

Chapitra 2: les outils de la génomique

La génomique englobe un ensemble de concepts et de techniques qui permettent d'étudier et de modifier les gènes et leurs fonctions, ou cœur même du fonctionnement des cellules et des organismes vivants.

1. Isoler l'ADN

Isoler l'ADN ou l'extraction d'ADN, c'est une technique fondamentale de toute l'analyse génomique qui permet de isoler et purifier l'acide désoxyribonucléique (ADN) des échantillons biologiques. Il y a plusieurs protocoles de l'extraction mais le processus suit toujours les mêmes étapes de base

A. Lyse des cellules

Pour libérer l'ADN des organes cellulaires (les chromosomes dans les noyaux, ADN de mitochondrie et ADN des chloroplastes), il faut d'abord casser la paroi et les membranes des cellules et des organes cellulaires; ceci est souvent réalisé par :

- Une action mécanique : broyage
- Chimique : utilisation de détergents ou de savon qui dissolvent les membranes lipidiques

Élimination des protéines liées avec l'ADN en utilisant des enzymes comme la Protéinase K pour dégrader les protéines. Pour éliminer l'ARN, il faut d'abord ajouter aussi l'ARNase.

B. Précipitation et Séparation

Dans cette étape l'ADN est séparé des autres composants cellulaires. On utilise des solutions salines et des solvants organiques (ex : Phénol, Chloroforme) ou des tampons à haute teneur en sel et en éthanol/isopropanol pour faire précipiter l'ADN hors de la solution.

L'ADN est une macromolécule très chargée négativement. Lorsqu'on ajoute des solutions salines, les ions positifs du sel (par exemple, le Na^+ ou NH_4^+) se lient aux charges négatives de l'ADN qui permettent de neutraliser la charge négative des groupes phosphate du squelette de l'ADN.

L'ajout de solvants organiques (Phénol, Chloroforme) permet de séparer l'ADN des protéines et lipides contaminants, souvent avant la précipitation à l'alcool. Le Phénol est un solvant très efficace pour dénaturer et dissoudre les protéines et les lipides.

Chapitre 2: les outils de la génomique

Le Chloroforme est ajouté pour séparer les phases et améliorer la densité du mélange, forçant les protéines dénaturées à rester dans la phase organique inférieure, tandis que l'ADN purifié reste dans la phase aqueuse supérieure. On obtient une phase aqueuse plus pure (contenant l'ADN) qui est ensuite récupérée pour la précipitation à l'éthanol.

C. Purification et Solubilisation

L'ADN est lavé avec de l'éthanol à haute concentration (souvent 70% à 90% et froid) pour éliminer les sels et les contaminants résiduels, puis il est solubilisé dans un tampon stérile (comme le tampon Tris-EDTA, ou l'eau) pour être conservé et utilisé dans des analyses futures (PCR, l'électrophorèse, ou le séquençage...)

Tampon Tris-EDTA (TE) :

- Le Tris (Tris (hydroxy méthyl) amino méthane) maintient le pH stable, ce qui est crucial pour la stabilité chimique de l'ADN.
- L'EDTA (acide éthylène diamine tétra acétique) est protégé l'ADN de la dégradation enzymatique.

2. Couper-coller l'ADN

a) Couper l'ADN

Les molécules d'ADN constituant les chromosomes sont beaucoup trop longs, complexes et fragiles pour être analysées facilement. Au laboratoire, pour étudier l'ADN il faut digérer le par des enzymes de restriction.

Les enzymes de restriction sont des protéines d'origine bactérienne. Capable de couper les deux brins d'ADN (l'ADN bicaténaire) à des sites spécifiques. Ce sont des endonucléases reconnaît et coupe une séquence nucléotidique donnée. Leur nom provient du nom de genre et d'espèce de la bactérie de laquelle elles ont été isolées.

Il existe trois types d'enzymes de restriction :

- **Les types I et III** sont des protéines complexes coupant l'ADN double brin en dehors de leur site de reconnaissance.
- **le type II** sont reconnaissent une séquence spécifique de 4, 6 ou 8 pb et coupent à l'intérieur de cette séquence, appelée site de restriction.

