

II.1. Définition et principe

Le terme clonage est issu du grec « klôn » : rejeton, il s'agit de l'opération qui à partir de cellules isolées, permet d'obtenir une lignée (plusieurs cellules similaires appelées clones) dérivant d'un seul ancêtre. Le principe du clonage est simple, il consiste à insérer le fragment d'ADN d'intérêt après son isolement et sa purification de l'organisme donneur dans une petite molécule d'ADN capable de se répliquer de façon autonome. Ce type de molécule d'ADN appelée « vecteur de clonage » est lui-même introduit dans un autre organisme unicellulaire (organisme récepteur = hôte, généralement une bactérie ou une levure) ou une cellule d'un organisme pluricellulaire en culture. La multiplication de cette cellule va produire plusieurs copies qui lui sont tout à fait identiques : c'est le **clone** recombinant. Toutes les cellules filles constituant le clone, contiennent une copie du gène ayant été introduit dans la cellule mère qui a donné naissance à ce clone. Ce **clonage** a donc permis d'obtenir un bon nombre de copies identiques à la séquence initialement insérée dans la cellule receveuse. Il est également possible de réaliser l'insertion des gènes d'un organisme donneur, mais chacun d'eux est inséré dans une cellule receveuse à part. A la fin on obtiendra plusieurs clones recombinants qui chacun d'eux porte une copie d'un gène inséré. L'ensemble de ces clones est appelé : **banque génomique** (ou librairie : *library*, plus de détail sur la construction d'une banque d'ADN génomique ou d'ADNc est présenté dans le chapitre IV). Après multiplication des cellules receveuses, on peut isoler un clone donné grâce à une propriété donnée : c'est le **criblage**.

II.2. Vecteurs de clonage

Les fragments d'intérêt obtenus sont dépourvus de moyens de se répliquer (origine de réplication) et d'extrémités protégées. Il est donc nécessaire de les inclure dans un vecteur qui assurera leur propagation et leur protection. A cet effet, deux notions existent :

- Vecteur natif (non recombiné) lorsqu'il ne contient que son ADN propre.
- Vecteur recombiné lorsqu'il a intégré un fragment d'ADN étranger, quel qu'il soit.

Les propriétés d'un bon vecteur sont :

- ✓ Réplication active et autonome dans la cellule hôte : Une origine de réplication de type eucaryote ne fonctionnera pas dans une bactérie et réciproquement. Lorsque l'on veut passer d'un type de cellule à l'autre il faut des vecteurs qui possèdent les deux types d'origine de réplication.
- ✓ Présence de marqueurs génétiques (gènes de sélection), comme des gènes de résistance aux antibiotiques ou le gène codant la β -galactosidase.
- ✓ Présence d'un site de polyclonage (*polylinker*) : site de clonage multiple « MCS ou SCM » (un court segment de l'ADN, qui contient de nombreuses sites de restriction) localisés dans les gènes de sélection.

- ✓ Sa présence ne doit pas perturber la physiologie de la cellule hôte et vice versa.
- ✓ Facile à isoler sous forme purifiée.

II.2.1. Plasmide : C'est un ADN bicaténaire (de 3 à 10 kb) circulaires (absence d'extrémités libres) extrachromosomique capables de se répliquer indépendamment du chromosome bactérien (possède une origine de réplication) dans une cellule bactérienne et d'être transférée dans une autre. Il est possible de les purifier en grande quantité car ils se multiplient dans des bactéries hôtes. Leur mise en évidence remonte à 1967 lorsque Watanabe fit la synthèse de certain nombre de fais, principalement la découverte suivante : l'agent causal du transfert de gènes entre des souches d'*Escherichia coli* était un facteur transmissible qui pouvait se disséminer dans la population bactérienne. Les plasmides ne contiennent généralement pas de gènes indispensables à la survie de la cellule. Ils apportent des gènes pouvant présenter un avantages dans des conditions de cultures particulières (antibiotique, métaux lourd, hyper salinité,...) où ils vont apporter un avantage pour la croissance de la cellule les possédant. Des molécules d'ADN exogènes de l'ordre de 4kb peuvent facilement être intégrées dans ces vecteurs plasmidiques. Par contre, des molécules de taille plus importante sont difficilement acceptées. Si la bactérie (hôte) n'est pas compétente (compétente : capable d'absorber l'ADN étranger), on peut l'introduire par méthode chimique (CaCl_2 ...) ou des méthodes physiques tel que l'électroporation.

