

Hybridation moléculaire, sondes, enzymes et vecteurs

I. Hybridation moléculaire

1. Définitions

* Dénaturation = Fusion : Rupture des liaisons hydrogènes entre les 2 brins. Il faut donc un apport d'énergie.

Ce phénomène est visualisable grâce à l'effet hypochrome : variation de la DO à 260 nm due au passage db --> sb.

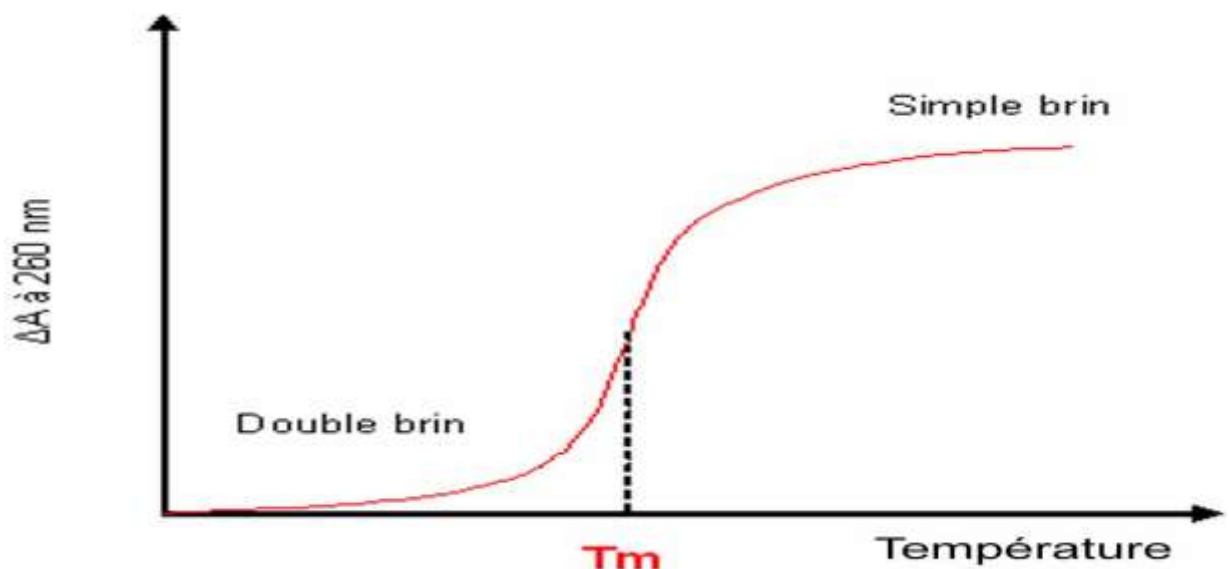
• Température de fusion T_m (T_f) : Température à laquelle se produit le passage de l'état bicaténaire à l'état monocaténaire.

T_m est fonction de plusieurs éléments :

- composition en bases
 - longueur des brins
 - les mésappariements
 - milieu environnant l'ADN :
- Une diminution de la force ionique diminue le T_m .
 - La présence de formamide diminue le T_m .

2. La température de fusion de l'ADN

C'est une température, appelée également Température de demi dénaturation (T_m) qui correspond à l'ouverture ou au déroulement de 50% de la chaîne de l'ADN chauffé. C'est l'effet hyperchromique qui correspond à la rupture des liaisons hydrogènes (liaisons faibles) et la séparation des 2 chaînes entraîne une augmentation dans l'absorption d'U.V de 40% à 260 nm. La valeur de T_m des ADN est une fonction linéaire du pourcentage de (G + C) de l'ADN :



Mise en évidence du phénomène coopératif de passage de l'ADN double brin à l'ADN simple brin et mesure de la T_m .

3. Formules utilisées pour calculer la Th

Formule de Wallace: $T_m = 4(C+G) + 2(A+T)$; $Th = T_m - 5$

La formule de Wallace permet de déduire la Th des petits oligonucléotides (Ex. Les amorces).

Pour les oligonucléotides avec $N < 100$ nucléotides on utilise la formule suivante en tenant compte de la concentration du sel, les mésappariements et la concentration de formamide:

$$T_m = 16.6 \log[Na^+] + 0.41(\%C+G) + 81.5 - (675/N) - \%mismatch - 0.65\% \text{ de formamide}$$

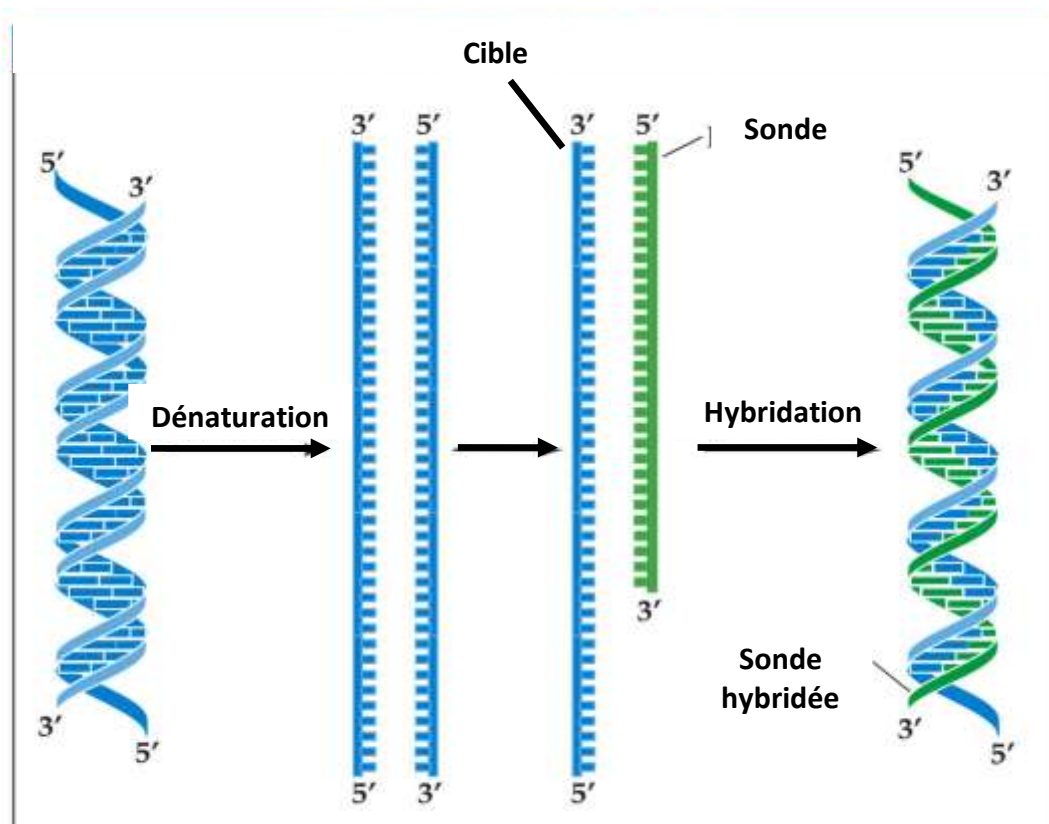
Dans les conditions réactionnelles standards la T_m pour les longs séquences (>100):

$$T_m = 69.3 + 0.41 (\%C+G)$$

La formule globale:

$$T_m = 81.5 + 16.6 \log [\text{sel}] + 0.41 [\%C+G] - \% \text{ mésappariement} - (500/N) - 0.65\%(\text{formamide}).$$

