

## COMMUNICATION

### **La puce à ADN de reséquençage : un outil rapide pour mieux identifier et comprendre une émergence virale et bactérienne**

MOTS-CLÉS : SÉQUENCE NUCLÉOTIDIQUE. SONDAS ADN. MALADIES TRANSMISSIBLES ÉMERGENTES

#### *Resequencing microarrays: a rapid tool for better identification and understanding of viral and bacterial emergence*

KEY-WORDS (Index medicus): BASE SEQUENCE. COMMUNICABLE DISEASES, EMERGING. DNA PROBES

Nicolas BERTHET \*

**L'auteur déclare ne pas avoir de liens d'intérêt en relation avec le contenu de cet article.**

## RÉSUMÉ

*L'introduction de la technologie des puces à ADN dans le monde de la microbiologie a transformé la détection et la caractérisation des agents pathogènes. Plusieurs plateformes de puces à ADN existent et peuvent être classées en différentes familles en fonction de leurs caractéristiques et applications. Les puces à ADN de reséquençage ont montré de nombreux avantages par rapport aux autres technologies dans la détection et la caractérisation des pathogènes.*

## SUMMARY

*The introduction of microarray technologies in microbiology has transformed the detection and characterization of microbial pathogens. Microarray-based platforms can be classified into different families according to their characteristics and applications. Resequencing*

---

\* Centre International de Recherches Médicales de Franceville (CIRMF), Département Zoonose et Maladies Émergentes, BP 769 — Franceville, Gabon. CNRS, UMR 3569, 25 rue du docteur Roux — 75724 Paris Cedex ; e-mail : nicolas.berthet@ird.fr

*Tirés-à-part* : Docteur Nicolas BERTHET, même adresse

*Article reçu le 15 novembre 2013*

*microarrays have several advantages over other technologies for pathogen detection and characterization.*

## **ABRÉVIATION**

PCR : Réaction de polymérisation en chaîne / *Polymerase Chain Reaction*  
ADNc : ADN complémentaire / *complementary DNA*  
ADN : Acide Désoxyribonucléique / *Deoxyribonucleic acid*  
ARN : Acide Ribonucléique / *Ribonucleic acid*  
BLAST : *Basic Local Alignment Search Tool*  
pdm : pandémie / *pandemic*  
RA : Amplification aléatoire / *Random amplification*  
QS : Score de Qualité / *Quality score*  
VGA : Virus grippaux de type A

## **INTRODUCTION**

Depuis de nombreuses années, les infections émergentes et réurgentes représentent un intérêt majeur aussi bien pour la communauté scientifique et les autorités de santé que pour le grand public. Ces maladies infectieuses ont toujours été une menace pour la population humaine. L'émergence et la réurgence des agents pathogènes et des infections associées résultent d'une interaction complexe entre l'agent infectieux, l'hôte et l'environnement. Ces agents possèdent d'importantes capacités d'adaptation leur permettant (i) de franchir les barrières d'espèces, (ii) d'étendre leur zone de répartition géographique et de coloniser des régions où ils étaient initialement absents et enfin (iii) d'être à l'origine de nouvelles pathologies humaines dans des contextes particuliers tels que l'immunosuppression, qu'elle soit liée à la transplantation d'organes ou à la chimiothérapie, l'immunodéficience acquise ou la sénescence du système immunitaire [1-3]. La détection de ces pathogènes, connus ou non, ainsi que leurs caractérisations génétiques sont devenues incontournables tant en microbiologie clinique qu'en santé publique.

La grande variabilité et diversité des agents pathogènes représentent une difficulté pour leur détection et leur identification empêchant ainsi certaines approches génériques. La nature polyphylétique des virus, qui se caractérise par l'absence d'un ancêtre commun, rend délicate l'identification de nouveaux agents ou variants viraux émergents. Les outils actuels de détection ciblent le plus souvent une, voire au mieux, plusieurs espèces ou variants. Bien que les outils sérologiques et la culture des pathogènes restent des méthodes de base pour le diagnostic, les méthodes s'appuyant sur la biologie moléculaire apportent de nombreux avantages en termes de rapidité, de spécificité, de coût et de capacité à détecter et caractériser un large éventail de pathogènes. Ces méthodes sont de plus en plus utilisées en microbiologie clinique, notamment l'utilisation des puces à ADN qui permettent de détecter en

parallèle une multitude de pathogènes. Cette puissance de détection repousse les limites engendrées par cette grande variabilité génétique. La technologie des puces à ADN a été utilisée dès le départ dans de nombreux domaines allant de la génomique humaine à l'expression des gènes et à la microbiologie. De façon minoritaire, les puces à ADN ont également été utilisées pour (i) la détection et l'identification de pathogènes bactériens et viraux, (ii) le typage et la mise en évidence de nouveaux variants, et (iii) la découverte de nouveaux gènes impliqués dans la résistance ou la pathogénicité.

Cette revue s'intéressera à la technologie des puces à ADN de reséquençage et à ses applications comme un outil potentiel pour mieux comprendre une émergence bactérienne et/ou virale. Les principaux objectifs seront de présenter (i) le principe général d'une puce à ADN de reséquençage, (ii) un bref aperçu non exhaustif des autres technologies existantes et (iii) quelques applications de cette technologie.

