

SEQUENÇAGE AUTOMATISE SELON LA METHODE DE SANGER

I) Objectifs

Cette méthode, aussi appelée “*dideoxynucleotide sequencing*”¹ ou “*chain termination sequencing*”¹, a pour but de séquencer des échantillons d’ADN, donc de connaître l’ordre dans lequel sont positionnés les 4 nucléotides : A (Adénine), C (Cytosine), G (Guanine) et T (Thymine).

II) Principe (Figure 1)

Le principe se base sur l’ADN polymérase et les di-déoxyribonucléotides (ddNTP). Les échantillons d’ADN à séquencer sont mélangés avec des amorces spécifiques, de l’ADN polymérase, des déoxyribonucléotides (dATP, dCTP, dGTP et dTTP) et des di-déoxyribonucléotides (sans groupement hydroxyle en position 3’) (ddATP, ddCTP, ddGTP et ddTTP). L’incorporation de ces ddNTP empêche la formation de la liaison phospho-diester, et donc la poursuite de l’élongation du brin d’ADN². Des fragments d’ADN de tailles variables se terminant par un ddNTP sont ainsi obtenus.

Dans la première version de cette méthode³, les amorces étaient marquées radioactivement et les bouts d’ADN migraient sur un gel dénaturant permettant l’obtention d’un profil de migration³. Actuellement, une amplification par PCR est réalisée en amont, et la méthode employée utilise l’électrophorèse capillaire et la fluorescence. L’électrophorèse permet de classer les fragments selon leur taille et ainsi obtenir la séquence nucléotidique. De plus, chaque ddNTP est marqué par un fluorochrome différent. Un rayonnement laser excite la molécule fluorescente qui réémet une longueur d’onde spécifique. Cette dernière est détectée par une caméra. A la fin du séquençage, cela permet d’obtenir un électrophorégramme.

