

Titre : Séquençage à haut débit pour le diagnostic en maladies infectieuses : exemple de la métagénomique *shotgun* dans les infections du système nerveux central.

Auteurs : Sarah Marchand^{1,2}, Christophe Rodriguez^{1,2,3}, Paul-Louis Woerther^{1,2,4}

Affiliations / Coordonnées :

- 1- Département de Microbiologie, Hôpital Henri Mondor, AP-HP, Créteil, France
- 2- Plateforme de Génomique, Hôpital Henri Mondor, AP-HP, Créteil, France
- 3- INSERM U955, Université Paris-Est Créteil, Créteil, France
- 4- EA 7380 Dynamyc, Université Paris-Est Créteil, Créteil, France.

Sarah Marchand : sarah.marchand@chu-nantes.fr

Séquençage à haut débit pour le diagnostic en maladies infectieuses : exemple de la métagénomique *shotgun* dans les infections du système nerveux central.

ABSTRACT

L'arrivée du séquençage à haut débit en microbiologie clinique ouvre la voie à de nouvelles approches diagnostiques et pronostiques dans le domaine des maladies infectieuses. La détection, l'identification et la caractérisation des microorganismes pathogènes constituent les étapes majeures dans le diagnostic et la mise en place d'un traitement antimicrobien adapté. Cependant, les méthodes standard de diagnostic microbiologique sont dans certains cas prises en défaut. De plus, l'émergence de nouvelles infections, favorisée par les voyages internationaux et le réchauffement climatique, rendent nécessaire la mise en place de méthodes de diagnostic innovantes. Parmi les différentes stratégies utilisées en microbiologie clinique et exposées dans cet article, la métagénomique *shotgun* est la seule technique permettant aujourd'hui une détection panpathogène et sans *a priori*, c'est-à-dire de l'ensemble des microorganismes potentiellement responsables d'une maladie infectieuse, incluant ceux encore inconnus. Cet article a pour objectifs d'exposer les différentes stratégies possibles du séquençage à haut débit utilisé dans le diagnostic microbiologique des maladies infectieuses, et de mettre en avant la contribution diagnostique de la métagénomique *shotgun* dans le domaine des infections du système nerveux central.

Mots clés : microbiologie clinique;NGS;méningo-encéphalites;métagénomique *shotgun*.

1- Introduction

Dans le domaine des maladies infectieuses, le diagnostic microbiologique (incluant détection, identification et caractérisation du microorganisme) est indispensable à la prise en charge optimale du patient. Afin de remplir ces objectifs, les laboratoires sont équipés d'un large éventail de méthodes dépendantes ou non de la culture. Dans l'ensemble, les techniques de routine comprennent d'une part les techniques de mise en évidence directe des microorganismes responsables, comme la mise en culture d'un échantillon biologique, la détection d'antigènes spécifiques ou la biologie moléculaire (PCR simplex ou multiplex, séquençage Sanger, séquençage à haut débit) et d'autre part les techniques indirectes comme la sérologie infectieuse.

Les maladies infectieuses se présentent sous la forme de syndromes la plupart du temps faiblement corrélés à la nature des agents pathogènes responsables. C'est pour cette raison que le diagnostic microbiologique, qui permet de choisir le traitement antimicrobien le mieux adapté au microorganisme responsable, revêt une si grande importance. En raison du très grand nombre de microorganismes potentiellement responsables, le bilan comprend souvent de nombreuses recherches effectuées dans plusieurs laboratoires différents et peut être limité par la quantité d'échantillon disponible, comme c'est le cas pour les LCS (liquide cébrospinal) dans le cadre des infections du système nerveux central (SNC). Parmi les méthodes disponibles, la culture reste le « gold standard » pour le diagnostic étiologique des infections bactériennes ou fongiques, bien qu'elle puisse être limitée par l'administration

préalable d'un traitement antimicrobien au patient ou prise en défaut en cas d'infection due à des microorganismes dont la croissance exige des conditions particulières. Les tests moléculaires tels que la PCR sont rapides et spécifiques, mais supposent une connaissance *a priori* des microorganismes suspectés d'être la cause de l'infection, qu'ils soient de nature bactérienne, virale, fongique ou parasitaire. Enfin, les PCR dites universelles (ciblant par exemple le gène *rrs* pour les bactéries ou les régions ITS pour les champignons) suivie d'un séquençage Sanger du produit d'amplification, bien que peu coûteuse, n'est contributive qu'en cas d'infection monomicrobienne et ne donne pas d'information sur la sensibilité aux antimicrobiens du microorganisme mis en évidence. De plus, la discrimination au niveau de l'espèce peut s'avérer difficile au sein de certains genres bactériens. Enfin, l'absence de gène viral universel empêche l'utilisation de ce type d'approche pour le diagnostic virologique.

Afin de simplifier les recherches et d'accélérer les délais de rendu des analyses microbiologiques, sont apparues les PCR multiplex dites « syndromiques ». Ces approches reposent, pour chaque syndrome infectieux (respiratoire, neuroméningé, intestinal etc...), sur la recherche simultanée des principaux microorganismes pathogènes incriminés à partir d'un échantillon unique. Bien que rapides et performantes, ces approches sont limitées par le nombre des pathogènes ciblés, et ne permettent pas la mise en évidence d'un microorganisme variant, rare ou émergent.

Le développement du séquençage de nouvelle génération (NGS pour « *Next Generation Sequencing* ») ouvre la voie à de nouvelles approches dans le domaine du diagnostic en maladies infectieuses. Les procédés techniques reposent sur deux stratégies différentes. La première approche est le *target* NGS, stratégie ciblée permettant l'amplification et le séquençage d'une ou plusieurs régions précises d'un (ou plusieurs) génome(s) donné(s). Cette approche est notamment utilisée pour l'identification microbienne (gène codant pour l'ARNr 16S pour l'identification bactérienne par exemple) ou la recherche de résistances (gène(s) associé(s) à des mutations de résistance virales par exemple), mais également pour le séquençage de génomes complets de petite taille comme ceux des virus (génome complet du SARS-CoV-2 pour l'étude épidémiologique de variants, par exemple). Il existe aussi à l'inverse une autre approche, non ciblée, qui est appelée *shotgun* NGS. Elle repose sur une première étape de fragmentation aléatoire des acides nucléiques présents dans un échantillon, suivie par le séquençage de l'ensemble du matériel ainsi obtenu. Elle est utilisée soit pour le séquençage complets de grands génomes comme ceux des bactéries ou des champignons à partir d'une culture (aussi appelé séquençage de souche), soit pour la métagénomique *shotgun* (SMg). Cette dernière approche repose sur le séquençage et l'analyse de l'ensemble du matériel génétique microbien (ADN + ARN) obtenu à partir d'un échantillon biologique et permet l'identification du ou des microorganismes qu'il contient à des fins de diagnostic microbiologique. Lorsque la quantité de séquences obtenues est suffisante, elle peut aussi permettre de reconstituer les génomes complets et de caractériser ainsi les microorganismes présents dans l'échantillon de départ, qu'ils soient connus ou non encore décrits. En effet, cette approche « à large spectre » sans *a priori* est capable de caractériser de nouveaux microorganismes, et devient alors incontournable dans le contexte d'émergence de nouvelles maladies infectieuses [1]. A titre d'exemple, dans le cadre de l'exploration des infections du SNC, les approches traditionnelles ne permettent d'obtenir un diagnostic que dans environ 50% des méningo-encéphalites aiguës [2]. Disposer de nouveaux outils de diagnostic microbiologique panpathogènes et sans *a priori* pourrait donc permettre d'augmenter les capacités diagnostiques à la disposition du clinicien et améliorer la prise en charge de ces infections.

