

## Harry Pot'TUT

### Génétique

## Séquençage Haut Débit : NGS

### P'tite intro :

Le séquençage haut débit est une technique utilisée dans le diagnostic de maladie génétique et permet de séquencer un très grand nombre de gènes. On parle de séquençage massif ! Il est très utilisé dans le diagnostic de maladies multigéniques ( avec plusieurs gènes en causes ) comme la maladie de Charcot Marie Tooth, pour analyser l'ADN de plusieurs patients en parallèle ou encore pour l'analyse de pathologies dont on ne connaît pas encore le ou les gènes responsables. Toutefois, même si techniquement parlant on peut l'utiliser pour le diagnostic de maladies monogéniques comme l'achondroplasie, il n'y en a pas une grande utilité puisque la maladie provient d'une mutation d'un gène précis. On va donc faire une PCR-RFLP plutôt que séquencer tout le génome ! Il existe actuellement deux plateformes ( sociétés ) qui commercialisent les technologies de NGS, Illumina et ThermoFisher / Life technologie.

**En bref, le séquençage haut débit est un séquençage massif en parallèle de molécules d'ADN individuellement séparées et amplifiées sous forme de clones ou de molécules uniques.**

### I/ Les étapes du NGS

Le NGS se déroule en 4 étapes que nous allons détailler par la suite :

- Préparation des échantillons → consiste à fragmenter l'ADN génomique ou les fragments de PCR pour avoir une taille entre 200 et 400 pb ( selon les plateformes ). Ensuite seront rajoutés sur ces fragments des adaptateurs et des barres-codes qui sont des séquences CONNUES d'oligonucléotides ayant pour rôle l'amplification PCR pour les adaptateurs et la reconnaissance des patients pour les BC ( chaque BC correspond exactement à un patient, on pourra donc mélanger les patients par la suite ).
- Enrichissement par PCR clonale → fixation et amplification de chaque fragment d'ADN sur un support solide, une lame de verre pour illumina et une sphère métallique pour ThermoFisher.

- Séquençage individuel des fragments avec des technologies différentes → utilisation de la fluorescence pour Illumina et des variations de PH pour ThermoFisher.
- Cette dernière étape est commune et très importante puisqu'il s'agit de l'analyse informatique. Il faut transformer les données ( fluorescence ou variation de PH ) en séquence.

## II / Préparation des échantillons

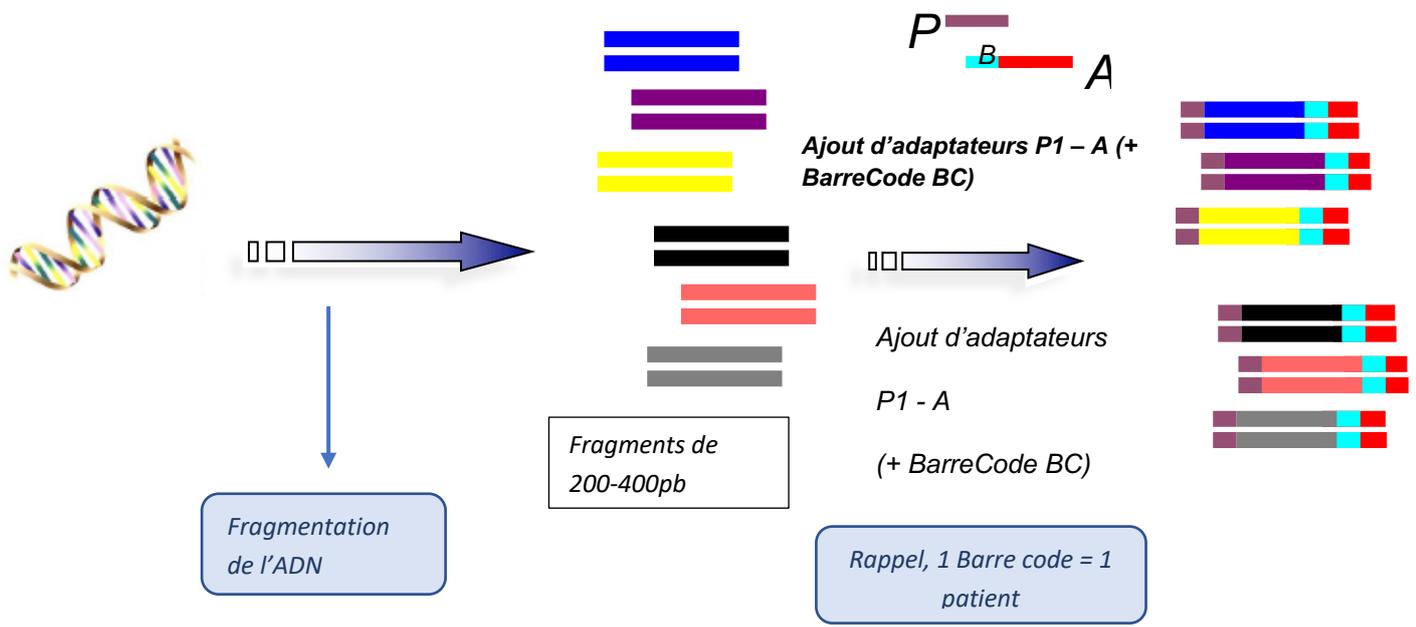
Attention !!!!! La préparation des échantillons est la MÊME pour les deux méthodes !!!!