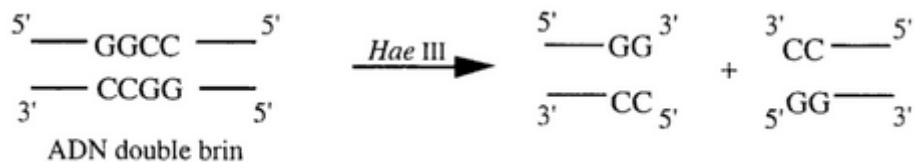
De très nombreuses enzymes de restriction sont disponibles dans le commerce. Elles sont généralement fournies avec leur tampon de réaction et la température optimale de

Chapitre 2: les outils de la génomique

fonctionnement est indiquée sur les catalogues. Il est possible de digérer un même fragment d'ADN par deux enzymes de restriction différentes. Si elles fonctionnent dans des conditions analogues (concentration en sels, pH, température...), les deux digestions se déroulent simultanément. Sinon, le fragment d'ADN doit être digéré par les deux enzymes successivement, en changeant entre chaque réaction enzymatique les conditions de l'expérience.

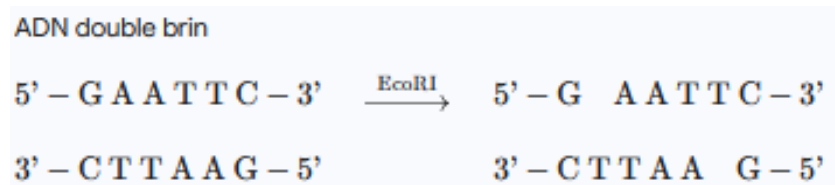
Les enzymes de restriction utilisées coupent dans la séquence d'ADN d'étude selon deux modes :

- **coupure franche** : sont obtenues des extrémités franches par coupure au même endroit sur les deux brins. Exemple de L'enzyme *Hae* III (isolée de *Haemophilus aegyptius*) coupe l'ADN double brin au niveau de la séquence GG/CC.



- **Coupure Cohésive** : dans ce type, L'enzyme est coupe les doubles brins d'ADN de manière décalée (en zigzag), laissant de courtes séquences d'ADN simple brin non appariées en extrémités (cohésives), ces cohésives est facilement hybrider (se coller) avec des extrémités complémentaires coupées par la même enzyme.

Exemple d'enzyme *Eco*RI coupe l'ADN de manière décalée, Le fragment produit a une extrémité simple brin AATT.



b) coller l'ADN

Le collage d'ADN se fait par L'ADN ligase qui catalyse la formation d'une liaison phosphodiester pour souder définitivement le fragment d'ADN étranger (le gène d'intérêt) au vecteur (le véhicule, souvent un plasmide). Cette étape permet d'assembler les fragments d'ADN découpés. (Voire fig1.)