2- Stratégies NGS utilisées pour le diagnostic microbiologique clinique et leurs principales applications

Le NGS regroupe des approches de séquençage basées sur deux principales stratégies, le *target* NGS ou le *shotgun* NGS, conduisant à applications différentes utilisées dans le diagnostic microbiologique.

A- Principes des stratégies *target* NGS et *shotgun* NGS

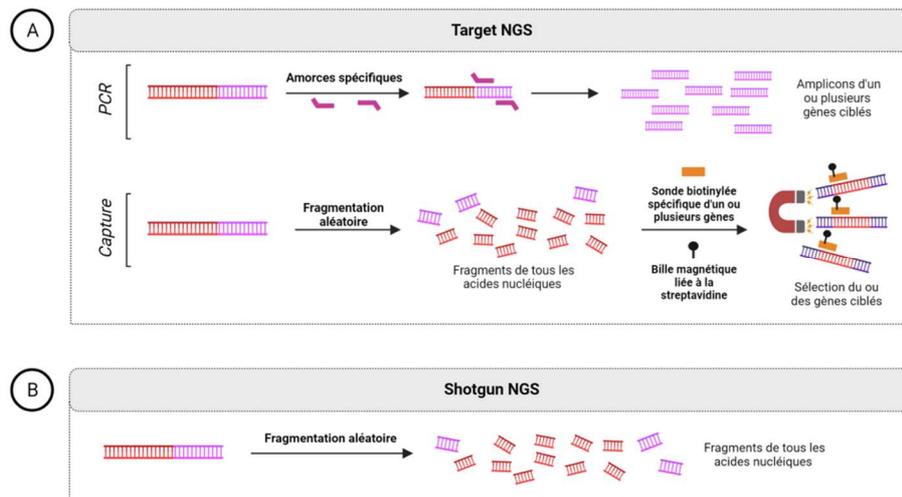


Fig. 1. Stratégies de séquençage à haut débit utilisées pour le diagnostic microbiologique clinique.

A. La technique *target* NGS est une stratégie ciblée de séquençage NGS. Deux méthodes peuvent être utilisées pour cibler une ou des régions précises ou un type de microorganismes particulier : la PCR et la capture.

B. La technique *shotgun* NGS est, contrairement au *target* NGS, une stratégie non ciblée de séquençage à haut débit. L'ensemble des acides nucléiques de l'échantillon biologique va subir une fragmentation *shotgun*, qui consiste à fragmenter de façon aléatoire l'ensemble des génomes présents. Tous les petits fragments générés sont ensuite séquencés.

“Created with BioRender.com”

Le *target* NGS est une technique qui permet de cibler par PCR ou capture une (ou plusieurs) région(s) d'intérêt (Fig. 1A). Concernant la PCR, grâce à des amorces spécifiques, la ou les régions ciblées sont amplifiées et le (ou les) amplicon(s) obtenu(s) est (sont) séquencé(s). L'enrichissement par capture consiste quant à elle à utiliser des sondes qui viennent s'hybrider sur une ou plusieurs région(s) ciblée(s) correspondant au(x) microorganismes recherché(s). Ces sondes sont biotinylées et sont utilisées sur des fragments d'acides nucléiques générés aléatoirement par fragmentation *shotgun*. Des billes magnétiques couplées à la streptavidine vont se lier aux sondes biotinylées et permettent ainsi l'isolement des acides nucléiques spécifiques des microorganismes ciblés grâce à un aimant magnétique. Seuls les fragments ciblés par capture seront ensuite séquencés. Ces deux stratégies permettent l'obtention de séquence(s) de haute qualité.

Contrairement à la stratégie *target* NGS qui cible une ou plusieurs régions génomiques, la stratégie de séquençage *shotgun* NGS repose sur une étape de fragmentation aléatoire *shotgun* de l'ensemble des acides nucléiques extraits, suivie d'une étape de séquençage de l'ensemble des fragments générés (Fig. 1B). Des procédés

mécaniques ou enzymatiques sont utilisés pour fragmenter aléatoirement le matériel génétique. Les fragments séquencés sont comparés à des bases de données, afin de permettre l'identification du (des) microorganisme(s) présent(s). Lorsque le nombre de séquences obtenues et la taille des génomes présents le permettent, il devient alors possible de procéder à leur reconstitution et donc à la caractérisation génomique précise des microorganismes présents.

B- Applications des méthodes basées sur le NGS en microbiologie clinique

Divers échantillons de patients ou des cultures microbiennes peuvent être analysés à l'aide des différentes stratégies de séquençage ciblées ou non ciblées selon les applications désirées.