- ✓ **Le plasmide pBR322 :** Ce plasmide est l'un des 1^{ers} plasmides artificiels (Fig.02) utilisés comme vecteur de clonage. Ce plasmide peut être présent à raison de 10 à 20 copies par cellule dans des conditions standards de culture. Si la culture est faite en présence de chloramphénicol qui bloque la réplication du chromosome bactérien mais pas celle du plasmide, ce dernier peut s'accumuler jusqu'à près de 1000 copies par cellule. L'origine de réplication (rep) provient du plasmide pMB1. Il porte également deux marqueurs génétiques ; des gènes de résistance à des antibiotiques. Le gène bla (Ap^R) code la β -lactamase qui dégrade l'ampicilline. Une bactérie portant pBR322 pourra donc pousser en présence de cet antibiotique : elle est devenue [AmpR] alors qu'une bactérie sans plasmide est tuée par l'ampicilline [AmpS]. Le gène tet (tc^R) confère la résistance à la tétracycline. Le plasmide pBR322 possède aussi plusieurs sites de restriction ; des sites uniques de restriction tels que les sites *EcoRI*, *EcoRV*, *BamHI*, *PstI*, *PvuI*, *Hind III*, parmi lesquels *BamHI* et *Pst I* sont respectivement localisés au niveau des gènes de résistance à la tétracycline et à l'ampicilline. L'introduction d'un ADN étranger dans l'un de ces sites se traduit donc par la perte de la résistance à l'antibiotique correspondant.
- ✓ **Les plasmides Puc18/19 :** Ce sont des plasmides de la 3^{eme} génération, de la famille des pUC (*plasmid of University of California*). Les plasmides pUC18 et pUC19 (Fig.03) sont des ADN circulaires double brin, ils ont été construits par recombinaison d'un fragment de pBR322 (2250 pb) et d'un fragment du phage M13 double brin (446 pb). Ces plasmides contiennent le gène de résistance à l'ampicilline (Amp^R) et l'origine de réplication issu de pBR322. L'ADN issu

de M13 contient le promoteur du gène (fragment de l'opéron lactose) qui code pour l'extrémité NH₂ de la β-galactosidase. Au sein de ce gène est placée une région renfermant un *polylinker*. L'interruption de ce gène (par insertion de l'ADN étranger) entraîne une perte de l'activité de la protéine.

