mammifères, selon la cellule hôte choisie pour cette expression.

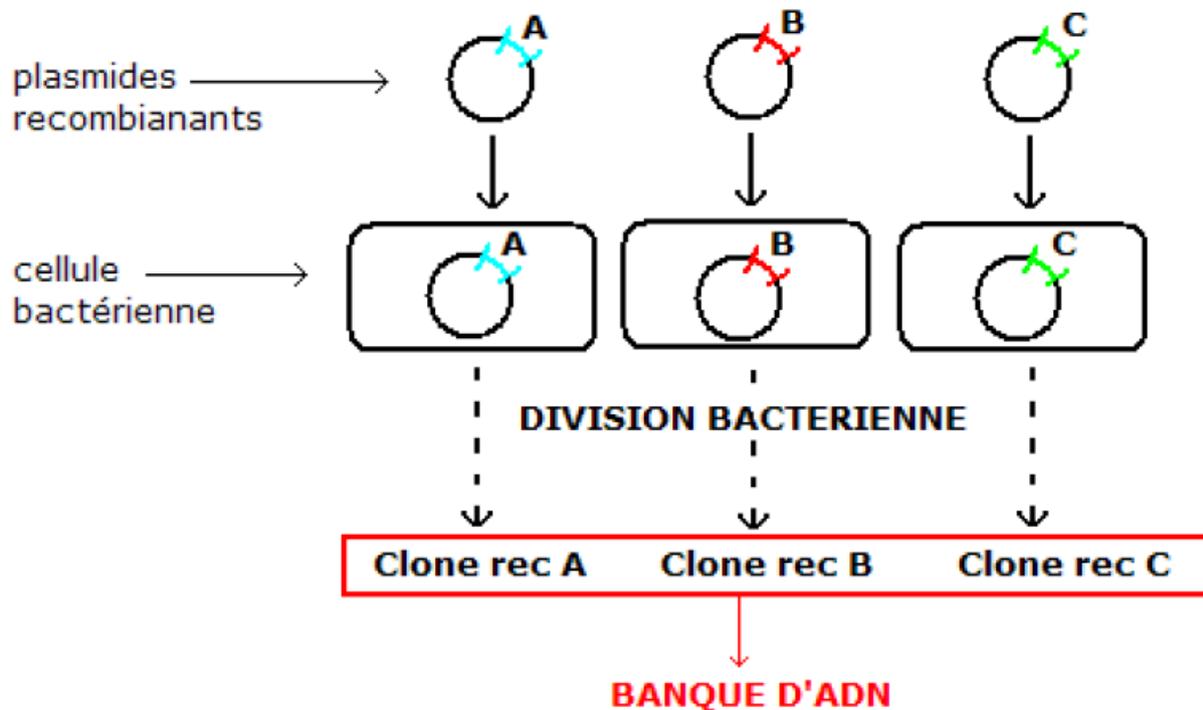
1-4- Notion de banque d'ADN

Une banque d'ADN (ou bibliothèque) est une collection de clones recombinants renfermant chacun un ADN étranger (insert) particulier.

✓ Banques d'ADNc

L'ADNc est une copie d'ADN double brin d'un ARNm, et ne possède donc pas d'introns. Une banque d'ADNc est une collection d'ADNc recombinants. Ainsi si on s'intéresse à une protéine X exprimée dans un tissu donné, on purifiera l'ARNm de ce tissu, et on fabriquera les ADNc correspondant à ces ARNm que l'on clonera dans un vecteur. Cette banque d'ADNc constituée représentera des séquences codant

pour les protéines de ce tissu. Cette information génétique se limitera donc aux gènes exprimés dans le tissu et aux séquences codant les protéines (pas d'introns).



✓ Banques génomiques

Une banque génomique s'oppose à une banque d'ADNc par le fait qu'elle représente la totalité du génome. Une telle banque est en effet obtenue par fragmentation d'ADN génomique.

Elle est constituée de toute une collection d'ADN recombinants correspondant aussi bien à des séquences codantes qu'à des séquences non codantes.

Chaque partie du génome y est théoriquement représentée au moins une fois. Pour connaître la structure d'un gène, il faut utiliser une banque génomique et non une banque d'ADNc. Il en est de même lorsqu'on s'intéresse à des régions d'ADN non transcrites.

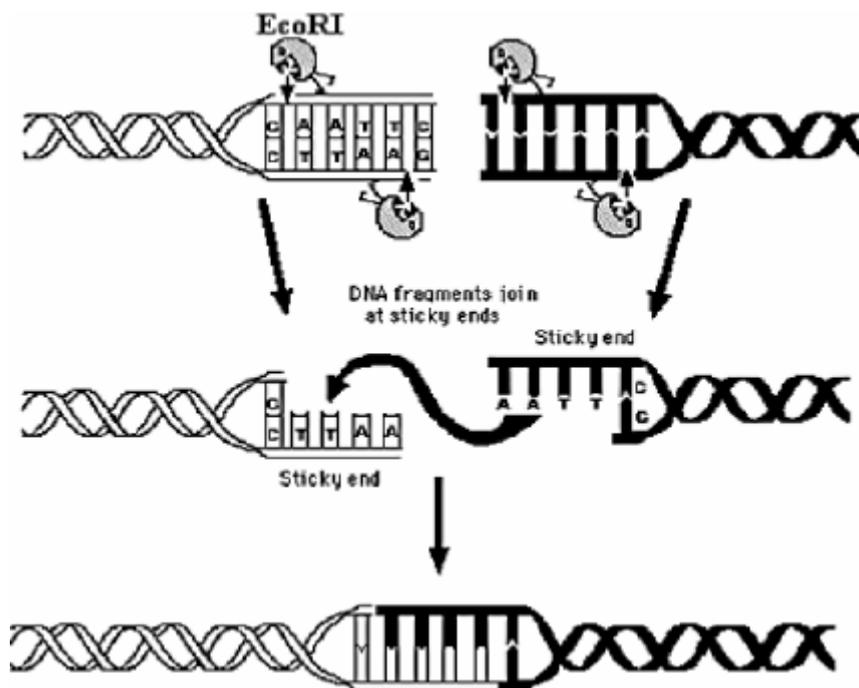
D'autre part, quand on ne sait pas dans quelle cellule est synthétisée la protéine étudiée, un moyen infallible de trouver son gène est d'utiliser une banque génomique. Mais, il est plus difficile de travailler avec une banque génomique dont la taille est bien plus élevée que celle d'une banque d'ADNc. Elle contient en effet environ cent fois plus d'information de séquences d'ADN qu'une banque d'ADNc.

2- Les principaux outils de la biologie moléculaire

2-1- Les enzymes

✚ Les enzymes qui coupent l'ADN (les enzymes de restriction)

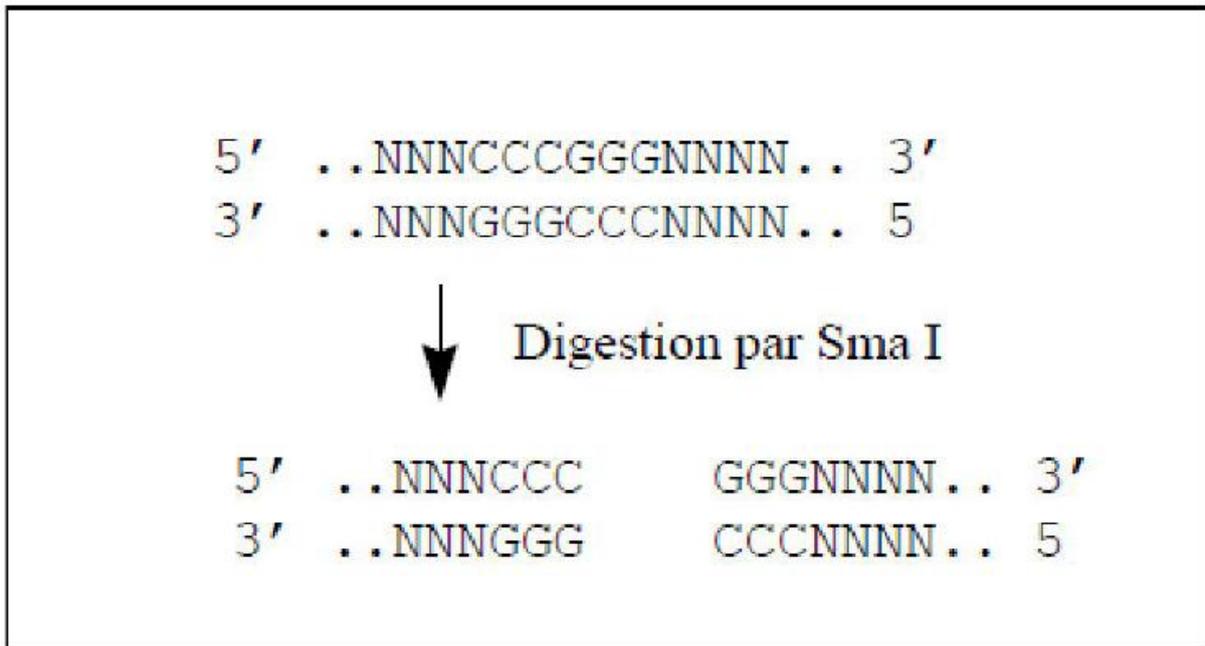
Les enzymes de restriction sont des enzymes qui vont modifier l'ADN au niveau de séquences spécifiques donc ils sont capables de **couper des molécules d'ADN double brin** en des sites **très spécifiques** de chaque enzyme et reconnus par lui. Ces enzymes sont classés en fonction de leur site de reconnaissance et leur lieu de coupure. Les séquences de nucléotides reconnues par les enzymes de restriction sont nommées « **palindrome** ».



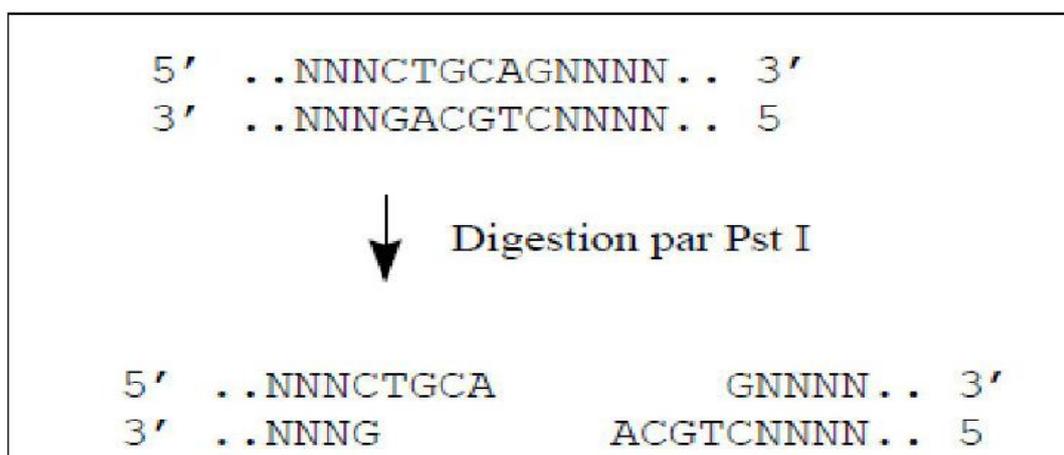
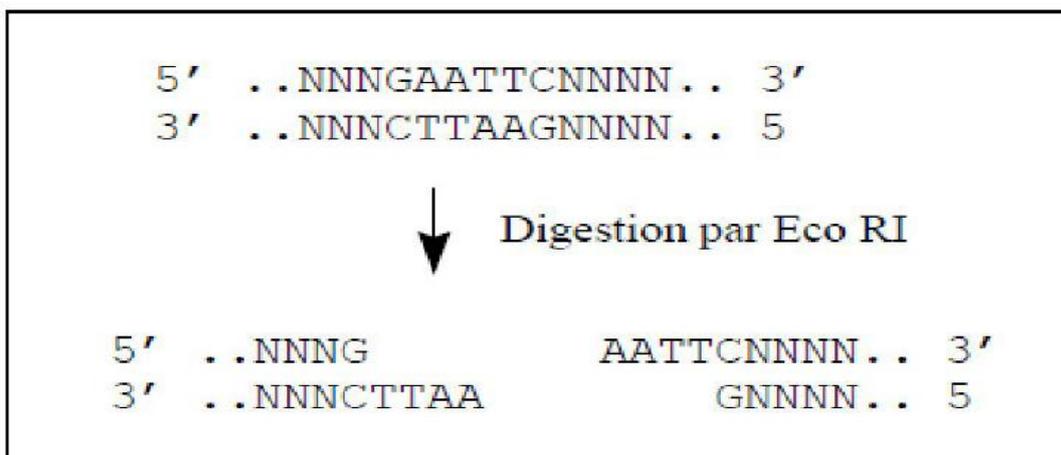
Les enzymes de restriction sont isolés de micro-organismes, bactéries le plus souvent. On donne à ces enzymes des noms à trois ou quatre lettres qui rappellent l'origine du micro-organisme d'où ils ont été isolés. Par exemple : EcoRI est isolé de *E. coli*, HpaI de *Haemophilus parainfluenzae* (R signifie une endonucléase, le chiffre romain est utilisé lorsque plusieurs enzymes ont été isolés à partir d'une même souche).