→ Coupure aléatoire de l'ADN double brin par des endonucléases ( enzymes différentes des enzymes de restrictions ) pour obtenir des fragments entre 200 et 400 pb

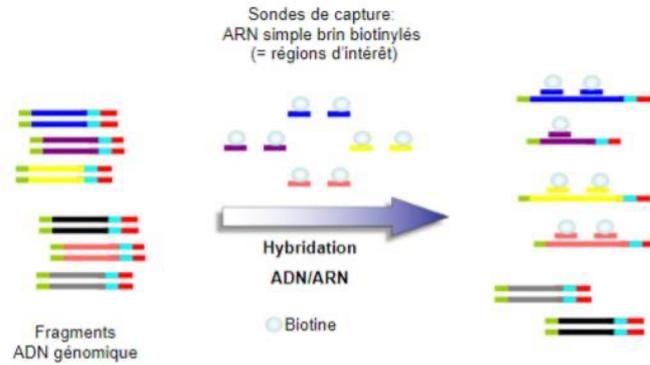
→ Ajout des adaptateurs et des BC grâce à l'ADN ligase pour les adaptateurs et l'ADN polymérase pour rendre les fragments à bords francs ( pas forcément le cas après coupure ) :

- . Les adaptateurs P1 se placent en 5'
- . Les adaptateurs A se placent en 3'
- . Entre les adaptateurs A et le fragment d'ADN se positionnent les BC pour reconnaître le patient

Attention !! Les adaptateurs P1 et A sont les mêmes pour les différents patients, ce ne sont que les BC qui sont spécifiques !!



→ Capture des séquences d'intérêts ( le but est de récupérer uniquement les régions d'intérêts comme les exons et éviter par-là de séquencer tout le génome ). On utilise des sondes de captures ( petits ARN simple brins ) biotinylées qui sont complémentaires des régions à capturer. Il va alors y avoir une hybridation ADN / ARN.

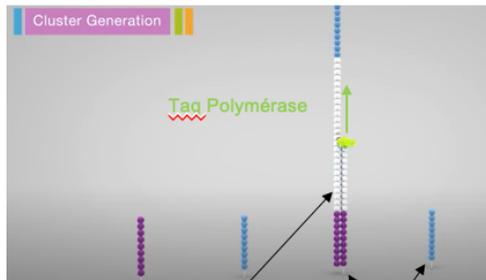


→ Il faut ensuite purifier le mélange pour ne récupérer que les fragments qui nous intéressent.

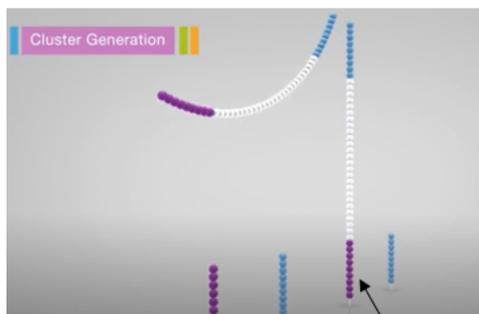
<p>1. Ajout de billes magnétiques</p>	<p>. Les billes sont recouvertes de streptavidine pour permettre l'hybridation avec la biotine</p>	<p><b>Les étapes du NGS</b></p> <p>Purification des fragments d'ADN capturés</p>
<p>2. Utilisation d'un aimant pour la récupération</p>	<p>. Aimantation des billes magnétiques sur les parois du tube . Les fragments non capturés vont être éliminés par lavage</p>	<p><b>Les étapes du NGS</b></p> <p>Purification des fragments d'ADN capturés</p>
<p>3. Dégradation des sondes de capture</p>	<p>. Utilisation de RNAses pour dégrader l'ARN de la liaison ADN/ARN . On peut alors ne récupérer que les fragments d'intérêt.</p>	<p><b>Les étapes du NGS</b></p> <p>Dégradation des sondes de capture (ARN)</p>