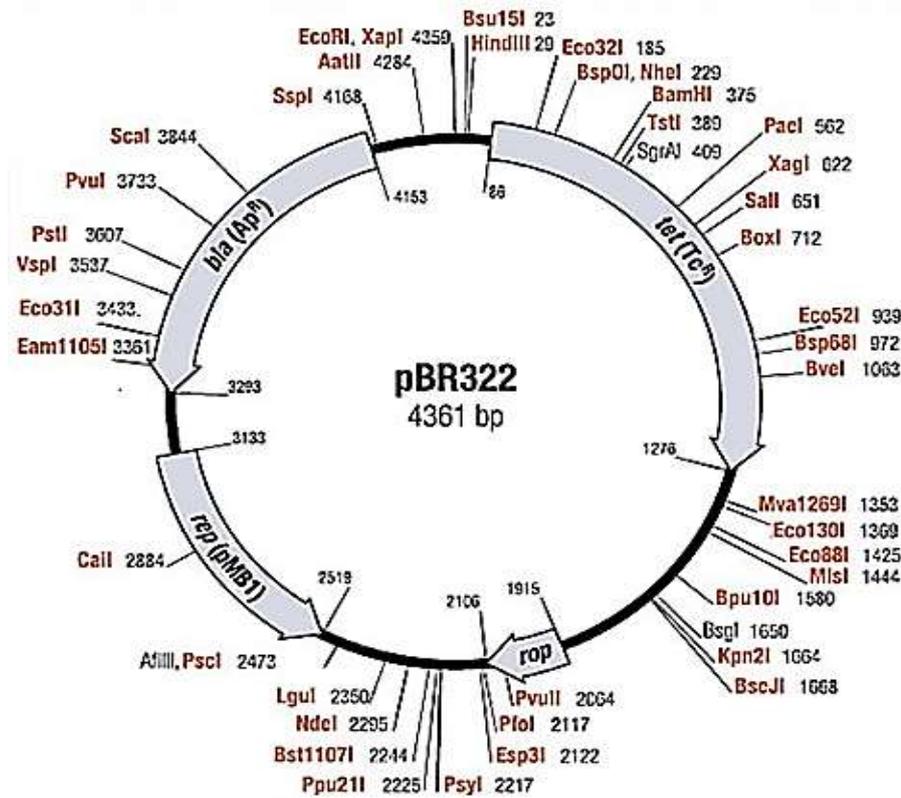


Fig.02 : Carte de restriction du plasmide pBR322.

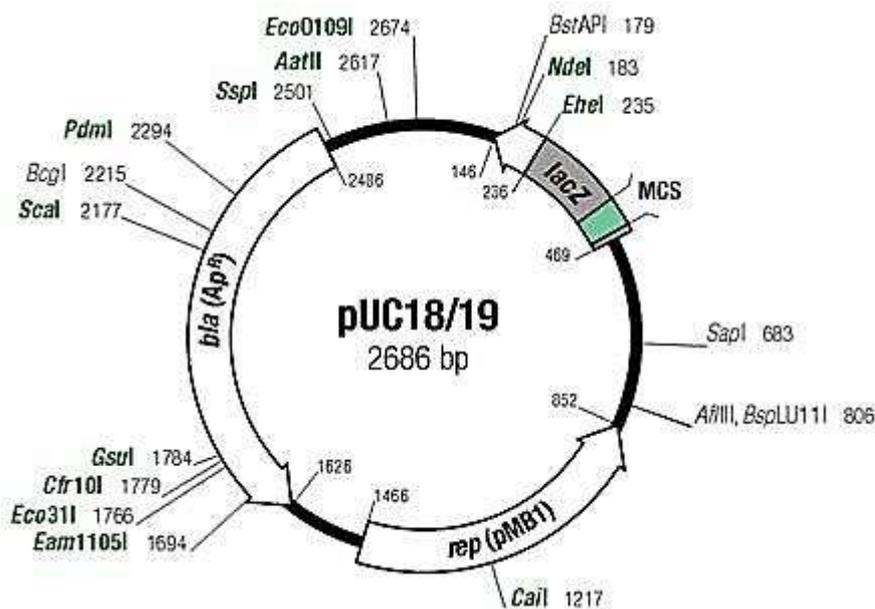


Fig.03 : Carte des plasmides Puc18/19

II.2.2. Phagemides : Les phagemides ou les phasmides sont des molécules hybrides entre un plasmide et un phage. Ce sont des molécules d'ADN bicaténares, circulaires, qui peuvent être obtenues sous forme monocaténaire dans certaines conditions. Ils possèdent une origine de réplication plasmidique, au moins un gène de résistance à un antibiotique, un site de polyclonage et une séquence provenant du phage M13 contenant l'origine de réplication qui permet d'obtenir la forme monocaténaire. En générale, des promoteurs reconnus par des ARN polymérases ont été introduits en amont et/ou en aval du site de polyclonage, afin de pouvoir produire des ARN par transcription *in vitro*. Des fragments d'ADN de l'ordre de 4 kb peuvent être introduits dans ces vecteurs, mais il est possible d'intégrer des inserts de l'ordre de 10 kb.

✓ **Les vecteurs pBluescript:** la figure 04 présente la carte de restriction de pBluescript II SK/KS (+/-), ce sont des phagemides dérivant de pUC19. Ces phagemides contiennent un gène de résistance à l'ampicilline (Ap^R). La séquence de *polylinker* est contenue dans un gène LacZ contrôlé conçu pour fournir une couleur bleue lorsqu'elle est exprimée dans des bactéries. Ils possèdent aussi une séquence *polylinker* avec 21 sites de reconnaissance uniques pour les enzymes de restriction sans les séquences des promoteurs T7 et T3. Ces vecteurs sont disponibles avec deux orientations *polylinker* KS et SK, selon cette convention :

(1) Orientation KS: Kpn I est le plus proche du promoteur lacZ et Sac I est le plus éloigné du promoteur lacZ

(2) Orientation SK: Sac I est le plus proche du promoteur lacZ et Kpn I est le plus éloigné.