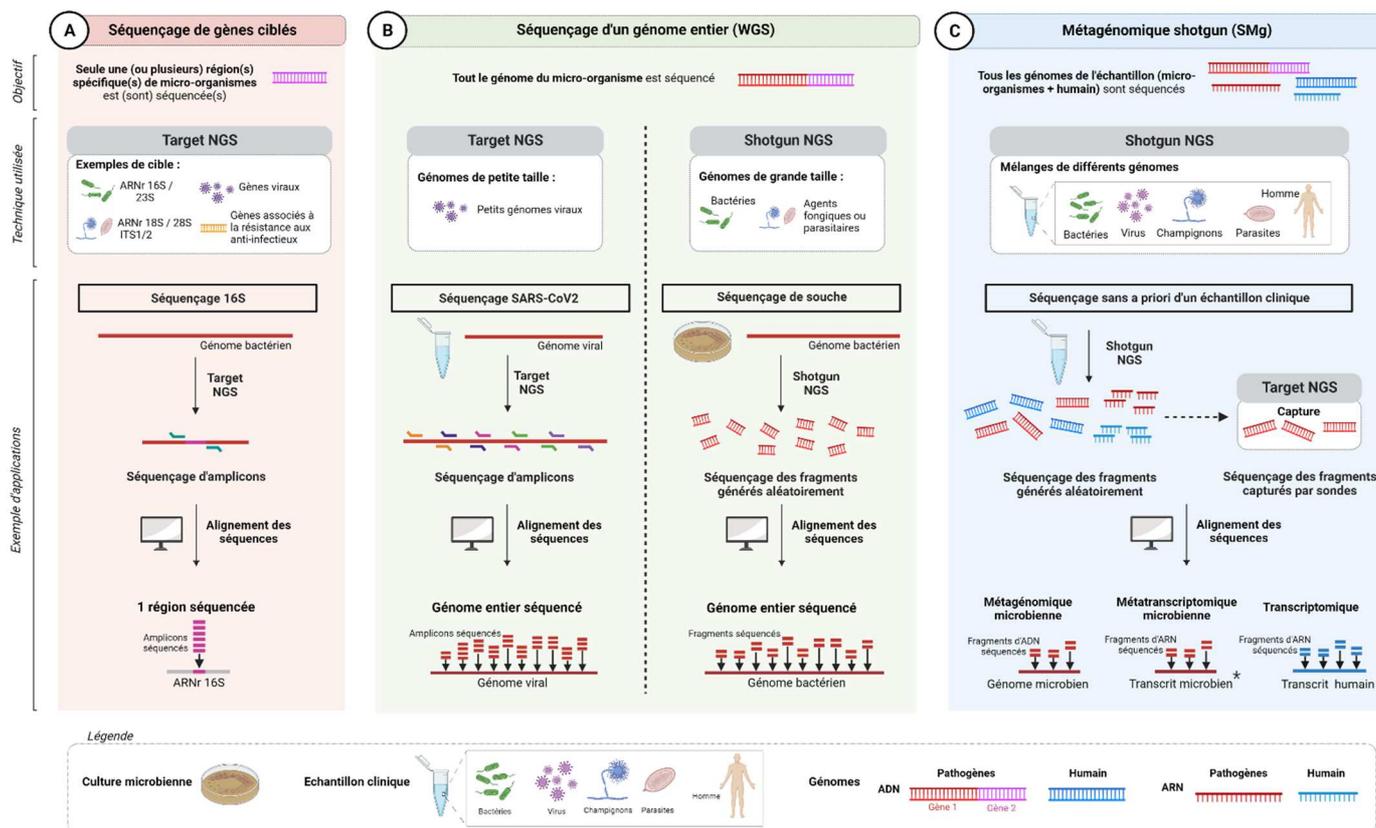


Fig. 2. Principales applications du séquençage à haut débit dans le diagnostic microbiologique clinique.

A. Le séquençage de gènes ciblés peut être réalisé à partir d'échantillons primaires ou sur des cultures microbiennes. Cette application repose sur la stratégie ciblée utilisant le target NGS avec diverses cibles possibles selon le gène à séquencer. L'exemple d'application présenté ici est le séquençage du gène codant pour l'ARNr 16S : grâce à l'utilisation d'amorces consensus ciblant ce gène et d'une réaction par PCR, des amplicons spécifiques de cette région sont générés et séquencés. L'analyse bioinformatique des séquences obtenues permet de les aligner sur une séquence de référence, et d'obtenir une identification bactérienne pour chaque amplicon.

B. Le séquençage d'un génome entier repose sur la stratégie target NGS pour le séquençage de petits génomes ou sur la stratégie shotgun NGS pour les grands génomes. L'exemple du séquençage du génome entier du SARS-CoV-2 est présenté sur cette figure et repose sur la stratégie target NGS. Des jeux d'amorces conçus pour se chevaucher sont utilisés et permettent de recouvrir l'ensemble du génome viral. L'analyse bioinformatique permet d'aligner les différents amplicons sur le génome de référence et d'obtenir un génome complet. L'autre application du séquençage de génome complet utilisant cette fois le shotgun NGS est par exemple le séquençage de souche bactérienne ou fongique. Il est réalisé sur des cultures bactériennes ou fongiques et permet de séquencer le génome entier d'un microorganisme isolé.

C. La dernière application est la SMg. Son utilisation permet le séquençage de l'ensemble des acides nucléiques d'un échantillon contenant éventuellement plusieurs microorganismes. La stratégie de séquençage SMg permet d'obtenir deux principales analyses bioinformatiques pour les séquences microbiennes : la métagénomique microbienne et la métatranscriptomique microbienne. En parallèle, les ARN humains peuvent être analysés, c'est la transcriptomique. La stratégie ciblée de capture peut être utilisée avant le séquençage des fragments afin de sélectionner uniquement les microorganismes ciblés comme par exemple des bactéries ou des virus.

"Created with BioRender.com"

1) Séquençage de gènes ciblés

a. Principe général

Le séquençage de gènes ciblés est réalisé par la stratégie *target* NGS (Fig. 2A). En Bactériologie et Mycologie, cette approche est fréquemment utilisée pour la détection et l'identification d'agents microbiens à partir d'un échantillon biologique. Pour chaque catégorie de microorganisme que l'on cherche à détecter et identifier, on choisira une cible commune dont la spécificité des séquences permettra ensuite de déterminer les espèces présentes. Une fois la cible choisie, elle sera amplifiée ou capturée grâce à l'utilisation d'amorces consensus ou de sondes encadrant des régions hautement variables. Cette approche permettra d'obtenir autant de séquences différentes qu'il y a d'espèces présentes au sein de l'échantillon. Les séquences obtenues permettront leurs identifications au niveau de l'espèce.

En Virologie, le *target* NGS permet la détection de mutations de résistance aux antiviraux. Une application courante en laboratoire est par exemple le génotypage de la résistance aux traitements antirétroviraux. Traditionnellement effectué par Sanger, ce dernier ne détecte que les populations virales majoritaires, représentant plus de 20% de l'ensemble des quasi-espèces virales présentes. Le *target* NGS améliore la sensibilité et permet de détecter les variants viraux minoritaires en se basant sur les séquences des gènes ciblés (>0,5-1%) [3].

b. Exemples d'application

Pour la plupart des bactéries, le gène *rrs* codant pour l'ARNr 16S est généralement choisi alors qu'en parasitologie/mycologie, ce sont les gènes codant pour l'ARN ribosomal 23S, 18S, 28S ou encore les régions intergéniques ITS1 ou ITS2 en. Par rapport au séquençage Sanger, cette méthode présente l'avantage de pouvoir identifier plusieurs microorganismes présents dans un même échantillon. En Virologie, le *target* NGS est utilisé pour le séquençage de régions associées à la résistance virale du VIH ou du VHC. Il peut permettre également une analyse précise de la diversité virale. Ses performances ont été démontrées par exemple pour le génotypage des HPV (Papillomavirus humains), notamment lors de co-infections par différents génotypes viraux ou lorsque les charges virales dans l'échantillon sont faibles [4].