On observe deux types de coupures par les enzymes de restriction :

-coupure donnant des « extrémités franches » : la coupure d'une molécule d'ADN peut se faire au milieu du palindrome ;



-coupure donnant des « extrémités adhésives » (ou « bouts collants ») ; d'autres types d'enzymes agissent en coupant de part et d'autre du centre de symétrie.



Un enzyme de restriction est donc capable de repérer sur toute la longueur d'un ADN certaines séquences caractéristiques. Si cette séquence se trouve cinq fois dans une molécule d'ADN, l'enzyme coupera cet ADN en ces cinq endroits, ce qui donnera cinq fragments d'ADN. Donc selon l'enzyme utilisé, un même ADN sera coupé différemment.

❖ La DNase

La DNase utilisée au laboratoire est généralement d'origine bovine (pancréas). C'est une endonucléase qui possède la propriété de **couper une molécule d'ADN double brin au hasard**, engendrant ainsi des fragments d'ADN double brin.

❖ La nucléase S1

La nucléase S1 (isolée d'un champignon, *Aspergillus oryzae*) n'attaque que l'ADN simple brin.

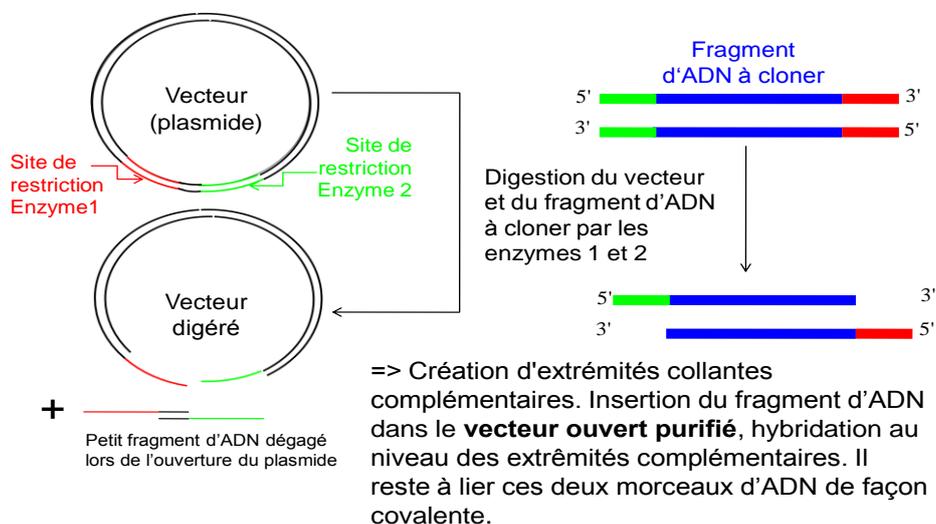
❖ Les exo-nucléases

Les exo-nucléases coupent les extrémités libres des molécules d'ADN en libérant des nucléotides. Par exemple, l'exonucléase III hydrolyse séquentiellement les extrémités 3' libres des molécules d'ADN dans le sens 3' vers 5'.

✚ Les enzymes qui ligaturent : les ligases

Pour **lier, souder** deux fragments d'ADN, on utilise une ligase, en présence d'ATP. Il est plus facile de liguer deux fragments à extrémités cohésives que deux fragments à coupure franche. La ligase choisie est en général la ligase du virus T4 (extraite de bactéries infectées par le virus T4). Lorsqu'il existe une coupure d'un des deux brins d'une molécule d'ADN double brin, la ligase est également capable d'effectuer une ligation entre les deux nucléotides adjacents séparés par une brèche.

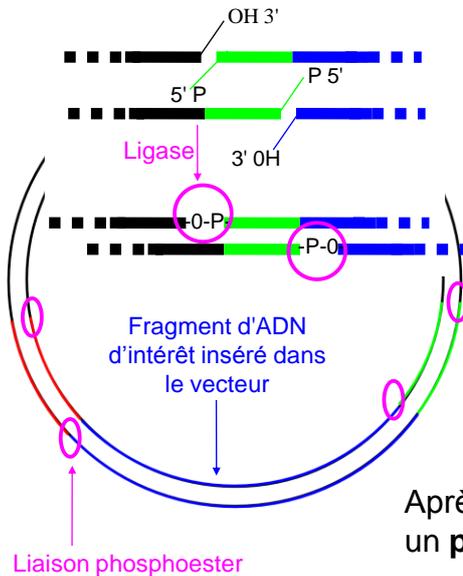
b) Ligation du fragment d'ADN dans le vecteur



37

4) Insertion de l'ADN dans le vecteur

b) Ligation du fragment d'ADN dans le vecteur



- La **ligase** est une enzyme capable de former des **liaisons covalentes** (formation de liaisons *phosphoester*) entre deux molécules d'ADN (au niveau d'une zone d'hybridation des molécules d'ADN)
- Intérêt d'utiliser deux enzymes différentes lors de la digestion :
 - insertion orientée de l'ADN à cloner dans le vecteur
 - limitation de la religation du vecteur sur lui-même

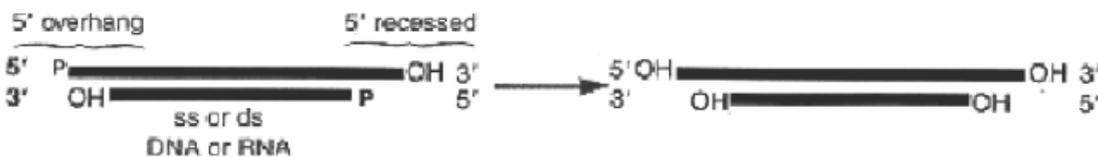
Après ligation, on obtient un **plasmide recombiné**

38

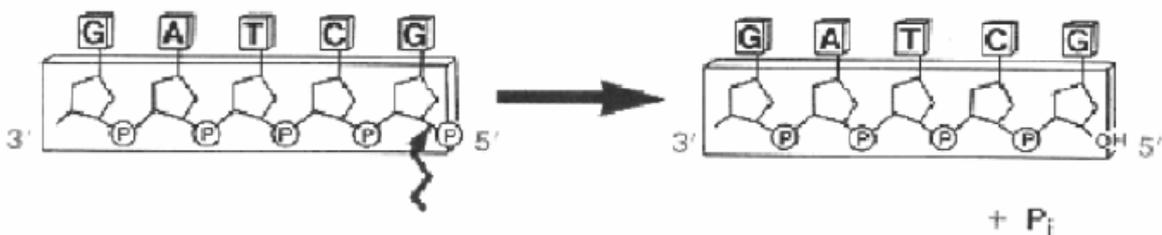
⚡ Les enzymes qui déphosphorylent : **les phosphatases**

Les phosphatases catalysent l'élimination d'un groupement phosphate en 5' d'une chaîne d'ADN. Il est courant de déphosphoryler un vecteur que l'on vient d'ouvrir par un enzyme de restriction, afin d'éviter une refermeture de ce vecteur (auto-ligation) qui empêcherait alors d'insérer le fragment d'ADN à étudier.

Eviter l'auto-ligation: phosphatase



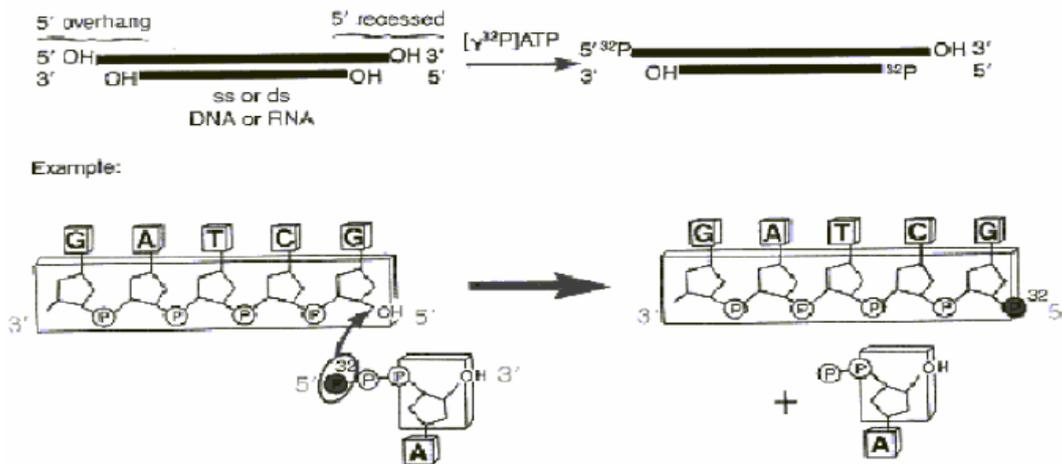
Example:



✚ Les enzymes qui phosphorylent : **les kinases**

Les **kinases** permettent le transfert d'un groupement phosphate à partir d'une molécule d'ATP. Il s'agit du phosphate en position gamma (le plus externe des trois groupements fixés sur l'Adénosine). Il sera transféré à l'extrémité 5' d'un ADN déphosphorylé.

Marquer: kinase



✚ Les enzymes qui « recopient » un acide nucléique

La copie d'une chaîne d'acides nucléiques, qu'il s'agisse de synthétiser une chaîne d'ADN ou d'ARN, s'effectue de manière complémentaire et antiparallèle (l'addition des nouveaux nucléotides se faisant dans le sens 5'phosphate vers 3' OH).

• ADN-ADN

Les **ADN polymérases** synthétisent une chaîne d'ADN à partir d'un modèle. En effet, une ADN polymérase n'est pas capable de commencer une chaîne d'acides nucléiques. Il est nécessaire d'apporter dans le milieu réactionnel, une amorce d'acide nucléique.

• ADN polymérase I et enzyme de Klenow

L'ADN polymérase habituellement utilisée est l'ADN polymérase d'*E.coli*. Elle est constituée d'une seule chaîne peptidique et possède trois activités enzymatiques :

- une activité de **synthèse ADN polymérase 5'-3'**;
- une activité de **dégradation exonucléase 3'-5'** expliquant ses propriétés d'enzyme « à **fonction d'édition** » ;
- une activité de **dégradation exonucléase 5'-3'**.

L'**enzyme de Klenow** est un enzyme très souvent utilisé au laboratoire, il est obtenu à partir de l'ADN polymérase I d'où l'on a éliminé, clivage avec une protéase, le petit fragment

responsable de l'activité exonucléase 5'-3'. Il ne reste donc plus que les activités polymérase 5'-3' et exonucléase 3'-5' (**fonction d'édition**).

La suppression de l'activité exonucléase 3'-5' permet d'augmenter la processivité de la polymérase, ce qui permet de synthétiser des fragments de grande longueur.

- **Les ADN polymérases thermorésistantes**

La **Taq polymérase** a été la première à être identifiée. C'est une ADN polymérase isolée des bactéries vivant dans les sources d'eau chaude. Elle agit à une température voisine de 65°C. Cette propriété trouvera son utilité dans les réactions où il faut utiliser une ADN polymérase thermostable, comme dans la **polymerase chain restriction PCR (réaction d'amplification d'ADN)**.

Elle peut aussi être utilisée dans **les techniques de séquençages de l'ADN** car elle permet de travailler à température élevée.

- **ARN-ADN**

- **La rétrotranscriptase « RT » (ou transcriptase reverse)**

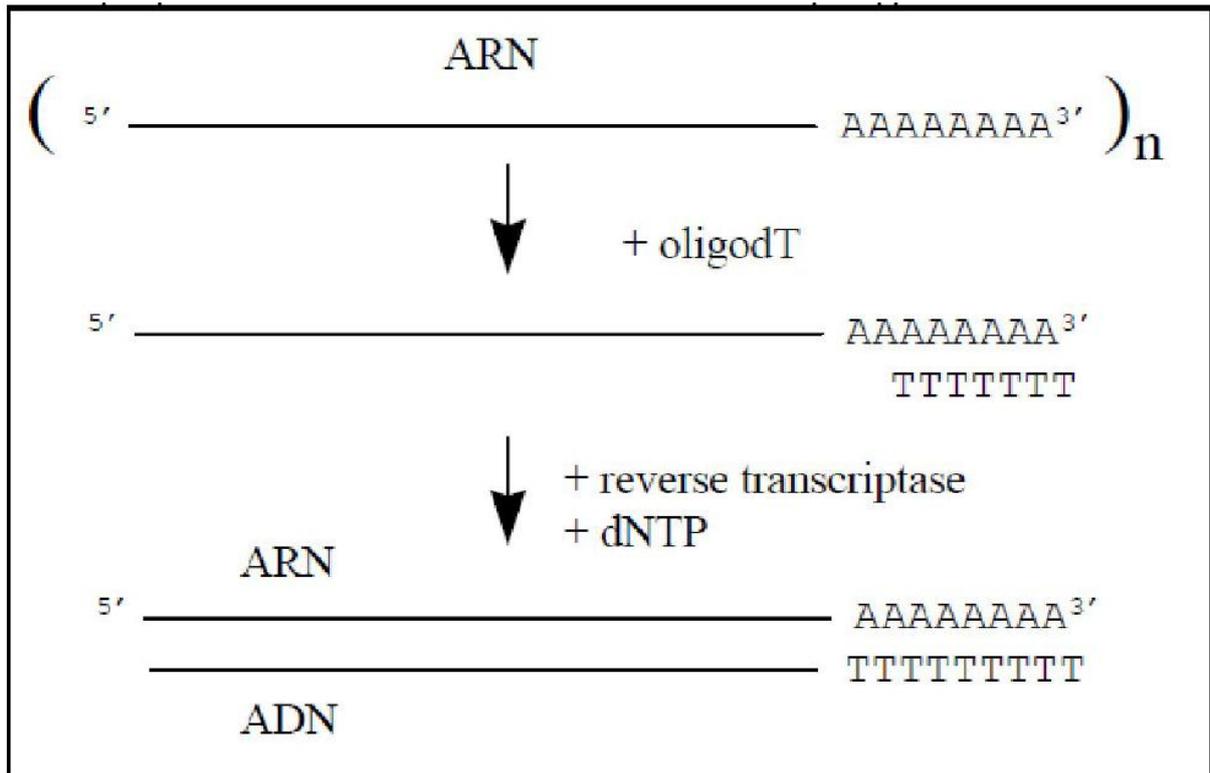
La RT est un enzyme isolé de **rétrovirus** comme HIV (virus de SIDA) et coronavirus (mais existant également ailleurs que dans les rétrovirus), qui permet de fabriquer un **ADNc à partir d'un ARNm**.