Chapitra 2: les outils de la génomique

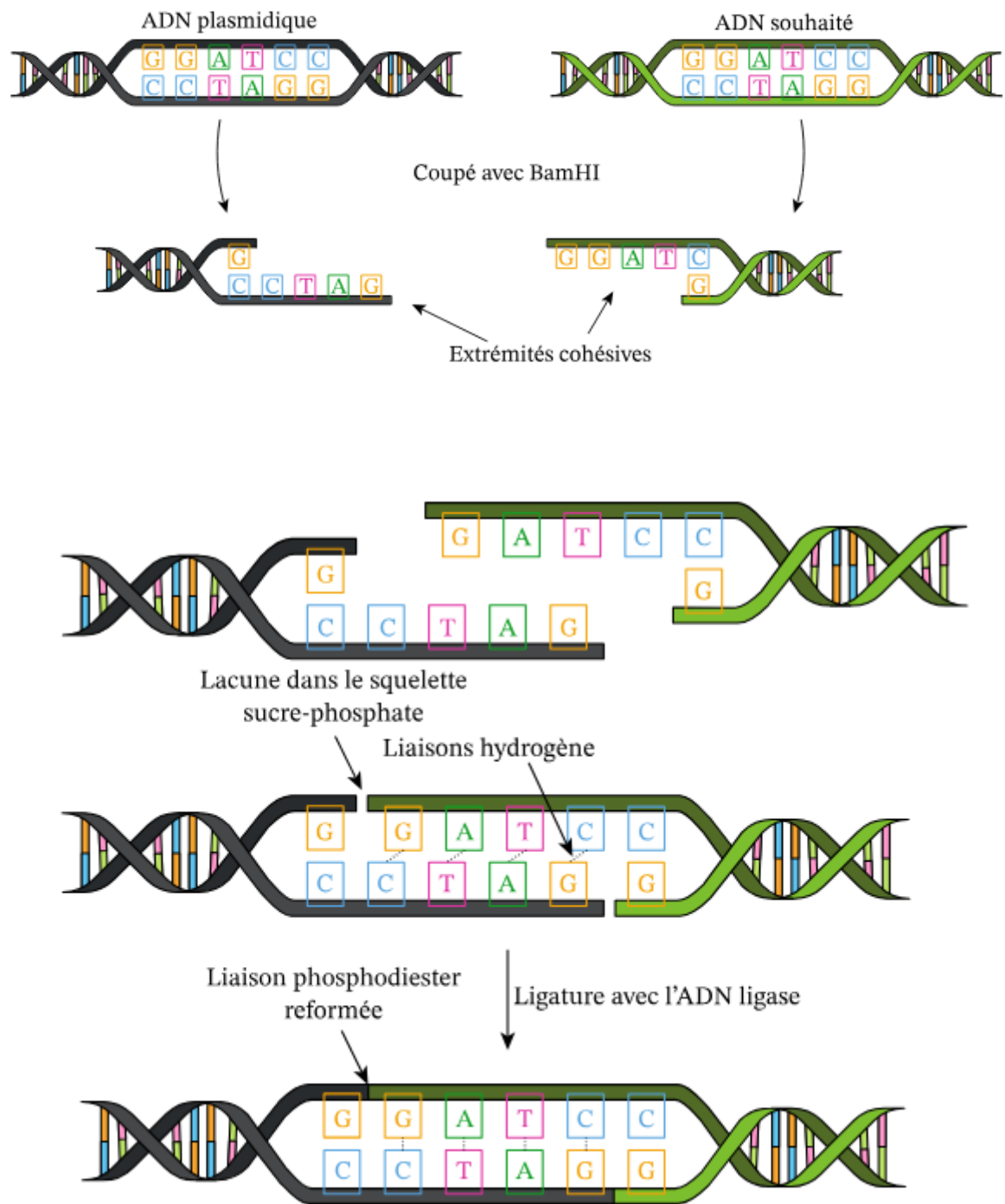


Fig. 1. La figure montre l'utilisation de la même enzyme de restriction (BamHI dans cet exemple) peut donner des extrémités cohésives compatibles. Deux extrémités cohésives compatibles peuvent être ligaturées à l'aide d'une ADN ligase. @ <https://www.nagwa.com>

Chapitra 2: les outils de la génomique

3. Recopier l'ADN

a) La PCR Conventionnelle (Réaction en Chaîne par Polymérase)

La PCR conventionnelle est une technique de biologie moléculaire base sur enzymes bactériennes, les polymérase qui permet d'amplifier (de copier) une séquence d'ADN spécifique in vitro, à partir d'une quantité infime d'ADN de départ. Elle base sur le principe de répllication d'ADN par la polymérase. Elle est facile à automatiser à grande échelle, et largement utilisée par la génomique.

Chaque cycle de PCR comporte trois étapes:

1 - dénaturation de l'ADN matrice à 95°C ;

2 - hybridation des amorces entre 55°C et 70°C (température fonction de la T_m des amorces, donc essentiellement de leur séquence et de leur longueur) ;

3 - extension des amorces par l'ADN polymérase à 72°C (température optimale de fonctionnement de la Taq polymérase) à partir de l'extrémité 3' des amorces.

Au début de chaque cycle, les segments double brin résultant de l'élongation du cycle précédent – constitués du brin matrice et du brin néosynthétisé – sont dénaturés et les brins néosynthétisés servent à leur tour de matrice. Ainsi, à chaque cycle, on a doublement de la quantité d'ADN, c'est-à-dire au total une réaction en chaîne qui conduit, à partir d'une très faible quantité d'ADN matrice, à une production exponentielle d'un fragment d'ADN, appelé fragment PCR (ou amplimère, ou amplicon), d'une taille correspondant au nombre de paires de bases comprises entre les extrémités 5' des amorces sens et antisens.

Les Composants Clés du PCR pour que la réaction ait lieu :

- La Matrice d'ADN (DNA Template)
- Les Amorces (Primers)
- L'ADN Polymérase Thermostable
- Les Désoxynucléotides Triphosphates
- Le Tampon de Réaction et les Ions Mg^{2+} (Le Tampon maintient le pH à un niveau stable, indispensable à l'activité enzymatique. Les Ions Mg^{2+} (Magnésium) sont des cofacteurs essentiels pour la fonction de l'ADN Polymérase. Sans ces ions, l'enzyme ne peut pas fonctionner.)

4. Séquencer d'ADN

Le séquençage de l'ADN est la détermination de l'ordre précis des nucléotides (Adénine, Thymine, Guanine, et Cytosine) dans une molécule d'ADN. Il permet d'étudier les gènes présents dans cette molécule et donc leurs expressions et leurs fonctions. C'est une technique fondamentale en génomique.

