

## CHAPITRE VII

### TECHNOLOGIES DE DIAGNOSTIC MOLECULAIRE

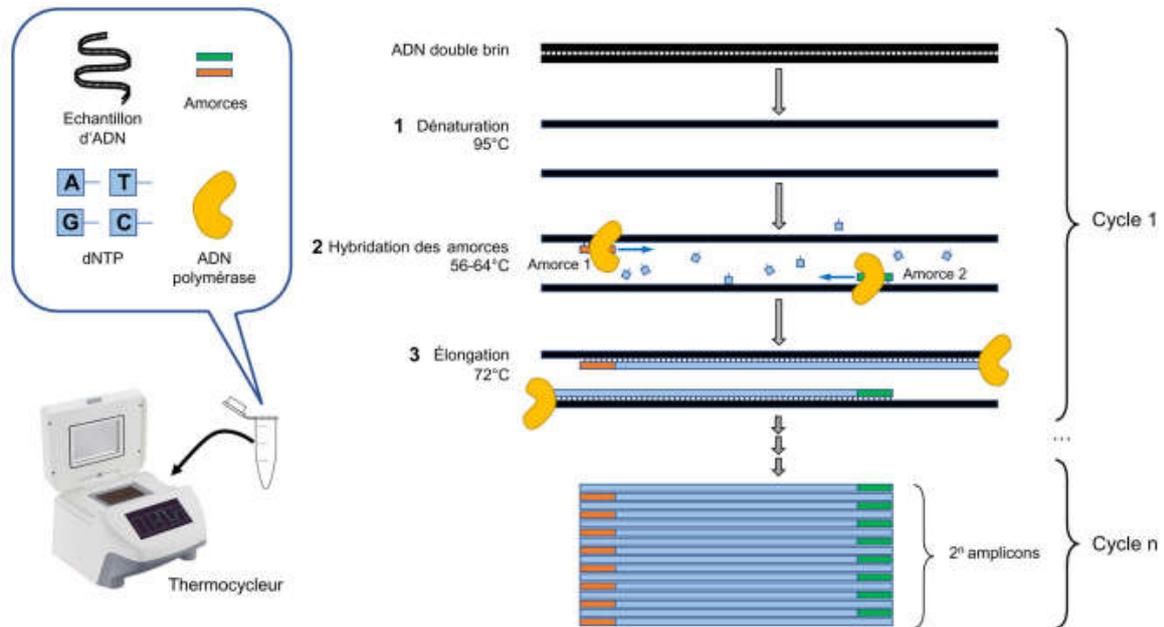
#### 1. Introduction

La biologie moléculaire désigne l'étude des acides nucléique, ribonucléique (ARN) et désoxyribonucléique (ADN). Ces techniques reposent essentiellement sur l'hybridation moléculaire, sur la réaction de polymérisation en chaîne (PCR) et sur le séquençage des acides nucléiques. Depuis leur apparition dans les années 1970, elles ont connu un développement exponentiel, et permettent aujourd'hui de recueillir en quelques heures les données à l'échelle du génome entier. Nous nous intéresserons ici aux avancées les plus récentes et à leurs applications cliniques ou en recherche translationnelle appliquée à la réanimation.