2) WGS (Whole Generation Sequencing) : séquençage de génomes complets

a. Principe général

Le séquençage de génomes complets (appelé aussi : WGS pour « Whole Generation Sequencing ») repose selon la taille du génome à séquencer sur l'approche *target* NGS pour les petits génomes (viraux par exemple) ou *shotgun* NGS pour les grands génomes (bactériens ou fongiques) (Fig. 2B). Cette dernière approche est aussi appelée séquençage de souche.

L'objectif de cette approche est de permettre d'analyser l'ensemble du génome d'un microorganisme, généralement à partir d'une culture ou parfois directement à partir d'un échantillon biologique. Le séquençage des génomes viraux complets utilise grâce à la stratégie *target* NGS plusieurs couples d'amorces conçues pour se chevaucher et permettant de recouvrir l'ensemble du génome viral. Cela facilite la caractérisation de petits génomes viraux directement à partir d'échantillons biologiques, même ceux ayant une charge virale faible.

Cette approche possède également de nombreuses applications en Bactériologie, principalement dans le domaine de l'épidémiologie et de l'hygiène hospitalière. En effet, en comparant les génomes d'isolats suspectés d'être responsables d'une épidémie, le WGS permet, en quantifiant le nombre de mutations différenciant les génomes entre eux (par détermination du nombre de « *single nucleotide polymorphisms* » ou SNP), d'affirmer ou non leur nature clonale. Le centre européen de prévention et de contrôle des maladies (ECDC) recommande l'utilisation accrue du WGS dans les enquêtes sur les épidémies et la surveillance en santé publique [5]. D'autres applications s'intéressant notamment à l'analyse des gènes codant pour de la résistance aux antibiotiques sont en cours de développement [6].

b. Exemples d'application

En Virologie, en raison de la petite taille des génomes étudiés, on recourt souvent aux approches *target* NGS. Grâce à l'utilisation de plusieurs couples d'amorces recouvrant tout ou partie du génome, l'approche *target* NGS permet la surveillance épidémiologique des variants du SARS-CoV-2 [7]. Le séquençage des virus Zika et Ebola repose également sur du *target* NGS. Cela facilite la caractérisation de petits génomes viraux entiers directement à partir d'échantillons biologiques, même en cas de charge virale faible. Cette méthode a été utilisée pour suivre l'évolution et la propagation du virus Ebola en 2015 en Guinée et du virus Zika en 2016 aux Etats-Unis [8,9].

En Bactériologie, les laboratoires mettent progressivement en œuvre des approches de *shotgun* NGS en remplacement de l'électrophorèse en champs pulsé qui permettait auparavant le typage de souches bactériennes. Il a par exemple été utilisé en Italie pour étudier l'émergence et la propagation régionale et nosocomiale d'un clone multi-résistant de *Klebsiella pneumoniae* (ST-147), responsable d'une épidémie en Toscane [10]. Le WGS est également appliqué dans la surveillance épidémiologique en Mycologie. Le CDC (*Centers for Disease Control and Prevention*) recommande par exemple d'utiliser cette stratégie pour étudier les épidémies de *Candida auris*, levure multirésistante émergente, en identifiant les souches dans différentes régions du monde [11]. Enfin, à une échelle

plus locale, le WGS peut être appliqué à l'étude d'épidémies de bactéries multirésistantes au sein d'établissements hospitaliers [12].

3) La SMg

a. Principe général

La SMg est une technique de diagnostic panpathogène, permettant la recherche non ciblée et exhaustive de l'ensemble des agents microbiens qu'ils soient bactériens, viraux, parasitaires ou fongiques, y compris ceux non encore découverts. Brièvement, cette technique consiste à séquencer l'ensemble des acides nucléiques (ADN et ARN) appartenant aussi bien à l'hôte et qu'aux microorganismes présents dans un échantillon biologique étudié (Fig. 2C). Les séquences ainsi obtenues sont analysées à l'aide d'outils bioinformatiques permettant dans un premier temps d'exclure les séquences d'origine humaines. Dans un second temps, les séquences restantes, lorsqu'elles sont présentes, sont comparées à des séquences contenues dans les bases de données qui correspondent aux génomes de l'ensemble des bactéries, virus, champignons ou parasites décrits. L'analyse bioinformatique prend en compte à la fois les séquences obtenues à partir des extractions d'ADN et d'ARN, l'ADN correspondant à la métagénomique microbienne (c'est-à-dire l'ensemble des génomes microbiens), tandis que l'ARN correspond à la métatranscriptomique microbienne (c'est-à-dire l'ensemble des transcrits microbiens ainsi que les génomes des virus à ARN). Ainsi, en plus de permettre l'identification de l'ensemble des microorganismes présents dans un échantillon, la SMg peut aussi se concentrer sur des aspects fonctionnels, en s'intéressant à l'analyse des transcrits d'origine microbienne. Comme cette approche repose sur la détection d'acides nucléiques indépendamment de la capacité des microbes présents à se multiplier, elle permet la mise en évidence de microbes de culture difficile ainsi que la documentation d'infection décapitées chez des patients préalablement exposés aux antibiotiques.

En analysant l'ensemble des acides nucléiques non-humains (ADN et ARN), la SMg permet la détection et l'identification de l'ensemble des microorganismes connus ou restant à décrire. Cependant, cette technique connaît de très nombreuses variantes selon le matériel utilisé et les objectifs fixés. En effet, comme la grande majorité des séquences obtenues à partir d'un échantillon biologique (généralement > 99%) appartiennent au génome humain et limitent la sensibilité analytique de détection des microorganismes ainsi que la reconstruction de génomes microbiens complets. Différentes stratégies peuvent alors être mises en place afin d'augmenter le rendement des séquences provenant des pathogènes d'intérêt, soit en déplaçant le matériel génétique provenant de l'hôte, soit en capturant spécifiquement les séquences microbiennes et en utilisant ainsi une stratégie ciblée de séquençage.