Le séquençage a connu une révolution technologique, classée en trois générations principales :

- **Première Génération (Méthode de Sanger, séquences environ 500 à 1000 paires de bases) :** C'est la méthode développée par Frederick Sanger, basée sur des nucléotides modifiés (didésoxyribonucléotides ; ddNTPs, qui n'ont pas de groupe OH en position 3'). Lorsque c'est un nucléotide qui se lie au brin complémentaire de la molécule d'ADN par l'ADN polymérase, ils sont incorporés, bloquant l'élongation de la copie. Cela produit une série de fragments d'ADN de toutes les longueurs possibles, chacun se terminant par un ddNTP spécifique. Ces fragments sont ensuite séparés par électrophorèse pour déterminer l'ordre des nucléotides de la séquence d'ADN.
- **Deuxième Génération : Séquençage de Nouvelle Génération (NGS) :** Il est réalisé de façon automatique par des robots de séquençages, utilise un nombre des méthodes (comme Illumina, Roche 454) qui permettent de séquençage massif et parallèle de millions de fragments simultanément. Il repose sur l'amplification de molécule d'ADN, ou des milliers de fragments d'ADN sont fixés à une surface solide et amplifiés. L'incorporation de chaque base est détectée par un signal fluorescent ou chimique.
- **Troisième Génération : Séquençage de Molécule Unique :** Ces technologies (comme Pacific Biosciences - PacBio ou Oxford Nanopore Technologies - ONT) permettent de séquencer des molécules d'ADN individuelles en temps réel, sans amplification par PCR. L'ADN est poussé à travers un nanopore (un trou nanoscopique) dans une membrane. Le passage des différentes bases (A, T, C, G) interrompt le courant électrique de manière unique, permettant de lire la séquence.

1. Séquençages de Première Génération

a. la technique de Maxam et Gilbert

la technique de Maxam et Gilbert est basée sur un procédé chimique. La technique de Maxam et Gilbert s'est peu développée car elle nécessite des réactifs chimiques toxiques, n'est pas facile à automatiser et reste limitée quant à la taille des fragments d'ADN.

Chapitre 2: les outils de la génomique

Le principe de cette méthode est :

- Après déphosphorylation, un fragment d'ADN linéaire est marqué en 5' par ^{32}P
- On cherche ensuite à ne conserver qu'un seul brin marqué. Pour cela, 2 méthodes sont possibles :
- l'ADN est dénaturé et les 2 brins séparés sur gel de polyacrylamide
- le double brin est clivé par une enzyme de restriction et les 2 fragments (marqué chacun sur un seul brin) sont séparés sur gel.
- Modification chimique et clivage au niveau de bases azotées spécifiques de telle sorte que la modification se produise aléatoirement au plus une fois par brin (1 tube par réaction)
- Migration sur gel de polyacrylamide et Autoradiographie et lecture (Fig. 2.)