#### 2. Amplification par réaction de polymérisation en chaîne (PCR)

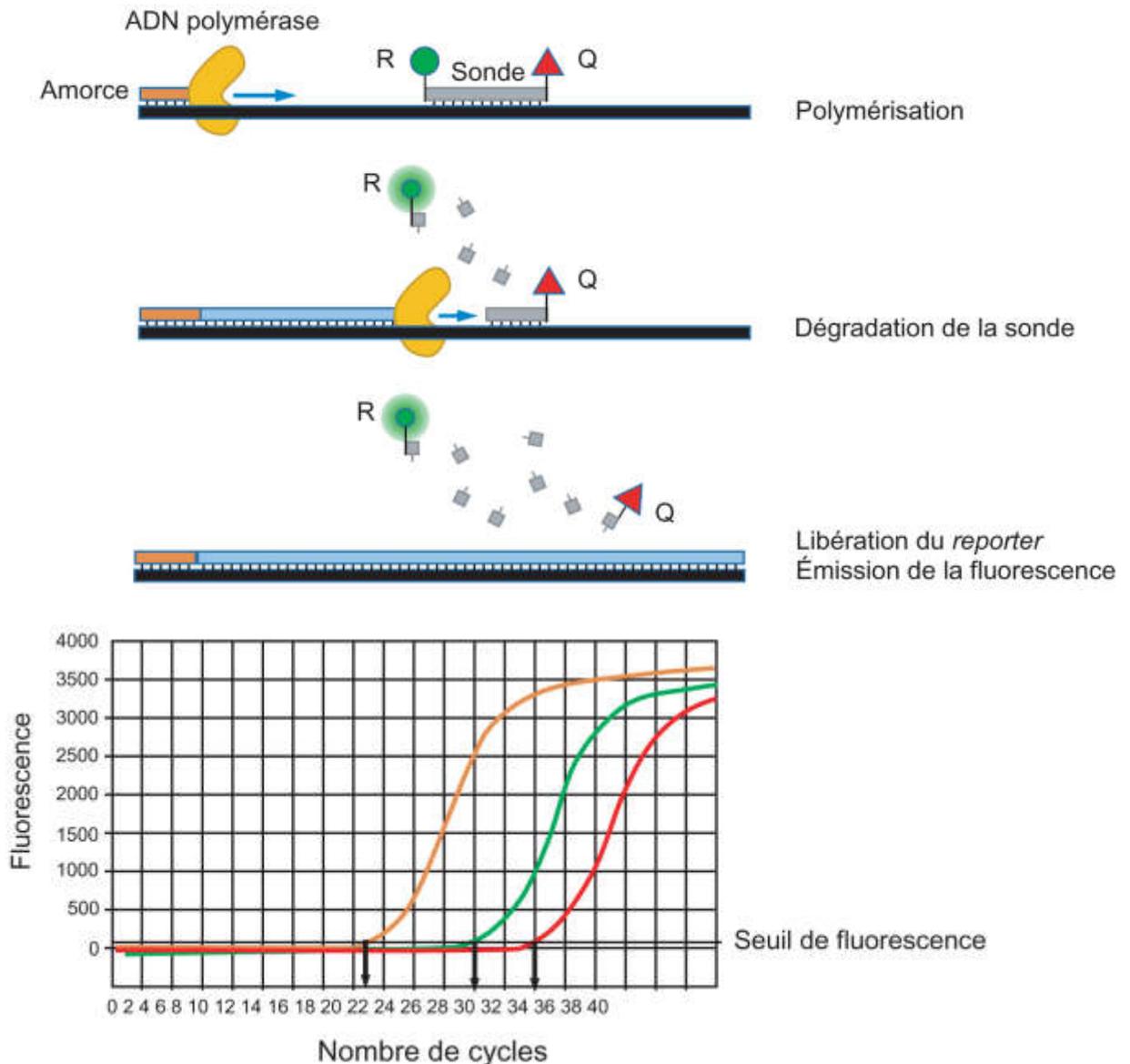
##### 2.1. Principe de la PCR

La PCR (polymerase chain reaction) est fondée sur une réaction enzymatique *in vitro* permettant d'amplifier plusieurs millions de fois des séquences d'ADN génomique ou d'ADN complémentaire (ADNc, généré par transcription inverse à partir d'une molécule d'ARN). Elle nécessite de connaître les séquences des extrémités de la région à amplifier, et de les utiliser pour élaborer des amorces nucléotidiques. Après fixation des amorces sur leurs séquences cibles par complémentarité, la région d'ADN cible est copiée par incorporation de désoxyribonucléotides libres grâce à une enzyme ADN polymérase thermostable (Fig. 1). Des variations cycliques de températures vont permettre successivement la dénaturation de l'ADN, l'hybridation des amorces et leur extension par l'ADN polymérase. La succession des cycles permet une augmentation quasi exponentielle de la quantité d'ADN. La réaction, réalisée de façon automatique grâce à un thermocycleur, permet de générer en moins d'une heure des millions de copies de la séquence cible à partir de très faibles concentrations initiales d'ADN (de l'ordre de 10–15 mol/l) [1]. La PCR, initialement dédiée à l'analyse de l'ADN, peut également s'appliquer à l'étude des ARN. Une étape initiale de transcription inverse (reverse transcription [RT]) par une ADN polymérase ARN dépendante (transcriptase inverse) permet de copier une molécule d'ARN en une molécule d'ADN complémentaire (ADNc) qui servira de matrice à la réaction de PCR.



