

LE CLONAGE MOLECULAIRE : Le gène isolé d'un organisme est isolé, purifié et introduit au sein d'un organisme unicellulaire d'une autre espèce réceptrice (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).

Dans certains cas, on réalise des banques génomiques de gènes qui s'expriment (c'est-à-dire qui produisent des protéines) dans des cellules différenciées. Les ARNm sont utilisés comme matrice pour fabriquer de l'ADN. Cet ADN qui provient des ARNm est appelé **ADNc**. A partir de cet ADNc, il est également de réaliser une banque génomique.

Après multiplication des cellules receveuses, on peut isolé un clone donné grâce à une propriété donnée : c'est le **criblage**.

Le choix du vecteur est une étape cruciale au moment de l'établissement de la banque, car en général, en travaillant sur les gènes d'organismes eucaryotes, on se trouve confronter au problème de la taille des gènes qui risquent de ne pas être acceptés par le vecteur.

La banque génomique doit contenir au moins une copie de l'ensemble du génome. Mais il est plus intéressant d'avoir des copies en nombre plus grand. La formule statistique qui donne le nombre de clones nécessaires est :

$$N = \frac{\log(1 - p)}{\log\left(1 - \frac{1}{n}\right)}$$

Avec :

p : probabilité de présence d'une séquence donnée

n : rapport entre la longueur du génome et la longueur moyenne des fragments insérés

Ou encore :

$$N = \frac{\ln(1 - p)}{\ln\left(1 - \frac{F}{G}\right)}$$

Avec :

P : Probabilité de présence d'une séquence au sein de la banque

F : Taille des fragments (inserts)

G : Taille du génome

Le tableau suivant résume quelques exemples pour le cas du génome humain qui est constitué de 3×10^9 pb :

Tableau N°6 : Nombre de clones (milliers) d'une banque génomique

<i>p</i>	Longueur de l'insert (Kb)		
	20	200	2000
0,99	640	65	6,5
0,95	415	40	4,2

5-1/ Construction d'une banque génomique : L'ADN génomique est dans un premier temps aléatoirement coupé par des enzymes de restriction. Les fragments d'intérêt issus de cette coupure possèdent l'équipement nécessaire pour leur propre réplication au sein de la cellule hôte.

Les fragments sont insérés individuellement dans un ADN vectoriel par liaisons covalentes, grâce à l'action d'une ligase, qui sera l'ADN recombiné à introduire dans une cellule hôte (par exemple *E. coli*) en présence de CaCl_2 . Les cellules ayant intégré l'insert sont ensuite sélectionnées par culture sur milieu spécifique (par exemple gélose à l'ampicilline). Les cellules n'ayant pas intégré le plasmide ne peuvent croître sur ce milieu. Seuls les cellules avec le vecteur donneront une pousse. Ce sont ces dernières qui constitueront la banque génomique.

On peut également utiliser les phages pour construire une banque génomique. Le principe est le même mais le criblage se fera en fonction des lysats obtenus après culture sur un milieu solide. Les bactéries qui échappent à l'infection par le phage donneront des colonies ordinaires et seront donc exclues car ne contiennent pas le phage hybridé avec le gène d'intérêt. La figure suivant donne le principe général de l'obtention d'une banque génomique par l'utilisation de phages comme vecteurs :