La stratégie de déplétion de l'hôte permet d'éliminer le fond génomique humain, tout en préservant les acides nucléiques appartenant aux microorganismes. Elle peut être effectuée grâce à des méthodes enzymatiques (tampon lysant les membranes des cellules humaines puis DNAses) ou grâce à des méthodes physiques (centrifugation, filtration), dégradant ou séparant ainsi le matériel de l'hôte. Gu *et al.* ont utilisé l'approche CRISPR (*clustered regularly interspaced short palindromic repeats*) pour cibler et appauvrir l'ARNr mitochondrial humain dans des échantillons cliniques de LCR, ce qui a permis d'améliorer la sensibilité de détection des agents pathogènes [13]. Bien que les méthodes de déplétion présentent l'avantage de conserver le côté impartial du séquençage

métagénomique par rapport à la capture qui nécessite de cibler des pathogènes connus, elles peuvent également entraîner une perte du matériel microbien limitant leur utilisation sur des échantillons faibles en acides nucléiques [14]. Par ailleurs, la déplétion compromet toute analyse impliquant l'utilisation de données humaines, dans le cadre de la recherche.

La stratégie de capture ou d'enrichissement positif permet d'augmenter le signal microbien plutôt que de réduire le fond génomique humain. Cette stratégie effectuée suite au *shotgun* NGS permet d'augmenter la sensibilité de détection pour les agents microbiens ciblés, mais entraîne également la perte de l'avantage « sans *a priori* » de la SMg et devient alors une stratégie ciblée. Des panels multiplex de capture virale [15] et bactérienne [16] existent. Ils peuvent couvrir jusqu'à plus de 200 espèces virales et 300 espèces bactériennes. Certains panels sont également composés de sondes ciblant certains gènes de résistance et de virulence bactériens. En utilisant une telle stratégie, on est capable de générer des génomes complets du virus respiratoire syncytial humain (VRS) directement à partir de prélèvements, permettant une surveillance continue du VRS au niveau de la population et ainsi, favoriser le développement de vaccins en temps réel [17].

En parallèle de l'analyse bioinformatique des séquences microbiennes, les séquences humaines, qui sont actuellement éliminées dans le cadre du soin, sont de plus en plus utilisées en recherche. En effet, l'analyse des ARN humains, analyse que l'on appelle transcriptomique humaine et qui reflète l'ensemble des transcrits présents à un moment donné, pourrait permettre d'identifier des profils de réponse immunitaires et permettre d'identifier de nouveaux biomarqueurs associés à différents types d'infection. Comme nous le verrons plus loin, cette approche pourrait dans l'avenir donner des informations sur la nature de l'infection en cours ou son pronostic [10].

b. Exemples d'application

La SMg a été appliquée à une large variété d'échantillons et pour le diagnostic de multiples syndromes infectieux [18,19]. Qu'il s'agisse des infections disséminées [20], des méningo-encéphalites [21], des infections nécrosantes de la peau et des tissus mous [22], des pneumonies [23,24] ou des infections ostéo-articulaires [25,26], la SMg a démontré des performances souvent supérieures aux méthodes de référence, notamment pour l'identification de microorganismes *a priori* non suspectés ou de croissance difficile comme les bactéries anaérobies strictes. Elle a également été contributive pour le diagnostic microbiologique des patients immunodéprimés, chez lesquels les infections sont souvent graves et complexes [27–29]. C'est par ailleurs la principale technique qui permet aujourd'hui de détecter et de décrire des microorganismes pathogènes émergents, comme le SARS-CoV-2 en 2019 [30]. Aujourd'hui, elle est principalement utilisée en deuxième intention, en cas d'impasse diagnostique [31].

3- Apports de la SMg en cas d'impasse diagnostique : exemple des infections du SNC

En raison de la diversité de leurs causes et de la difficulté à les explorer, les infections du SNC, comme les encéphalites et les méningo-encéphalites, sont parmi les indications les plus fréquentes à la SMg. Dans l'ensemble, on estime que le diagnostic étiologique des atteintes du SNC n'est identifié que chez environ 50 % des patients [32], avec une prévalence des encéphalites d'origine auto-immune équivalente aux encéphalites d'origine infectieuse

(13,7/100 000 et 11,6/100 000, respectivement en 2014) [33]. Dans ce contexte, et en raison des options de prise en charge thérapeutique radicalement différentes, disposer d'un outil diagnostique permettant à lui seul d'identifier une étiologie infectieuse et de débiter précocement un traitement adapté correspond à un réel besoin. En raison de ses performances dans cette indication, des publications recommandent aujourd'hui son utilisation en routine dans le diagnostic des infections du SNC [34]. Ainsi, Wilson et Ramachandran ont élaboré un algorithme pour la prise en charge des infections du SNC dans lequel la SMg est intégrée [35]. Après un bilan de première intention, le prélèvement de LCS est analysé par SMg, permettant au clinicien de rechercher les infections faisant partie des diagnostics différentiels classiques mais aussi les infections non suspectées, y compris les causes rares ou potentiellement inconnues de méningo-encéphalites.

La première utilisation de la SMg dans l'exploration d'une méningo-encéphalite remonte à 2013, chez un enfant de 14 ans atteint d'un déficit immunitaire combiné sévère (déficit en adénosine désaminase). L'utilisation de cette approche avait alors permis d'identifier *Leptospira santarosai* sur LCS, permettant une prise en charge adaptée et un rétablissement progressif du patient quatre mois après le début de son hospitalisation et alors qu'aucune étiologie infectieuse n'avait été documentée par les approches classiques [36].

Depuis, l'intérêt de la SMg dans l'exploration des encéphalites et méningo-encéphalites en impasse diagnostique n'a jamais été démenti. Des études ont démontré que la SMg avait une sensibilité de 73 à 92 % et une spécificité de 96 à 99 %, selon l'agent pathogène détecté dans le LCS [37]. Wilson et al ont inclus dans une étude 204 patients atteints de méningite, avec ou sans encéphalite associée [21]. Sur 58 infections diagnostiquées, 13 (22%) ont été diagnostiquées par SMg, 26 (45%) par des méthodes conventionnelles et 19 (33%) par les deux approches. Le diagnostic par SMg seule a notamment permis à 9 patients sur 13 une adaptation du traitement antimicrobien.

Schématiquement, la SMg possède un réel intérêt face à des situations complexes comme une présentation clinique atypique ou la présence d'un microorganisme pathogène inhabituel, émergent ou non encore décrit. Grâce à son approche panpathogène sans *a priori*, la SMg permet de résoudre des situations où les signes cliniques peuvent parfois être atypiques, comme l'illustre le cas d'une patiente atteinte par le VIH et présentant un syndrome encéphalitique. Malgré la réalisation d'un bilan étiologique négatif, c'est finalement l'approche SMg qui a permis de détecter le virus de la rougeole à partir d'une biopsie cérébrale, chez une patiente qui n'avait présenté ni éruption cutanée, ni aucun autre signe clinique classiquement associé à la rougeole [38]. Le même type d'approche a permis d'identifier *Nocardia cyriacigeorgica* à partir du LCS d'un patient sans comorbidité identifiée qui présentait une méningo-encéphalite [39], alors que toutes les méthodes standard étaient restées négatives. Ce diagnostic inattendu avait alors permis non seulement une adaptation du traitement antibiotique mais avait aussi conduit à l'exploration et la découverte d'un déficit immunitaire acquis lié à une production d'anticorps anti-GM-CSF.

