

## 1. Introduction

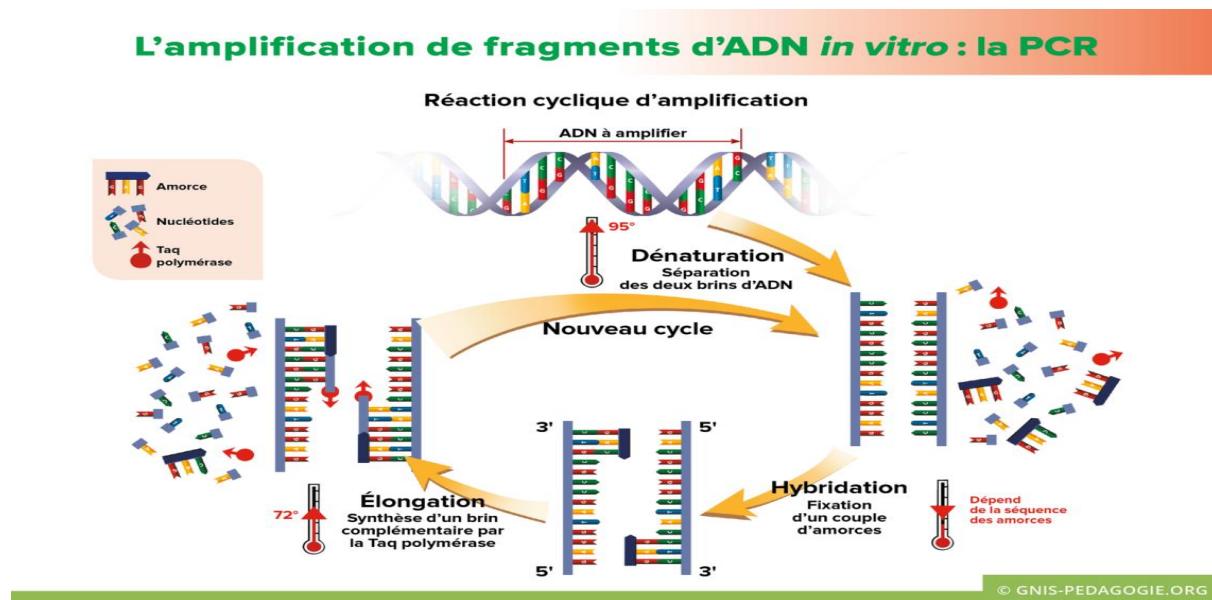
La génétique moléculaire a de nombreuses applications en médecine et dans des domaines variés : -Médecine légale (reconnaissance de paternité, criminologie) -Recherche de prédispositions génétiques -Diagnostic de maladies génétiques, épigénétique, maladies infectieuses (quel est l'agent infectieux impliqué). -Thérapie génique. -Expression recombinante de protéines d'intérêt thérapeutique -Nouvelles cibles thérapeutiques (recherche de marqueurs de thérapies ciblées en cancérologie). -Pharmacogénétique : on recherche sur le génome des marqueurs qui permettent d'évaluer la réponse au traitement ou les éventuels effets secondaires.

### 1.1. Principales applications diagnostiques.

**1.1.1. Recherche et analyse de génome étranger** : microbiologie. Pour faire le diagnostic, on recherche l'ADN ou l'ARN de l'agent infectieux soupçonné. Mais l'ADN dans le germe peut rester présent trois semaines après une antibiothérapie efficace (c'est la limite de la méthode). On peut également rechercher la résistance à un agent infectieux. On fait une quantification de la charge virale (dans l'hépatite, on compte le nombre de copies/ml). La biologie moléculaire sert aussi pour l'épidémiologie moléculaire.

**1.1.2. Analyse du génome humain.** Tout échantillon biologique peut être utilisé : -sang total -prélèvement buccal -liquide amniotique -villo-sités choriales -cheveux... . Le sang doit être prélevé dans un tube EDTA. Il contient un anticoagulant. Quand on fait une analyse ADN, on s'intéresse aux cellules nucléées et la récupération d'ADN est plus difficile si ces cellules sont coincées au milieu d'un caillot de sang. De plus, l'hémoglobine est un inhibiteur puissant de la Taq polymérase (enzyme de la PCR). On analyse le génome à la recherche d'une maladie génétique ou à la recherche de facteurs de risques associés à une maladie (cancer du sein ou Alzheimer). Le diagnostic d'une maladie génétique amène une consultation pour un conseil génétique.