**Fig. 1 Principe de la PCR 1.** Dénaturation : une première étape de chauffage à 95 °C permet de séparer les deux brins d'ADN, d'homogénéiser le milieu réactionnel et d'inhiber d'autres enzymes potentiellement présentes dans la solution. La Taq ADN polymérase est en revanche stable à cette température. 2. Hybridation des amorces (56–64 °C) : les amorces s'apparient spécifiquement avec leur brin complémentaire d'ADN dénaturé. 3. Élongation (72 °C) : l'ADN polymérase synthétise le brin complémentaire de chaque brin d'ADN matrice à partir des désoxyribonucléotides (dNTP) libres présents dans le milieu réactionnel. 4. Lors du cycle suivant, la nouvelle étape de dénaturation sépare les brins d'ADN matrices des brins néosynthétisés qui deviennent à leur tour matrice de la réaction suivante. 5. La répétition du cycle de variation de température environ 25 à 35 fois permet d'amplifier la quantité totale d'ADN correspondant à la séquence cible de façon quasi exponentielle

La technique de PCR quantitative en temps réel (qPCR) permet, parallèlement à l'amplification, la détection et la quantification de l'ADN. Elle repose sur l'utilisation d'un marqueur fluorescent (par exemple SYBR Green ou sonde Taqman®) dont le signal augmente proportionnellement à la quantité d'ADN synthétisée au cours de la réaction (Fig. 2).



**Fig. 2 PCR en temps réel avec sonde spécifique** Parallèlement à la paire d'amorces, une troisième sonde spécifique (ici de type Taqman®) se fixe au centre de la région à amplifier. Cette sonde possède un fluorophore rapporteur (reporter, R) et un désactivateur (quencher, Q) qui inhibe l'émission de fluorescence par le rapporteur lorsqu'il est situé à proximité. Au cours de l'élongation, l'activité exonucléase de l'ADN polymérase dégrade la sonde hybridée au brin matrice, permettant la libération du rapporteur et l'expression de sa fluorescence. La cinétique d'émission de la fluorescence est proportionnelle à la quantité d'amplicons produits au cours de la réaction. Plus une séquence est abondante, plus elle sera détectée précocement. Parallèlement au gène d'intérêt, la mesure sur un même échantillon du niveau d'expression d'un gène de ménage (dont l'expression est constante au sein du tissu étudié et selon les conditions physiologiques) permet de s'assurer de la qualité de l'échantillon et de quantifier de façon relative l'expression du gène d'intérêt.

## 2.2. Applications de la PCR en clinique et en recherche

Par sa simplicité et sa rapidité, la PCR a rapidement trouvé de nombreuses applications cliniques, en particulier dans le domaine des maladies infectieuses [2] :

- diagnostic précoce des infections bactériennes, virales ou fongiques : détection de *Mycobacterium tuberculosis* dans les prélèvements respiratoires [3], diagnostic et suivi thérapeutique des infections virales (VIH, hépatites, CMV...);
- diagnostic des endocardites infectieuses à hémocultures négatives. La PCR sur tissu valvulaire (en particulier la PCR dite « universelle » du gène codant pour l'ARN ribosomal 16S) permet d'obtenir une documentation microbiologique dans plus de 60 % des cas [4] (Fig. 3);
- la détection de plusieurs séquences en une seule réaction (PCR multiplex) constitue la dernière avancée majeure de la technique. Le panel FilmArray méningite/encéphalite permet par exemple de rechercher en moins d'une heure 14 pathogènes responsables de méningoencéphalites sur un volume de 200  $\mu$ l (soit quatre gouttes) de liquide céphalorachidien [5].

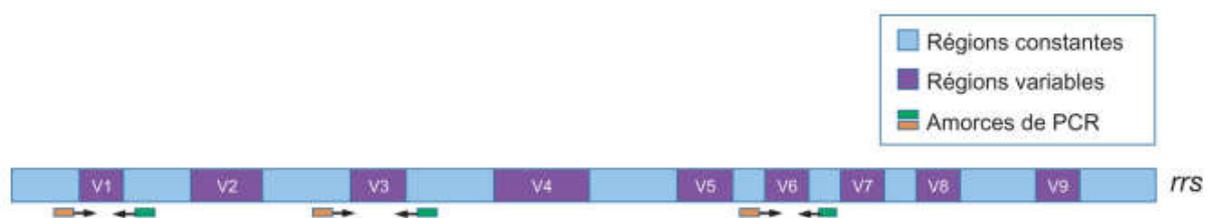


Fig. 3 PCR du gène codant pour l'ARN ribosomal 16S Le gène *rrs* qui code l'ARNr 16S est présent chez l'ensemble des espèces bactériennes, en un nombre variable de copies. Il est constitué de séquences conservées chez toutes les eubactéries, et de séquences variables. L'amplification de l'ARNr 16S à l'aide d'amorces universelles complémentaires des régions conservées permet de détecter la présence de n'importe quelle espèce bactérienne. Le séquençage des régions variables permet, après comparaison des séquences avec des banques de données, d'identifier l'espèce bactérienne