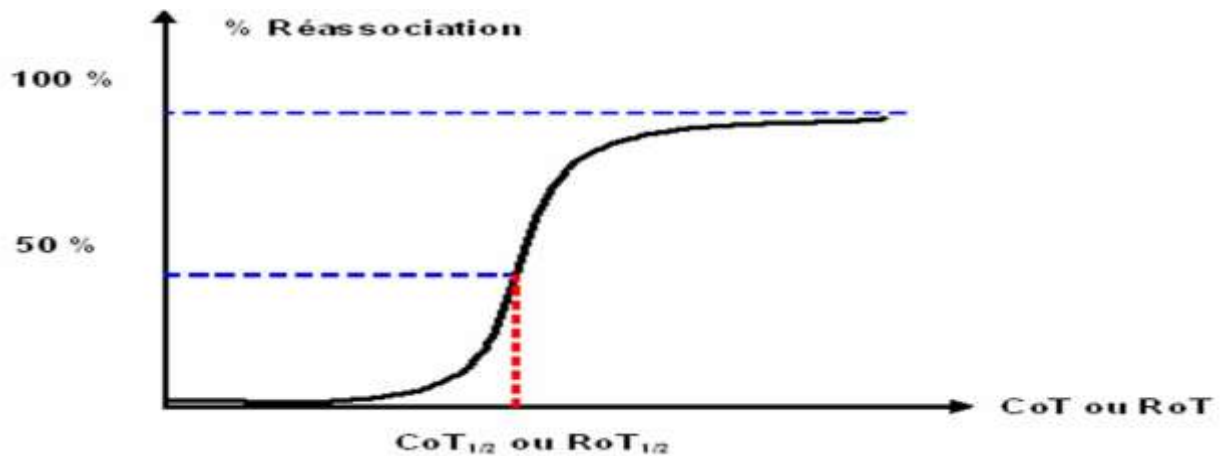
4. Principe de l'hybridation



5. Notion de Cot et Rot

Concentration de l'ADN et temps :

L'hybridation des séquences nucléiques est aléatoire. Cependant, si la concentration de l'ADN est grande, le nombre de copies hybridées sera grand. Il en résulte donc que la vitesse d'hybridation augmente lorsque la concentration de l'ADN augmente. Il en est de même pour le facteur temps : la probabilité d'association des brins complémentaires est importante lorsque le temps est long. L'hybridation est quantifiée en fonction de ces deux variables prises ensemble. Dans le cas des ADN, cette variable est nommée le CoT, pour le cas des hybrides ADN/ARN, elle est dite le RoT :



6. Types d'hybridation

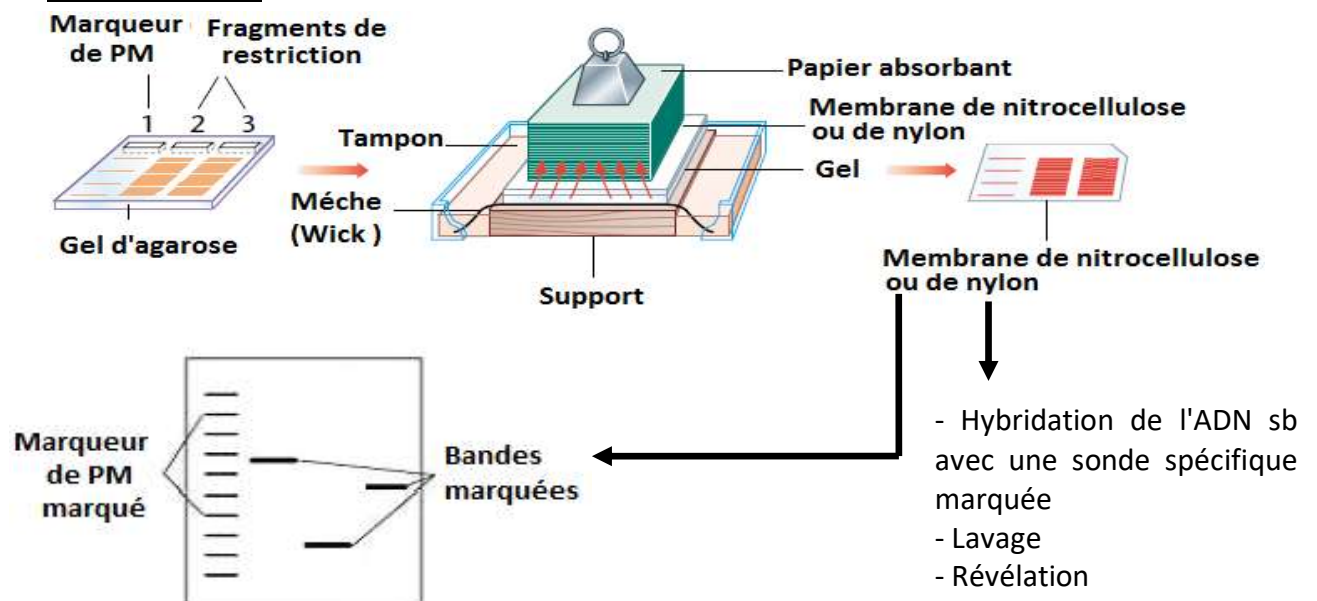
L'objectif de l'hybridation est la détection de la présence d'un acide nucléique d'une séquence donnée par l'utilisation d'un fragment d'ADN complémentaire = sonde. Cela peut avoir lieu en solution ou sur support solide (immobilisation de la cible sur une membrane [nitrocellulose, nylon], sur verre, colonies bactériennes, chromosomes, plage de lyse ...etc.).

6. 1. Hybridation d'ADN sur support solide : Southern blot

La procédure de ce type d'hybridation est résumée en sept étapes

- 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

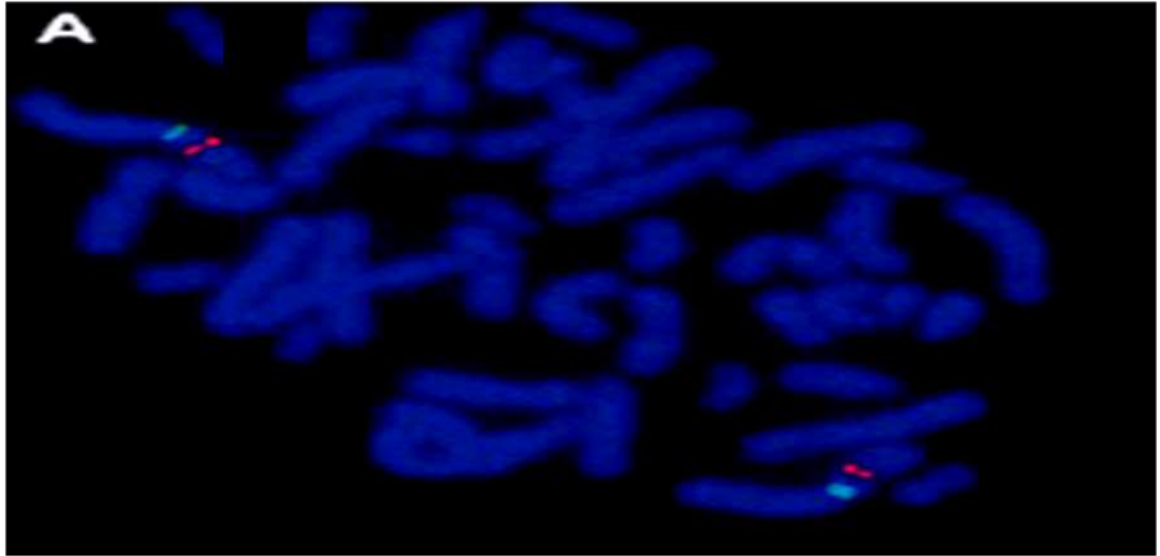
7. Southern Blot



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

8. Hybridation *in situ* sur chromosome

Cette technique permet de déterminer sur quel chromosome et dans quelle région se trouve le gène d'intérêt. Les chromosomes fixés sur lame sont hybridés avec une sonde marquée (H^3). L'examen microscopique révèle les signaux positifs sous forme de grains noirs disposés sur le chromosome. Comme exemple d'application est la détection des gènes viraux intégrés dans les chromosomes. Examine



Rouge : sonde gène *BCL11A*

Vert: sonde centromère chromosome

⇒ *Localisation en 2q13*

9. Hybridation sur colonie de bactéries, sur plage de lyse

Ce type d'hybridation permet la détection parmi un grand nombre de bactéries ou de phages recombinants (plage de lyse) celle ou celui qui contient le fragment d'ADN cible. La réalisation de cette technique passe par les étapes suivantes:

- Culture pure sur boîte de Pétri (colonies ou plage de lyse)
- Transfert sur membrane de nylon ou de nitrocellulose
- Lyse alcaline des bactéries et dénaturation de l'ADN
- Fixation par la chaleur ou les UV
- Hybridation avec une sonde marquée
- Lavage
- Autoradiographie

II. Les sondes

1. Définition

* **Sonde** : Séquence nucléotidique, généralement marquée, complémentaire d'une séquence d'ADN ou d'ARN avec laquelle elle va s'hybrider.