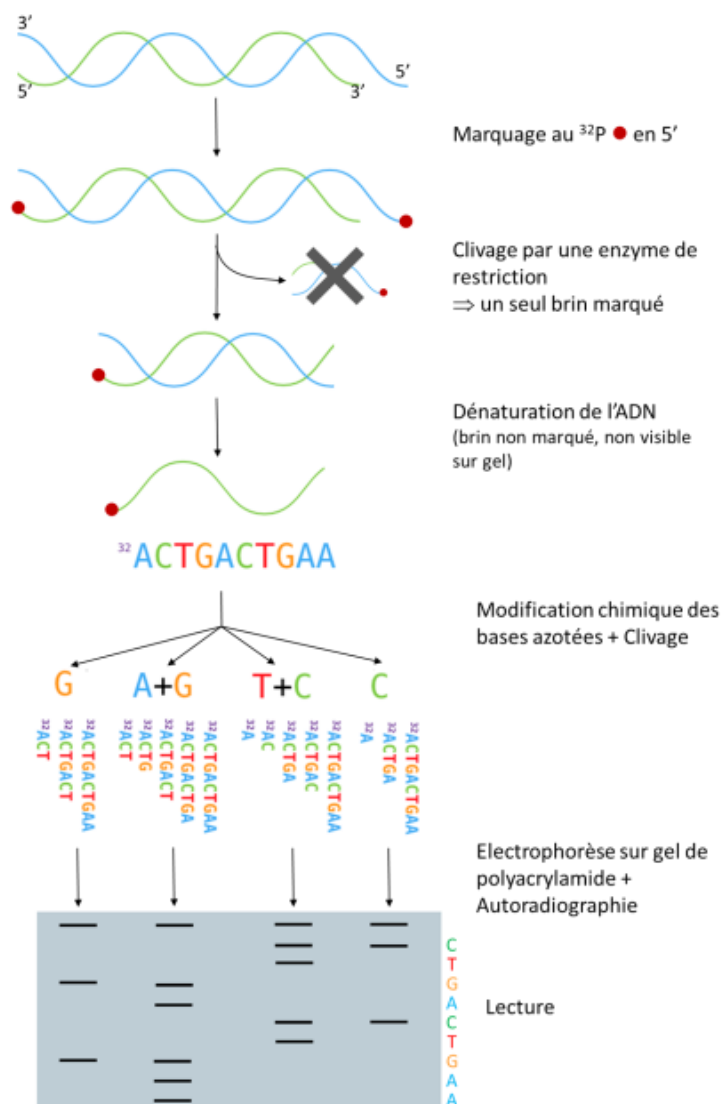


Fig. 2. Séquençage d'ADN par méthode de Maxan et Gibbert

Chapitre 2: les outils de la génomique

b. La technique de Sanger

La technique de Sanger base sur le principe de la réplication d'ADN, mais il utilise des 2'.3' di-déoxynucléotides (ddNTP) qui inhibent l'élongation de l'ADN polymérase. Les ddNTP ne présentent pas de groupement hydroxyle sur le carbone 3'. La liaison phosphodiester entre cet atome et l'atome de carbone 5' d'un autre nucléotide ne peut pas s'établir. Par conséquent, l'incorporation d'un ddNTP à la place d'un dNTP interrompt la polymérisation, le ratio entre la concentration des ddNTP et dNTP est tel que statistiquement. Il y a l'incorporation d'un ddNTP à chaque portion ainsi la réaction enzymatique va générer un mélange de fragments d'acide nucléiques ayant tous la même extrémité 5' et avec des résidus ddNTP à l'extrémité 3'. Les composants de cette réaction sont :

- ADN séquencé : de petites tailles (quelque centaines de nucléotides), amplifié par PCR ou clonale.
- Les amorces capables de s'hybrider à un des brins de la matrice et qui permettent le démarrage de la polymérisation d'ADN.
- Du ddNTP correspondant à une des quatre bases (ddATP, ddGTP, ddCTP, ddTTP)
- polymérase Taq
- des dNTP correspondant à une des quatre bases

Les différents fragments de chaque réaction enzymatique séparés par électrophorèse sur un gel de polyacrylamide coulé entre deux plaques de verre peuvent. (Fig. 3.)