## 3. Microarrays

### 3.1. Concept général

La technologie des microarrays, ou biopuces, est une méthode permettant d'analyser simultanément plusieurs milliers de gènes au sein d'un seul échantillon. Elle se décline sous de nombreuses formes permettant l'analyse de l'ADN (génomique) ou de l'ARN (transcriptome) (Tableau 1).

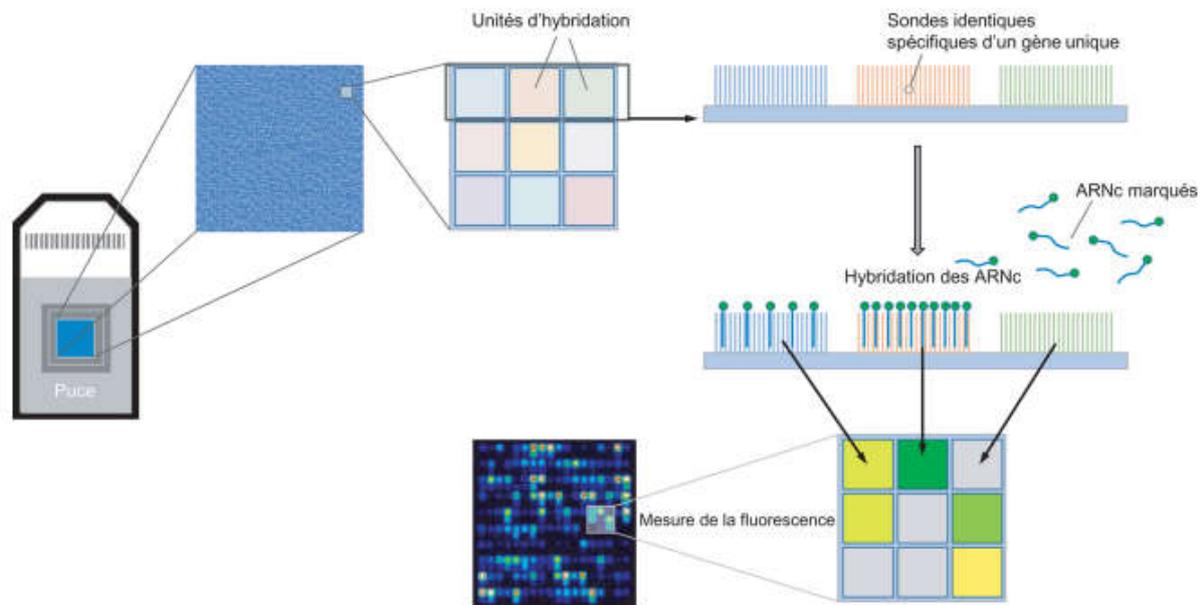
Type de puce	Cible	Description
Transcriptomique ( <i>gene expression profiling microarrays</i> )	ARN	Mesure du profil d'expression des gènes. Étude simultanée de l'ensemble des ARN messagers ± ARN non codants
Génotypage	SNPs	Génotypage des polymorphismes nucléotidiques ( <i>single-nucleotide polymorphisms</i> [SNPs])
Méthylation <i>Chromatin immuno-precipitation</i> (ChIP)	Ilots CpG Sites de fixation de l'ADN	Quantification de la méthylation de l'ADN Étude de l'interaction entre protéines et ADN (ex. : facteurs de transcription et promoteurs)
Hybridation génomique comparative (CGH)	ADN	Diagnostic des duplications anormales de l'ADN

Ce type d'analyses à haut débit a fait la preuve de son intérêt dans la classification des cancers et la prédiction de la réponse aux traitements [6]. Exemple des puces transcriptomiques Contrairement à l'ADN, dont la séquence est largement similaire entre les cellules d'un même organisme, le contenu en ARN est hautement variable selon le tissu étudié et le contexte physiopathologique. Alors que la PCR ciblée permet d'étudier les fonctions de quelques gènes, les puces transcriptomiques permettent de savoir, à l'échelle du génome dans un compartiment cellulaire donné, quels gènes sont actifs, et dans quelle proportion ils sont transcrits. Les puces sont des matrices de verre sur lesquelles sont fixées de courtes séquences nucléotidiques appelées sondes (probes), organisées en unités d'hybridation (Fig. 4).