* **Sondes directes** : exploration d'un gène ou d'une région génique dont on a déjà quelques informations.

* **Sondes indirectes ou sondes anonymes** : exploration de gènes peu ou pas connus

2. Les différents types de sondes

- **sonde génomique** : fragment obtenu par coupure de l'ADN génomique

- **sonde ADNc** : sonde ADN obtenue par transcription réverse d'un ARNm messager (= séquence codante)

- **sonde oligonucléotides** : ADN (ou ARN) simple brin de 18 à 50 nucléotides synthétisé chimiquement

- **Les ribosondes**

3. Obtention des sondes

- clonage

- PCR

4. Marquage des sondes

4. 1. Agent de marquage :

- **marquage radioactif**

- ^{32}P , ^{35}S , ^3H

1. nécessité de se protéger contre le rayonnement émis, maniement des sondes inconfortable

2. décroissance rapide du ^{32}P , d'où besoin de marquer les sondes fréquemment

Avantage: grande sensibilité

- **marquage froid**

- direct - nucléotide modifié par un fluorophore

- indirect - nucléotide marqué par un reporter qui sera repéré par une molécule affine

4. 2. Stratégies de marquage

- Nick translation

- multi-amorçage au hasard

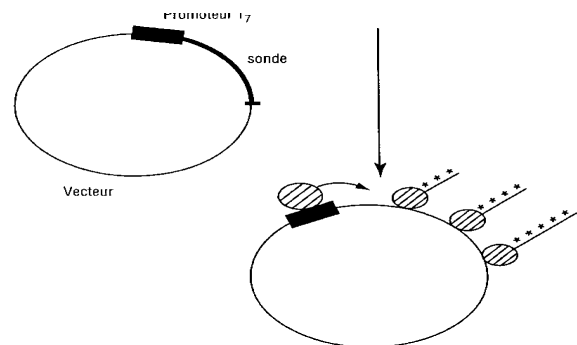
- marquage des oligonucléotides

5. Les sondes froides

Marquage direct avec un fluorophore

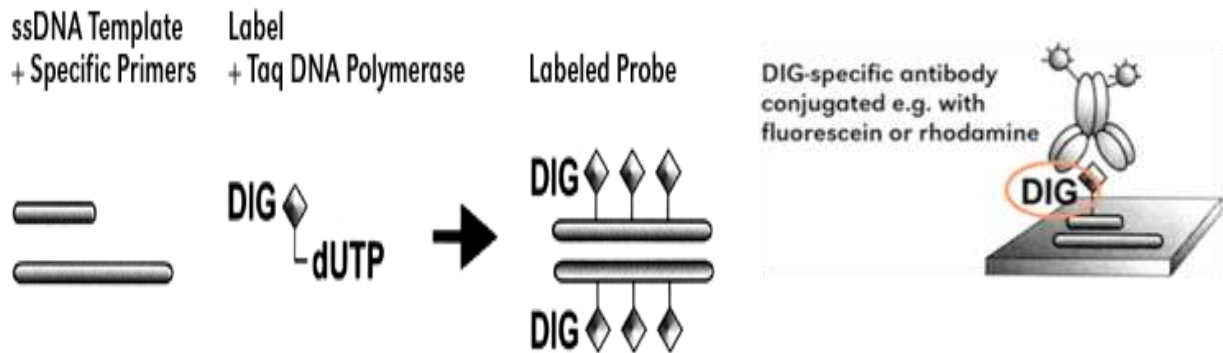
Le nucléotide portant le marqueur est incorporé directement.

Groupe chimique qui fluoresce quand il est exposé à une longueur d'onde donnée: Fluorescéine, Cy5, Cy3, Rhodamine, Texas red, TAMRA, TET.