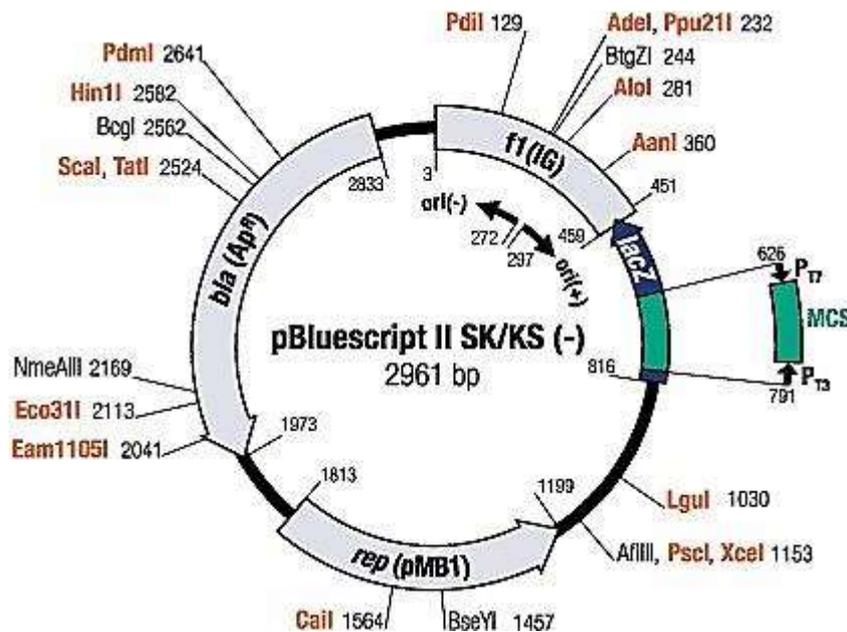


Fig.04 : Carte de restriction du phagemide pBluescript.

II.2.3. Phages : Un bactériophage (ou phage) est un virus bactérien. L'ADN phagique est enveloppé dans une coque protéique. Lors de l'infection d'une bactérie, la particule phagique se fixe à l'extérieur de la bactérie et libère alors sa molécule d'ADN à l'intérieur de la cellule bactérienne. La machinerie enzymatique de la bactérie est alors utilisée pour transcrire et traduire les protéines codées par l'ADN du phage, notamment les polypeptides constituant sa coque protéique. L'ADN phagique va également se répliquer un grand nombre de fois. Les molécules d'ADN nouvellement synthétisées seront encapsidées dans les nouvelles coques phagiques à l'intérieur de la cellule bactérienne. On parle alors de virions. Puis la cellule bactérienne se lyse et libère ainsi des milliers de particules phagiques identiques au phage qui a infecté la bactérie (cycle lytique). Ces nouveaux phages vont à leur tour infecter des bactéries avoisinantes. L'avantage d'un vecteur phagique sur un vecteur plasmidique est que son ADN accepte de plus grands fragments d'ADN exogène (20 kb en moyenne contre 4 kb pour un plasmide). Cependant, à cause de la grande taille de l'ADN, il est difficile de le manipuler *in vitro*. Dans la pratique, les phages sont utilisés comme vecteurs de construction de banques d'ADNc ou génomiques. Leur multiplication est rapide et le nombre de copies / cellule est élevé. Le rendement d'une infection phagique est supérieur à ce qui est obtenu lors de la transformation par un plasmide. Lorsque les clones d'intérêt ont été repérés, l'ADN phagique est purifié et les inserts extraits sont intégrés dans un vecteur plasmidique de manipulation plus aisée. On parle alors de sous-clonage. Plus simplement, ce sous-clonage peut également être effectué par PCR.