Chaque unité contient plusieurs millions d'exemplaires de la même sonde. Les ARN extraits d'une population cellulaire sont marqués à la biotine puis déposés sur la puce. Ils se fixent dans les unités d'hybridation contenant des sondes qui leur sont spécifiques. Les ARN sont ensuite révélés à l'aide d'un fluorochrome, et l'intensité de fluorescence émise par chaque unité est mesurée par un module de criblage. L'intensité de fluorescence émise dépend du nombre de molécules d'ARN spécifiquement appariées.

### 3.2. Applications en clinique/recherche

Les données générées par biopuces sont particulièrement adaptées à la découverte de nouveaux processus physiopathologiques ou à la recherche de biomarqueurs. Par exemple, en comparant les profils transcriptomiques leucocytaires sanguins de patients atteints de pneumonie à ceux de patients admis en réanimation avec une présentation clinique semblable mais sans infection, il a été possible de découvrir une signature de 78 gènes spécifiques du diagnostic de pneumonie et d'identifier le ratio d'expression génique FAIM3: PLAC8 comme biomarqueur d'infection plus sensible que la procalcitonine [7]. Une approche similaire a permis de mettre au point le test moléculaire SeptiCyte LAB qui mesure simultanément le niveau d'expression des gènes CAECAM4, LAMP1, PLA2G7 et PLAC8, permettant de distinguer les patients admis en réanimation pour un sepsis des patients admis pour un syndrome de réponse inflammatoire systémique d'origine non infectieuse [8].



**Fig. 4 Puces transcriptomiques.** Les ARN extraits d'une population cellulaire sont transcrits en ADNc par transcription inverse (les ADNc, plus stables, peuvent être conservés plus longtemps) puis les ADNc sont à nouveau transcrits en ARN complémentaires (ARNc) marqués à la biotine. Les ARNc sont déposés sur la puce et s'hybrident uniquement dans les régions contenant des sondes qui leur sont spécifiques. L'ARNc hybridé est coloré à l'aide d'une molécule fluorescente (streptavidine–phycoérythrine) qui se lie à la biotine. Les puces sont scannées avec un laser. Un module de criblage mesure l'intensité de fluorescence émise par chaque unité d'hybridation. Chaque unité est spécifique d'un gène. Seules les unités contenant de l'ARN émettent de la fluorescence, et l'intensité de fluorescence de chaque unité est directement proportionnelle au nombre de molécules d'ARN présentes dans l'échantillon, c'est-à-dire au niveau d'expression de chaque gène.

## 4. Séquençage de l'ADN

### 4.1. Principe général

Depuis la description par Sanger en 1977, les techniques de séquençage ont fait l'objet d'innovations permettant d'en améliorer le rendement, la fiabilité et la rentabilité [11]. Regroupées sous le nom générique de séquençage haut débit ou next generation sequencing (NGS), les stratégies disponibles reposent sur trois étapes (Fig. 5) :

- réalisation d'une librairie : l'ADN est découpé en fragments ligaturés par des adaptateurs ;
- amplification clonale : une PCR dont les amorces ciblent les adaptateurs permet de multiplier le nombre de fragments matriciels ;

- le séquençage parallèle : les fragments d'ADN sont lus et traduits en séquences de nucléotides.