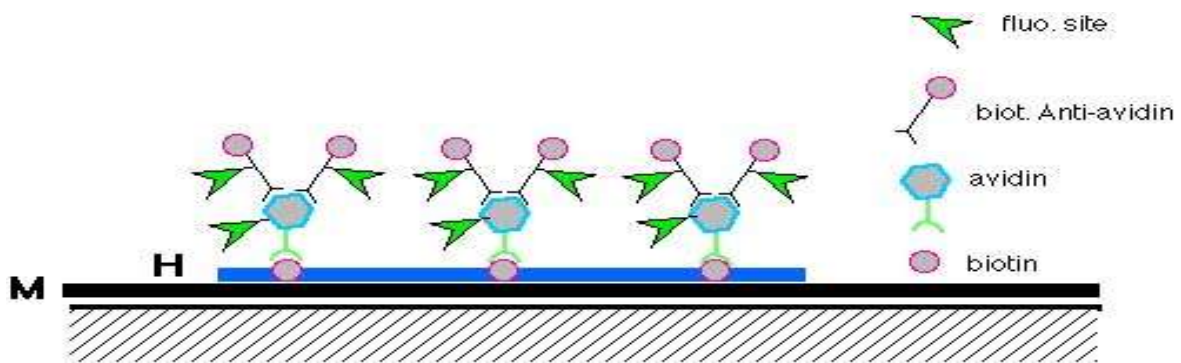


Marquage indirect

- digoxigénine

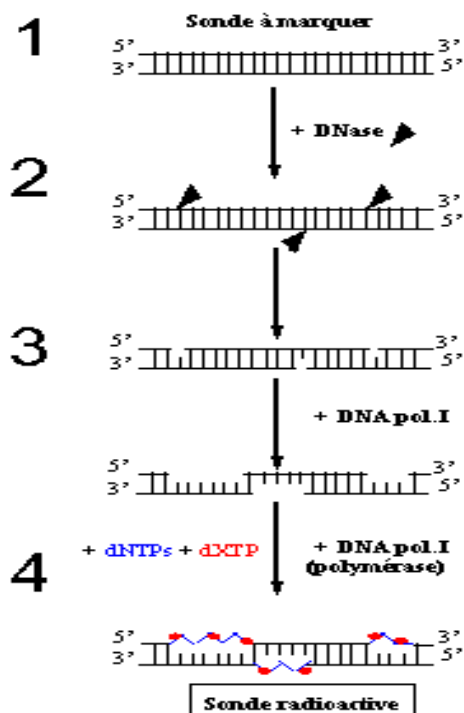


- marquage par la biotine-streptavidine



6. Marquage par Nick-translation

Marquage par Nick-translation



Endonucléase utilisée dans des conditions ménagées

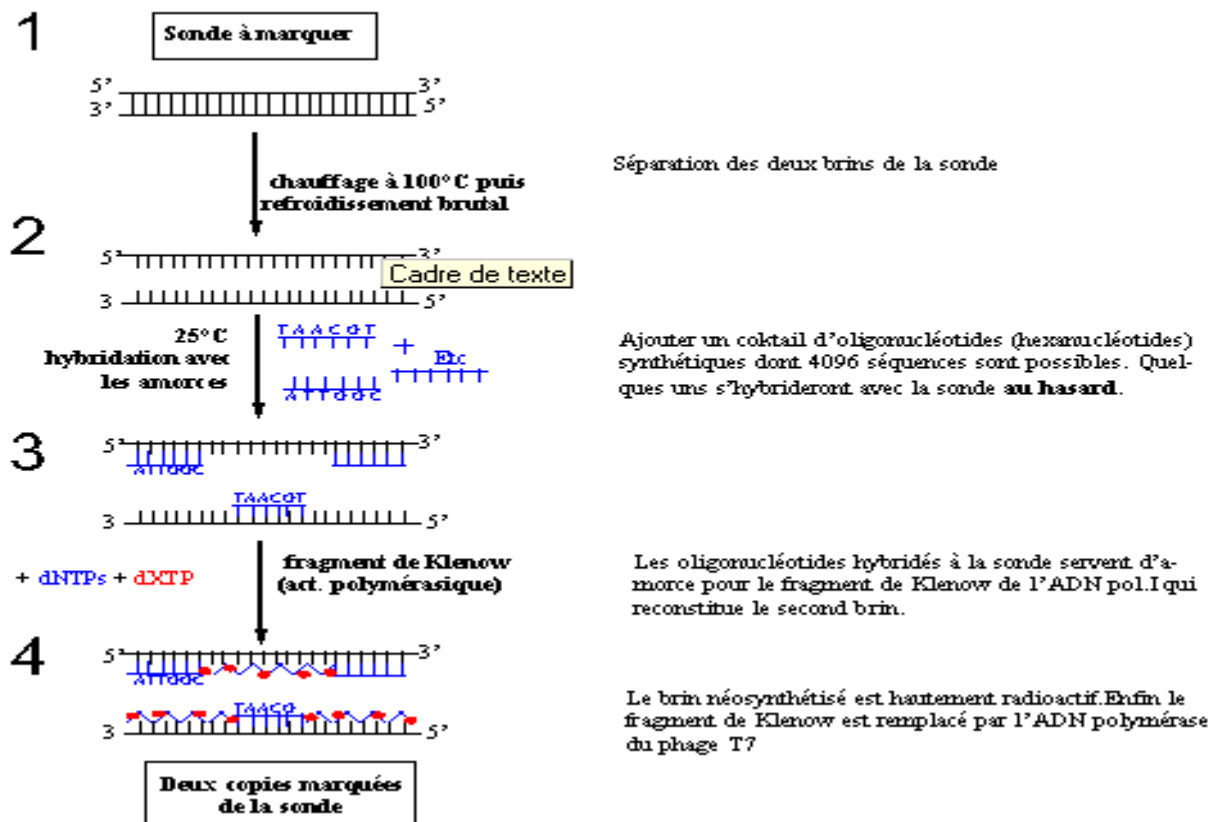
La DNase génère quelques cassures aléatoires simple brin

Au niveau des cassures la DNA polymérase I détruit l'ADN par son activité exonucléasique (5'-3')....

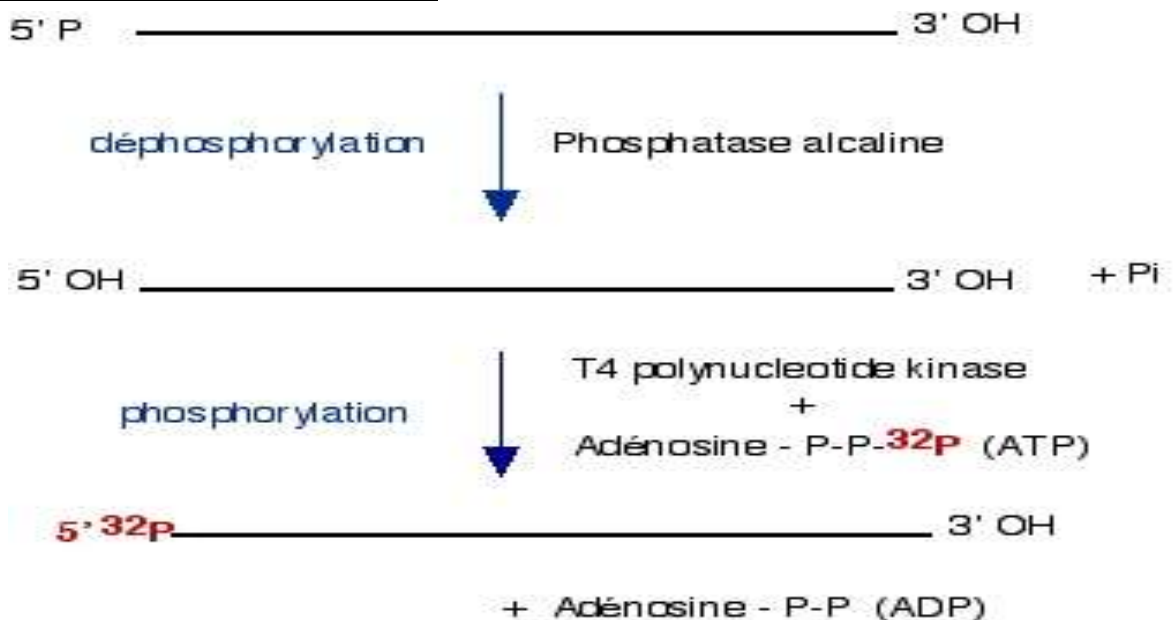
... et le synthétise par son activité polymérase en Présence de nucléotides dont un est marqué

7. Marquage par random priming

Marquage par random priming



8. Marquage d'un oligonucléotide



Différentes Applications

- Hybridation ADN/ADN ou ARN/ADN sur membrane
- Sondes oligonucléotides
- Séquençage
- Hybridation in situ

III. Enzymes de restriction et de modification de l'ADN

1. Enzymes de restriction

Les enzymes de restriction sont des endonucléases spécifiques, leur utilisation permet d'obtenir des fragments d'ADN aux extrémités bien caractérisées, dont certains peuvent contenir des gènes.

Il existe trois types d'enzymes de restriction:

Les enzymes de **type I** et **III** ne sont pas utilisées en pratique, car elles coupent l'ADN de façon aléatoire.

Les enzymes de restriction de **type II**, clivent spécifiquement les deux brins d'ADN au niveau d'une séquence bien définie. Elles ont la particularité de reconnaître et de couper des séquences palindromiques.

Exemple de quelques endonucléases et leurs origines

Enzyme	microorganisme	SITE DE RESTRICTION
<i>EcoRI</i>	<i>Echerichia coli</i>	↓ GAATTC CTAAG ↑
<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 ↑

2. Nomenclature des enzymes de restriction

Les trois premières lettres de chacune d'entre elle concernent

- La première lettre du nom de genre, par exemple E (*Escherichia*)
- Les deux premières lettres du nom de l'espèce, par exemple co (*coli*).
- La quatrième concerne la souche bactérienne d'où est extraite l'enzyme en question.
- Le chiffre romain: indique l'ordre de la caractérisation de l'enzyme chez la même souche.

HindIII signifie que c'est la troisième (**III**) Enzyme de restriction isolée et caractérisée de la souche bactérienne ***Haemophilus influenzae Rd.***

3. Exemples des enzymes de restriction de type II

Enzyme	Source Organism	Recognition Sequence
<i>HpaII</i>	<i>Haemophilus parainfluenzae</i>	C/CGG GGC/C
<i>MboI</i>	<i>Moraxella bovis</i>	/GATC CTAG/
<i>NdeI</i>	<i>Neisseria denitrificans</i>	/GATC CTAG/
<i>EcoRI</i>	<i>Escherichia coli</i> RY13	G/AATC CTTAA/G
<i>EcoRII</i>	<i>Escherichia coli</i> RY13	/CC(A or T)GG GG(T or A)CC/
<i>EcoRV</i>	<i>Escherichia coli</i> J62/pGL74	GAT/ATC CTA/TAG
<i>BamHI</i>	<i>Bacillus amyloliquefaciens</i>	G/GATCC CCTAG/G
<i>SauI</i>	<i>Staphylococcus aureus</i>	CC/TNAGG GGANT/CC
<i>BglI</i>	<i>Bacillus globigii</i>	GCCNNNN/NGGC CGGN/NNNNCCG
<i>NotI</i>	<i>Nocardia otitidis-caviarum</i>	GC/GGCCGC CGCCGG/CG
<i>DraI</i>	<i>Deinococcus radiophilus</i>	RG/GNCCY YCCNG/GR

N: n'importe quelle base **R:** n'importe quelle purine **Y:** n'importe quelle pyrimidine **/:** Position de coupure.