- ✓ **Phage λ :** C'est un virus qui infecte la bactérie *Escherichia coli* (bactériophage), et c'est le phage le plus utilisé pour le clonage moléculaire. Le phage λ peut se multiplier selon deux modes : selon le **cycle lytique** et le **cycle lysogénique**. Lorsque ce phage est utilisé comme vecteur de clonage, c'est le cycle lytique qui nous intéresse, puisqu'il entraîne la production d'un grand nombre de nouveaux phages qui vont à leur tour infecter d'autres bactéries et ainsi de suite. Son ADN bicaténaire linéaire, de 48502 pb, présente des extrémités cohésives (séquence *cos*) lui permettant de se circulariser dans la bactérie infectée. Les gènes de structure du phage sont regroupés aux extrémités de son génome. Les séquences médianes non indispensables pour la réplication du phage peuvent être éliminées et remplacées par des sites de restriction facilitant le clonage. Les extrémités 5' des 2 brins sont plus longues (donc simple brins) d'environ 12 nucléotides et sont complémentaires l'une avec l'autre. Ces extrémités cohésives (L et R) qui sont nécessaires pour l'intégration de l'ADN du phage dans le génome de son hôte (donc nécessaires à la réplication et à l'encapsidation). L'utilisation de phage λ non modifié (λ gt10) présente des problèmes, parmi lesquels l'existence de plusieurs sites de restriction pour des enzymes identiques, d'où la nécessité de construire des phages λ modifiés possédant un seul site de restriction pour que l'insertion de l'ADN étranger se fasse à un endroit précis du génome du vecteur. Le vecteur « λ gt11 » qu'est le phage λ modifié le plus utilisé (puisque'il est considéré comme un bon vecteur de clonage) possède un génome de 43,7 Kb dans lequel a été inséré dans la partie centrale un marqueur génétique

(gène de sélection) à la place du gène répresseur du cycle lytique. Le gène placé est une partie de l'opéron lactose (codant la β -galactosidase) et possédant un seul site de restriction *EcoRI* situé dans l'opéron lactose. La carte génétique du phage λ est présentée dans la figure 5.

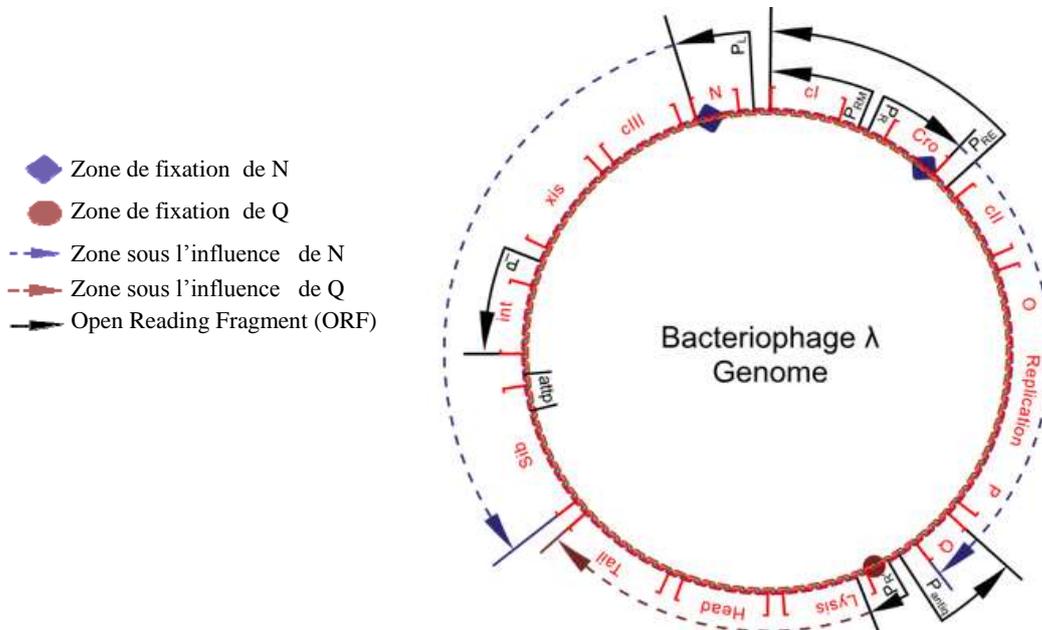


Fig.05 : Carte génomique circulaire du phage λ . La protéine N : Anti-terminateur, se lie à l'ARN polymérase, permet à cette dernière d'ignorer les sites de terminaison. La protéine Q se lie à l'ARN et aux cofacteurs de l'ARN polymérase, les sites de terminaison sont ignorés, seule la séquence de terminaison de Q est efficace.

