

## Le séquençage de l'ADN :

### Introduction :

La séquence d'ADN contient l'information nécessaire aux êtres vivants pour survivre et se reproduire. Déterminer cette séquence est donc utile aussi bien pour les recherches visant à savoir comment vivent les organismes que pour des sujets appliqués. En médecine, elle peut être utilisée pour identifier, diagnostiquer et potentiellement trouver des traitements à des maladies génétiques et à la virologie. En biologie, l'étude des séquences d'ADN est devenue un outil important pour la classification des espèces.

### 1. Historique :

Les premières techniques de séquençage ont été développées en parallèle au milieu des années 1970. Les méthodes de **Sanger** (Grande-Bretagne) et **Gilbert** (Etats-Unis) ont toutes deux été récompensées d'un prix Nobel de chimie en 1980.

Le premier organisme a été séquencé en 1977. Il s'agissait du virus bactériophage  $\phi$ X174, possédant un ADN simple brin ne nécessitant donc pas l'étape de dénaturation utilisé dans les méthodes de Sanger et Maxam et Gilbert.

Depuis une trentaine d'année, l'amélioration de la technique de production des amorces, de l'amplification des brins et la généralisation des traceurs fluorescents ont permis d'améliorer considérablement les techniques de séquençage. L'apparition des **séquenceurs automatiques** a notamment permis l'automatisation de ces technologies et ont contribué à la finalisation du génotypage humain en 2003.

### 1.1 Définition de séquençage :

Le **séquençage de l'ADN** est une technique d'analyse de l'ADN, consiste à déterminer l'ordre d'enchaînement des nucléotides pour un fragment d'ADN donné. Actuellement la technique de séquençage repose majoritairement sur la méthode enzymatique de Sanger.

### 1.2 Les différents acteurs du séquençage :

□ l'ADN provient des organismes dont on souhaite séquencer le génome. L'ADN, le plus souvent sous la forme double-brin, est dénaturé afin de séparer les deux brins. Celui qui sera séquencé s'appelle le brin « matrice »

les **nucléotides**, ou plutôt désoxynucléotides, sont les bases de l'ADN (A, C, G ou T).

Ils sont attachés les uns aux autres grâce à des liaisons chimiques nécessitant la présence d'un groupement OH particulier (sur le carbone 3' du ribose)

□ les **didésoxynucléotides** sont des nucléotides privés du groupement OH en position 3'.

Leur incorporation dans une chaîne d'ADN interrompt définitivement la synthèse de l'ADN (voir figure 1).

- une **amorce** est un brin d'ADN très court (une vingtaine de nucléotide), qui peut s'hybrider à une séquence complémentaire spécifique
- l'**ADN polymérase** est une enzyme qui a pour rôle de copier l'ADN, en synthétisant un brin complémentaire au brin matrice. L'ADN polymérase ne fonctionne qu'en ajoutant des nucléotides à **une amorce** déjà présente, en fonction de la succession de nucléotides du brin matrice (un A est toujours positionné en face d'un T et un C toujours en face d'un G).

### 1.3 Méthode de Sanger :

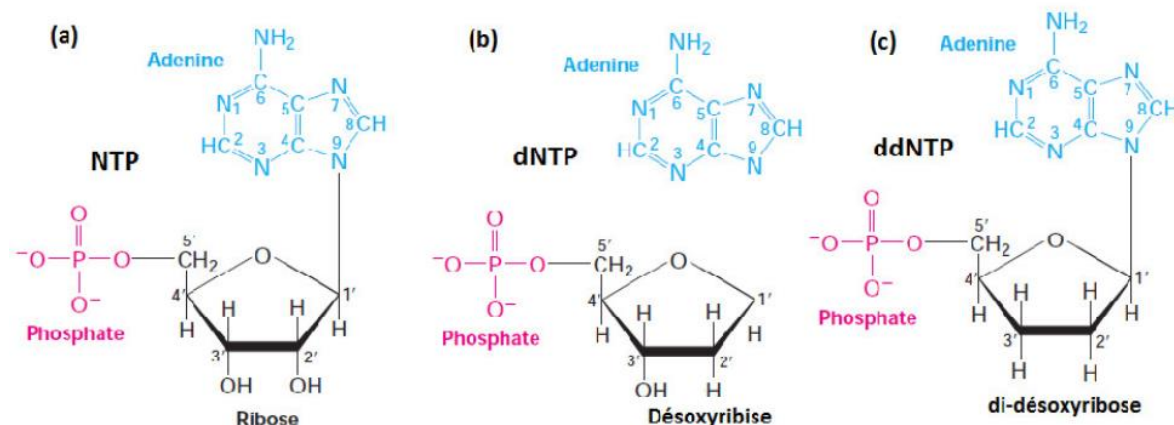


Figure : les différentes formes des nucléotides.