### III / Méthode Illumina

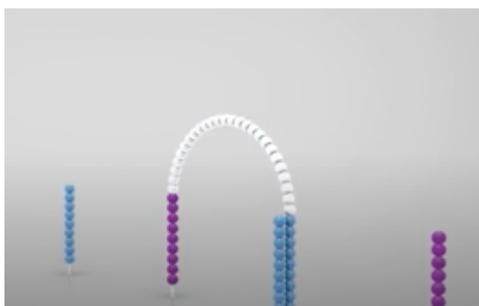
Pour rappel avec cette méthode, on utilise une lame de verre ( aussi appelé Flow Cell ) recouverte d'oligonucléotides complémentaires avec les séquences des adaptateurs. Il y en a donc deux types sur la lame. Les échantillons d'ADN sont ensuite ajoutés sur la lame et il y a hybridation de nos fragments avec les oligonucléotides de la lame par complémentarité ( comme d'hab quoi ;). Ensuite plusieurs étapes vont se répéter, c'est le PCR « bridge » :



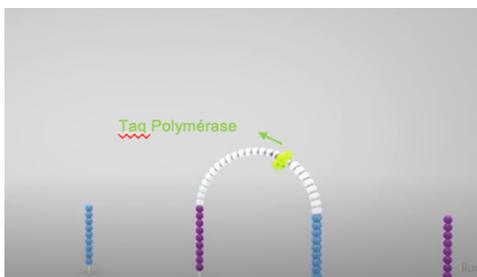
*Amplification par PCR avec synthèse par une Taq polymérase de 5' en 3'*



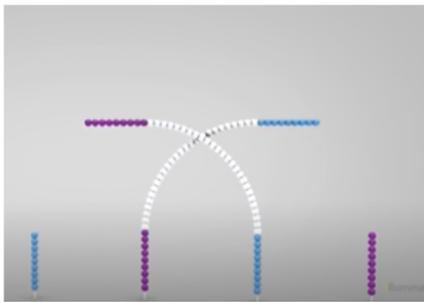
*Dénaturation par lavage, le premier fragment posé sur la lame est libéré ( on garde seulement celui nouvellement formé qui est fixé à la lame de verre )*



*Formation d'un pont avec hybridation en 3' du brin d'ADN avec un 2<sup>e</sup> oligonucléotide*



*PCR « bridge » en cluster, synthèse du brin complémentaire = le brin sens par une Taq polymérase*



Dénaturation et lavage pour obtenir 2 brins d'ADN fixés à la lame de verre, le brin sens et le brin reverse

Ces étapes vont être répétées un grand nombre de fois pour amplifier les fragments d'ADN et obtenir la formation de clusters = groupes de produits PCR. A la fin des cycles il y a clivage des brins reverse ( le brin complémentaire du brin initialement déposé sur la lame ) puis lavage pour les éliminer. Il y a conservation uniquement des brins sens !

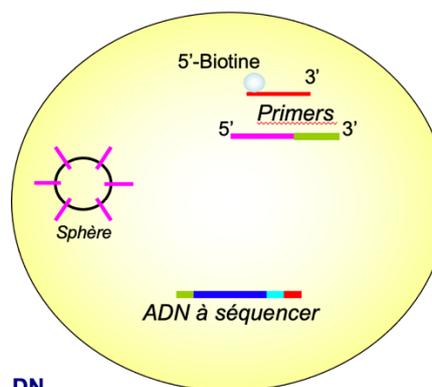
Pour finir avec cette méthode ( allez courage il reste pas grand-chose !!! ) il faut faire le séquençage directement sur la lame par détection de fluorescence. Pour cela on rajoute un primer ( oligonucléotides complémentaires de l'adaptateur ), de l'ADN polymérase ( qui synthétise de 5' en 3' ) et des dNTP fluorescents. Il va y avoir incorporation des nucléotides par la polymérase et libération de fluorescence par excitation qui va être mesurée par un automate ( la couleur dépend du nucléotide ajouté ). Des milliers de clusters sur la Flow Cell sont analysés en même temps. On parle donc de séquençage massif !!!