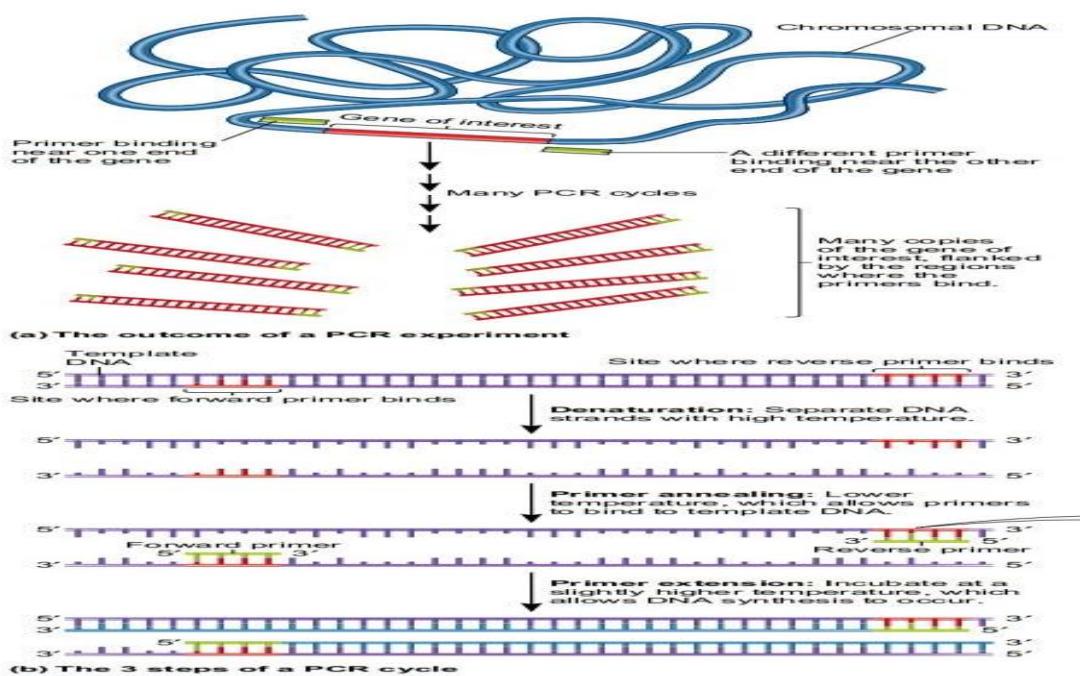
**1.2. PCR.** Ce diagnostic en biologie moléculaire peut être très délicat du fait de l'hétérogénéité génétique. En effet pour un seul gène on a plusieurs mutations et pour une maladie génétique plusieurs gènes peuvent être impliqués (**Fig.01**).