Chapitre 2: les outils de la génomique

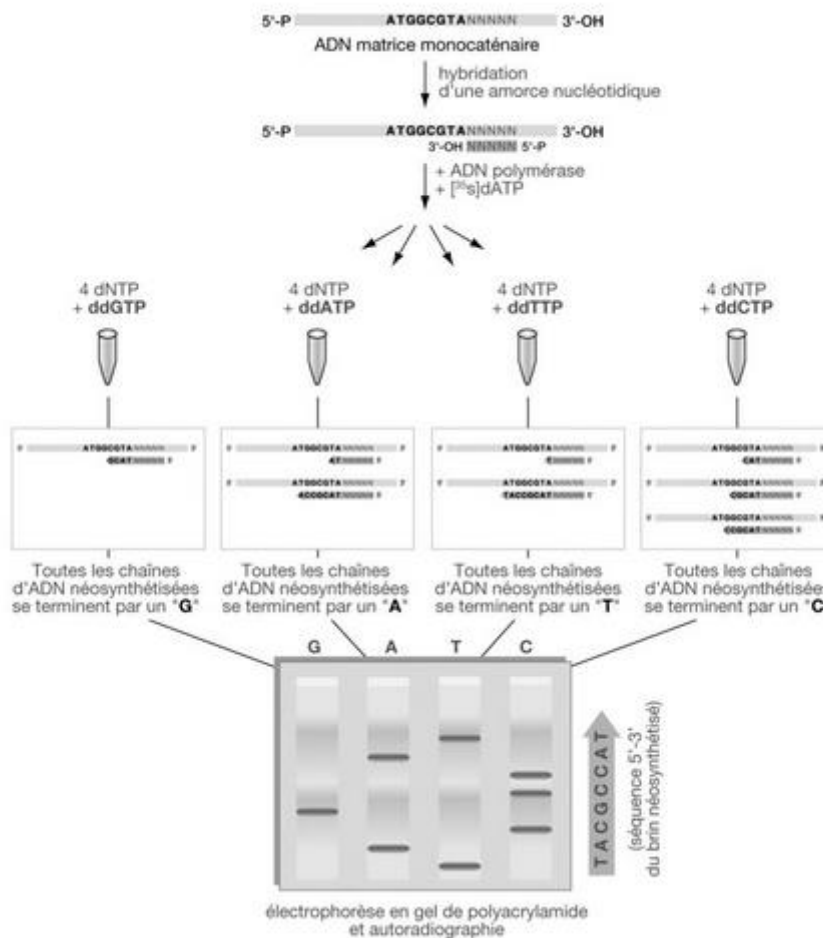


Fig. 3. la technique de Sanger@ Denis Tagu et al. 2018.

Donc être visualisés par autoradiographie. La longueur des fragments identifie la position de base. La résolution électrophoretique de ces quatre réactions produit un profil de bande des radioactives qui reflète la séquence complémentaire de l'ADN matrice.

Le technique de Sanger s'est développée avec succès car les produits utilisés sont maintenant toxiques et la longueur d'une séquence est plus importante (jusqu'à tout les nucléotides avec la technique originelle) par conséquent jusqu'à dernières années. Elle constituait la technique courante de séquençage.

2. Deuxième Génération : Séquençage de Nouvelle Génération

Le séquençage de Nouvelle Génération (NGS) se distingue par sa capacité à séquencer des millions, voire des milliards, de fragments d'ADN simultanément et en parallèle (d'où le terme "massif parallèle").

Chapitre 2: les outils de la génomique

Plusieurs plateformes existent, mais la majorité des technologies NGS reposent sur le principe du Séquençage par Synthèse (Sequencing by Synthesis - SBS), où une base est identifiée lors de son incorporation à un brin d'ADN en cours de construction.

a. Séquençage par Synthèse (Illumina)

C'est la technologie dominante actuellement, responsable de la majorité des données de séquençage générées.

Des fragments d'ADN sont fixés et amplifiés sur une lamelle (flow cell) pour former des clusters clonaux. Chaque base est ajoutée séquentiellement en utilisant des nucléotides marqués par fluorescence et terminés de manière réversible. À chaque cycle, une seule base est incorporée, le signal fluorescent est capturé, le terminateur et le marqueur sont clivés, et le cycle recommence. Parmi l'Avantages de Séquençage par Synthèse est Très haut débit, grande précision et coûts faibles par base. Longueur des lectures relativement courte ($\approx 50-300$ paires de bases). (Fig. 4)

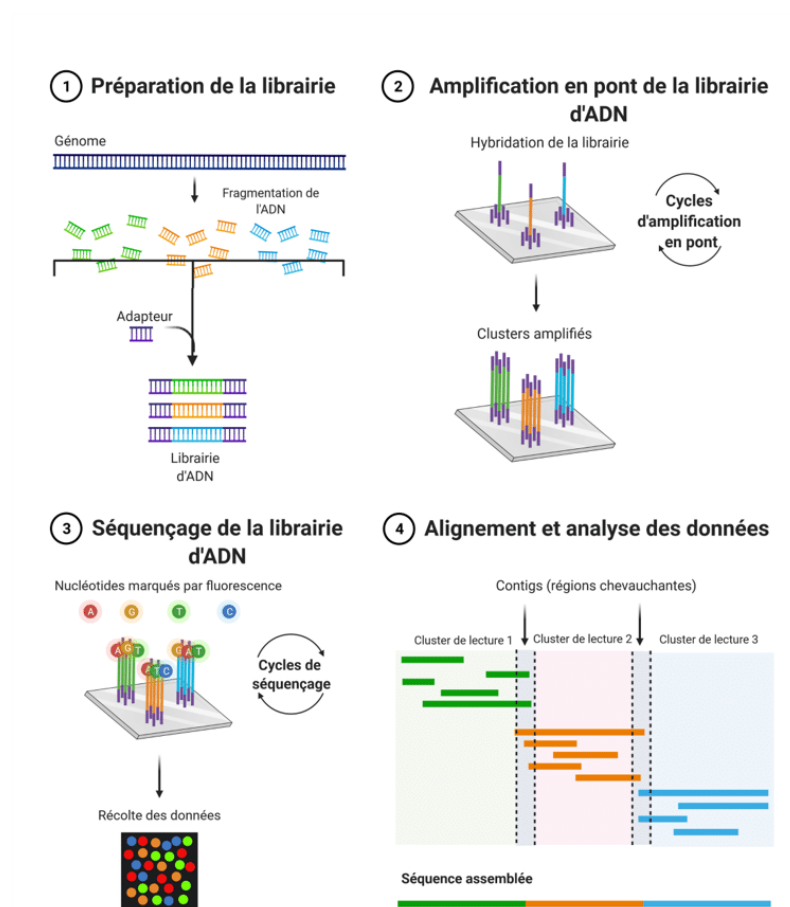


Fig.4. Principes et étapes du séquençage Illumina. BioRender.com (2021).