La plupart des enzymes de restriction de type II ont un site de reconnaissance de 4 ou 6 paires de bases, et un certain nombre d'enzymes de restriction reconnaissent des séquences à 5 ou à 7, 8 ... paires de bases.

4. Les séquences reconnues : palindromiques

Le site de coupure, compris dans le site de reconnaissance est représenté par une flèche verticale, et les produits de la digestion d'un génome par les enzymes de restriction sont appelés les fragments de restriction. Selon la position de la coupure (dans l'axe ou loin de l'axe) on distingue deux types de coupure:

- Coupure franche ex : **Sma I** : Dans l'axe \longrightarrow Extrémités franches

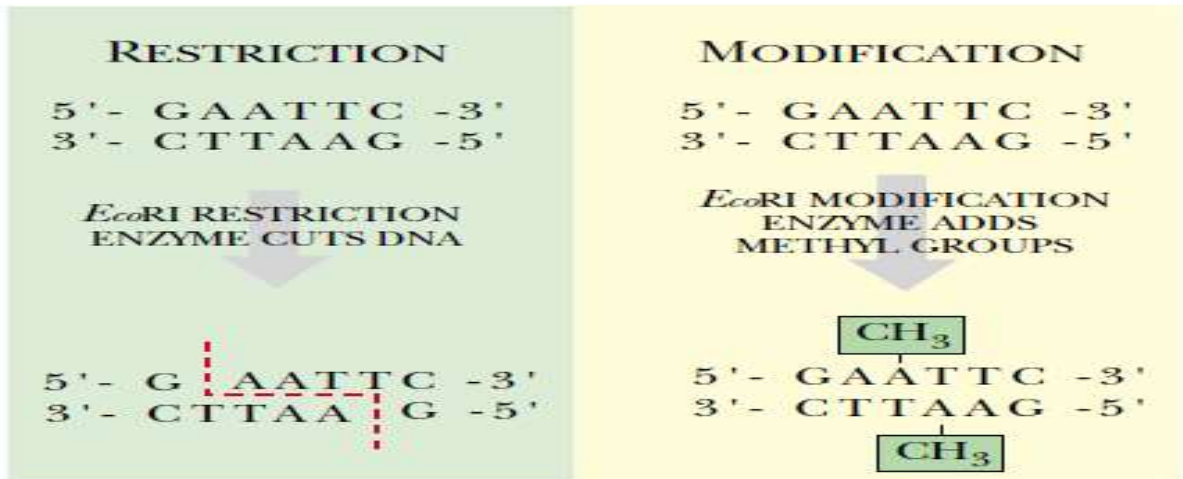


- Coupure cohésive (collant) ex : **EcoR I** : N'est pas dans l'axe. \longrightarrow Extrémités cohésives



5. Systèmes de restriction et de méthylation

Ces enzymes participent à un mécanisme de défense des bactéries vis-à-vis des virus "système de restriction méthylation". Les endonucléases de restriction coupent l'ADN étranger à des sites spécifiques. Et afin de protéger l'ADN bactérien de l'hydrolyse par les enzymes, une méthylase, codé par le gène de méthylation, va méthyler l'ADN au niveau des sites de coupure pour qu'ils ne soient plus reconnus par l'enzyme de restriction.



6. Isoschizomères et famille d'enzymes

En général, des enzymes de restriction différentes reconnaissent des séquences différentes. Cependant, plusieurs enzymes isolées d'organismes différents reconnaissent des séquences identiques. On les appelle Isoschizomères. Ceux-ci ne coupent pas toujours la séquence reconnue au même endroit.

Exemples:

Asp 718

5' g gatcc 3'

Kpn I

3' cctag g 5'

5' ggatc c 3'

Bam HI

3' c ctagg 5'

5' g gatcc 3'

3' cctag g 5'

Une famille d'enzymes est caractérisée par le fait que tous les membres de cette famille produisent les mêmes extrémités, même si elles ont des séquences de reconnaissance différentes.

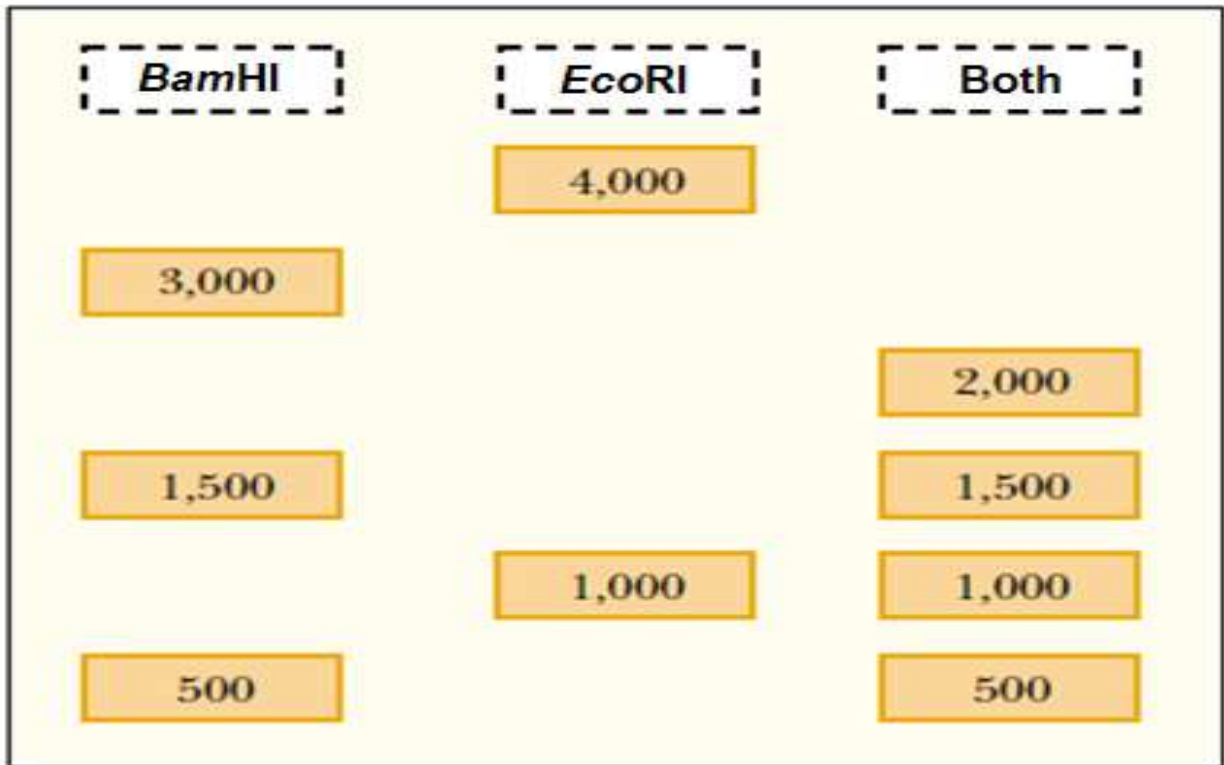
Exemples:

- gatc : *Sau* 3AI, *Bgl* I, *Bam* HI, *Bcl* II, *Xho* II.