Dans le ddATP, le **groupement 3'-OH** est remplacé par un **hydrogène**. Cette modification empêche la poursuite de la synthèse de l'ADN qui continue normalement sur le 3-OH

### 1.4 Principe du séquençage par la méthode de Sanger :

Le principe de cette méthode consiste à initier la polymérisation de l'ADN à l'aide d'un petit oligonucléotide (amorce) complémentaire à une partie du fragment d'ADN à séquencer.

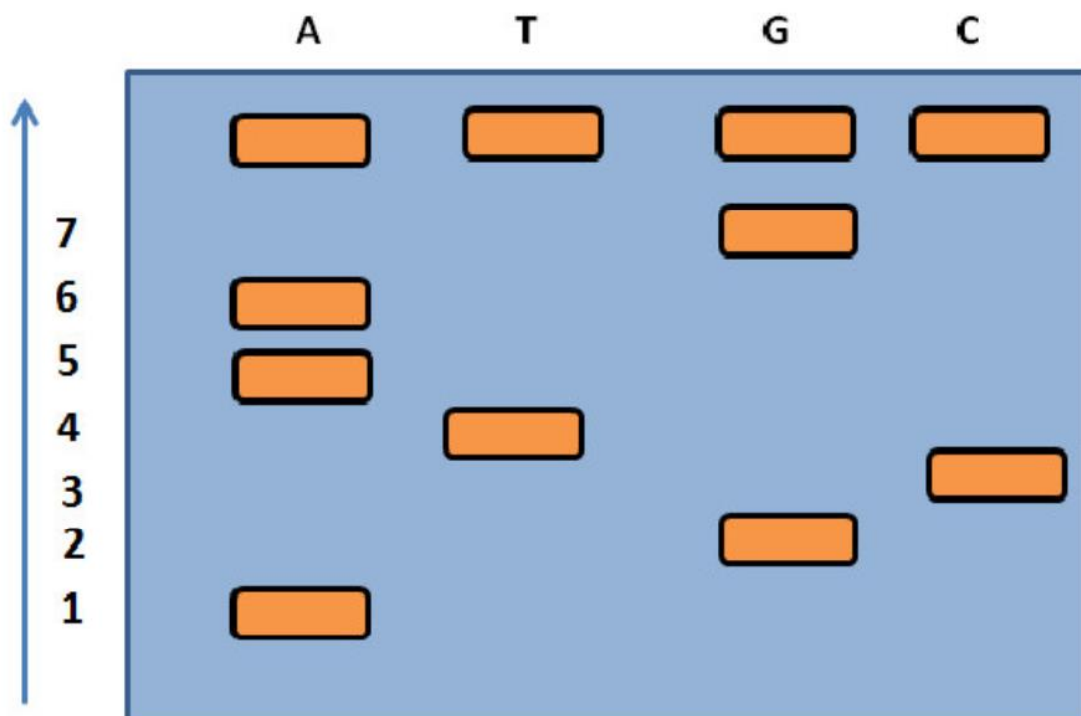
L'élongation de l'amorce est réalisée par une ADN polymérase I dépourvue d'activité exonucléase 5' → 3' et maintenue par des ADN polymérases thermostables, celles qui sont

utilisées pour la PCR. On utilise quatre tubes séparés contenant les quatre désoxyribonucléotides (dATP, dCTP, dGTP, dTTP), ainsi qu'une faible concentration de l'un des quatre didésoxyribonucléotides (ddATP, ddCTP, ddGTP ou ddTTP).

Ces didésoxyribonucléotides agissent comme des « poisons » terminateurs de chaîne : une fois incorporés dans le nouveau brin synthétisé, ils empêchent la poursuite de l'élongation.

Cette terminaison se fait spécifiquement au niveau des nucléotides correspondant au didésoxyribonucléotide incorporé dans la réaction. Pour le séquençage complet d'un même fragment d'ADN, on répète cette réaction quatre fois en parallèle, avec les quatre didésoxyribonucléotides différents.