3. Troisième Génération : Séquençage à Molécule Unique

Cette génération est la plus récente et vise à surmonter la limitation des lectures courtes du NGS. Ces technologies lisent de très longues sections d'ADN.

Le Principe de Séquençage à Molécule Unique est la Lecture directe de l'ADN ou de l'ARN d'une seule molécule, sans étape d'amplification (PCR) préalable. Les deux principales plateformes sont :

A. PacBio (Pacific Biosciences) - Séquençage SMRT

Séquençage en Temps Réel à Molécule Unique (SMRT). L'ADN polymérase est fixée au fond d'un puits (Zero-Mode Waveguide - ZMW) et incorpore des nucléotides marqués par fluorescence. Cette technique est Caractérisé par Très longues lectures ($\approx 10\text{--}30$ kb ou plus), ce qui est excellent pour les régions répétées et l'assemblage de génomes.

B. Oxford Nanopore Technologies (ONT)

Dans cette Technique, L'ADN ou l'ARN passe à travers un nanopore (un trou nanoscopique dans une membrane). Chaque base perturbe le courant électrique d'une manière unique, permettant son identification. il Caractérise par une Lectures ultra-longues (potentiellement des mégabases), temps réel, et l'appareil est portable (ex. : MinION).

5. Reconnaître son semblable

L'Hybridation Moléculaire est Une technique abondamment utilisée dans les expériences de génétique moléculaire repose sur la complémentarité des deux brins de la molécule d'ADN ou sur la complémentarité entre une molécule d'ADN et une molécule d'ARN complémentaire.

Lors de hybridation moléculaire, des fragments d'ADN ou d'ARN, marqués par un traceur radioactif ou fluorescent sont employés pour détecter la présence d'une séquence proche dans un mélange complexe, par exemple un génome entier. (fig. 5)