#### IV / Méthode Thermofisher

Cette fois pas de lame de verre, mais l'ADN est sur une sphère métallique. L'amplification clonale se fait dans un microréacteur créé par un système d'émulsion ( mélange d'huile avec le milieu réactionnel de la PCR qui correspond à une solution aqueuse ). Un microréacteur est composé de :

- 1 sphère porteuse d'un primer
- 1 fragment d'ADN à purifier
- Le milieu réactionnel de la PCR ( primers, ADN polymérase, dNTP et tampon )

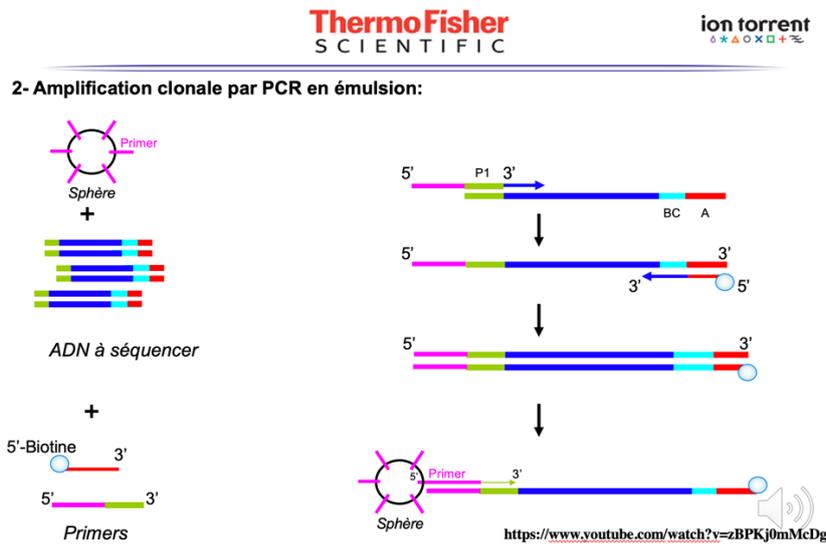
Pour la PCR clonale, les étapes sont les mêmes que pour la PCR classique ( je vous renvoie à mon cours précédent 🤔 ).



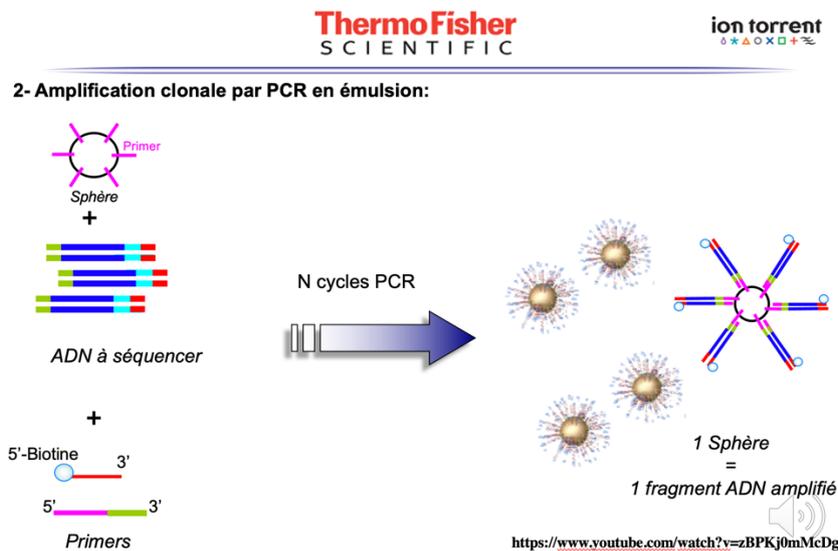
DN

Les primers utilisés sont spécifiques des extrémités 5' et 3'. Il y en a un complémentaire de l'adaptateur P1 portant en 5' une séquence complémentaire du primer porté par la sphère. Le deuxième est complémentaire à l'adaptateur 1 et est porteur de biotine en 5' du primer.