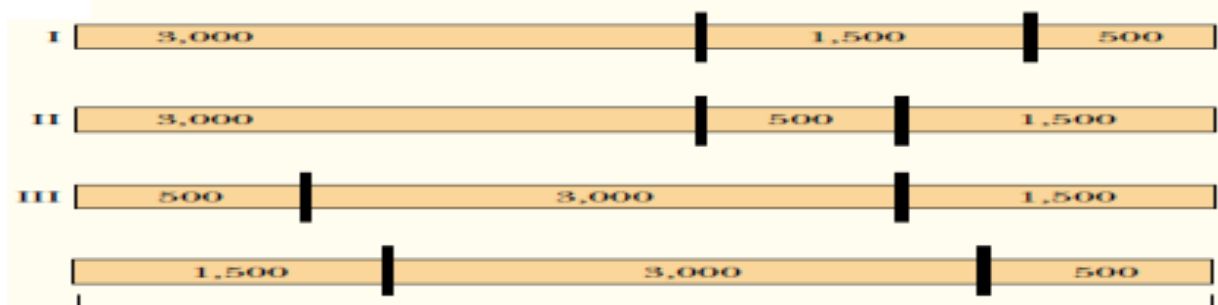
- ctac : *Mae* I, *Spe* I, *Nhe* I, *Avr* I, *Xba* I.

7. Utilisation des endonucléases de restriction pour établir la mappe de restriction

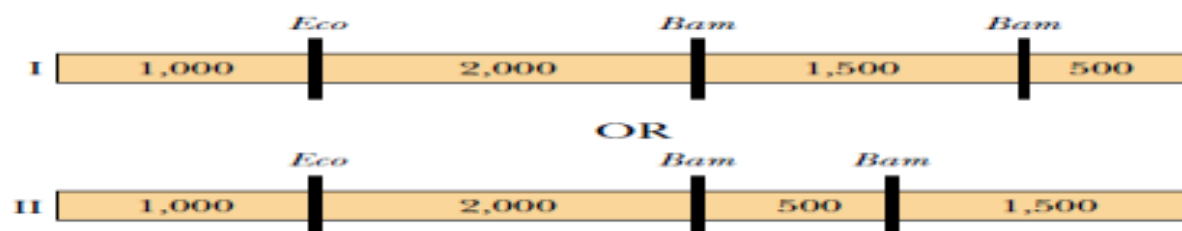
La figure au dessous montre la digestion d'un fragment d'ADN de 5 kb par deux enzymes de restriction (*Bam*HI et *Eco*RI). L'analyse des résultats de la digestion nous permet de déterminer les différentes possibilités de la localisation des sites de restriction pour chaque enzyme



Gel d'agarose des produits de digestion d'un fragment de 5kb par deux enzymes de restriction



Les différentes possibilités de la localisation des sites de coupure par *Bam*HI



Les différentes possibilités de localisation des sites de restriction des deux enzymes (*Eco*RI, et *Bam*HI)

12. Enzymes de modification des acides nucléiques

Les polymérases

- ADN polymérase I
- Le fragment de Klenow de la Pol I
- La Taq polymérase
- Les ARN polymérases
- La transcriptase réverse
- Phosphatase alcaline
- Kinases

Les ligases

- ADN ligase, T4 DNA ligase

Nucléases

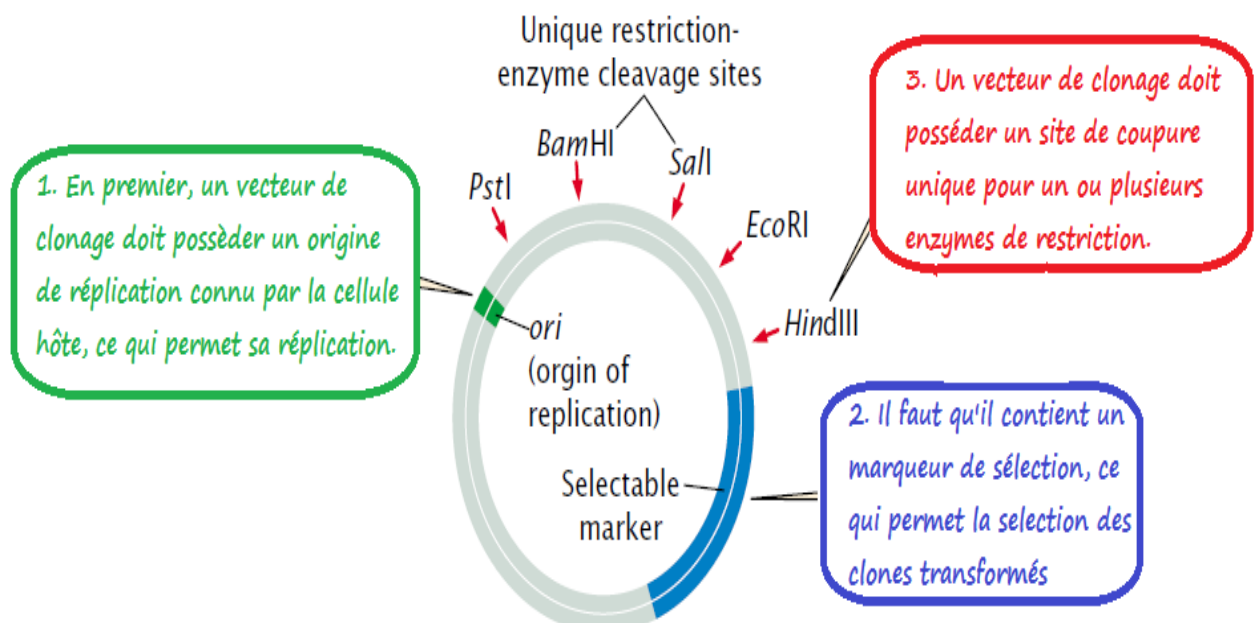
- Rnase A (U et C)
- DNase I (C et T)
- Nucléase S1
- Exonucléase de phage λ
- Exonucléase III
- Ribonucléase T1
- Ribonucléase T2
- Nucléase Bal31

IV. Vecteurs de Clonage

1. Définitions

- **Vecteur** est une séquence d'ADN permettant la propagation, la sélection, la modification d'une séquence d'ADN d'intérêt. En bref l'étude et la manipulation d'une séquence d'ADN isolée.
- **Clonage** - pour amplifier et conserver une séquence d'ADN
 - pour exprimer une séquence d'intérêt
 - pour introduire un gène dans des cellules ou des organismes
 - pour produire une protéine d'intérêt
- **ADN génomique** est le support physique de l'ensemble des gènes de la cellule.
- **ADN complémentaire (ADNc)** est la copie en ADN des ARNm.