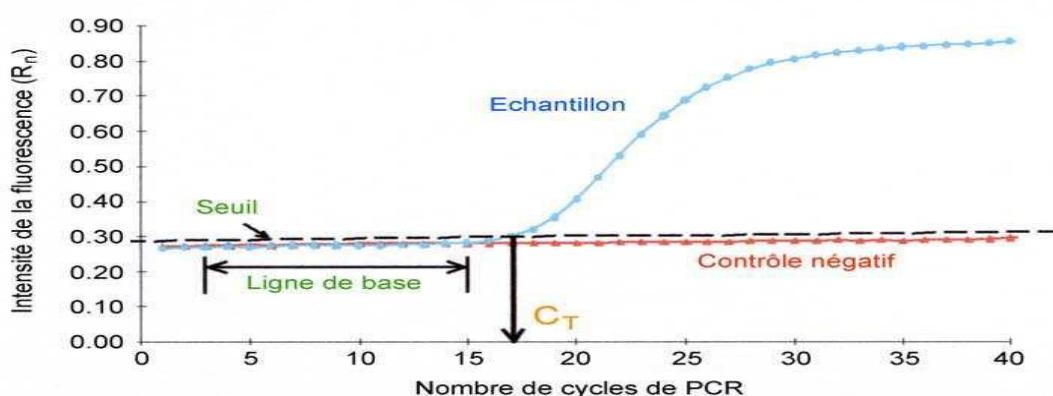
**Fig.01.** Amplification génique PCR

**1.2.1. Modalités de prélèvement.** Il faut faire très attention aux inhibiteurs d'enzymes (Taq polymérase+++) = hémoglobine et urée. Attention aussi aux nucléases qui dégradent les acides nucléiques et qui sont présentes partout (sur les mains++ +). La PCR est un outil de base en biologie moléculaire. Les prélèvements doivent être le plus stérile possible ! On peut prélever sur du sang, des cultures, les urines, le LCR, des expectorations, écouvillonage, ponctions, biopsies... ADN : robuste peut voyager par courrier à température ambiante. ARN : très fragile et sensible aux RNases. Il doit être congelé à -80 degrés et expédié en carboglace dans les 3h. On recueille les acides nucléiques dans des tubes spéciaux du type PAXgene. Les RNases sont ubiquitaires, difficiles à inactiver (renaturation après choc thermique), n'ont pas besoin de cofacteur pour fonctionner et sont présentes dans toutes les cellules.

**1.2.2. Les étapes de la PCR.** 3 étapes : -Dénaturation (94°C) -Hybridation des amorces (50-60°C) -Elongation (72°C) Dans le tube de réaction on trouve : 2 amorces, des dNTP, la Taq polymérase, un buffer (Mg<sup>2+</sup>) et le morceau d'ADN qu'on souhaite amplifier. L'amplification génique de la PCR peut être représentée par une équation mathématique exponentielle :  $Y=Ax2^n$  A=quantité initiale. n=nombre de cycles Y=produit de réaction En général n=30 cycles. Au bout de 30 cycles, on obtient plus d'un milliard de copies à partir de deux brins de départ. En PCR, on travaille dans trois pièces différentes par peur de contamination : une pièce pour la préparation de la PCR, une pour la PCR et une pour l'analyse post-PCR (**Fig.02**).

**Fig.02.** Etapes de la PCR

**1.2.3. PCR quantitative en temps réel.** La cassure avec le plateau dans la courbe de la PCR signifie que les amorces et les enzymes ont fini par s'épuiser. La formule  $Y=Ax2^n$  est donc valable au début de la réaction de PCR. L'équation réelle est  $Y=A(1+R)^n$  (avec  $R$ =efficacité de la réaction de PCR). En modifiant  $R$ , on modifie le rendement de la réaction de PCR. En PCR en temps réel, on se place le plus tôt possible dans la réaction pour détecter un signal (**Fig.03**). Car plus la réaction avance, plus on prend le risque que les valeurs soient fausses (à cause de la présence de possibles contaminants). On joue sur la longueur du produit de PCR : plus le produit est grand, plus l'efficacité de la PCR est inférieure à 1. Pour la PCR en temps réel on choisit des petits morceaux d'ADN. On évite aussi les séquences riches en GC (elles sont plus difficiles à amplifier).

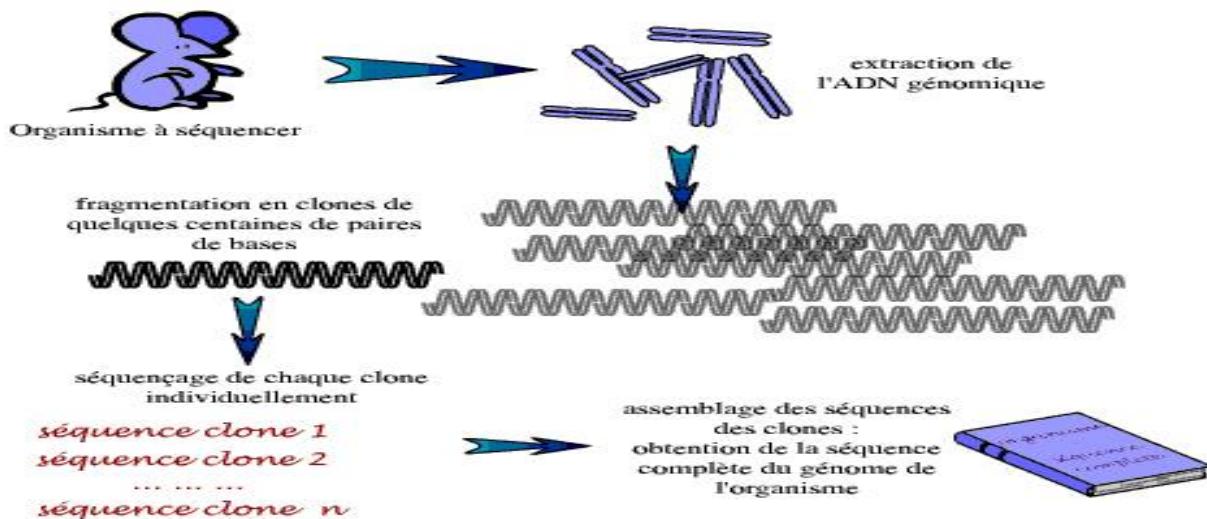
**Fig.03.** Principe de la quantification par PCR en temps réel