La RT est une ADN polymérase 5'-3' ayant les caractéristiques suivantes :

- elle est ARN dépendante ;
- elle est dépourvue d'activité exonucléase 3'-5' et est donc dénuée de fonction d'édition ; elle est ainsi encline aux erreurs ;
- elle possède une activité RNase.

Elle permet de synthétiser le premier brin de l'ADNc (ADN double brin) à partir d'un ARNm. Une amorce **oligo (dT)** qui s'hybrideront au poly (A) de l'ARNm et permettra d'initier l'action de la RT, qui synthétisera une copie complémentaire de l'ARNm jusqu'à son extrémité.

Une réaction très utile se produit à la fin de la copie de l'ARNm. La RT provoque le retournement de l'extrémité 3' du transcrit sur lui-même. Elle « tourne le coin » et est prête à recopier son propre travail, ce qui crée une épingle à cheveu de 10 à 20 bp. Seules quelques bases à la toute extrémité de la boucle ne sont pas appariées.



- **ADN-ARN**

- **L'ARN polymérase**

L'ARN polymérase est utilisé pour effectuer des transcriptions *in vitro*.

2-2- Les vecteurs de clonage

On appelle « vecteur » l'ADN dans lequel on insère l'ADN à étudier et qui sert en fait de support.

Les vecteurs sont des petites molécules d'ADN dans lesquels on insère le fragment d'ADN à étudier. Ces petits ADN sont généralement des virus bactériens (bactériophages) ou des plasmides. Ils possèdent dans leur génome les signaux nécessaires pour leur répllication, mais ils ne savent pas se multiplier seuls. Ils doivent être introduits dans des cellules hôtes (bactéries par exemple).

2-2-1- Plasmides

Les plasmides sont des petits éléments génétiques extra-chromosomiques doués de la répllication autonome, ce sont typiquement des molécules d'ADN double brin, circulaires, leur taille varie de 1 kb à deux ou trois centaines de kb (< 5% de la taille du chromosome bactérien). Ils contiennent des gènes codants souvent pour des protéines qui donnent un ou des avantage(s) à la cellule hôte. Comme par exemple :

- Résistance aux antibiotiques

- Résistance aux métaux lourds
- Dégradation de composés aromatiques

Les plasmides sont de bons vecteurs de clonage chez les bactéries parce qu'ils se multiplient en nombre de copies important et qu'ils sont aisément purifiables. Les marqueurs de sélection (Ex. gènes de résistance aux antibiotiques) permettent l'identification des bactéries recombinantes (transformées) qu'ils les portent.

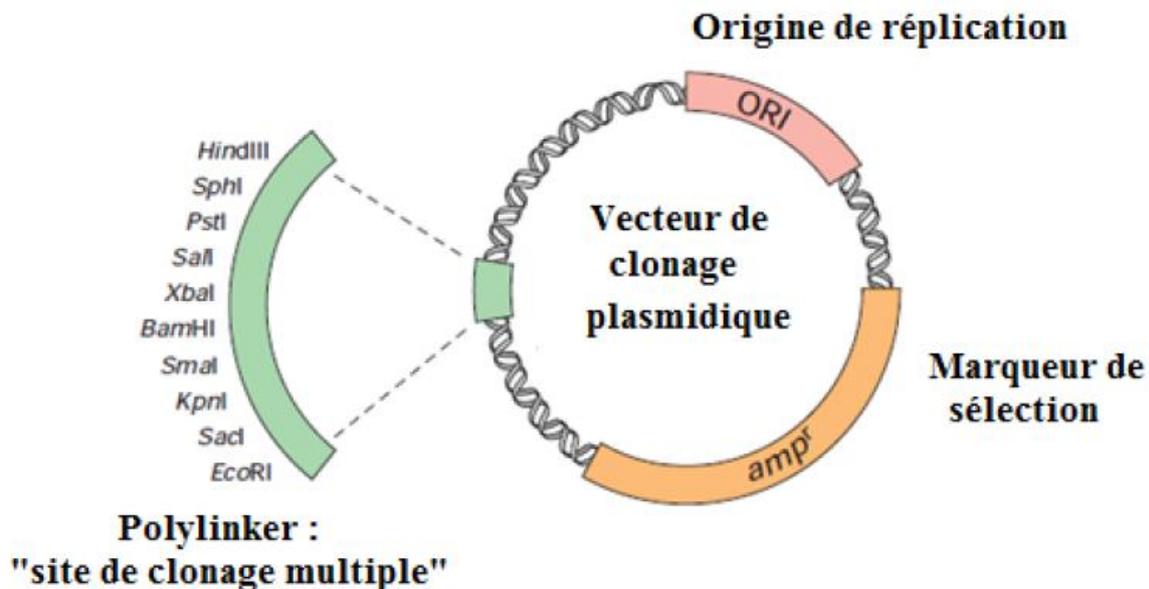


Figure 1. Composants basiques d'un vecteur plasmidique qui peut se répliquer chez *E. coli* (Lodish *et al.*, 2003).

➤ Plasmide pBR322

Le pBR322 appartient à une série de vecteurs de clonage de première génération, partiellement construit par génie génétique.

✓ Caractéristiques du plasmide pBR322

- pBR322 est un petit plasmide constitué de 4361 paires de bases, dont la séquence nucléotidique est complètement connue.
- Il est possible d'y insérer un fragment d'ADN de bonne dimension sans toutefois dépasser la taille de 10kpb sous peine de l'instabilité plasmidique.
- Il possède deux gènes de résistances aux antibiotiques : l'un pour l'ampicilline (Amp^R), l'autre pour la tétracycline (Tc^R), l'expression de l'un de ces deux gènes facilite la sélection des clones recombinants.
- Il possède également vingt sites uniques pour des enzymes de restriction.
- Il est facilement transférable par transformation ou par électroporation.

➤ Plasmide PUC 19 (plasmid of University of California)

De nouvelles générations de plasmides plus puissants ont été développées depuis pBR322 et ces dérivés. C'est le cas de la famille pUC appelés (les vecteurs de clonage de seconde génération).

Les vecteurs de clonage de seconde génération ce sont des petits plasmides d'environ 2700pb. Le plasmide pUC19 contient le gène de résistance à l'ampicilline de pBR322, mais en plus il possède une partie du gène *lacZ* dans lequel a été introduit un site multiple de clonage (Polylinker), contenant toute une série de sites de coupure unique.

Le fait que ce polylinker soit inséré dans le gène *lacZ* qui intervient dans le catabolisme du lactose permet de révéler facilement l'intégration d'un insert par « inactivation insertionnelle ». L'utilisation d'un inducteur coloré comme le X-gal (5-bromo-4chloro-3-indonyl- β -D-galactoside), l'hydrolyse de ce dernier en dibromo-5,5-dichloro-4, 4-indigo (de couleur bleu), indique la production de β -galactosidase.

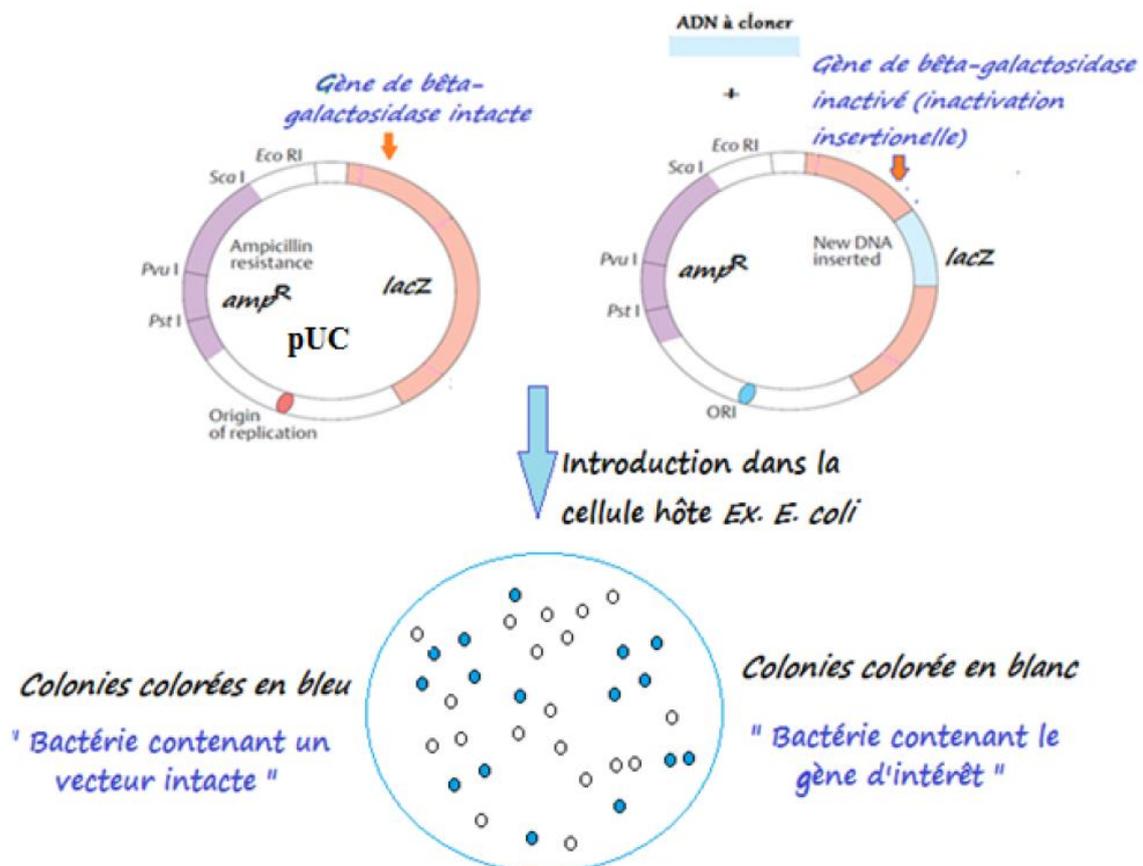


Figure 2. Insertion d'un fragment d'ADN (< 10 kb) dans un vecteur (type pUC). L'inactivation insertionnelle du gène *lacZ* permet de révéler facilement les colonies contenant le gène d'intérêt (sélection positive: colonies blanche). Le milieu de sélection est supplémenté par l'ampicilline.

➤ Plasmide Ti (Tumor inducing Plasmide)

La bactérie *Agrobacterium tumefaciens* renferme de grand plasmides (140 à 235 kb), mais les cellules de plante transformées intègrent seulement un petit fragment spécifique du plasmide d'une taille d'environ 23kb appelé l'ADN-t, ce fragment spécifie le type d'opine synthétisé dans le tissu végétal. Des souches d'*A.tumefaciens* porteuse du plasmide Ti sont maintenant disponibles pour produire des plantes transgénique. Exemple: l'obtention des plantes résistantes aux insectes (BT), aux herbicides, etc.