- ✓ **M13 :** C'est un bactériophage de 6.4 kb, ADN simple brin, qui contient une dizaine de gènes. Il infecte seulement les bactéries qui expriment le pilus sexuel codé par le facteur F. Il pénètre par le pilus, la protéine de structure *g8p* est retirée et l'ADN est transféré dans le corps principal de la bactérie. Ce brin d'ADN (appelé brin +) est alors converti en double brin circulaire appelé ADN RF (forme répliquative). Cette conversion est due à des enzymes bactériennes, une ARN polymérase initie la répliquative qui est effectuée par une ADN polymérase. Ce virus ne provoque pas la lyse de la cellule bactérienne, toutefois une diminution de la croissance cellulaire chez les cellules infectées est observée. Les vecteurs dérivés du phage M13 ont été développés pour leur capacité de produire de l'ADN monocaténaire qui constituait la matrice utilisés pour les réactions de séquençage par la méthode de Sanger. L'ADN simple brin était également le substrat pour la technique de mutagenèse dirigée, permettant de changer à volonté la séquence d'acides aminés du produit d'un gène cloné.

II.2.4. Cosmide : Un cosmide est un plasmide (avec son origine de répliquative, ses gènes de sélection et son site de polyclonage) dans lequel on a inséré les séquences *cos* nécessaires à l'encapsidation dans des particules de bactériophage λ . Les fragments d'ADN exogènes introduits dans les cosmides doivent avoir une taille comprise entre 35 et 45 kb. Ainsi, après ligature d'un cosmide préalablement

linéarisé et d'un fragment d'ADN étranger de taille convenable, l'ADN recombinant peut être encapsidé comme celui d'un phage, puis utilisé pour infecter des bactéries. Une fois dans la bactérie, le cosmide se réplique comme un plasmide, car il ne possède pas les gènes du phage nécessaire, à la production de nouvelles particules phagiques. Les cosmides sont des vecteurs utilisés pour construire des banques d'ADN génomique.

II.2.5. Chromosomes artificiels : Plusieurs chromosomes peuvent être utilisés comme vecteur. Ils se caractérisent par leur capacité d'insertion de grands fragments d'ADN.

✓ **Yeast Artificial Chromosome (YAC):** Ils ont été décrits pour la première fois par Murray et Szostak en 1983, utilisé pour cloner de très gros fragments d'ADN. En effet, les recherches visant à caractériser la structure de génomes entiers nécessitent l'utilisation de vecteurs de clonage permettant l'insertion de grands fragments d'ADN. Cela permet de manipuler un nombre restreint de vecteurs recombinants tout en couvrant le génome entier. Les YACs sont des vecteurs construits à partir de séquences d'ADN chromosomiques de la levure *Saccharomyces cerevisiae* ligaturé dans un plasmide bactérien, dont les quatre principaux types de plasmides de levure sont : YIp, YRp, YCp et YEp. Les chromosomes artificiels de levure (Fig.06) possèdent :

- Une origine de réplication (ARS : *Autonomous Replicating Sequence*).
- Un site centromérique (CEN).
- deux séquences télomériques (TEL)
- Deux gènes de sélection codant pour des enzymes permettant de sélectionner les levures ayant incorporé un vecteur viable (TRP1 et URA3).
- Un site de clonage unique, situé dans le gène SUP4.
- Un gène de résistance à l'ampicilline (Amp^R).