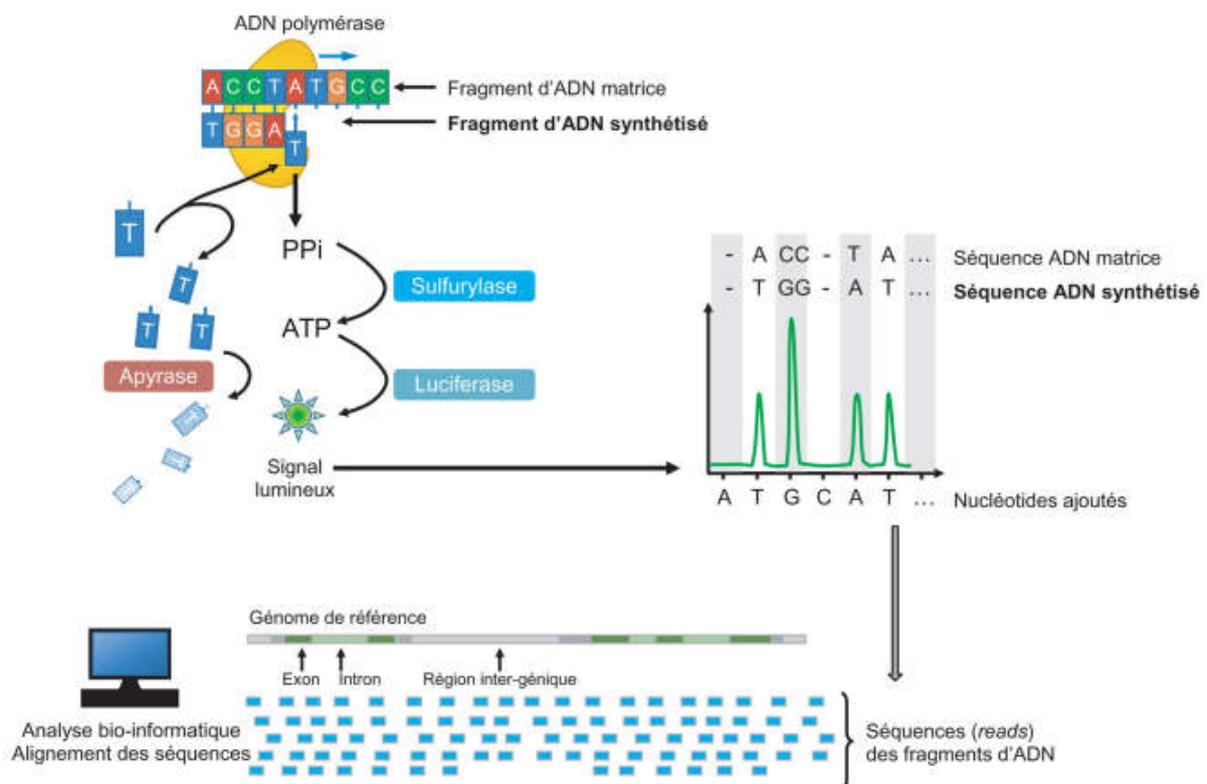


Fig. 5 Séquençage de l'ADN à haut débit (exemple du pyroséquençage) Les méthodes de séquençage à haut débit (next generation sequencing [NGS]) nécessitent une fragmentation du génome en petites séquences de quelques centaines de bases qui sont amplifiées par PCR. Le séquençage est réalisé à l'aide d'un ADN polymérase qui va synthétiser un brin complémentaire du fragment à séquencer. Contrairement à une PCR conventionnelle, les désoxyribonucléotides sont ici apportés séparément de façon séquentielle dans le milieu réactionnel. Si le nucléotide présent est celui attendu par la polymérase (nucléotide complémentaire de l'ADN matrice), il s'apparie en libérant un pyrophosphate inorganique (PPi). La libération de PPi est détectée par une cascade de réactions enzymatiques aboutissant à l'émission de lumière par bioluminescence. Si le nucléotide présenté n'est pas celui attendu par la polymérase, il ne s'incorpore pas, et aucun PPi n'est libéré. Les nucléotides non incorporés sont détruits par une apyrase, puis on apporte successivement les nucléotides suivants (A, C, G, T...). L'enregistrement du signal lumineux émis ou pas à chaque nucléotide permet de déduire la séquence de la matrice. La taille du pic d'émission est proportionnelle au nombre de nucléotides consécutivement incorporés. Les systèmes de NGS permettent de séquencer simultanément plusieurs centaines de fragments d'ADN. Les fragments de séquences sont ensuite réalignés sur un génome de référence à l'aide d'outils bio-informatiques

## 4.2. Applications cliniques et en recherche

Le NGS est utile au diagnostic des maladies d'origine génétique. Selon la pathologie ciblée, il est possible de séquencer un groupe de gènes précédemment identifiés (maladies épileptiques héréditaires par exemple), l'exome (génome codant pour des protéines) ou bien le génome entier [12]. Le reséquençage des génomes humains permet d'étudier l'impact des polymorphismes génétiques sur la santé à l'échelle des populations (projet 1 000 genomes [13]), ou sur la réponse immunitaire à l'échelle cellulaire (projet Human Functional Genomics [14]). Le NGS constitue enfin la pierre angulaire de la métagénomique appliquée à la clinique (séquençage des micro-organismes présents dans un échantillon biologique dans un but diagnostique) ou l'étude des microbiotes (écosystème des micro-organismes qui colonisent différents compartiments de l'organisme) et est en passe de révolutionner la microbiologie [15,16].