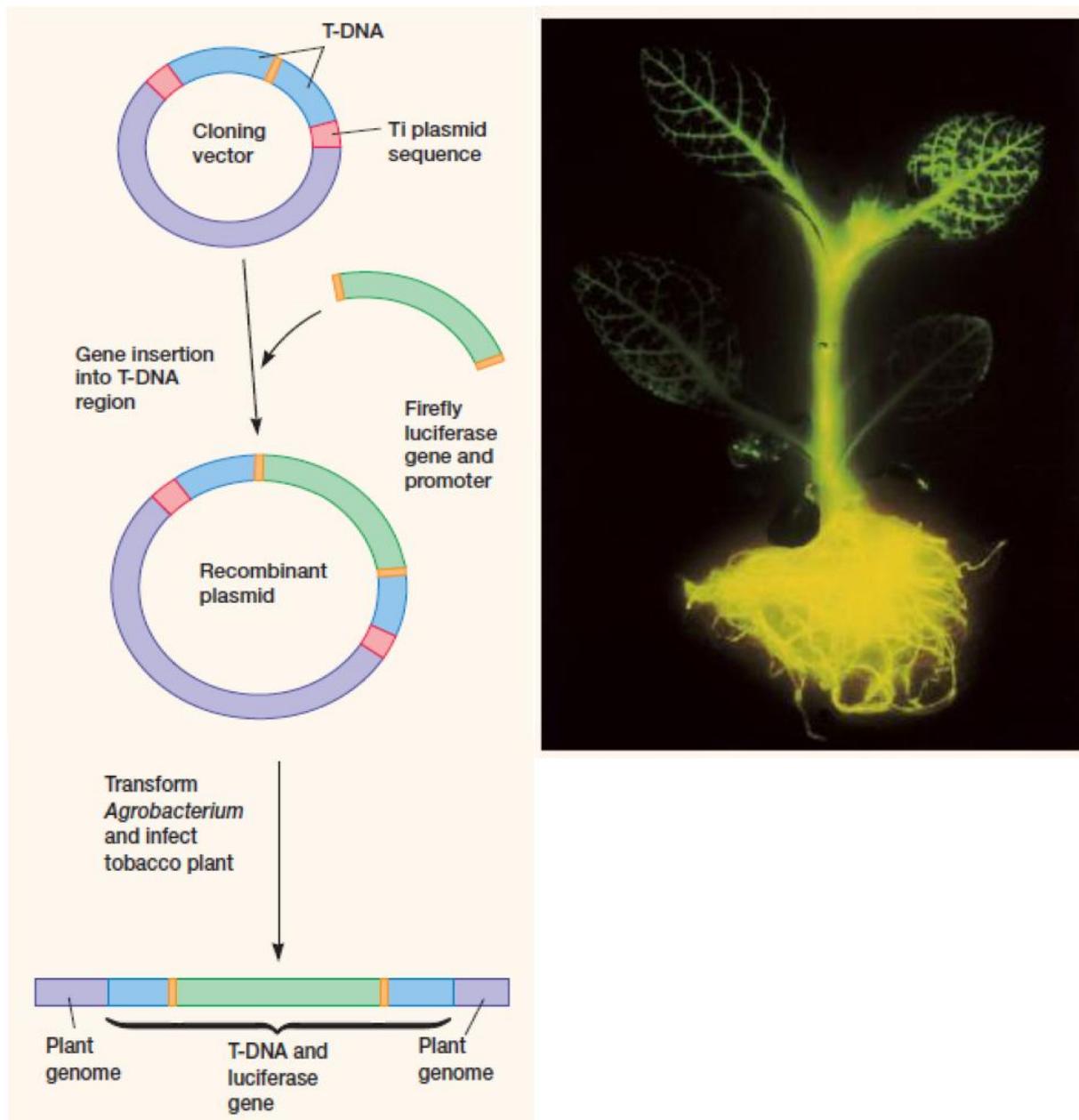


Figure 3. L'utilisation du vecteur Ti pour transformer la plante *Nicotiana tabacum* par le gène d'une luciférase. Le gène est introduit dans la région de l'ADN-T. Cette transformation rend les bactéries bioluminescentes (Prescott, 2002).

2-2-2- Les bactériophages

Deux phages ont été très utilisés comme vecteurs dans les premiers temps de la biologie moléculaire. Ce sont les phages λ (lambda) et le phage M13, ou plus exactement des dérivés de ces deux phages qui doivent en effet subir divers types de modifications pour pouvoir être utilisés comme vecteurs de clonage.

➤ Le phage λ (lambda)

Plusieurs vecteurs sont dérivés du phage λ et peuvent être utilisés dans des constructions génétiques, essentiellement pour la constitution de banques d'ADNc ou génomiques.

On distingue deux parties dans le phage λ :

- la tête du phage qui renferme l'ADN ;
- la queue du phage qui permettra au virus de se fixer sur la cellule hôte bactérienne

Le phage λ est un phage à ADN double brin, linéaire, d'une longueur de 48,5 Kb. A chaque extrémité 5' se trouve une région monocaténaire de 12 nucléotides, l'une complémentaire de l'autre et leur association donne une structure circulaire à l'ADN dans la cellule hôte. L'association de ces extrémités cohésive naturelles forme le site *cos* (figure 4) [*COS*: Des éléments important pour la réplication et l'encapsidation de bactériophage λ].

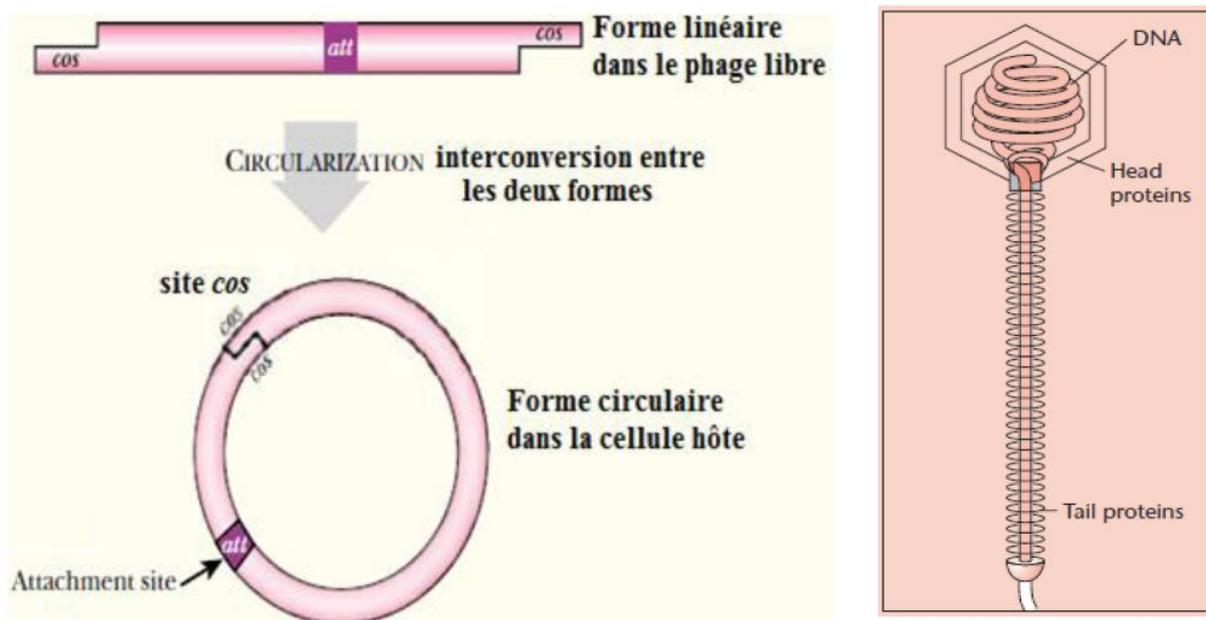


Figure 4. Différentes formes du génome du bactériophage λ , et sa structure (Primose *et al.*, 2002).

Le bactériophage λ a donné naissance aux premiers vecteurs de types phagiques car:

- Ça biologie est bien connue
- La quantité d'ADN qu'il peut intégrer est plus importante que celle véhiculée par les vecteurs plasmidiques. Il est possible d'insérer jusqu'au 22kb, après élimination de la partie indispensable au cycle de vie du phage (figure 5).

Polycopie de cours de Biologie moléculaire

- La possibilité d'empaqueter *in vitro* l'ADN phagique nu recombinant dans les têtes de phage.
- Il a une capacité d'infection (transfection) de l'hôte très rapide.
- Le nombre de copies par cellule étant considérable.
- Le rendement de cette transfection est très supérieur à ce qui est obtenu lors de la transformation de la bactérie par les plasmides.

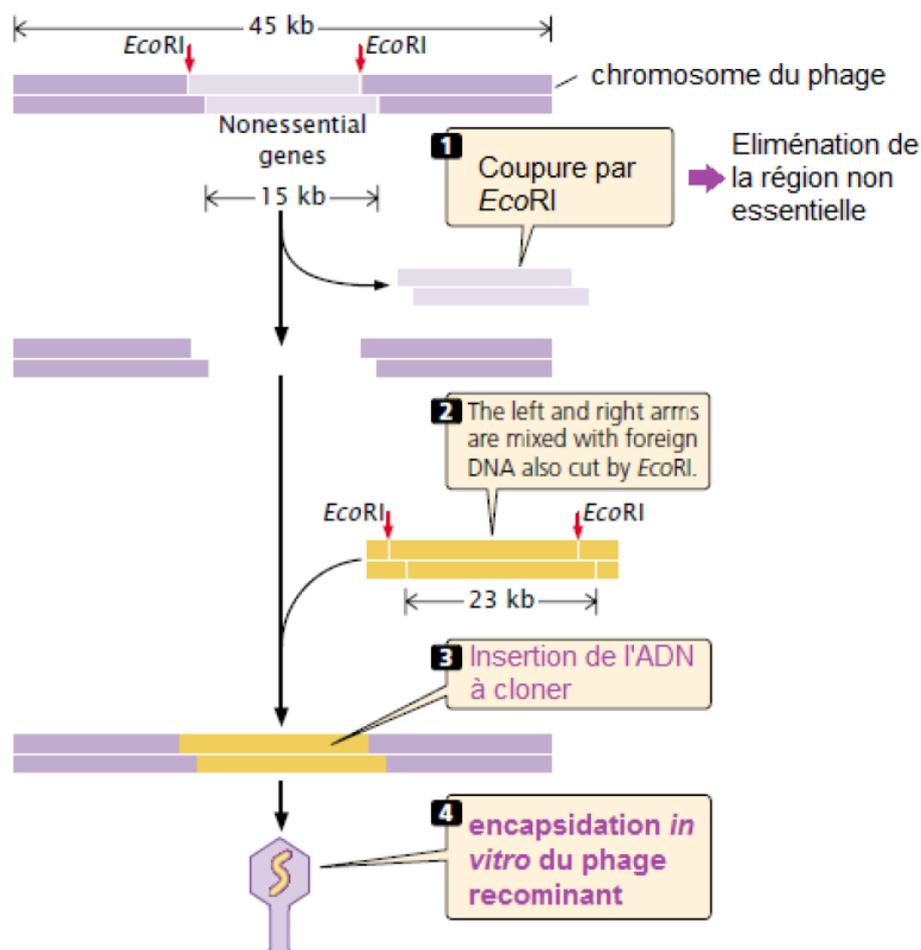
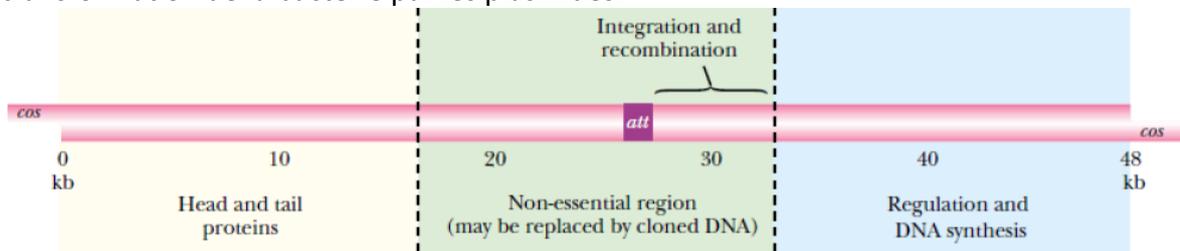


Figure 5: L'élimination de la partie indispensable (non essentielle) au cycle de vie du bactériophage λ augmente la capacité d'insertion de grand fragment (22kb) (Passarge, 2007).

2-2-3- Les cosmides

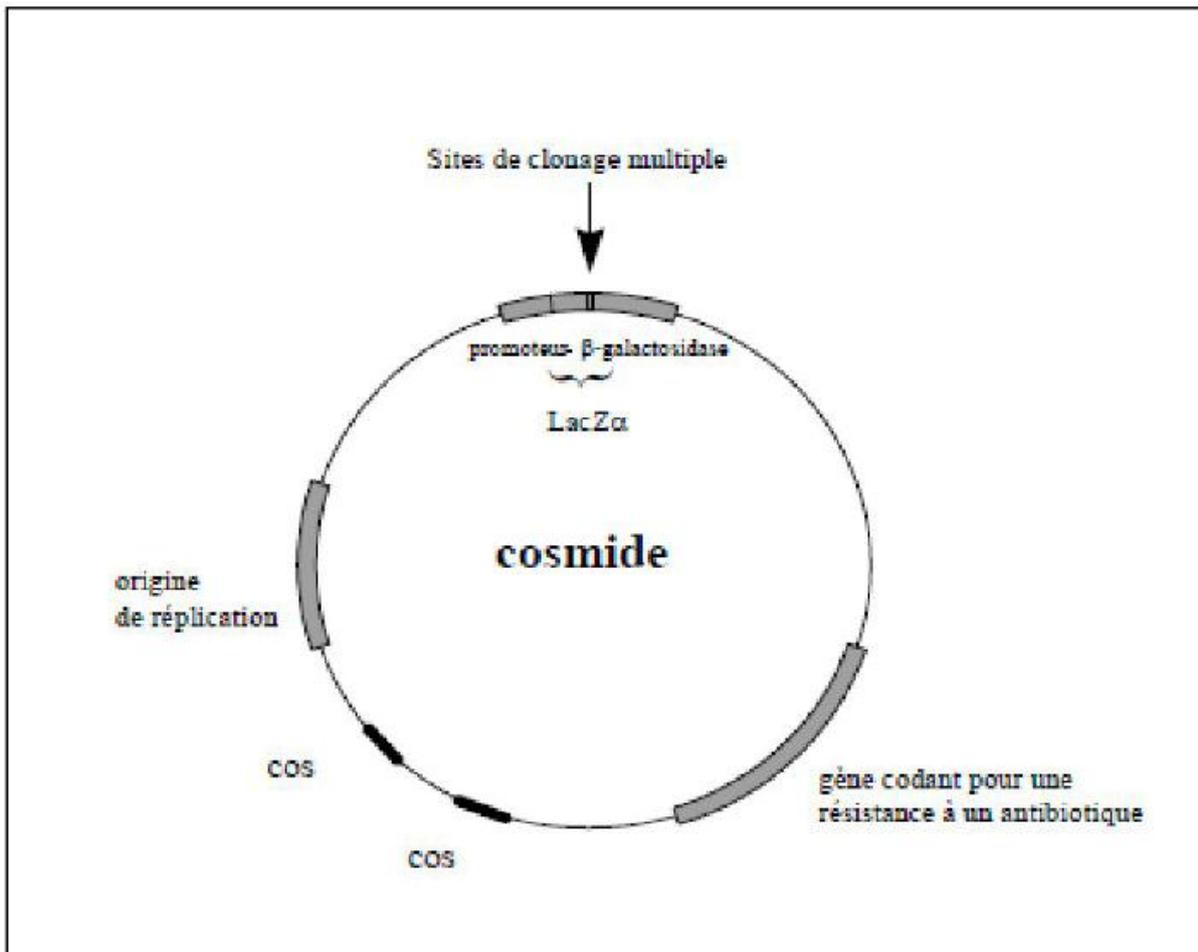
Les cosmides sont des vecteurs artificiels hybrides : plasmide-phage λ . L'avantage des cosmides est qu'il est possible de cloner de plus grands fragments d'ADN étranger (35 à 45 Kb) qu'avec les vecteurs dérivés du phage λ (23 Kb maximum). Ils servent à fabriquer des

Polycopie de cours de Biologie moléculaire

banques génomiques et sont utilisés plus spécialement lorsque le gène étudié s'étend par exemple sur une longueur de 30 à 40 Kb.