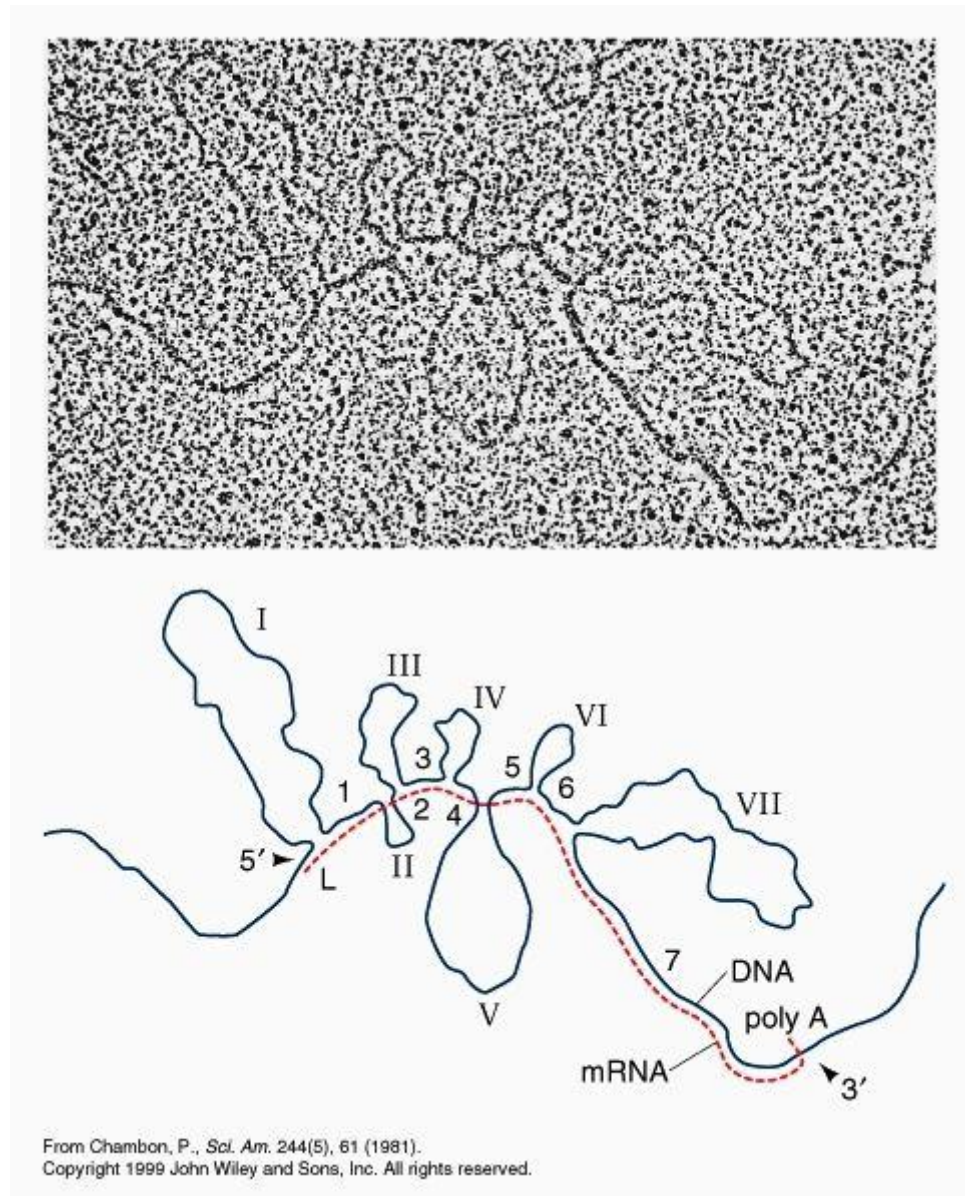


Fig. 5. L'hybridation de molécules d'ADN avec une molécule d'ARN.@

<https://svtmarcq.over-blog.com>

6. Stocker l'ADN: clonage et banques

La méthode de stockage de l'ADN par le clonage et la création de banques d'ADN, est essentielle en biologie moléculaire

Le Clonage est une technique d'isoler et de copie d'une séquence d'ADN au but d'obtenir plusieurs copie identique. Le gène d'intérêt et le vecteur (généralement un plasmide bactérien, une petite molécule d'ADN circulaire) sont découpé avec la même d'enzymes de restriction, pour créant des extrémités cohésives compatibles. Ces extrémités sont collées ensemble par l'ADN ligase, formant un ADN recombinant. Qu'il introduit dans un organisme hôte (ex :

Chapitra 2: les outils de la génomique

bactérie). lorsque cette hôte est multiplié, elle réplique l'ADN recombinant et créant des milliers de copies.

Les Banques d'ADN (Bibliothèques Génétiques) est une méthode de stockage systématique qui collection complète de tous les fragments d'ADN clonés d'un organisme donné. Le génome est découpé en milliers de fragments, chacun est cloné dans un vecteur, et les bactéries hôtes sont stockées dans des milliers de puits ou boîtes, représentant l'ensemble du matériel génétique. Les Banques d'ADN permettent aux chercheurs de trouver et d'étudier n'importe quel gène d'un organisme sans avoir à ré-isolé l'ADN de l'organisme d'origine à chaque fois. Il présente par :

- Banque Génomique qui contient l'intégralité du génome d'un organisme.
- Banque d'ADNc (ADN complémentaire) : Elle représente uniquement les gènes qui sont exprimés.

7. La transgénèse

La transgénèse est une technique de transformation, qui Réintroduire des gènes modifiés ou non dans un organisme vivant afin de vérifier leur fonction. Pour les plantes, deux méthodes principales de transfert d'ADN sont utilisées :

- La méthode indirecte qui utilise *Agrobacterium tumefaciens*. c'est une bactérie vas une capacité de transforme un gène tumoraux de l'ADN-T dans le génome des cellules végétales (la maladie de la galle de collet). cette capacités permette au scientifiques insérant le gène souhaite par le remplacement de ces gènes avec le gène souhaité ; lorsque la bactérie modifie est infectant la cellule végétale, il introduit le gène interé dans leur génome. Cette méthode est la plus courante pour de nombreuses plantes.
 - La méthode directe qui utilise La Biobalistique. C'est des micro-particules métalliques (d'or ou de tungstène) recouvertes d'ADN. Il transfert l'ADN direct au tissu végétal par bombarder le avec microbilles couvertes de l'ADN intérêt à une vitesse importante, qu'elle permet à ces billes de pénétrer dans la cellule en traversant sa paroi. L'ADN solubilisé dans le cytoplasme peut alors dans certains cas, aller s'insérer dans le génome de la plante. Les cellules transformées des tissus pourront également donner naissance à une plante entière. , cette méthode est utilisée pour les plantes récalcitrantes à l'infection par l'*agrobacterium*.