## 5. Séquençage de l'ARN (RNA-seq)

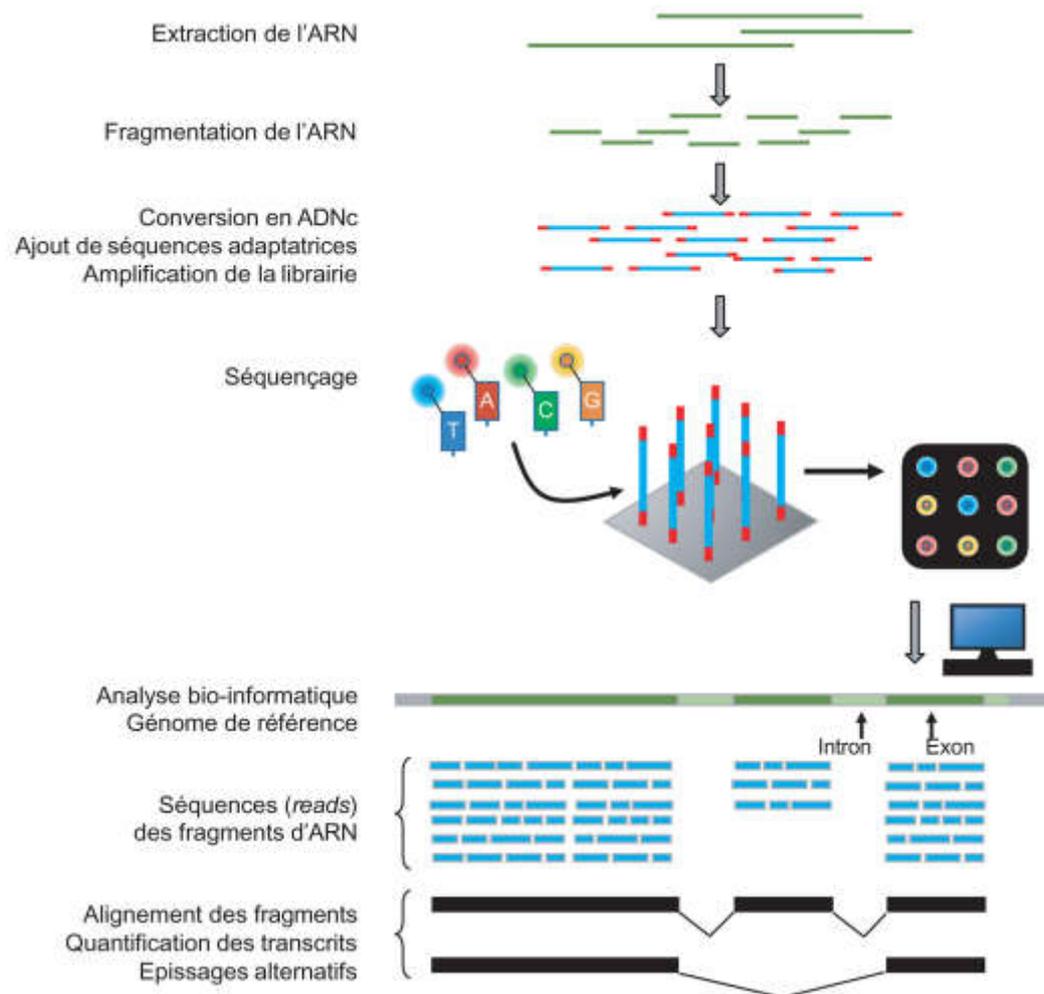
### 5.1. Principe du RNA-seq

L'application des méthodes de séquençage à l'étude des ARN apporte une dimension nouvelle à l'étude du transcriptome. Alors que l'utilisation des puces repose sur la conception de sondes nucléotidiques, le RNA-seq permet de déterminer la séquence des transcrits sans connaissance préalable du gène d'intérêt. Le RNA-seq repose sur trois étapes (Fig. 6) :

- l'extraction et le fractionnement des ARN, suivis de la constitution d'une librairie d'ADNc ;
- séquençage proprement dit ;
- analyse bio-informatique (alignement, assemblage et quantification) des millions de séquences générées.