Dans d'autres cas, la SMg permet la mise en évidence de microorganismes de culture difficile et non réalisable en routine et dont la recherche n'est pas habituellement réalisée dans les laboratoires. C'est le cas de la neuroborréliose à *Borrelia miyamotoi*, maladie décrite mais exceptionnelle qui atteint exclusivement les individus immunodéprimés. Nous avons eu l'occasion de faire ce diagnostic par SMg chez un patient présentant une méningo-encéphalite chronique en errance diagnostique. Ce diagnostic a conduit d'une part à la guérison complète du patient

grâce à l'administration d'un traitement adapté, et d'autre part à l'identification, pour la première fois en France, de cette bactérie dans le cadre d'une méningo-encéphalite (données non publiées). Ainsi, la SMg peut se révéler contributive dans des cas d'infections rares ou méconnues, qu'elles soient d'origine bactérienne comme nous venons de le voir, mais aussi parasitaires comme l'angiostrongylose du SNC [40].

Enfin, la SMg est actuellement le seul test permettant de détecter et de décrire, grâce à son approche sans *a priori*, des microorganismes non encore décrits. C'est ainsi que des virus ont pu être découverts et leurs génomes reconstitués à partir de biopsies cérébrales, comme le Cristolivirus appartenant à la famille des Orthobunyavirus et mis en cause dans une encéphalite mortelle, ou le Cache Valley virus dans une méningo-encéphalite [41,42]. Elle a aussi permis d'identifier un nouveau clade dans la famille des Astrovirus chez un patient présentant une encéphalite [43]. Il est important de rappeler dans ces cas que lier la présence d'un nouveau microorganisme à une pathologie humaine nécessite de réunir un faisceau d'arguments microbiologiques, cliniques mais également histologiques.

4- Perspectives et place de la SMg en pratique clinique

Comme nous venons de le voir, en plus de pouvoir identifier les microorganismes détectés, la SMg peut également contribuer à leur caractérisation. Sur la base de l'analyse des données génomiques, des informations sur le typage ou la présence de gènes de résistance ou de virulence peuvent être révélées, permettant de mieux comprendre l'épidémiologie d'un microorganisme ou prédire sa virulence ou sa sensibilité aux anti-infectieux. Ainsi, lors de l'identification du Cristolivirus, le séquençage du génome complet a révélé l'expression à haut niveau d'un facteur de virulence, ce qui pourrait expliquer le tropisme neurologique du virus ainsi que la gravité de l'infection chez le patient [41]. Cependant, ces approches WGS nécessitent encore des développements et restent aujourd'hui du domaine de la recherche.

Une autre perspective possible de la SMg concerne la transcriptomique humaine, c'est-à-dire les informations que pourraient donner l'analyse des transcrits humains en cas d'infection. Face à un agent pathogène, l'expression génique des cellules du système immunitaire de l'hôte peut varier. Ainsi, des profils transcriptomiques distincts ont pu être mis en évidence en fonction de la nature du microorganisme responsable d'infection pulmonaire ou de sepsis [44–46] mais aussi sur le pronostic de certaines infections, comme cela a été démontré pour le COVID [47], ou encore le caractère approprié ou non d'un traitement antimicrobien [48]. Malgré ces perspectives encourageantes, la diversité et la complexité des situations requièrent encore d'autres études afin de préciser les performances de cette approche.

Face à une situation infectieuse comme dans le cas des infections neuro-méningées, les analyses permises par les approches NGS, combinant l'étude des agents infectieux et la signature transcriptomique de l'hôte, contribuent ainsi à l'amélioration du diagnostic et à une prise en charge personnalisée. L'identification de transcrits associés à des microorganismes pathogènes pourrait également aboutir à la détermination de panels de biomarqueurs spécifiques [49]. Pour finir, le développement de l'intelligence artificielle combinée aux méthodes OMICS pourraient constituer l'avenir du diagnostic et de la prise en charge des maladies infectieuses, en particulier dans le domaine du système nerveux central.

Plusieurs obstacles ralentissent la mise en place de la métagénomique en microbiologie clinique. A ce jour, quelle que soit l'approche de séquençage, les technologies NGS ont un coût élevé (automates, consommables et équipements bioinformatiques) et nécessitent du personnel de laboratoire qualifié. Le manque de validation/standardisation des étapes (analytiques et post analytiques) et de recommandations posent des problèmes réglementaires pour l'implantation du séquençage à haut débit en microbiologie et limitent son utilisation à grande échelle. Enfin, l'analyse des données nécessite la formation d'équipes multidisciplinaires (bioinformaticien, biologiste, clinicien) afin d'interpréter les résultats selon chaque contexte clinique. Ainsi, seuls quelques laboratoires de référence proposent la métagénomique en technique de recours lors d'impasse diagnostique, avec un délai de rendu de résultats variant entre 10 à 21 jours. En 2023, cette technique est principalement utilisée en support aux méthodes traditionnelles. En effet, la culture traditionnelle ou les techniques de PCR ciblées sont moins coûteuses que les différentes stratégies NGS et restent le premier choix en matière de diagnostic microbiologique. Pour simplifier les comparaisons de coût réactif, il apparaît assez simple de garder en tête qu'une approche *target* NGS coûte en général entre une et deux fois le coût d'un panel multiplex PCR contre deux à trois fois pour la stratégie *shotgun* NGS. Cependant, le surcoût des technologies à haut débit pourrait être compensé en partie par un nombre plus restreint de tests diagnostiques totaux effectués, une thérapie davantage ciblée voire une sortie d'hospitalisation plus précoce pour un patient. Des études médico-économiques, avec des essais cliniques prospectifs démontrant l'amélioration de la prise en charge des patients, doivent être menées. Ces données permettront également d'ouvrir la voie à l'approbation réglementaire et aux conditions de remboursement.