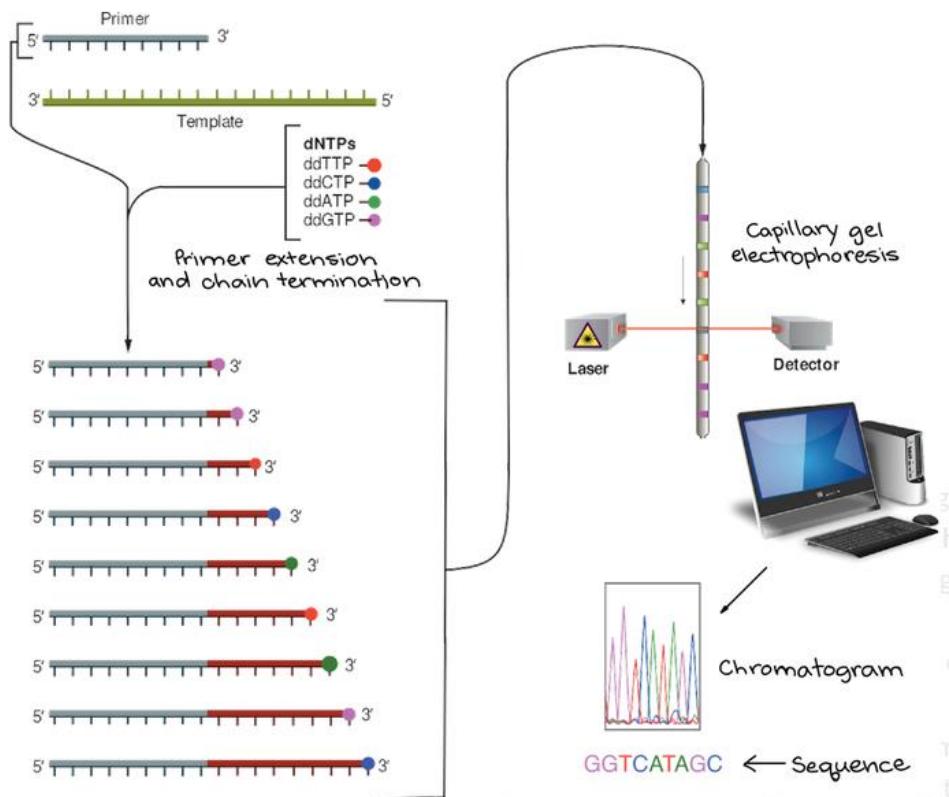


Figure 1 : Principe du séquençage Sanger⁴

III) Mode opératoire

Tout d'abord, une amplification par PCR de séquençage est réalisée. L'incorporation des ddNTP marqués par les fluorochromes va permettre l'obtention de nombreux fragments de différentes tailles de l'ADN à séquencer. L'extrémité 3'OH des dNTP est remplacée par un 3'H d'un ddNTP, empêchant la formation de la liaison phosphodiester entre ce dernier et le nucléotide suivant. Le rapport ddNTP/ dNTP, ainsi que l'affinité de la Taq polymérase pour chaque nucléotide sont optimisés pour que statistiquement, un ddNTP soit incorporé à chaque position possible. La technologie BigDye™ est généralement utilisée. Le BigDye™ contient les dNTPs, ddNTPs, de l'ADN polymérase et du MgCl₂. Les tubes à PCR sont donc incubés dans un thermocycleur, un temps d'activation et d'extension finale ne sont pas nécessaires. Les échantillons sont par la suite purifiés.

Pour cela, du sodium d'acétate (pH= 4,6) est d'abord utilisé avec de l'éthanol à 100% pour précipiter l'ADN au fond des puits de la plaque (avec incubation à RT et centrifugation). Puis l'ADN est lavé avec de l'éthanol à 70% qui solubilise les impuretés (+ centrifugation).

Les échantillons sont finalement resuspendus dans du Hi-Di (Formamide hautement désionisée) permettant de re-suspendre les échantillons pour l'électrophorèse capillaire. L'ADN est alors dénaturé 1 min à 95°C dans un thermocycleur (pour obtenir de l'ADN simple brin), puis la plaque est prête pour le séquençage selon Sanger. Avant de lancer le séquençage, il faut programmer le séquenceur en précisant les puits utilisés.

Une électrophorèse capillaire en conditions dénaturantes est réalisée par le séquenceur. Les produits de la réaction de séquence sont séparés selon leur taille, et la fluorescence spécifique des ddNTP est enregistrée et analysée pour réaliser l'électrophorégramme.

IV) Présentation des résultats

Le séquenceur donne les résultats sous forme de séquence nucléotidique. Afin de mieux présenter et interpréter les résultats du séquençage, l'utilisation du logiciel Chroma (ou similaire) est nécessaire (Figure 2).

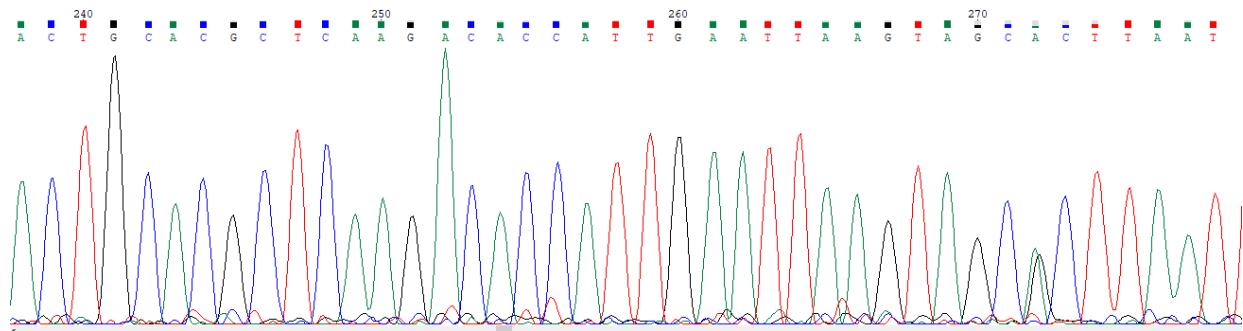


Figure 2 : Exemple de résultats du séquençage sur Chromas, communication personnelle

V) Interprétation des résultats

Un logiciel permet de visualiser l'électrophorégramme (Figure 2), et ainsi modifier les nucléotides N que l'analyseur n'a pas su déterminer en A, T, C ou G. Pour cela, les pics de l'électrophorégramme permettent de déterminer le nucléotide N et ainsi le remplacer pour obtenir la séquence finale.