Ces vecteurs rassemblent à la fois les propriétés intéressantes des plasmides comme :

- L'origine de réplication
 - Gène de résistance à un antibiotique
- Et celles du bactériophage
- Encapsidation *in vitro* de grand fragment d'ADN.



2-2-4- Phagemides ou phasmides

Sont des vecteurs qui combinent des éléments d'origine plasmidiques et phagiques. Le phagmide le plus utilisé est **pBluescriptII KS**, c'est un dérivé du plasmide pUC19, il contient un polylinker, interrompu par deux promoteurs (T3 et T7) se lisent en sens opposés, il contient aussi un promoteur *lac* inductible avec une partie du gène *lacZ* (blanc-bleu sélection). Une origine de réplication dérivée de M13 (phage filamenteux). Un *ori ColEI* pour permettre la réplication du phage comme un plasmide. Ils sont utilisés pour cloner de grand fragment d'ADN et la manipulation des gènes.

2-2-5- Autres types de vecteurs

➤ Les banques YAC et BAC (*yeast or bacterial artificial chromosome*)

Il est maintenant possible de cloner de grands fragments d'ADN, longs de 200 à 500 Kb, voire plus, sous forme de chromosomes artificiels, dans la levure ou dans la bactérie.

Le principe de cette technique consiste à doter les grands fragments d'ADN à cloner des éléments nécessaires pour que la levure ou la bactérie les réplique comme s'il s'agissait de ses propres chromosomes.

Les YACs (Yeast Artificial Chromosomes) doivent avoir :

- Une origine de réplication
- Des télomères pour la réplication de l'ADN aux extrémités du chromosome
- Un centromère (ségrégation lors de la mitose).
- Site de clonage multiple (MCS : multiple cloning site or polylinker)
- Marqueur de sélection.

2-3- Les cellules-hôtes

La cellule-hôte destinée à recevoir le vecteur, phage ou plasmide (ou cosmide), doit répondre à certaines conditions :

➤ Choix de la souche

Une des premières conditions est que la souche d'E coli choisie ne soit pas pathogène pour l'homme. C'est le cas par exemple des souches E coli. K12. Ces souches ont la particularité de ne pas coloniser l'appareil digestif humain. Mais elles synthétisent l'enzyme de restriction EcoK (enzyme de type I qui reconnaît le site : AAC et GTGC). Il faut utiliser des souches dérivées ne synthétisant pas l'enzyme de restriction EcoK (ou qui modifient les sites EcoK) sinon tout ADN à étudier possédant un site EcoK serait clivé.

➤ Croissance en milieu liquide

Après une phase initiale de latence, la croissance d'E coli est exponentielle, le nombre de bactérie double toutes les 20 minutes. Lorsque le substrat et l'oxygène ne sont plus en conditions optimales, la croissance ralentit. Puis la culture entre en phase stationnaire, les bactéries meurent aussi vite qu'elles se divisent. Finalement c'est la mort cellulaire.

❖ Distinction entre infection, transformation et transfection

- Infection

Dans l'infection, la bactérie est mise en contact avec le virus.

L'infection par le phage λ en phase lytique se traduit par l'apparition de plages de lyse. L'infection par le phage M13 (ou par le virus λ en phase lysogénique) aboutira à des colonies bactériennes contenant ces virus.

- Transformation

La transformation consiste à mettre en contact un ADN double brin (plasmide) et les bactéries prétraitées (on dit compétentes). Dans ces conditions, l'ADN pénétrera dans les bactéries et pourra s'y répliquer.

- Transfection

La transfection, enfin, concerne essentiellement les cellules de mammifères et se rapproche de la transformation des bactéries. Elle consiste à faire entrer dans les cellules un ADN double brin, soit en fragilisant la membrane plasmique (électroporation, agents chimiques), soit en incluant l'ADN dans des gouttelettes lipidiques (liposomes) qui seront internalisées par les cellules (lipofection), soit encore en formant avec l'ADN un précipité de phosphate de calcium qui entrera dans les cellules.

2-4- Les sondes nucléiques

▪ Caractéristiques des sondes

Les sondes nucléiques sont des segments d'acides nucléiques que l'on utilise pour repérer de manière spécifique, dans une réaction dite « **d'hybridation moléculaire** », le fragment d'acide nucléique auquel on s'intéresse.

Une sonde possède les caractéristiques suivantes :

- c'est un segment d'acide nucléique (ADN ou ARN) n'ayant obligatoirement qu'un brin ; elle peut être plus ou moins longue selon les cas et peut ne posséder qu'une vingtaine de nucléotides (sonde oligo-nucléotidique) ou plusieurs centaines de nucléotides ;
- elle est complémentaire et antiparallèle du segment d'acide nucléique à reconnaître, cette reconnaissance pouvant se faire entre ADN-ADN (ou ADN-ARN ou ARN-ARN) ; bien évidemment, dans cette réaction d'hybridation moléculaire, il faut que les acides nucléiques à reconnaître soient également sous forme monobrin ;
- la sonde peut couvrir tout ou une partie du segment d'acide nucléique à reconnaître ; elle est capable de détecter sa copie complémentaire parmi des milliers de fragments d'ADN ou d'ARN différents.
- une sonde doit être repérable : des sondes radioactives sont le plus souvent utilisées ; il est ainsi facile de repérer l'emplacement de cette sonde et par conséquent le clone à identifier qui se sera hybridé avec cette sonde.

▪ **Hybridation moléculaire : Stringence**

L'utilisation de sondes d'acides nucléiques repose sur la complémentarité des deux brins d'ADN, qui peuvent être dans certaines circonstances dissociés, puis réassociés avec la sonde repérable.

Si l'on chauffe de l'ADN, il se produit, à une certaine température « température de fusion », une rupture des liaisons hydrogène entre les bases. La double hélice se défait, les deux brins se séparent : on dit que l'ADN est dénaturé. Plus il y a de bases C : G dans un ADN, plus la température de fusion sera élevée. Ceci se comprend aisément, étant donné que les bases C : G sont reliées par trois liaisons hydrogène, alors que les bases A : T ne sont reliées que par deux liaisons hydrogène.

La dénaturation de l'ADN est réversible : quand la température est abaissée progressivement, les deux brins d'ADN peuvent s'hybrider à nouveau c'est-à-dire se rapprocher selon les règles de complémentarité de bases (renaturation). D'où la nécessité de refroidir brusquement pour maintenir une dénaturation souhaitée. Cette hybridation est mise à profit pour identifier la séquence d'intérêt : l'introduction de molécules d'ADN marqué (sonde) dans la mixture de séquences d'ADN dénaturées, permettra l'hybridation de la sonde avec la séquence recherchée.

▪ **Obtention d'une sonde**

Toute la difficulté consiste bien évidemment à obtenir une sonde spécifique de l'ADN à étudier. Plusieurs possibilités existent. Peuvent être utilisées comme sonde les molécules ci-dessous :

✓ **ADNc**

L'ADNc constitue une excellente sonde. En pratique on utilise le plus souvent un fragment de restriction cloné couvrant la plus grande partie possible de l'ADNc. Encore faut-il avoir pu construire l'ADNc correspondant, ce qui n'est pas toujours le cas, surtout lorsque l'on débute l'étude d'un gène.

✓ **ARNm**

L'ARNm peut également servir de sonde, mais cette fois encore, on ne dispose pas toujours de l'ARNm spécifique.

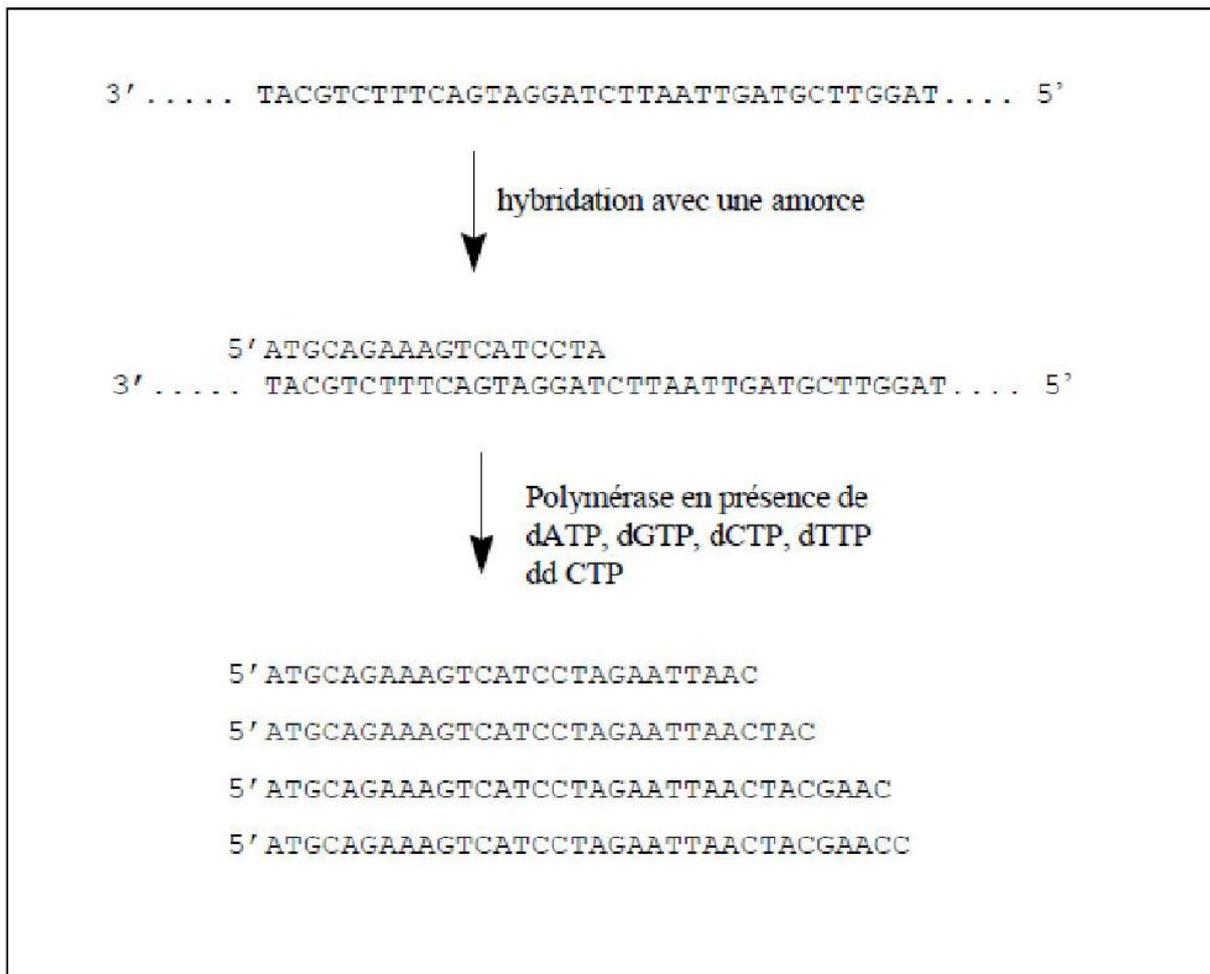
✓ **Oligo-nucléotides de synthèse**

- Si la séquence de l'ADN n'est pas connue, il faut purifier une petite quantité de la protéine correspondant au gène étudié, déterminer la séquence des acides aminés d'un court segment de cette protéine, on déduit d'après le code génétique la séquence de segment d'ADN

Polycopie de cours de Biologie moléculaire

correspondant et synthétiser alors de court sondes d'ADN. On peut imaginer les difficultés rencontrées lorsque le segment protéique séquencé contient des acides aminés tels que Leu, Ser ou Arg codés par six codons différents ! Il faut alors synthétiser tout un ensemble de ces courtes sondes d'ADN afin de couvrir toutes les possibilités.

- Si la séquence d'une partie d'ADN à repérer est connue, il devient alors beaucoup plus facile de synthétiser une sonde d'une vingtaine de nucléotides. Il existe des appareils automatiques (« synthétiseurs d'oligonucléotides ») qui permettent d'assembler jusqu'à une centaine de nucléotides dans l'ordre souhaité.