5- Conclusion

Aujourd'hui, il est admis que l'utilisation des technologies de séquençage à haut débit, et notamment de la SMg, jouera un rôle croissant dans le diagnostic microbiologique des maladies infectieuses. Malgré des obstacles liés à la nécessité de disposer du personnel de laboratoire qualifié et à l'acquisition du matériel indispensable, les performances de la SMg et son caractère universel en font un outil très prometteur. A l'avenir, les développements de cette méthode permettront en plus de documenter une infection, de caractériser de façon très précise les microorganismes aussi bien sur leur capacité à présenter des mécanismes de résistance aux anti-infectieux, que sur leur contenu en gènes de virulence. Par ailleurs, la SMg pourrait dans l'avenir exploiter les données de transcriptomique humaine, ouvrant la voie à l'analyse de la réponse inflammatoire de l'hôte. Des travaux encore préliminaires ont permis d'identifier des profils de réponses inflammatoires pouvant correspondre à une infection spécifique ou au pronostic d'une infection en cours [47]. Ainsi, par son approche globale, cette technique pourrait ouvrir la voie, en grande routine, à la médecine personnalisée en maladies infectieuses.

Conflit d'intérêts : aucun.

Références

1. Marani M, Katul GG, Pan WK, Parolari AJ. Intensity and frequency of extreme novel epidemics. *Proceedings of the National Academy of Sciences*. 2021;118(35):e2105482118.
2. Wilson MR, Sample HA, Zorn KC, Arevalo S, Yu G, Neuhaus J, et al. Clinical Metagenomic Sequencing for Diagnosis of Meningitis and Encephalitis. *N Engl J Med*. 2019;380(24):2327-40.
3. Ávila-Ríos S, Parkin N, Swanstrom R, Paredes R, Shafer R, Ji H, et al. Next-Generation Sequencing for HIV Drug Resistance Testing: Laboratory, Clinical, and Implementation Considerations. *Viruses*. 2020;12(6):E617.
4. Molet L, Girlich D, Bonnin RA, Proust A, Bouligand J, Bachelerie F, et al. Identification by high-throughput sequencing of HPV variants and quasispecies that are untypeable by linear reverse blotting assay in cervical specimens. *Papillomavirus Res*. 2019;8:100169.
5. European Centre for Disease Prevention and Control. ECDC strategic framework for the integration of molecular and genomic typing into European surveillance and multi-country outbreak investigations – 2019–2021. 2019; Stockholm.
6. Ruppé E, Cherkaoui A, Charretier Y, Girard M, Schicklin S, Lazarevic V, et al. From genotype to antibiotic susceptibility phenotype in the order Enterobacterales: a clinical perspective. *Clin Microbiol Infect*. 2020;26(5):643.e1-643.e7.
7. European Centre for Disease Prevention and Control. Guidance for representative and targeted genomic SARS-CoV-2 monitoring. Stockholm: ECDC; 2021 mai.
8. Grubaugh ND, Ladner JT, Kraemer MUG, Dudas G, Tan AL, Gangavarapu K, et al. Genomic epidemiology reveals multiple introductions of Zika virus into the United States. *Nature*. 2017;546(7658):401-5.
9. Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature*. 2016;530(7589):228-32.
10. Martin MJ, Corey BW, Sannio F, Hall LR, MacDonald U, Jones BT, et al. Anatomy of an extensively drug-resistant *Klebsiella pneumoniae* outbreak in Tuscany, Italy. *Proc Natl Acad Sci U S A*. 2021;118(48):e2110227118.
11. Chow NA, Gade L, Tsay SV, Forsberg K, Greenko JA, Southwick KL, et al. Multiple introductions and subsequent transmission of multidrug-resistant *Candida auris* in the USA: a molecular epidemiological survey. *Lancet Infect Dis*. 2018;18(12):1377-84.
12. Royer G, Fourreau F, Boulanger B, Mercier-Darty M, Ducellier D, Cizeau F, et al. Local outbreak of extended-spectrum β -lactamase SHV2a-producing *Pseudomonas aeruginosa* reveals the emergence of a new specific sub-lineage of the international ST235 high-risk clone. *J Hosp Infect*. 2020;104(1):33-9.
13. Gu W, Crawford ED, O'Donovan BD, Wilson MR, Chow ED, Retallack H, et al. Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. *Genome Biology*. 2016;17(1):41.
14. Maljkovic Berry I, Melendrez MC, Bishop-Lilly KA, Rutvisuttinunt W, Pollett S, Talundzic E, et al. Next Generation Sequencing and Bioinformatics Methodologies for Infectious Disease Research and Public Health: Approaches, Applications, and Considerations for Development of Laboratory Capacity. *J Infect Dis*. 2020;221(Suppl 3):S292-307.
15. Wylie KM, Wylie TN, Buller R, Herter B, Cannella MT, Storch GA. Detection of Viruses in Clinical Samples by Use of Metagenomic Sequencing and Targeted Sequence Capture. *J Clin Microbiol*. 2018;56(12):e01123-18.
16. Allicock OM, Guo C, Uhlemann AC, Whittier S, Chauhan LV, Garcia J, et al. BacCapSeq: a Platform for Diagnosis and Characterization of Bacterial Infections. *mBio*. 2018;9(5):e02007-18.
17. Lin GL, Golubchik T, Drysdale S, O'Connor D, Jefferies K, Brown A, et al. Simultaneous Viral Whole-Genome Sequencing and Differential Expression Profiling in Respiratory Syncytial Virus Infection of Infants. *J Infect Dis*. 2020;222(Suppl 7):S666-71.