### 5.2. Applications cliniques

La méthode RNA-seq est utile à l'identification de transcrits rares, à l'étude des épissages alternatifs ou encore des multiples formes d'ARN (ARN codants pour des protéines et ARN non codants qui régulent le fonctionnement cellulaire) [17,18]. En microbiologie, le RNA-seq permet d'étudier des agents pathogènes émergents tels que le virus Zika [19]. En cancérologie, la technique de RNA-seq a permis d'établir des signatures capables de prédire la réponse au traitement [20]. Ces signatures sont en train de trouver une place essentielle dans la stratégie de prise en charge des cancers du sein ou du côlon, et pourraient prochainement trouver leur place dans la prise en charge des malades de réanimation [21].



**Fig. 6 Étapes du RNA-seq** 1. Préparation de la librairie d'ADNc : après purification, les ARN sont fragmentés en courtes séquences (les procédés de séquençage ne peuvent prendre en charge que de courts fragments, de l'ordre de 200 à 300 bases alors que les transcrits d'ARN peuvent contenir plusieurs milliers de bases). Les fragments d'ARN sont convertis en ADNc double brin (plus stable et plus faciles à amplifier et manipuler que l'ARN). Les fragments sont ligaturés de part et d'autre par des adaptateurs de séquençage qui permettent à l'automate de reconnaître les fragments et de séquençer plusieurs échantillons en même temps. Les ADNc sont ensuite amplifiés. 2. Séquençage : la plateforme Illumina™ utilise des nucléotides marqués avec des fluorochromes différents. Les fluorochromes incorporés au brin synthétisé sont visualisés un par un à l'aide d'un lecteur optique. 3. Analyse bio-informatique des séquences. Les millions de séquences de fragments d'ARN sont réalignés et assemblés. Il est ensuite possible de quantifier les niveaux d'expression des transcrits, de détecter des altérations génomiques, de relever des variations de séquence non détectables à l'échelle génomique, telles que l'expression de transcrits alternatifs.

## 6. Conclusions et perspectives

La rapidité, la sensibilité et la spécificité des techniques de biologie moléculaire en ont fait des outils de choix dans le domaine des maladies infectieuses et de la réanimation. Alors que

le premier séquençage du génome humain a nécessité 13 années de travaux et coûté trois milliards de dollars, les techniques de nouvelle génération permettent aujourd'hui d'obtenir en moins de 24 heures et pour quelques centaines d'euros les séquences complètes des ADN ou ARN d'un échantillon. L'intégration de ces nouvelles informations permet désormais d'appréhender plus finement la réponse de l'hôte à diverses situations d'agression, de découvrir de nouveaux biomarqueurs diagnostiques ou thérapeutiques ou de mieux comprendre les susceptibilités individuelles dans des contextes cliniques particuliers (Tableau 2). La mise en application de ces techniques soulève toutefois des questions éthiques intéressantes autant la propriété des données issues du séquençage que leur utilisation dans la prise en charge médicale. Dans un futur proche, ces nouvelles techniques de biologie moléculaire pourraient permettre une prise en charge encore plus rapide et personnalisée, adaptée autant à la situation clinique qu'à la réponse spécifique en temps réel de chaque patient.

Technique	Principe	Applications courantes
<i>Polymerase chain reaction (PCR)</i>	Amplification/détection d'une séquence ciblée d'ADN ou d'ARN (RT-PCR)	Détection d'agents infectieux
Biopuces ( <i>microarrays</i> )	Analyse simultanée de plusieurs milliers de gènes	Recherche de biomarqueurs Analyses de polymorphismes Analyses de signatures transcriptomiques
Séquençage ADN à haut débit ( <i>DNA next generation sequencing</i> )	Séquençage nucléotide par nucléotide de séquences ou de l'intégralité du génome	Diagnostic de maladies d'origine génétique Détection d'agents pathogènes Étude des microbiotes Étude des pathogènes émergents
Séquençage ARN ( <i>RNA-seq</i> )	Séquençage nucléotide par nucléotide des ARN codants et non codants (pour des protéines)	Classification moléculaire des cancers Étude du rôle des ARN non codants Étude des transcrits rares et épissages alternatifs Classification moléculaire des cancers