La PCR en temps réel permet d'observer les choses au fur et à mesure de l'amplification. On quantifie une fluorescence. En effet, on ne calcule pas la quantité Y de produit créé, mais le nombre de cycles qu'il aura fallu pour atteindre un seuil de fluorescence : le Ct. Ce seuil est déterminé par l'expérimentateur. Le Ct est inversement proportionnel à la quantité dans l'échantillon de départ. Sonde utilisée. La sonde d'hydrolyse (chimie Taq-Man). En plus des deux amorces, on rajoute une sonde fluorescente entre ces deux amorces qui contient deux fluorochromes : le quencher et le rapporteur. L'émission de base du rapporteur est coincée par le Quencher lors de la dénaturation. Lors de l'elongation, la Taq polymérase « grignote » le quencher et libère le rapporteur → émission de fluorescence.

**1.2.4. Application au virus de l'hépatite C.** Prévalence : 1% de la population. Transmission : transfusions toxicomanes Clinique : aigue ou chronique Traitement : interferon alpha et ribovirus. Il existe une variabilité génétique du VHC. Tests virologiques : test sérologique test par biologie moléculaire : détection des ARN, quantification des ARN (mesure charge virale), analyse quantitative du génome viral. Diagnostic. Inconvénients de la sérologie : l'apparition des AC est retardée. Pour les hémodialysés et les sujets immunodéprimés le taux d'AC est indétectable. Enfin, certains résultats sont douteux et ininterprétables. De plus, les AC peuvent être le reflet d'une infection ancienne ou guérie ou avoir été transmise passivement (transmission mère-enfant).

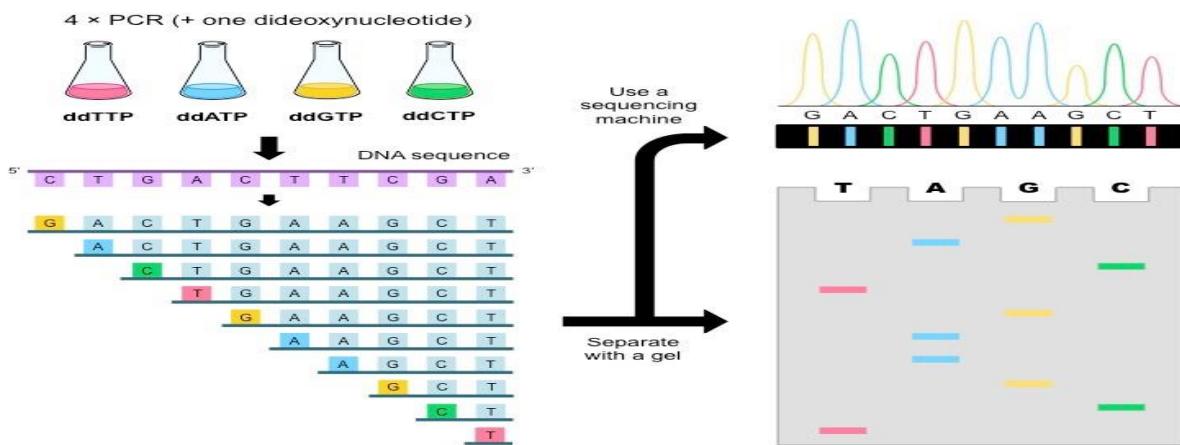
En biologie moléculaire, la détection d'ARN viral prouve qu'il existe une réplication du virus et donc qu'il s'agit d'une infection en cours. Indications de la biologie moléculaire : -diagnostic de patients immunodéprimés -tests sérologiques douteux -sérologie positive mais bilan hépatique normal -recherche de transmission mère-enfant chez une mère positive - détermination du rôle du VHC au cours d'hépatopathies associées à plusieurs causes possibles. La prise en charge thérapeutique et la durée du traitement dépend ainsi de la charge virale et du génotype.

**1.3. Séquençage.** Recherche d'un variant allélique, d'une mutation ou d'un polymorphisme. Il s'agit d'une technique qualitative. Le séquençage est une technique qui prend du temps, on ne fait pas quand la mutation est connue (il existe alors d'autres techniques plus rapides) (*Fig.04*).