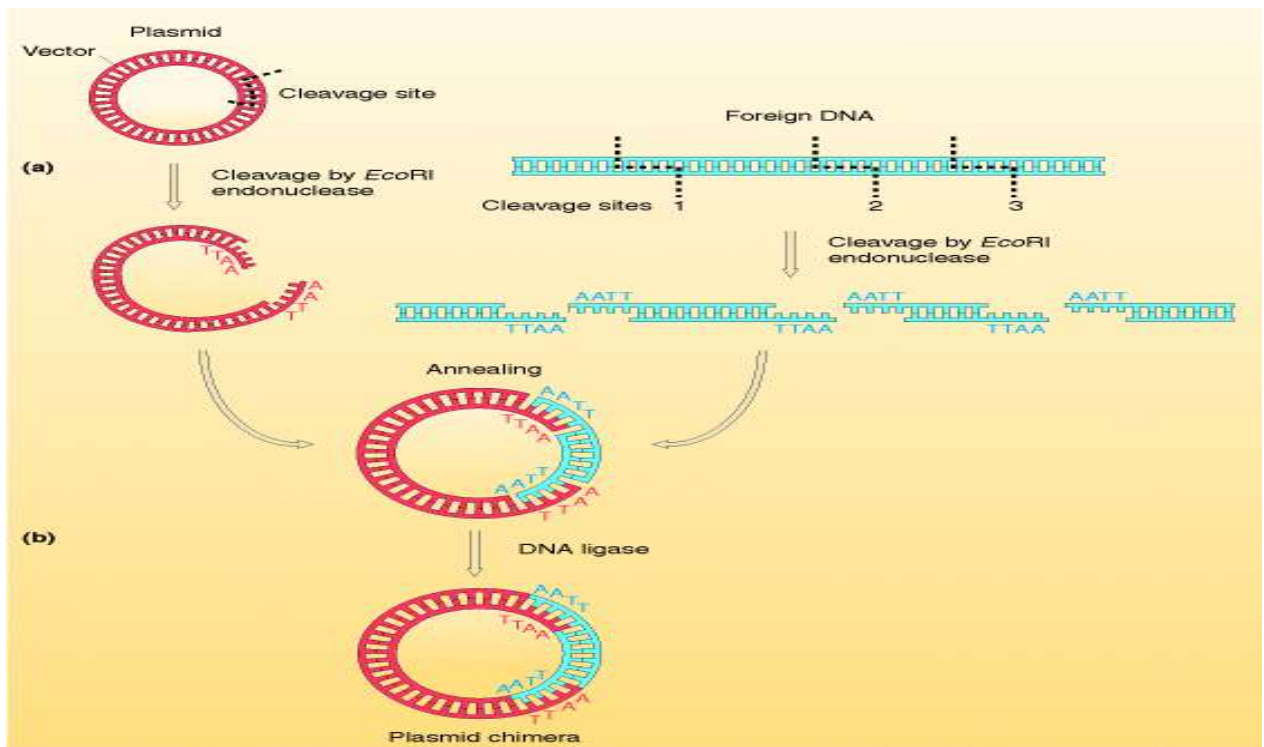
2. Les caractéristiques d'un vecteur de clonage idéal



3. Principes généraux d'utilisation d'un vecteur

1. préparation de l'insert à cloner (fragment d'ADN d'intérêt) et du vecteur
2. Intégration - ligation de l'insert dans le vecteur
3. transformation de l'hôte par le vecteur recombiné
4. sélection des hôtes recombinants

4. Obtention de l'ADN recombinant



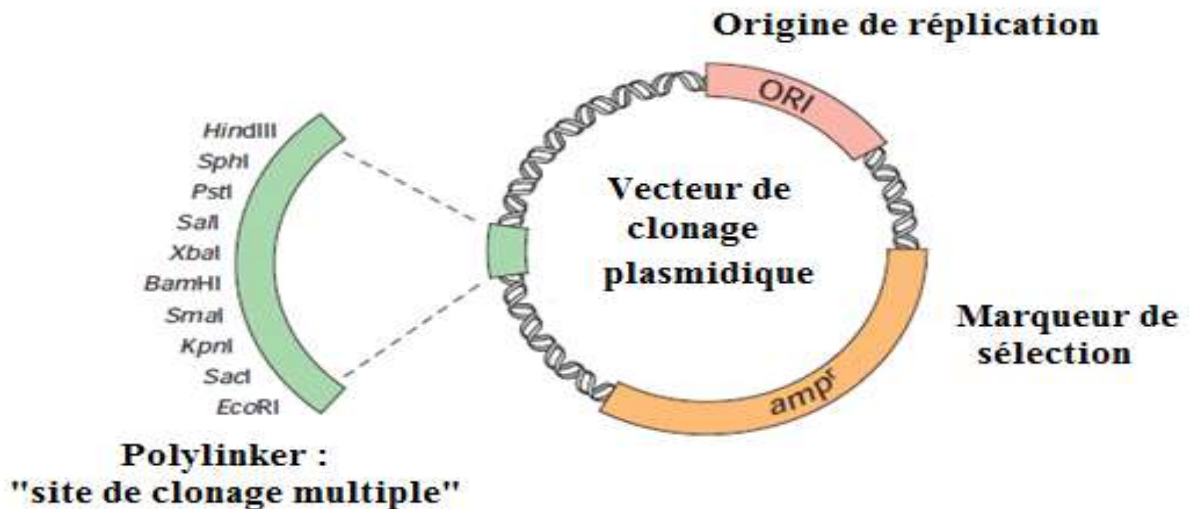
5. Différents vecteurs pour différentes utilisations

Taille maximum approximative du fragment d'ADN qui peut être cloné dans chaque vecteur

VECTEURS	HOTE	INSERT (Kb)
Plasmides	Bactérie	10
Phage lambda	Bactérie	25
Cosmide	Bactérie	45
Phage P1	Bactérie	100
BAC (bacterial artificial chromosome)	Bactérie	300
YAC (yeast artificial chromosome)	Levure	1000

6. Les plasmides comme vecteurs de clonage

- Molécule d'ADN de taille réduite, d'origine bactérienne
- Molécule d'ADN circulaire – taille 2kb à 5kb
- Peut accepter jusqu'à 10kb d'ADN exogène
- Peut être considérée comme un minichromosome capable de réplication autonome
- Confère la résistance à un ou plusieurs antibiotiques = sélection



7. Les classes de plasmides

7.1. Les plasmides de première génération :

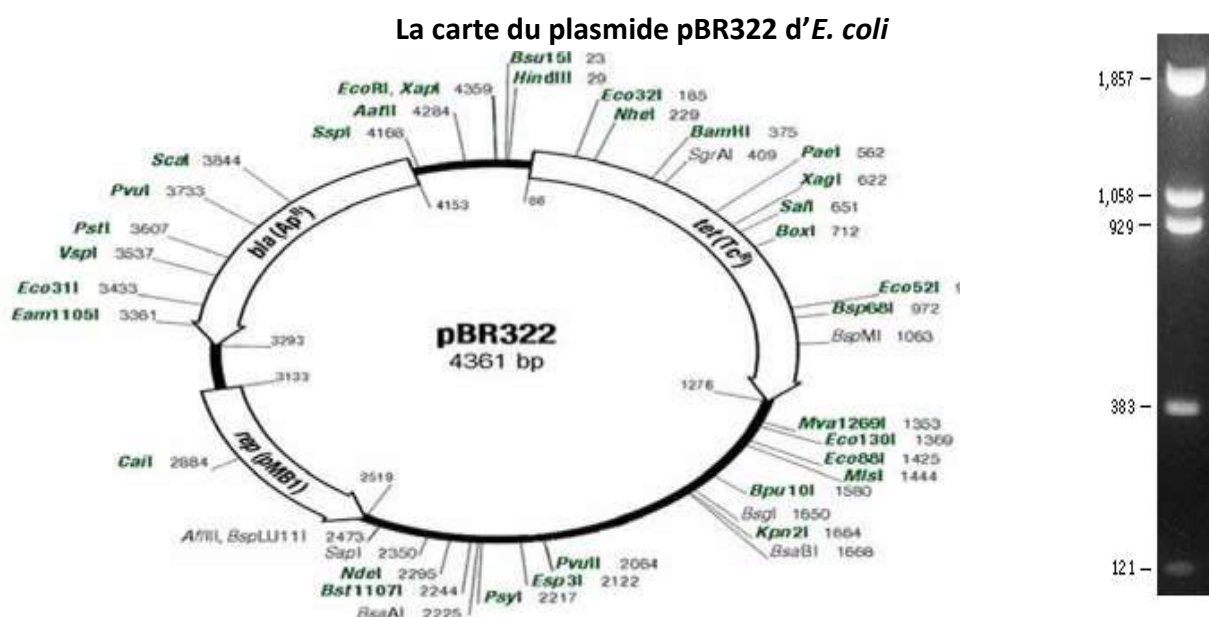
Ce sont les premiers à avoir été utilisés en génie génétique. Ce sont des plasmides à l'état naturel, non modifiés au laboratoire. Il s'agit des plasmides suivants :

- ColE1
- RSF 2124
- pSC 101

7.2. Les plasmides de deuxième génération :

Ce ne sont pas des plasmides naturels mais résultent de plusieurs transformations : plasmides "artificiels".