VI) Intérêts et limites

Une fois la séquence obtenue, il y a de nombreuses utilisations différentes. En effet, connaître la séquence d'ADN permet de réaliser des études d'arbres phylogénétiques, diagnostiquer et étudier des maladies ou encore contrôler des pathogènes². Par exemple, pour aligner la séquence obtenue à une autre séquence d'intérêt, des logiciels d'alignements peuvent être utilisés, comme le logiciel MAFFT. Pour identifier et comparer la séquence, le logiciel *Nucleotide blast* de NCBI est utile. Cela permet de vérifier un produit de PCR et/ou de plasmide. En conclusion, grâce à la bio-informatique, les possibilités à partir de la séquence obtenue sont nombreuses.

Néanmoins, cette méthode ne permet de séquencer que des fragments allant de 600 à 1 000 paires de bases par réaction⁵. Elle est encore utilisée de nos jours en diagnostic clinique lorsque la recherche est orientée vers un petit nombre de gènes, ou bien pour confirmer les résultats de NGS (« New Generation Sequencing ») lorsque la mutation et le gène ont été identifiés⁵. Toutefois, le séquençage Sanger ne s'adapte pas pour un séquençage massif de gènes, et dans ces cas-là, le NGS est largement utilisé de nos jours. Le choix d'une méthode ou de l'autre dépendra donc de l'objectif du séquençage (Figure 3).

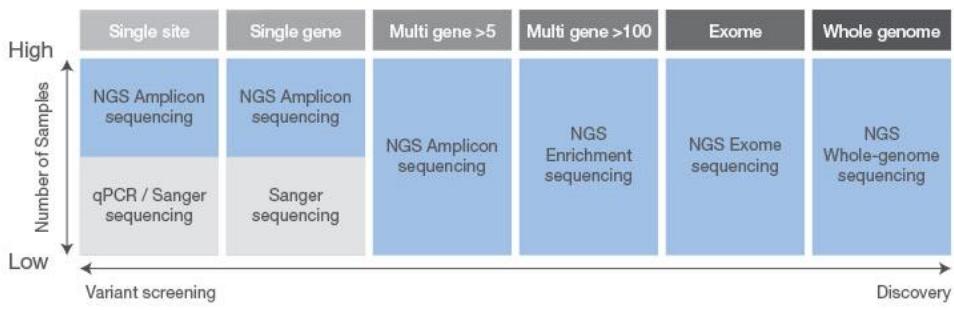


Figure 3 : Options d'utilisation de la méthode Sanger ou NGS⁶

De plus, si une amplification par PCR est réalisée avant le séquençage, des parties de la séquence du vecteur d'amplification peuvent être retrouvées. Finalement, il peut également y avoir une mauvaise reconnaissance de l'amorce et donc des erreurs en début de séquence⁷.

VII) Références bibliographiques

- 1-Carpente L, Cerdeira-Pena A, Lorenzo-Freire S, Places ÁS (2019). Optimization in Sanger sequencing. *Computers & Operations Research* **109**: 250–262
- 2-Garrido-Cardenas J, Garcia-Maroto F, Alvarez-Bermejo J, Manzano-Agugliaro F (2017). DNA Sequencing Sensors: An Overview. *Sensors* **17**: 588
- 3-Sanger F, Nicklen S, Coulson AR (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* **74**: 5463–5467
- 4-Khan Academy (modifié en 2020), Dna sequencing, Method of Sanger sequencing, Khan Academy (en ligne). <https://www.khanacademy.org/science/high-school-biology/hs-molecular-genetics/hs-biotechnology/a/dna-sequencing> (consulté le 18/03/2020)
- 5- Carole André, Paloma Moreno-Elgard (2018). Les nouvelles techniques de séquençage. *Savoir & Comprendre*. 3-4.
- 6- Illumina® (2020). Key differences between next-generation sequencing and Sanger sequencing, Comparison of Sanger Sequencing and NGS, Illumina (en ligne). <https://www.illumina.com/science/technology/next-generation-sequencing/ngs-vs-sanger-sequencing.html?langsel=/us/> (consulté le 21/03/2020)
- 7- Martina Kujanová (2020). Higher-quality sequence data at the beginning of electropherogram - Modified sequencing primers, SEQme (en ligne) (consulté le 21/03/2020)

7- **Martina Kujanová** (2020). Higher-quality sequence data at the beginning of electropherogram - Modified sequencing primers, SEQme (en ligne) (consulté le 21/03/2020)