3- Techniques de Biologie moléculaire

3-1- Criblage de banque

Banques d'ADN génomique procaryote:

Prenons l'exemple du clonage d'un gène bactérien G. Comme l'ADN chromosomique de la bactérie est un petit ADN, nous pouvons utiliser les plasmides comme vecteurs. On suit les étapes suivantes:

Il faut sortir ce gène et l'amplifier dans des cultures bactériennes.

1) Digestion de l'ADN chromosomique et de l'ADN plasmidique par la même enzyme de restriction:

- L'ADN chromosomique possède plusieurs sites de restriction spécifiques à l'enzyme utilisée: on obtient alors plusieurs fragments d'ADN dont un seul comprend le gène G.
- L'ADN plasmidique ne contient qu'un seul site de restriction spécifique, au sein d'un gène de sélection. L'ADN est ouvert à ce niveau.

2) Ligature: On réunit l'ADN chromosomique fragmenté et l'ADN plasmidique dans le même tube et on ajoute de la T4 ADN ligase (ligatures d'extrémités adhésives ou non). Après cette étape, on peut obtenir 3 types de vecteurs:

- vecteurs recombinants:
 - (1) Plasmide recombinant avec de l'ADN chromosomique ne comprenant pas le gène
 - (2) Plasmide recombinant contenant le gène G
- vecteurs non-recombinants:
 - (3) Plasmide s'étant refermé sans intégrer de fragment d'ADN étranger.

3) Transformation: Il s'agit de vecteurs plasmidiques, donc il faut les faire pénétrer dans les cellules bactériennes (E. coli) grâce à du CaCl₂ ou à l'Electroporèse. Il faut également faire en sorte qu'un seul ADN plasmidique au mieux par cellule.

4) Sélection I (clones transformés): On étale les clones sur un milieu ajouté de l'antibiotique pour lequel le vecteur plasmidique est résistant et dans lequel il n'y avait pas le site de restriction.

- seuls les clones transformés (par (1) (2) ou (3)) survivront car ils auront acquis la résistance à cet antibiotique par le vecteur plasmidique. (attention: il faut que la souche bactérienne soit sensible à cet antibiotique)

5) Sélection II (clones recombinant):

- Si le plasmide utilisé est de type pBR322: On utilise le gène de sélection dans lequel se situait le site de restriction de l'enzyme utilisée. On utilise les clones transformés par (1) (2) ou (3) avec un velours stérile
 - ceux qui ne survivront pas sur le milieu comprenant l'antibiotique pour lequel le plasmide utilisé est résistant seront des clones recombinants. En effet, si le gène de résistance comprend un insert, il est altéré et donc il ne s'exprimera pas et les clones seront sensibles à cet antibiotique.

- Si le plasmide utilisé est de la famille pUC: Coloration des colonies (lac Z', cf. chapII: Les Vecteurs).
- les colonies colorées ne seront pas recombinantes (lac Z' fonctionnel).
 - les colonies blanches seront recombinantes (lac Z' altéré).

nb: l'ensemble des clones recombinants forme une banque d'ADN génomique.

- 6) Criblage de la banque: Il s'agit de repérer les clones recombinants possédant le gène G. On utilise souvent l'hybridation sur colonies:

- Transfert des bactéries sur filtre de nylon (ou nitrocellulose).
- Lyse des bactéries et dénaturation de l'ADN par NaOH
- Fixation de l'ADN bactérien monobrin par chauffage (= ADNs reste ADNs)
- Hybridation par sonde radioactive spécifique du gène étudié.
- Lavage: élimination des sondes non hybridées
- Autoradiographie:
 - si rayonnement β = sonde hybridée donc présence du gène étudié = recombinant voulu
 - si pas de rayonnement β = pas de sonde hybridée = recombinant sans le gène voulu

2. Banques d'ADN génomique eucaryote:

Les ADN génomiques étant de grande taille, on ne peut pas utiliser les plasmides comme vecteur (trop petit insert). Nous devons choisir des phages ou des cosmides. Dans l'exemple présenté, le vecteur est un phage.

Fig 4 p 26:

- 1) Digestion: L'ADN phagique et l'ADN génomique sont digérés par la même enzyme de restriction.
 - L'ADN phagique est ouvert au niveau du site de restriction.
 - L'ADN génomique est fragmenté.
- 2) Traitement: L'ADN phagique ouvert est traité par la phosphatase alcaline de manière à ce qu'il ne puisse se refermer sur lui-même sans avoir intégré un ADN étranger (fragment d'ADN génomique).
 - Action de la phosphatase alcaline: Elle est capable d'hydrolyser les ADN, ARN et nucléotides libres: elle retire le P en 5' et libère donc:
 - un P inorganique
 - une extrémité 5'OH
 - Elle empêche donc l'action des ligases sur les acides nucléiques (car elles ont besoin d'une extrémité 5'P libre sur le brin à lié).
- 3) Ligation: On réunit les deux digestions en présence de T4 ADN Ligase.
 - Formation d'ADN phagique recombinant
- 4) Encapsulation
- 5) Infection
- 6) Sélection sur plage de lyse: Chaque plage correspond à 1 phage recombinant avec un insert particulier. L'ensemble des plages de lyses correspond à la banque d'ADN génomique eucaryote.
- 7) Criblage: Hybridation de plage de lyse

3-2- purification des acides nucléiques

3-2-1- Préparation de l'ADN à partir du sang total

Les globules blancs (lymphocytes, macrophages...etc.) représentent la source majeure d'ADN en médecine. A partir de 10 à 30 ml de sang en présence d'un anticoagulant (ex. EDTA), il est possible de récupérer une quantité d'ADN ($\approx 100\mu\text{g}$) suffisante, pour des études qualitatives et quantitatives.

➤ Protocole expérimental

- 1- Prélèvement de 10 à 30 ml de sang dans un tube contenant un anticoagulant (ex. EDTA).
- 2- Éclatement des globules rouges par une solution hypotonique
- 3- Récupération des globules blancs par centrifugation.
- 4- Préparation du lysat cellulaire en utilisant la solution de lyse (détergent comme le SDS ou sarcosyl + Protéinase K). Cette étape permet la libération de l'ADN nucléaire dans le milieu et la digestion des protéines qui lui étaient associées.
- 5- Extraction phénolique : cette étape permet de récupérer l'ADN dans la phase aqueuse après la centrifugation, car le phénol est un déprotéinisant puissant dans lequel les acides nucléiques ne sont pas solubles.
- 6-Extraction au chloroforme ou à l'éther : cette étape complète toujours une extraction phénolique, car elle permet d'éliminer les traces de phénol qui aurait pu être importé par la phase aqueuse (le phénol peut inhiber les enzymes utilisées dans l'amplification, la restriction...etc.).
- 7- Précipitation : cette étape permet la précipitation de l'ADN pour faciliter sa récupération. Il y'a deux types largement utilisés :
 - a**-Précipitation à l'alcool éthylique : Elle doit être effectuée à haute force ionique (2.5 volumes d'éthanol 95° contre un volume d'échantillon). Pour les faibles concentrations, il faut prolonger le temps de précipitation ($> 10\text{h}$), comme il est possible d'accélérer la précipitation par le froid (-20°C à -70°C). L'ADN est récupéré par centrifugation.
 - b**-Précipitation à l'isopropanol : Le principe est le même que la précipitation éthanolique. Ce type de précipitation se fait volume à volume. Mais deux caractéristiques majeures la différencient de la précipitation éthanolique.
 - Le sel n'est pas nécessaire

- Les très petits fragments de DNA même à hautes concentrations ne sont pas précipités, ce qui permet de les éliminer.

8- Lavage : l'utilisation de l'éthanol 70° permet d'éliminer les sels, et les traces de l'isopropanol. L'ADN est récupéré par centrifugation.

9- Séchage: cette étape permet l'élimination de l'éthanol (s'évapore).

10- Resuspendre dans 10mM tris EDTA, l'eau purifiée de nucléases...etc.

3-2-2- Dosage des acides nucléiques

Les pyrimidines et les purines absorbent fortement les UV à 260 nm. Une unité de densité optique à 260 nm correspond à:

- Une solution de DNA double brin à 50 µg/ml

- Une solution de DNA simple brin ou RNA à 25 µg/ml

$$C = A_{260} * DF * 100$$

Ces valeurs s'appliquent à des acides nucléiques parfaitement purs et en solution homogène.

Pour vérifier la pureté de l'ADN il faut calculer :

$$P \text{ (pureté)} = A_{260} / A_{280}$$

Une solution d'ADN est considérée pure si : $1.7 \leq P \leq 2$

3-3- Southern blot

La procédure de ce type d'hybridation est résumée en sept étapes (figure 6)

1- Extraction de l'ADN

2- Digestion enzymatique (par les enzymes de restriction)

3- Electrophorèse du produit de la digestion

4- Transfert sur membrane

5- Hybridation avec une sonde spécifique marquée.

6- Lavages

7- Révélation

On peut aussi révéler la présence des séquences spécifiques dans les ARN (Northern Blot) et des protéines (Western Blot) utilisant le même principe.

3-4- Northern blot

Elle dérive du Southern blot, sauf qu'au lieu d'étudier de l'ADN, on étudie de l'ARN. L'ARN va être analysé par électrophorèse et détecté par une sonde. Une différence du procédé par rapport au transfert de Southern est l'utilisation de formaldéhyde dans le gel d'électrophorèse

comme dénaturant, parce que le traitement d'hydroxyde de sodium utilisé en transfert de southern dégraderait l'ARN. Comme dans le transfert de Southern, la sonde d'hybridation peut être faite à partir d'ADN ou d'ARN.

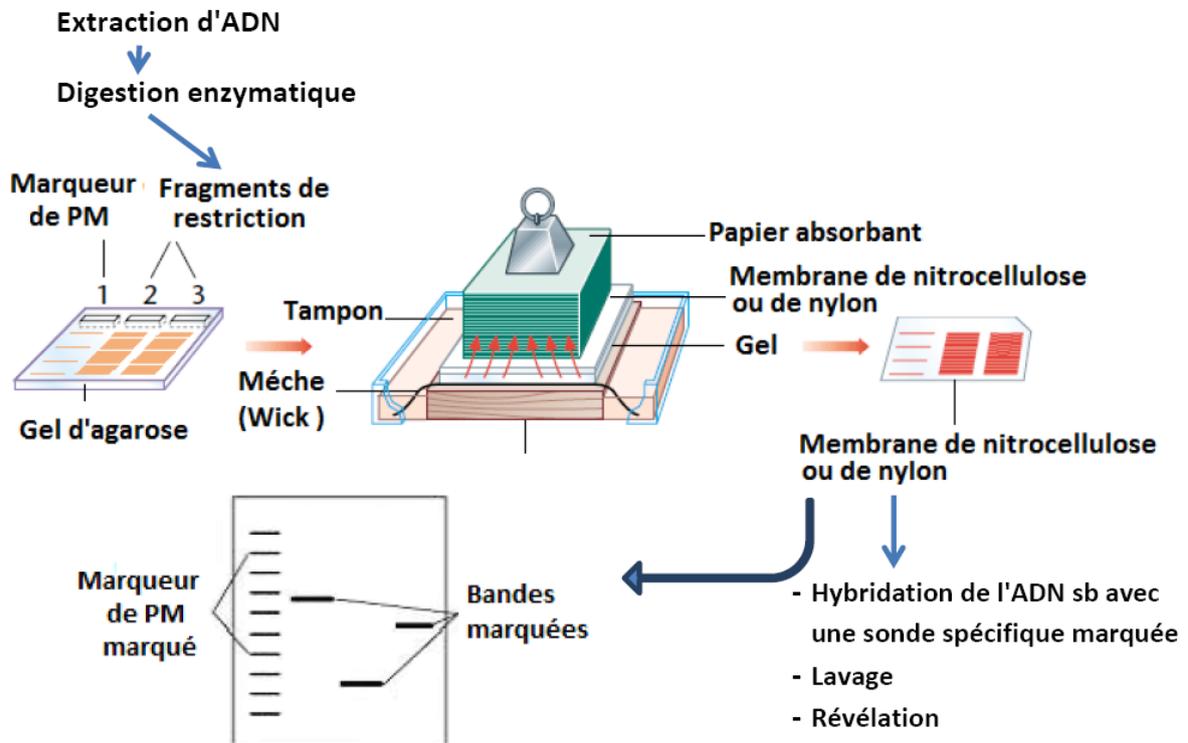


Fig.6: Southern Blot.

3-5- Réaction de Polymérisation en Chaîne "PCR"

La synthèse et le séquençage d'ADN ont permis l'émergence d'une méthode d'amplification de l'ADN appelée PCR (**P**olymerase **C**hain **R**eaction), à partir d'un gène (ou fragment) spécifique, de grand quantité d'ADN sont obtenues *in vitro*. L'ADN polymérase copiant jusqu'à un milliard de fois cette région cible « quantité suffisante pour être révélée ».