*Fig.04.* Séquençage des génomes

**1.3.1. Principe du séquençage.** Le principe est le même que pour la PCR sauf qu'il n'a qu'une seule amorce, pas de sonde et qu'on ajoute au milieu réactionnel des ddNTP (didésoxynucléotides). Sur chaque ddNTP, on met un fluorochrome différent (cela permet de bien différencier les quatre ddNTP). La présence d'une seule amorce empêche la superposition des brins. La lecture se fait sur un système de capillaire (sorte de fil de pêche très fin) (*Fig.05*).



*Fig.05.* Principe de séquençage

**1.3.2. Interprétation.** Problème des variants : différence par rapport à la séquence d'origine enregistrée dans la base de données. Est-ce un polymorphisme ou une mutation ? Les altérations possibles sont des délétions, des insertions... Elles peuvent entraîner un décalage du code de lecture. Le décalage du cadre de lecture entraîne des protéines tronquées (apparition prématuée d'un codon stop).

**N.B** : si on a une délétion ou une insertion de 3 bases, on dit qu'elles ont équilibrées car elles n'entraînent pas de décalage du cadre de lecture. Dans l'interprétation, on différencie : -les mutations tronquantes (non-sens, décalage du cadre de lecture) -les mutations non-tronquantes ou faux-sens. Pour progresser dans l'interprétation des variants, on peut utiliser les données de la littérature, la corrélation génotype-phénotype, les études fonctionnelles, l'analyse *in silico* (à partir de logiciels de prédiction : cette mutation altère-t-elle la fonction de la protéine ?)

Aujourd'hui on classe les variations de génotype comme ceci : -Polymorphisme -Variant non délétère -Variant probablement non délétère -Variant de signification inconnue -Variant probablement délétère -Variant délétère -Mutation

**1.3.3. Application** : Néoplasie endocrinienne multiple de type 2 (NEM2). C'est une affection multiglandulaire. Dans 100% des cas, il y a une atteinte des cellules C de la thyroïde qui sécrètent la calcitonine. On a alors un cancer médullaire thyroïdien (CMT) avec hypersécrétion asymptomatique de calcitonine (diagnostic de la maladie très tardif et au stade de métastases).

On peut avoir en association avec le CMT un phéochromocytome et une hyperparathyroïdie. C'est une maladie héréditaire autosomique dominante avec une pénétrance de 100%. La prévalence est de 0,2/1000. 3 catégories de NEM (et donc 3 phénotypes). NEM2A (60%) : CMT, phéochromocytome, hyperparathyroïdie et notalgia. NEM2B (5%) : CMT, phéochromocytome, phénotype marfanoïde, neurones sous-muqueux, ganglioneuromatose digestive. (Forme très agressive). FMTC (35%) : CMT isolé, peu agressif. Il y a une corrélation génotype-phénotype ! Le gène impliqué dans la maladie est le gène RET. Il code pour un récepteur transmembranaire à activité tyrosine-kinase. Un de ses ligands est le GDNF. Il agit par l'intermédiaire d'un co-récepteur GDNF-alpha. Ce récepteur agit sous forme d'hétérodimère, avec phosphorylation de la tyrosine-kinase et activation de voies de transduction.

2 types de mutation : l'homodimérisation du récepteur (mutation inactivatrice) ou une transautoactivation du domaine TK (mutation activatrice). Une fois muté, RET devient un oncogène. Cas index= 1er cas de la maladie dans la famille. Devant tout CMT, analyse systématique de RET. On recherche un diagnostic syndromique et une forme familiale. Mais

certains CMT sont sporadiques et nontransmissibles. Traitement du cas index : On fait une thyroïdectomie totale et un curage ganglionnaire. La chimiothérapie est peu efficace. On recherche des pathologies associées (phéochromocytome, HPT...). Pronostic : la survie des patients non guéris est de 81,5% à 5 ans et de 43,7% à 10 ans. Traitement des sujets génétiquement à risque : si présence de la maladie, idem que pour le cas index si maladie pas encore déclarée, thyroïdectomie prophylactique.

**N.B** : en oncogénétique on fait 2 prélèvements indépendants. Si la mutation de RET est identifiée, on fait un dépistage génétique familial avec toujours deux prélèvements (évite les faux positifs ou faux négatifs).