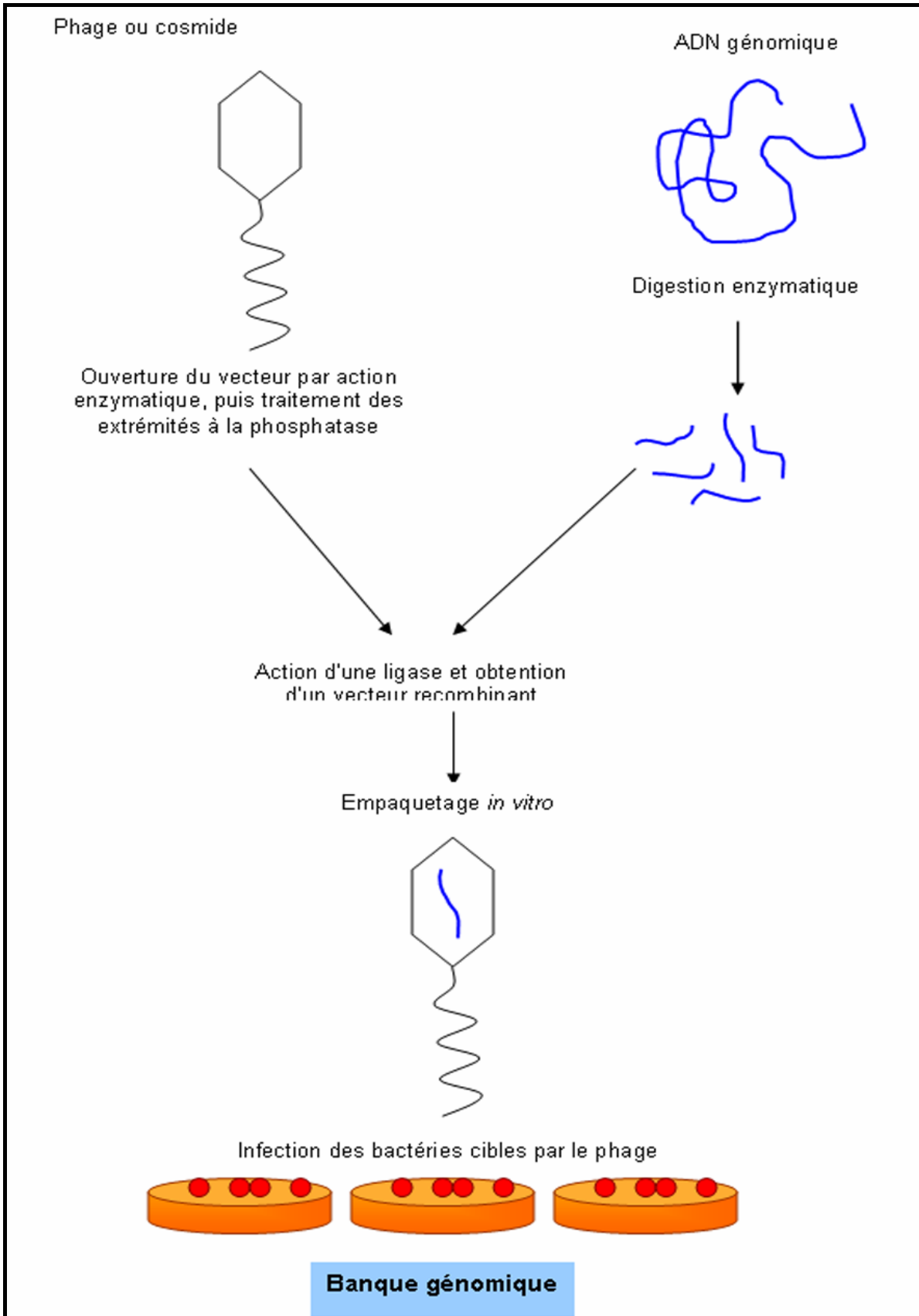


Figure N°19: Principe de construction d'une banque génomique

5-2/ Construction d'une banque ADNc : L'ADNc est un ADN complémentaire obtenu à partir de l'ARN messager suite à l'action d'une transcriptase réverse d'origine virale. Dans ce cas les banques obtenues sont spécifiques du type cellulaire d'où ont été isolés les ARNm : Une banque de pancréas contient des ADNc de l'insuline mais pas les ADNc de la myéline qu'on retrouve dans une banque d'ADNc de cerveau par exemple.

5-3/ Le criblage : Il s'agit de rechercher les clones ayant intégré le recombinant d'intérêt. Plusieurs techniques permettent cette recherche.

5-3-1/ Criblage par autoradiographie : L'ADNc étant synthétisé à partir d'un ARNm en présence de trinucleotides phosphates marqués, sera utilisé comme sonde afin d'hybrider les bactéries de la banque. Les images de l'autoradiographie montrent les bactéries qui possèdent le plasmide recombinant d'intérêt.

5-3-2/ Criblage par oligonucléotide de synthèse : Le principe de cette technique est d'utiliser une sonde oligonucléotidique préalablement synthétisée à partir d'une partie d'une séquence protéique de 20 acides aminés au plus. Grâce au code génétique, on peut établir les séquences possibles (*dégénérescence du code génétique*) relatives à ce "morceau" de protéine : c'est le microséquençage.

Dans une seconde étape, les séquences oligonucléotidiques possibles sont traitées par des programmes informatiques pour déterminer quelles sont les sondes susceptibles de donner un meilleur résultat.