3-5-1- Etapes de PCR

La dénaturation de la double hélice nécessite une haute température (autour de 95°C), donc l'enzyme utilisée doit être résistante aux hautes températures ou bien elle devait être ajoutée lors de chaque cycle. Ce problème a été résolu en utilisant une ADN polymérase thermiquement stable d'une bactérie thermophile : *Thermus aquaticus*, isolée d'une source chaude. Cette enzyme nommée ADN Taq polymérase est stable à 95°C et n'est pas affectée par l'étape de dénaturation. Avec l'utilisation du Taq polymérase, l'ADN étant copié à 72°C et non plus 37°C, son utilisation augmente la spécificité de la PCR. Car aux hautes températures l'hybridation non spécifique d'amorces à l'ADN **non** cible est rare. Les

produits de l'ADN Taq polymérase sont donc plus homogènes que ceux obtenus avec l'ADN polymérase III d'*E. coli*. En résumé il y'a quatre étapes :

1- **Dénaturation** (autour de 95°C) : Sert à dénaturer l'ADN pour obtenir des matrices simple brin.

2- **Hybridation** (autour de T_h): Lors du refroidissement du mélange, les amorces étant en excès, la plupart des brins d'ADN cibles se fixent à celles et non entre eux.

3- **Extension** (72°C) : L'ADN polymérase « prolonge les amorces : côté 3'-OH libre » en utilisant les brins cibles comme matrice.

4- Après une période appropriée d'incubation, le mélange est chauffé de nouveau pour séparer les brins. Il est alors refroidi pour que les amorces s'hybrident aux régions complémentaires de l'ADN nouvellement synthétisé. Le processus complet est alors répété « cycle ».

Cette technique permet entre autres de détecter la présence du virus VIH, coronavirus (c'est la technique utilisée dans l'institut Pasteur pour révéler les cas suspects) ou de mesurer la charge virale (concentration du virus dans le plasma), des OGM (organismes génétiquement modifiés), des virus des hépatites B, C et D. De plus en plus utilisée en criminalistique.

3-5-2- Thermocycleur

Les trois étapes, constituant un cycle de PCR, sont effectuées à des températures différentes permettant de contrôler l'activité enzymatique. Pour effectuer ces transitions de températures, les microtubes contenant le mélange réactionnel sont placés dans un appareil programmable : un thermocycleur (figure 7). Cet appareil permet d'exposer les tubes à des températures choisies et pour des durées déterminées par l'expérimentateur (effectuer automatiquement les cycles de chauffage) (figure 8).

Le secret de la PCR est que le produit d'un cycle d'extension sert de matrice pour le suivant (fig. 8), ainsi la technique PCR est remarquable par le fait que l'ADN cible se double à chaque cycle. En pratique 20 à 40 cycles sont habituellement réalisés, ils génèrent 10^6 à 10^9 fois la séquence cible. « Elle repose sur la succession de plusieurs cycles » (figure 9).

La progression géométrique de raison 2 (tableau 1) permet d'établir une relation pour calculer le nombre des copies totale et cibles.

Polycopie de cours de Biologie moléculaire

Tableau 1. Progression géométrique de raison 2 des cycles de PCR et nombre de copies.

Cycle	1	2	3	4	5	6	-----n	n : nombre de cycle
Copies totale (N)	2	4	8	16	32	64	-----	$N=2^n$
Copies parasites	2	4	6	8	10	12	-----	$2n$
Copies cibles	0	0	2	8	22	52	-----	2^n-2n

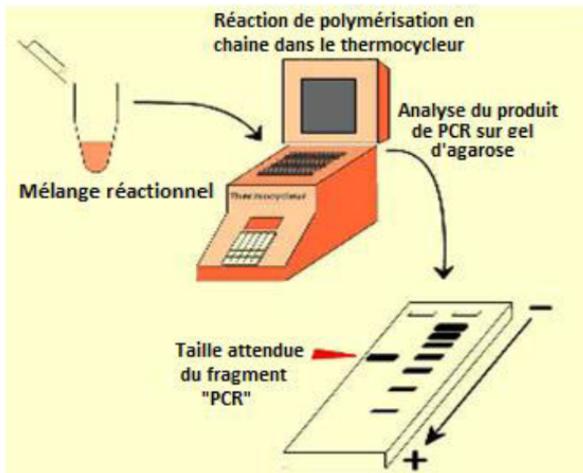


Fig.7:Réaction d'amplification d'ADN dans le thermocycleur et visualisation du produit de PCR par électrophorèse.

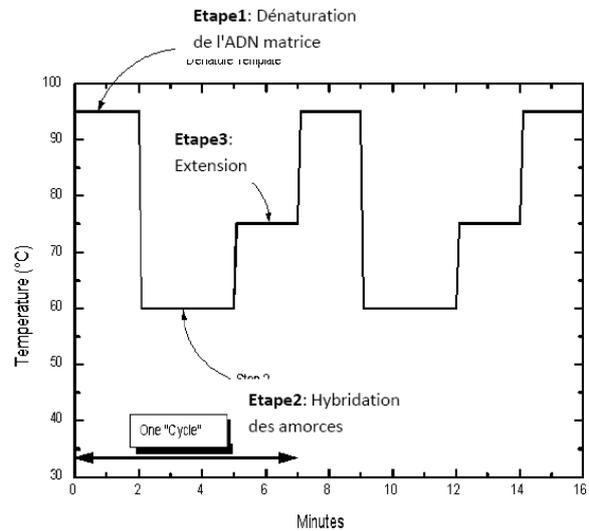


Fig.8:Profile de cycles de la température de la PCR

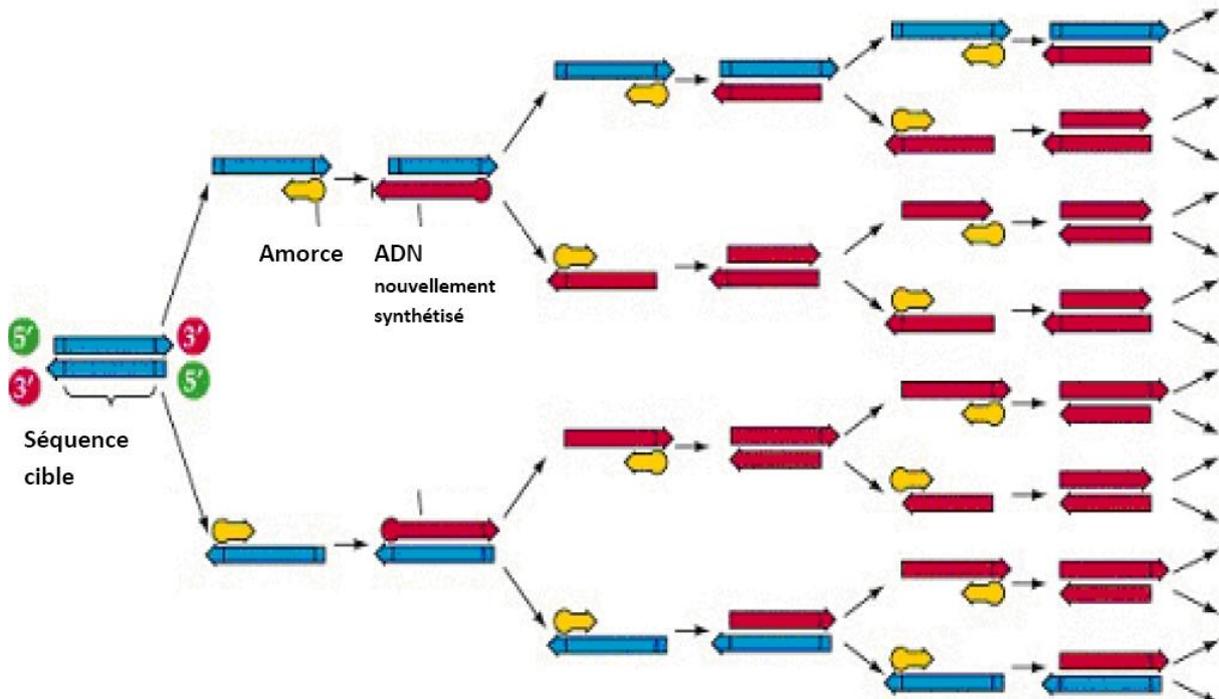


Figure 9. Amplification d'ADN par PCR. Le produit d'un cycle sert de matrice pour le cycle qui suit.

3-6- Séquençage

Le séquençage est une technique d'analyse de l'ADN, il permet de déterminer l'ordre des nucléotides. Actuellement la technique de séquençage repose majoritairement sur la méthode enzymatique de Sanger.

3-6-1- Méthode de Sanger

Cette méthode génère des fragments d'ADN terminés par l'une des quatre bases marquées au préalable, car l'élongation de la chaîne s'arrêtera après l'incorporation d'un didésoxynucléotide (ddNTP) (fig. 10). C'est le principe sur lequel baser cette méthode.

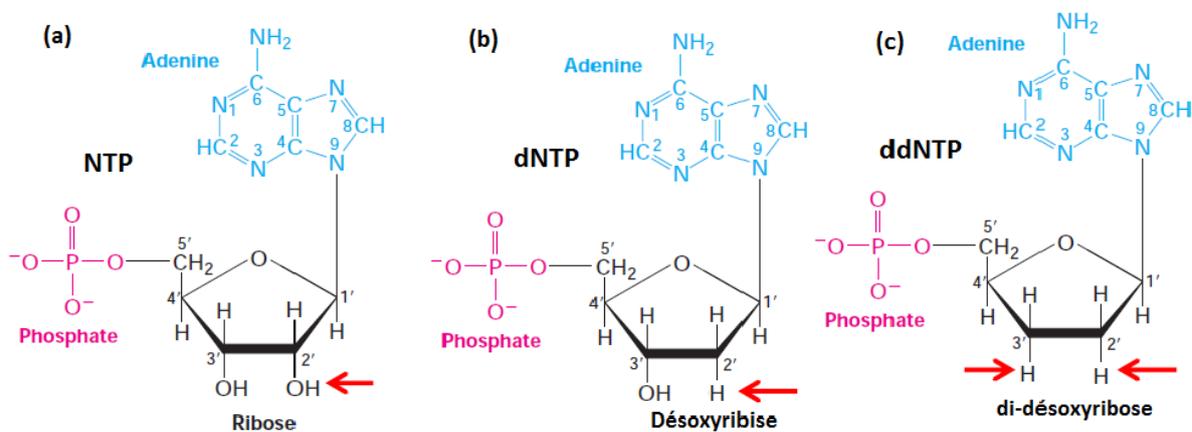


Fig.10: Les différentes formes des nucléotides.

La méthode de Sanger permet de déterminer une séquence inconnue par l'intermédiaire d'une copie d'ADN simple brin. Réalisé grâce à l'ADN polymérase et une amorce spécifique. Pour mieux comprendre le principe de séquençage, on utilise des ddNTP non marqués. Alors on utilise quatre tubes séparés contenant l'un des ddNTPs (ddATP, ddCTP, ddGTP, ddTTP). Chaque tube contient: ADN polymérase (Ex. Taq Polymérase); amorce, dNTPs (dATP, dTTP, dCTP, dGTP); eau purifiée de nucléases, MgCl₂ et l'ADN à séquencer.

La détermination de l'enchaînement des nucléotides dans un fragment d'ADN simple brin par la méthode de Sanger comporte deux étapes:

1- Synthèse du brin complémentaire d'un brin matrice à l'aide d'une polymérase à partir d'une amorce, en présence des 4 dNTPs et d'un ddNTP qui joue le rôle de terminateur de chaîne (absence de 3'-OH). Quatre réactions en parallèle sont nécessaires, chacune avec l'un des terminateurs présents en faible quantité.

2- Séparation par électrophorèse (PAGE) des fragments synthétisés, le nombre de ces fragments dépend du nombre de résidus nucléotidiques. Donc la position des terminateurs dans la séquence (figure 10).