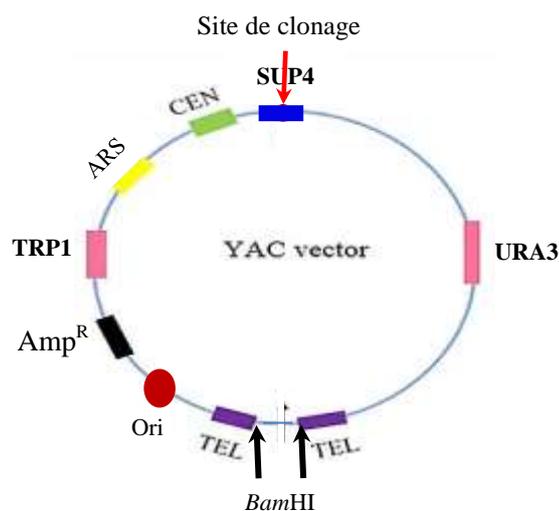


Fig. 06: Carte génétique de YAC. Le gène URA3 Impliqué dans la biosynthèse d'uracile. Le gène TRP1 : Impliqué dans la biosynthèse de tryptophane. Le gène SUP4 : code pour l'ARN-tyrosine.

- ✓ **Bacterial Artificial Chromosome (BAC)** : sont des dérivés du facteur F (plasmide de fertilité). Leur taille est d'environ 100 kb, et ils se maintiennent chez *E. coli* comme simple copie. Les BACs se comportent comme des plasmides de très grande taille. Ils sont de plus en plus utilisés par rapport aux YACs, le clonage de l'ADN y est plus aisé et efficace, et les clones sont plus stables dans *E. coli* dans la levure. Cependant, la taille des fragments intégrés ne dépasse pas 300kb.
- ✓ **P1-derived Artificial Chromosome (PAC)** : sont des dérivés du phage P1. C'est un type de vecteur utilisé pour cloner des fragments d'ADN au sein de la bactérie *Escherichia coli*.

Le tableau ci-dessous présente la capacité d'insertion de chaque vecteur et leurs applications majeures

Tableau 01 : Différents types de vecteurs de clonage et leurs applications majeures

Vecteurs	Forme du vecteur	Hôte	Taille de l'insert (kb)	Usage majeur
plasmide	AND db circulaire	<i>E. coli</i>	0.1-10	Banque d'ADNc, sous clonage
Bacteriophage λ	Virus, AND db linéaire	<i>E. coli</i>	10-20	Banque génomique et d'ADNc
cosmide	AND db circulaire	<i>E. coli</i>	35-45	Banque génomique
Phagemides	Hybride plasmide/phage	<i>E. coli</i>	4-10	Banque d'ADNc, sous clonage
YAC	Levure, chromosome artificiel	levure	200-1000	Banque génomique
BAC	Chromosome artificiel de bactérie	<i>E. coli</i>	100-500	Banque génomique
PAC	Virus, AND circulaire	<i>E. coli</i>	80	Banque génomique

II.3. Hôte

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

II.3.1. Hôtes bactériens

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

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

II.3.2. Hôtes eucaryotes

Les hôtes eucaryotes peuvent être des cellules animales en culture, des levures ou des plantes. L'avantage de ces systèmes réside dans leur niveau d'expression élevé, l'absence de corps d'inclusion et l'existence de modifications post-traductionnelles. Cependant, leur manipulation est complexe et coûteuse et ils ne sont utilisés que :

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

II.4. Transformation génétique des bactéries et sélection

Le plasmide pUC19 est à utilisation fréquente en génie génétique, car il montre un système de sélection des bactéries ayant été transformées par les plasmides recombinants. Un système enzymatique appartenant à l'opéron lactose peut être utilisé à la place d'un second antibiotique. L'insertion du fragment d'ADN dans le plasmide pUC19 aboutit à l'inactivation du gène qui code pour la β -galactosidase. Pour vérifier la présence ou l'absence de l'activité enzymatique, on utilise le galactoside X-gal, dont la couleur passe du bleue au blanchâtre quand 'il est clivé par la β -galactosidase. Pour pouvoir métaboliser le X-gal, la cellule doit être exposée à un inducteur, qu'est l'IPTG (isopropylthio-b-D-galactoside). En culture et en présence d'IPTG et X-gal, les bactéries résistantes à l'ampicilline et transformées par les plasmides recombinants se présentent sous forme de colonies blanchâtres car elles ont perdu la capacité de clivage de l'équivalent coloré du lactose (le X-gal) par la β -galactosidase. Par contre, les bactéries résistantes à l'ampicilline et non transformées par les plasmides recombinants se présentent sous forme de colonies bleues. La figure 7 nous montre les

trois principaux cas que l'on peut obtenir et les résultats attendus sur boîtes, où on peut sélectionner les bactéries transformées par les plasmides recombinants visuellement.