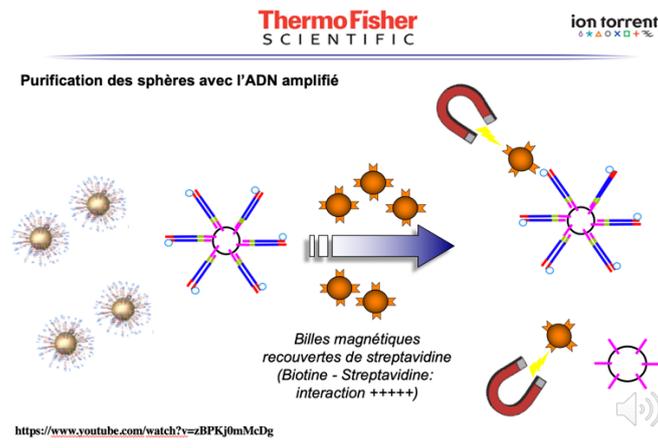
→ Dans un premier temps il y a hybridation du primer d'abord avec la séquence P1 puis lors du deuxième cycle, c'est le deuxième primer ( le biotinylé ) qui s'hybride à l'adaptateur.



Après plusieurs cycles, la sphère sera recouverte de fragments d'ADN double brin biotinylé.



Par la suite il y a purification des sphères avec l'ADN amplifié pour éliminer celles qui sont non recouvertes d'ADN. Le principe est le même que pour la méthode Illumina, on utilise des billes magnétiques recouvertes de streptavidine qui vont interagir avec la biotine en créant des liaisons. Grâce à un aimant les billes vont pouvoir être récupérées.



Chaque sphère va être ensuite séquencée individuellement puis être placée dans un puit sur une puce. Le séquençage se fait directement dans les puits. Pour la réaction il faut rajouter le milieu réactionnel composé :

- D'une amorce
- D'un ADN polymérase
- Des DNTPs

*Attention, on n'utilise pas de ddNTP puisqu'on ne cherche pas à avoir de la fluorescence mais on étudie les variations de pH !!!! Lorsque des nucléotides sont fixés il y a libération d'ions H+ donc si le pH change le nucléotide a été incorporé sinon il ne l'est pas.*

## V / Analyse informatique

*Bon là je vais pas vous mentir c'est le moment du cours bien relou mais faut quand même le faire ....*

Une fois qu'on a récolté les informations ( donc fluorescence ou pH ) il faut les transformer en séquence. C'est cette analyse qui va permettre de mettre en évidence des variants qu'il faudra par la suite interpréter en fonction du gène concerné, de la protéine, des différentes répercussions ...

1. Détermination des bases ou base calling	. On transforme les signaux lumineux ou les variations de pH en séquence
2. Étape de contrôle de la qualité	. On élimine les lectures de séquences de mauvaises qualités
3. Alignement	. On essaye d'aligner notre séquence sur une séquence de référence qui est entièrement connue ( le génome humain quoi ! ) ce qui va permettre de mettre en évidence les variants.
4. Analyse	. On vérifie que les gènes qu'on a ont été séquencés correctement ( est ce qu'il en manque ? Nombre de lecture pour chaque base ? ). Pour cela il y a deux notions hyper importantes : <ul style="list-style-type: none"> <li>- Profondeur de lecture → nombre de lecture indépendante de chaque base</li> <li>- Couverture → % de bases séquencées par lecture indépendante ( pour une profondeur donnée )</li> </ul>

5. Annotation des variants	. Identification des variants par rapport à la séquence de référence ce qui permet de déterminer la localisation chromosomique, la fréquence du variant dans la population ...
6. Interprétation	. Expertise clinico---biologique très importante car dépend du contexte clinique, de la fonction du gène ...

*Et voilààà c'en est enfin fini pour ce cours !!!!! Courage pour les révisions, je sais que c'est pas le meilleur cours et que les techniques ne sont pas forcément simples à comprendre, mais avec de la persévérance ( des fois même beaucoup ) ça va . Et si j'ai un conseil à vous donner c'est de faire tout pleins de QCM c'est comme ça qu'on apprend !!!*

*Et juste en passant comme ça petite dédi à Justine et Maéva qui sont en Las 2, vous allez tout défoncer !!!*