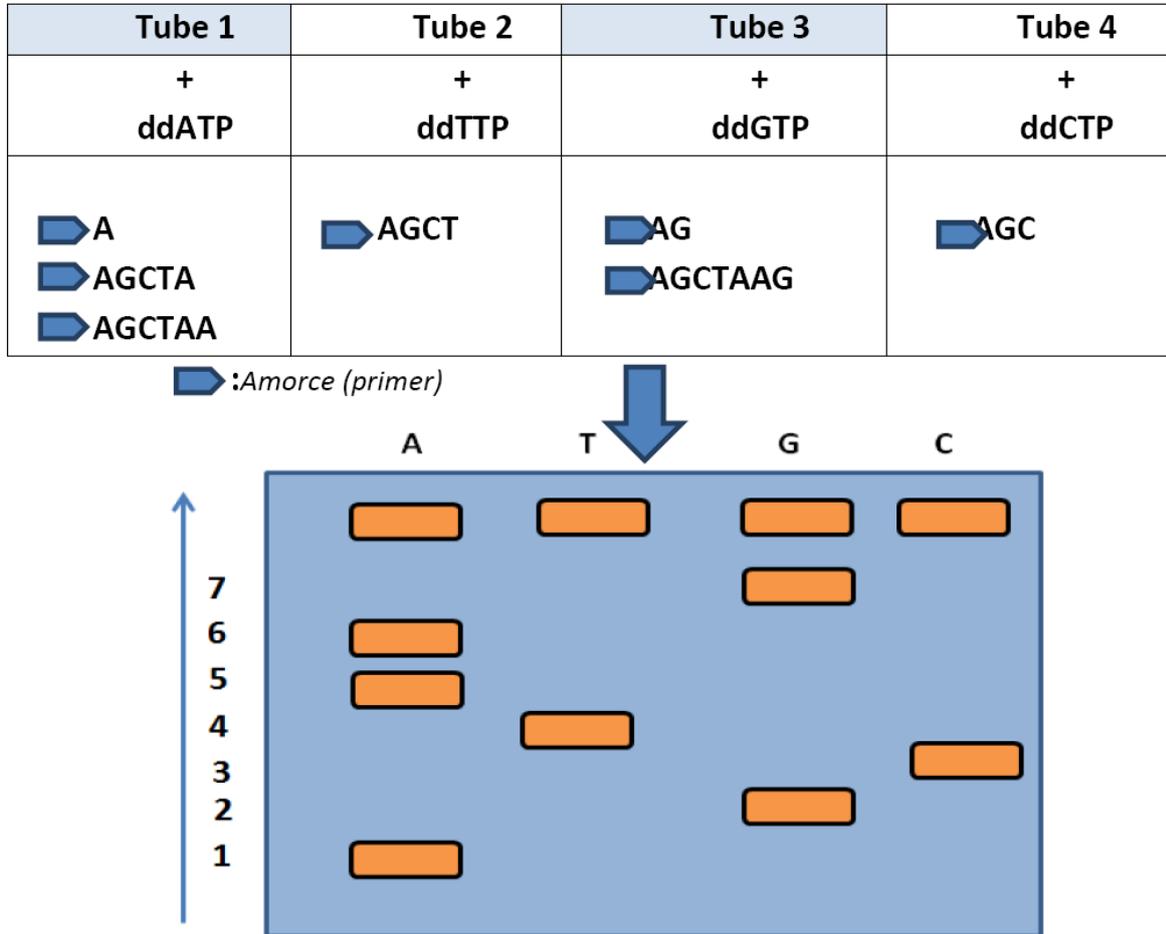


Fig.10: Séquençage de l'ADN par la méthode de Sanger

4- ADN C " ADN complémentaire"

4-1- Caractéristiques de l'ARNm eucaryote

Chez les eucaryotes la région codante peut être interrompue par 1 ou plus d'introns. Donc si on utilise l'ARNm on peut réaliser les avantages suivants :

- 1- Les introns sont éliminés par épissage (splicing).
- 2- Les tissus spécifiques permettent d'avoir des quantités considérables d'ARNm (plus de copies par rapport au gène d'intérêt).

Il faut noter que l'ARNm chez les eucaryotes est caractérisé par la queue poly (A), donc on peut utiliser poly (T) (12 à 18) comme amorce pour des ARNm eucaryotes différents et aussi pour leur extraction (fig.11).

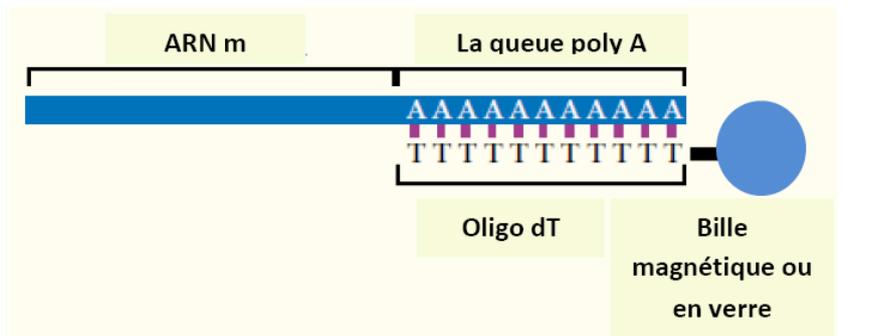


Fig.11: L'utilisation d'une queue poly (T) fixée sur un support pour l'extraction des ARNm eucaryotes.

4-2- Synthèse de l'ADNc

La synthèse d'ADNc est catalysée par des transcriptases inverses (RT). Ces enzymes sont des ADN polymérases ARN dépendantes, capables d'utiliser un brin d'ARN comme matrice pour catalyser la synthèse d'un brin d'ADN complémentaire. Cela correspond effectivement à l'«inverse» d'une réaction de transcription de l'ADN en ARN. Les transcriptases inverses sont issues de rétrovirus dont elles sont une des principales caractéristiques. Par exemples, la transcriptase inverse du Virus du Myéloblastome Aviaire (AMV), celle du Virus de la Leucémie Murine (Mo-MLV [Molony Murine Leukemia Virus] ou Mu-LV).

Comme toutes les ADN polymérases, les transcriptases inverses ne peuvent pas initier seules la synthèse d'un brin d'ADN. Elles ont besoin d'une amorce possédant une extrémité 3'-OH libre. Lorsque les ARN à amplifier sont polyadénylylés en 3' (ARNm eucaryotes par exemple), l'amorce choisie peut être simplement une séquence polyT (12 à 18) constituée d'une succession de désoxythymidines (comme sur le schéma ci-dessous en rouge). Dans ce cas, tous les ARNm sont a priori copiés en ADNc.

La ribonucléase H ou RNase H (E.C.3.1.26.4) est une enzyme qui hydrolyse le brin d'ARN dans un double brin hybride ADN:ARN. Elle libère des extrémités 3'-OH et 5'-phosphate. La RNase H n'hydrolyse pas l'ARN simple ou double brin (fig.12).

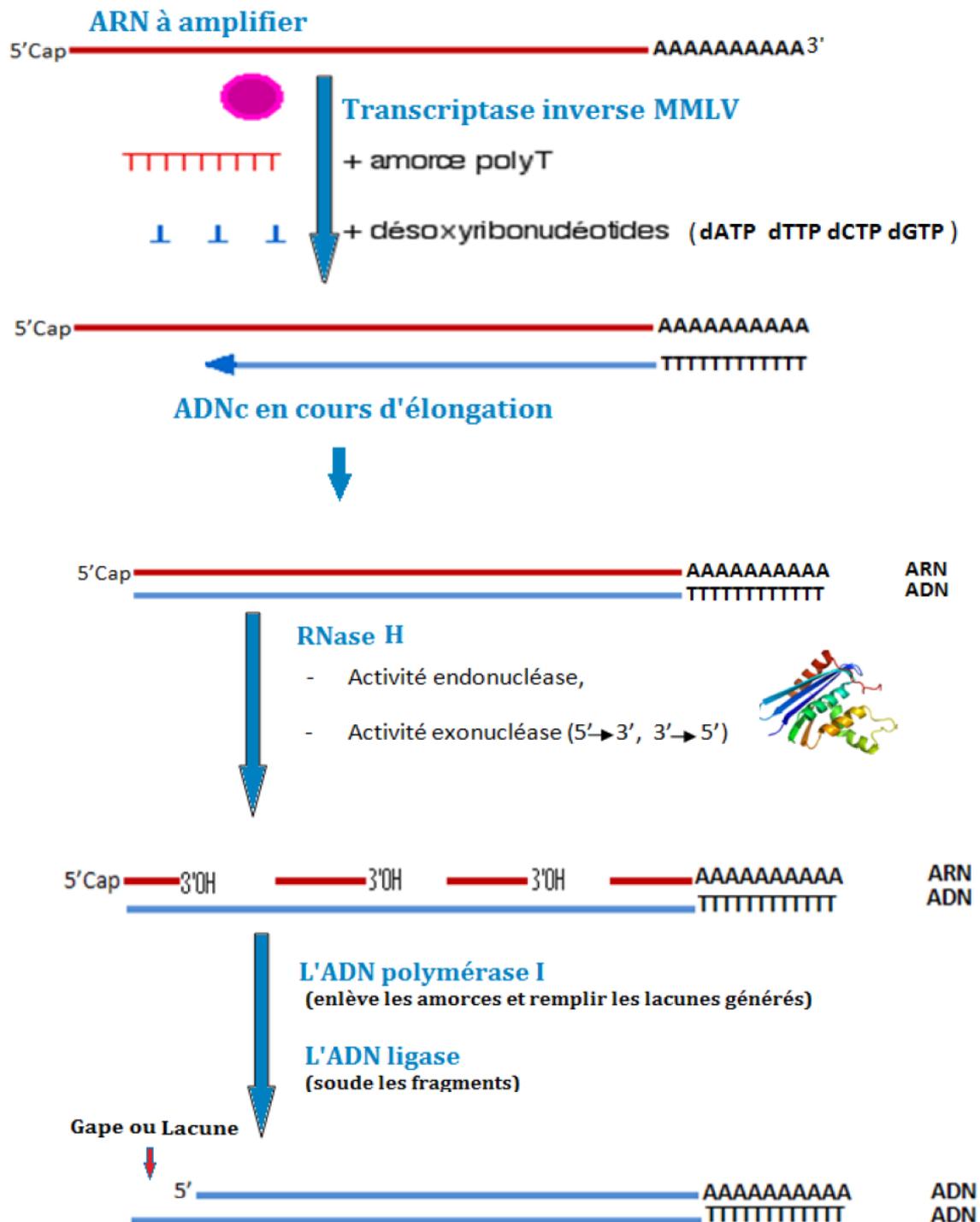


Figure12. Les étapes de la synthèse d'ADNc.

4-3- Insertion de l'ADNc dans un vecteur approprié

Il y'a deux méthodes d'insertion de l'ADNc dans les vecteurs:

- 1- Addition d'une queue homopolymérique.
- 2- Addition d'un joint (linker).

Polycopie de cours de Biologie moléculaire

L'addition d'une queue homopolymérique fait par une ADN polymérase inhabituelle trouvée chez des types de cellules eucaryotes appelées « pré-lymphocytes ». Cette enzyme est appelée transférase terminale, elle est capable d'ajouter un désoxynucléotide dans la partie 3'-OH d'un brin d'ADN (ne nécessite pas d'une amorce et d'une matrice).

à la présence de cations divalents (Mg^{++} , Co^{++}), l'enzyme catalyse l'addition de dNTPs (Purines : Mg^{++} ; Pyrimidines : Co^{++}) à l'extrémité 3' de l'ADN. Selon les conditions de réaction on ajoute trois à quelques milles nucléotides.

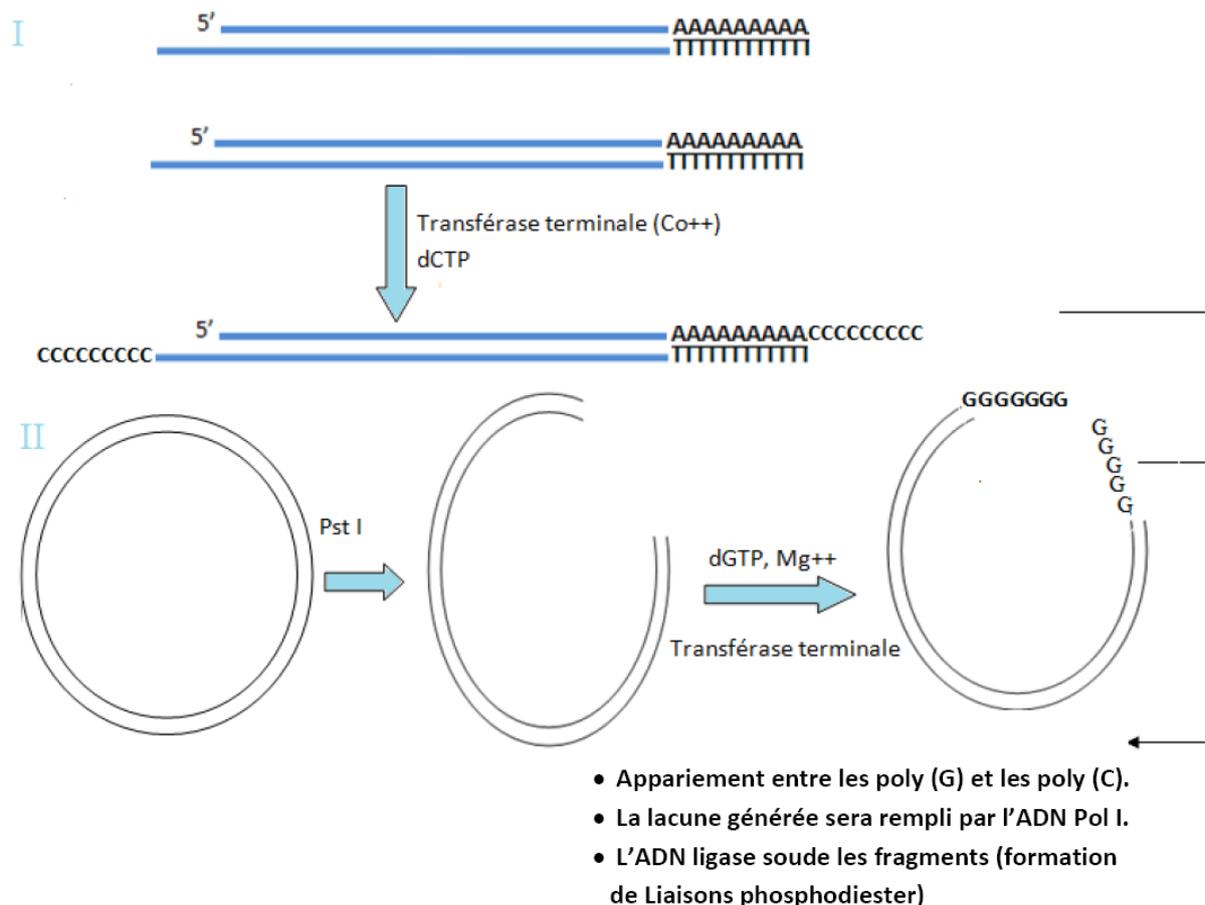


Figure 13. Insertion de l'ADNc dans un vecteur.