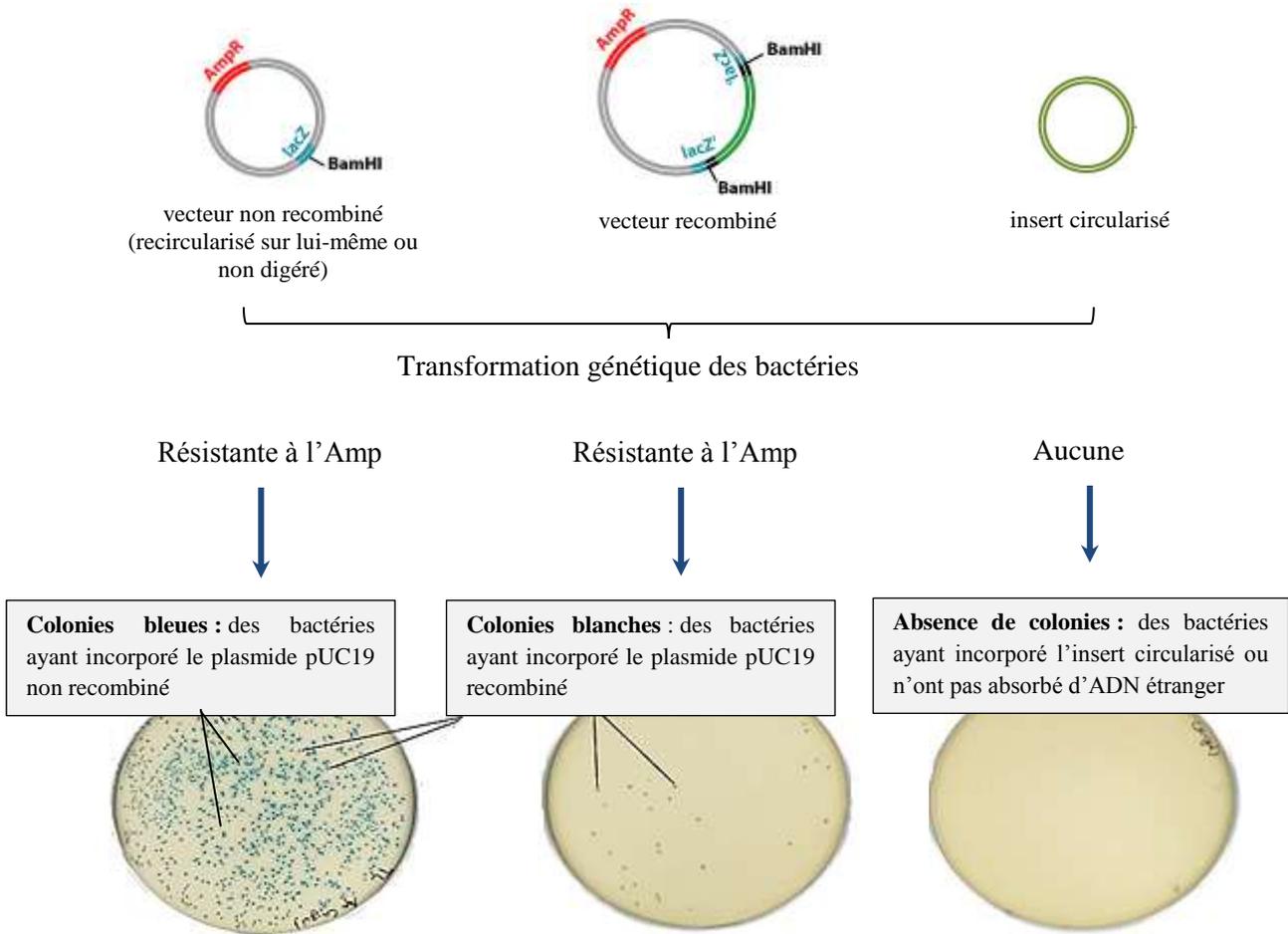


Fig.07 : Schéma illustratif de la transformation génétique des bactéries (*E. coli*) par le plasmide pUC19.