La série la plus importante de ces plasmides est la série **pBR 312 à pBR 322**. Le plasmide pBR 322 est constitué de 4,4 Kb et possède deux gènes de résistance : un pour la tétracycline (**Tc^R**), l'autre pour l'ampicilline (**Ap^R**). Il possède, en plus, 20 sites uniques pour les endonucléases de restriction dont 11 localisés sur les deux gènes de résistance.



La localisation des sites de coupure par certaines enzymes de restriction est montrée. Ce vecteur possède deux gènes de résistance (Apr: Résistance à l'ampicilline, Tetr: Résistance à la tétracycline).

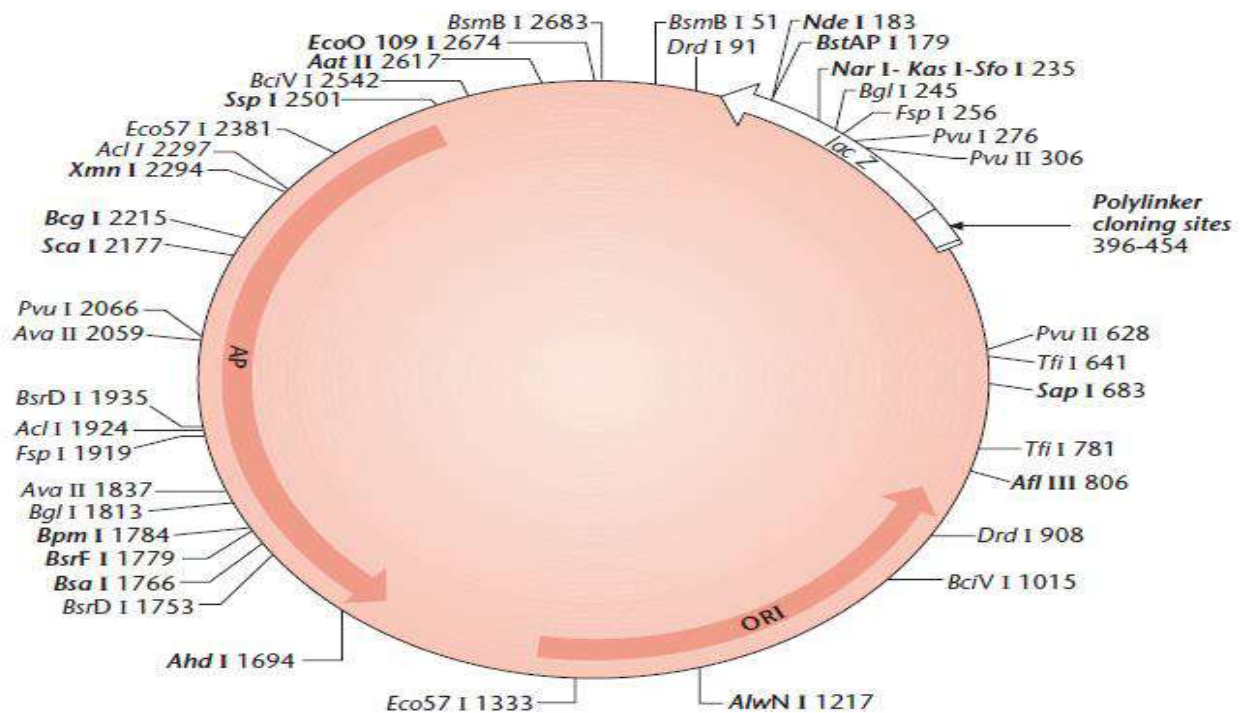
Polylinkers de pUC8 et pUC9

pUC8	ACGAATTCCCGGGGATCCGTCGACCTGCAGCCAAGCTTGGCACTG
Polylinker	
pUC9	ACGCCAAGCTTGGGCTGCAGGTCGACGGATCCCCGGGAATTCAGT
Polylinker	

7.3. Les plasmides de troisième génération

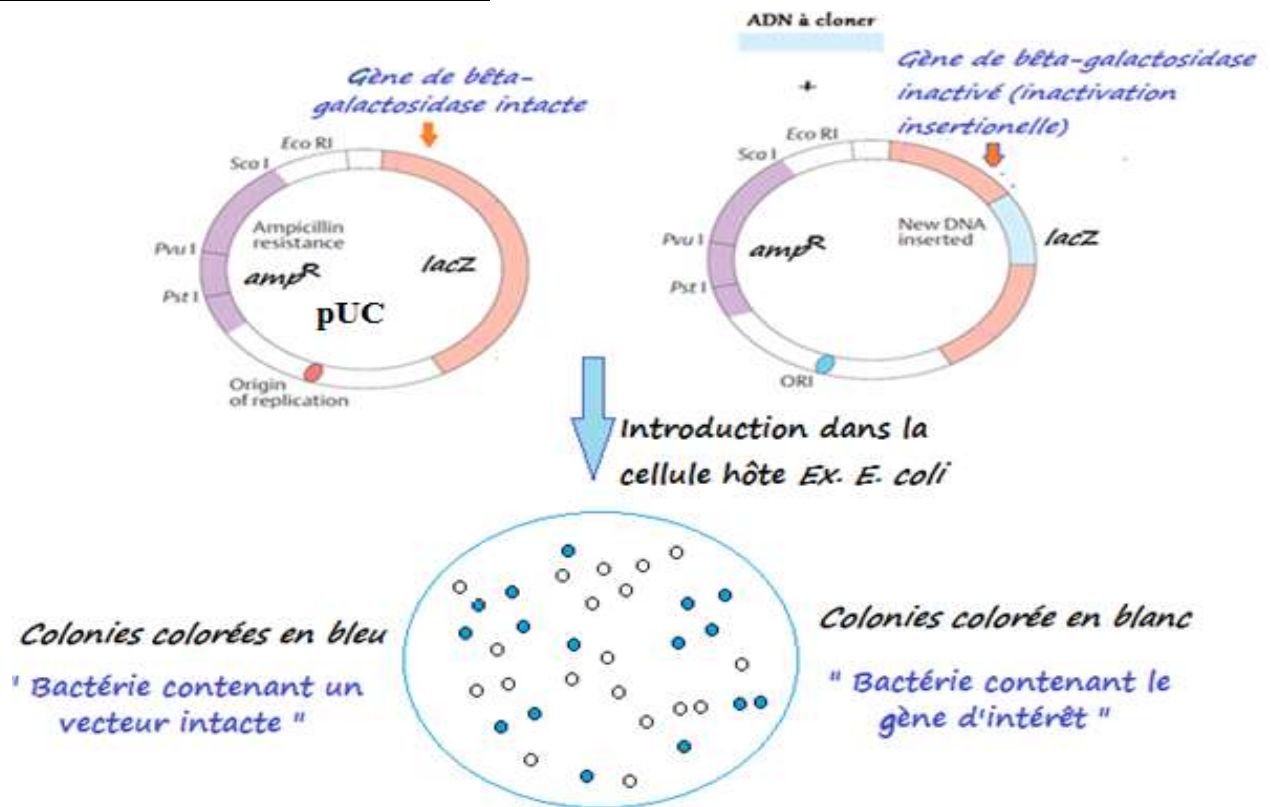
La famille pUC : Ont une taille qui avoisine 2,6 Kb et ayant intégré les gènes de résistance à l'ampicilline (Ap^R) et *lacZ*. Un polylinker identique à celui du phage M13 est associé à *lacZ*. Les différents pUC (de pUC8 à pUC19) ne diffèrent que par le nombre de nucléotides et l'emplacement du polylinker

La carte génétique d'un vecteur de type pUC dérivés de pBR322



Le site de clonage multiple (polylinker) est introduit dans le gène *lacZ*, sans interrompre la fonction du gène (Primose et al. 2002).

Criblage des clones intéressants



Insertion d'un fragment d'ADN (< 10 kb) dans un vecteur de deuxième génération (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.