Par exemple, dans la réaction où on a ajouté du ddGTP, la synthèse s'arrête au niveau des G. Le mélange réactionnel contenant, à la fois du dGTP et un peu de ddGTP, la terminaison se fait de manière statistique suivant que l'ADN polymérase utilise l'un ou l'autre de ces nucléotides. Il en résulte un mélange de fragments d'ADN de tailles croissantes, qui se terminent tous au niveau d'un des G dans la séquence. Ces fragments sont ensuite séparés par électrophorèse sur un gel de polyacrylamide, ce qui permet ainsi de repérer la position des G dans la séquence.



**Figure :** Séquençage de l'ADN par la méthode de Sanger

La détermination de l'enchaînement des nucléotides dans un fragment d'ADN simple brin par la méthode de Sanger comporte 2 étapes ;

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

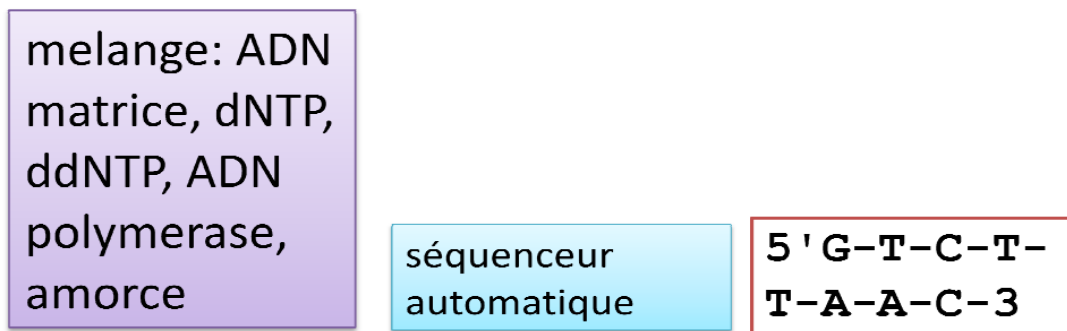
(absence de 3'-OH). 4 réactions en parallèle sont nécessaires, chacune avec l'un des terminateurs présents en faible quantité .

- Séparation par électrophorèse des fragments synthétisés, le nombre de ces fragments dépend du nombre de résidus nucléotidiques. Donc la position des terminateur dans la séquence.

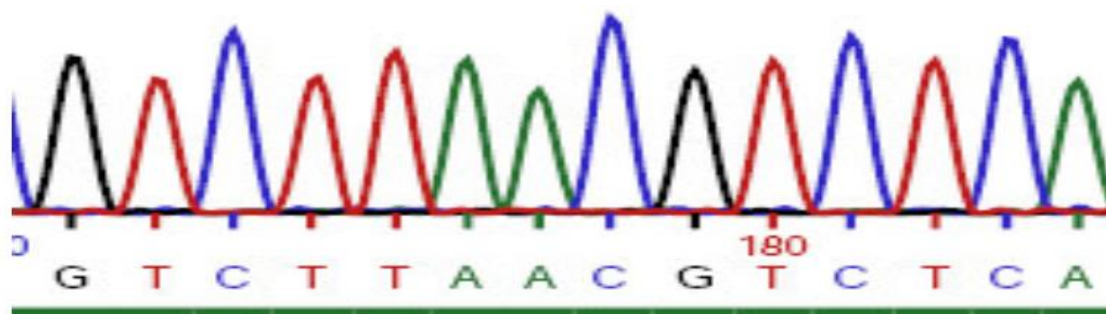
La détection des fragments ainsi synthétisés se fait en incorporant un traceur dans l'ADN synthétisé. Initialement ce traceur était radioactif, aujourd'hui, on utilise des traceurs fluorescents, attachés soit à l'oligonucléotide, soit au didésoxyribonucléotid

### 1.5 Automatisation du séquençage

Des séquenceurs automatique sont capables de réaliser les réactions de séquence puis de les lire.



• Les didésoxynucléotides sont marqués avec des fluorescences différentes , cela permet de faire une seule réaction et une seule lecture . Le séquenceur détecte la fluorescence et repère la taille des fragments d'ADN.



On obtient des courbes de fluorescences qui sont interprétées en terme de nucléotides