18. Chiu CY, Miller SA. Clinical metagenomics. *Nat Rev Genet.* 2019;20(6):341-55.
19. Simner PJ, Miller S, Carroll KC. Understanding the Promises and Hurdles of Metagenomic Next-Generation Sequencing as a Diagnostic Tool for Infectious Diseases. *Clin Infect Dis.* 2018;66(5):778-88.
20. Morales M. The Next Big Thing? Next-Generation Sequencing of Microbial Cell-Free DNA Using the Karius Test. *Clin Microbiol Newsl.* 2021;43(9): 69-79.
21. Wilson MR, Sample HA, Zorn KC, Arevalo S, Yu G, Neuhaus J, et al. Clinical Metagenomic Sequencing for Diagnosis of Meningitis and Encephalitis. *N Engl J Med.* 2019;380(24):2327-40.
22. Rodriguez C, Jary A, Hua C, Woerther PL, Bosc R, Desroches M, et al. Pathogen identification by shotgun metagenomics of patients with necrotizing soft-tissue infections. *Br J Dermatol.* 2020;183(1):105-13.
23. Zhou H, Larkin PMK, Zhao D, Ma Q, Yao Y, Wu X, et al. Clinical Impact of Metagenomic Next-Generation Sequencing of Bronchoalveolar Lavage in the Diagnosis and Management of Pneumonia: A Multicenter Prospective Observational Study. *J Mol Diagn.* 2021;23(10):1259-68.
24. Schlaberg R, Chiu CY, Miller S, Procop GW, Weinstock G, the Professional Practice Committee and Committee on Laboratory Practices of the American Society for Microbiology, et al. Validation of Metagenomic Next-Generation Sequencing Tests for Universal Pathogen Detection. *Arch Pathol Lab Med.* 2017;141(6):776-86.
25. Goswami K, Shope AJ, Tokarev V, Wright JR, Unverdorben LV, Ly T, et al. Comparative meta-omics for identifying pathogens associated with prosthetic joint infection. *Sci Rep.* 2021;11(1):23749.
26. Ruppé E, Lazarevic V, Girard M, Mouton W, Ferry T, Laurent F, et al. Clinical metagenomics of bone and joint infections: a proof of concept study. *Sci Rep.* 2017;7(1):7718.
27. Parize P, Muth E, Richaud C, Gratigny M, Pilimis B, Lamamy A, et al. Untargeted next-generation sequencing-based first-line diagnosis of infection in immunocompromised adults: a multicentre, blinded, prospective study. *Clin Microbiol Infect.* 2017;23(8):574.e1-574.e6.
28. Pérot P, Bielle F, Bigot T, Foulongne V, Bolloré K, Chrétien D, et al. Identification of Umbre Orthobunyavirus as a Novel Zoonotic Virus Responsible for Lethal Encephalitis in 2 French Patients with Hypogammaglobulinemia. *Clin Infect Dis.* 2021;72(10):1701-8.
29. Palacios G, Druce J, Du L, Tran T, Birch C, Briese T, et al. A new arenavirus in a cluster of fatal transplant-associated diseases. *N Engl J Med.* 2008;358(10):991-8.
30. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. *Nature.* 2020;579(7798):265-9.
31. Woerther PL, Surgers L, Lamoureux C, Lepeule R, Demontant V, Gricourt G, et al. Diagnostic microbiologique pan-pathogène par métagénomique clinique, retour d'expérience en routine. *Médecine et Maladies Infectieuses.* 2020;50(6, Supplement):S26.
32. Granerod J, Tam CC, Crowcroft NS, Davies NWS, Borchert M, Thomas SL. Challenge of the unknown. A systematic review of acute encephalitis in non-outbreak situations. *Neurology.* 2010;75(10):924-32.
33. Dubey D, Pittock SJ, Kelly CR, McKeon A, Lopez-Chiriboga AS, Lennon VA, et al. Autoimmune encephalitis epidemiology and a comparison to infectious encephalitis. *Ann Neurol.* 2018;83(1):166-77.
34. Kanjilal S, Cho TA, Piantadosi A. Diagnostic Testing in Central Nervous System Infection. *Semin Neurol.* juin 2019;39(3):297-311.
35. Ramachandran PS, Wilson MR. Metagenomics for neurological infections — expanding our imagination. *Nat Rev Neurol.* 2020;16(10):547-56.

36. Wilson MR, Naccache SN, Samayoa E, Biagtan M, Bashir H, Yu G, et al. Actionable Diagnosis of Neuroleptospirosis by Next-Generation Sequencing. *New England Journal of Medicine*. 2014;370(25):2408-17.
37. Miller S, Naccache SN, Samayoa E, Messacar K, Arevalo S, Federman S, et al. Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. *Genome Res*. 2019;29(5):831-42.
38. Rodriguez C, Gouilh MA, Weiss N, Stroer S, Mokhtari K, Seilhean D, et al. Fatal Measles Inclusion-Body Encephalitis in Adult with Untreated AIDS, France. *Emerg Infect Dis*. 2020;26(9):2231-4.
39. Courbin V, Riller Q, Amegnizin JL, Gricourt G, Demontant V, Fihman V, et al. Case Report: Cerebral Nocardiosis Caused by *Nocardia cyriacigeorgica* Detected by Metagenomics in an Apparently Immunocompetent Patient. *Front Immunol*. 2022;13:719124.
40. Feng L, Zhang A, Que J, Zhou H, Wang H, Guan Y, et al. The metagenomic next-generation sequencing in diagnosing central nervous system angiostrongyliasis: a case report. *BMC Infect Dis*. 2020;20(1):691.
41. Rodriguez C, Gricourt G, Ndebi M, Demontant V, Poiteau L, Burrel S, et al. Fatal Encephalitis Caused by Cristoli Virus, an Emerging Orthobunyavirus, France. *Emerg Infect Dis*. 2020;26(6):1287-90.
42. Wilson MR, Suan D, Duggins A, Schubert RD, Khan LM, Sample HA, et al. A novel cause of chronic viral meningoencephalitis: Cache Valley virus. *Ann Neurol*. 2017;82(1):105-14.
43. Naccache SN, Peggs KS, Mattes FM, Phadke R, Garson JA, Grant P, et al. Diagnosis of neuroinvasive astrovirus infection in an immunocompromised adult with encephalitis by unbiased next-generation sequencing. *Clin Infect Dis*. 2015;60(6):919-23.
44. Langelier C, Kalantar KL, Moazed F, Wilson MR, Crawford ED, Deiss T, et al. Integrating host response and unbiased microbe detection for lower respiratory tract infection diagnosis in critically ill adults. *Proc Natl Acad Sci USA*. 2018;115(52):E12353-62.
45. Cheng H, Tan S, Sweeney T, Jeganathan P, Briese T, Khadka V, et al. Combined use of metagenomic sequencing and host response profiling for the diagnosis of suspected sepsis [Internet]. 2019 [cité 8 sept 2022]. Disponible sur: [bioRxiv 854182](https://doi.org/10.1101/854182); doi: <https://doi.org/10.1101/854182>
46. Kalantar KL, Neyton L, Abdelghany M, Mick E, Jauregui A, Caldera S, et al. Integrated host-microbe plasma metagenomics for sepsis diagnosis in a prospective cohort of critically ill adults. *Nat Microbiol*. 2022;7(11):1805-16.
47. Rodriguez C, de Prost N, Fourati S, Lamoureux C, Gricourt G, N'debi M, et al. Viral genomic, metagenomic and human transcriptomic characterization and prediction of the clinical forms of COVID-19. *PLoS Pathog*. 2021;17(3):e1009416.
48. Poissy J, Terrier O, Lina B, Textoris J, Rosa-Calatrava M. La modulation de la signature transcriptomique de l'hôte infecté : une nouvelle stratégie thérapeutique dans les viroses graves ? Exemple de la grippe. *Reanimation*. 2016;25(Suppl 2):53-61.
49. Sweeney TE, Wong HR, Khatri P. Robust classification of bacterial and viral infections via integrated host gene expression diagnostics. *Sci Transl Med*. 2016;8(346):346ra91.