## **PRINCIPE D'UNE PUCE À ADN DE RESÉQUENCAGE**

### **Principe général**

Le principe général d'une puce à ADN repose sur une hybridation par complémentarité des bases (A-T et G-C) entre l'ADN d'un échantillon biologique et un ensemble de sondes immobilisées et organisées sur un support solide. Les fragments nucléiques déposés sur le support solide sont appelés « sondes » tandis que les acides nucléiques marqués présents dans l'échantillon sont les « cibles ». La réaction d'hybridation moléculaire est spécifique et permet de détecter et d'identifier la ou les séquence(s) présentes dans l'échantillon. Ces réactions d'hybridation sont réalisées en phase solide, ce qui permet de travailler simultanément avec un nombre considérable de cibles, dont les positions sur le support sont parfaitement connues. Les différentes étapes de lavage de la puce permettent l'élimination des cibles non hybridées ou hybridées de façon non spécifique. L'analyse de la surface de la puce permet de localiser les hybridations spécifiques entre la cible et sa sonde correspondante, grâce à l'émission d'un signal de fluorescence ou de radioactivité selon la technique de marquage utilisée. Depuis l'avènement de la première puce à ADN sur membrane de cellulose, il apparaît qu'une très large diversité de technologies est dorénavant disponible pour la détection et la caractérisation de pathogènes. Les puces à ADN peuvent être classées en différentes familles, en fonction du type de support (solide ou liquide), de la densité et de la taille des sondes, de la méthode de révélation de l'hybridation, des coûts relatifs et des applications potentielles [4].

La technologie des puces à ADN de reséquençage se base sur une synthèse des sondes directement sur le support. La synthèse des sondes par photolithographie utilise des masques successifs ainsi qu'une illumination ultra-violette (UV) contrôlée à la surface du support solide composé d'une plaque fine et lisse de silice (Fig. 1) [5-7]. Le premier nucléotide est donc fixé de façon covalente au groupement

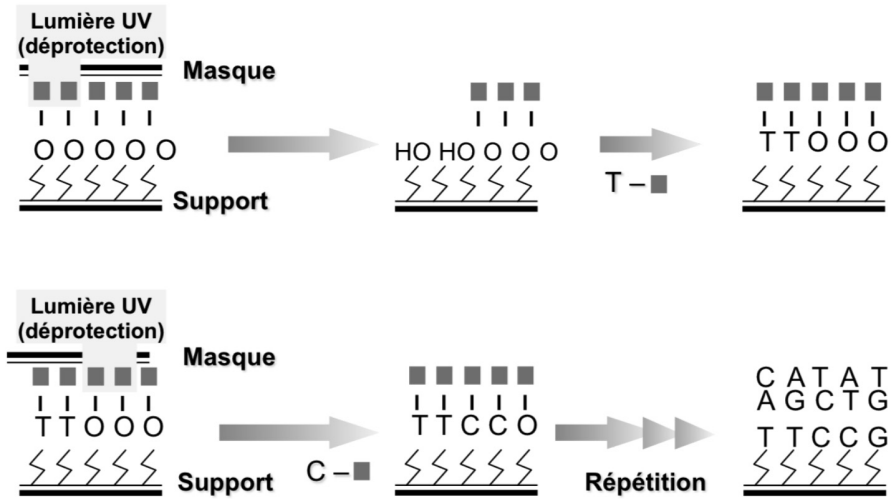


FIG. 1. — Description de la synthèse des sondes *in situ* par photolithographie. Cette synthèse se base sur un ensemble de cycles de déprotection contrôlée à la surface du support solide grâce à l'utilisation de masques physiques et l'addition de nucléotides à ces positions. À chaque étape, des nucléotides sont incorporés jusqu'à obtenir la synthèse complète des sondes désirées. Le masque est une surface plane composée d'une multitude de trous. L'association de plusieurs masques entre eux permettra de laisser passer l'illumination UV uniquement aux positions désirées pour l'ajout du nucléotide.

hydroxyle. Des cycles alternant la déprotection spécifique et l'addition de nucléotides sont répétés afin de synthétiser les sondes désirées [8-10]. La plaque de silice est découpée en petits dés et chaque élément est incorporé dans une cartouche de plastique. L'utilisation de sondes courtes spécifiques, présentes à haute densité (> 1 million) offre une résolution au nucléotide près et donc la détection simultanée d'un grand nombre de variations ponctuelles, ainsi que le « reséquençage » de gènes d'intérêt. Celui-ci est réalisé par hybridation de la cible sur des sondes de 25 nucléotides synthétisées à partir d'une séquence de référence. La première sonde couvre la région 1 à 25 de la séquence de référence. La seconde sonde la région 2 à 26 et ainsi de suite pour les sondes suivantes jusqu'à la fin de la séquence de référence avec un décalage d'une base à chaque fois. Afin de détecter d'éventuelles mutations de l'échantillon par rapport à la séquence de référence, chaque sonde ainsi dessinée verra sa base centrale (position 13) déclinée selon les 4 nucléotides possibles : A, T, C et G. Une sonde sur les 4 s'hybridera parfaitement avec la cible (*match*) tandis que les trois autres présenteront une différence de séquence (*mismatch*) en position centrale. Ainsi, la mesure des intensités de fluorescence pour chaque sonde au niveau des quatre nucléotides centraux permet de déterminer la base dont l'intensité est la plus forte à une position déterminée donc, par complémentarité, la base présente dans l'échantillon. La succession des lectures permet de déduire la séquence finale de l'agent présent dans l'échantillon (Fig. 2).



Cependant, cette approche d'amplification est particulièrement efficace depuis des matrices acellulaires telles que le sérum par rapport à des matrices tissulaire où la proportion en acides nucléiques de l'hôte est beaucoup plus importante. Puis l'ADN amplifié est fragmenté puis marqué à la biotine en position terminale. Ces fragments d'ADN simple brin sont hybridés sur les puces à ADN et plusieurs lavages sont réalisés afin d'éliminer les hybridations non spécifiques. Les réactions d'hybridations positives sont révélées ainsi que la visualisation par scan de la puce est obtenue. L'analyse des données permet d'identifier le ou les pathogènes présents dans l'échantillon (Fig. 3).

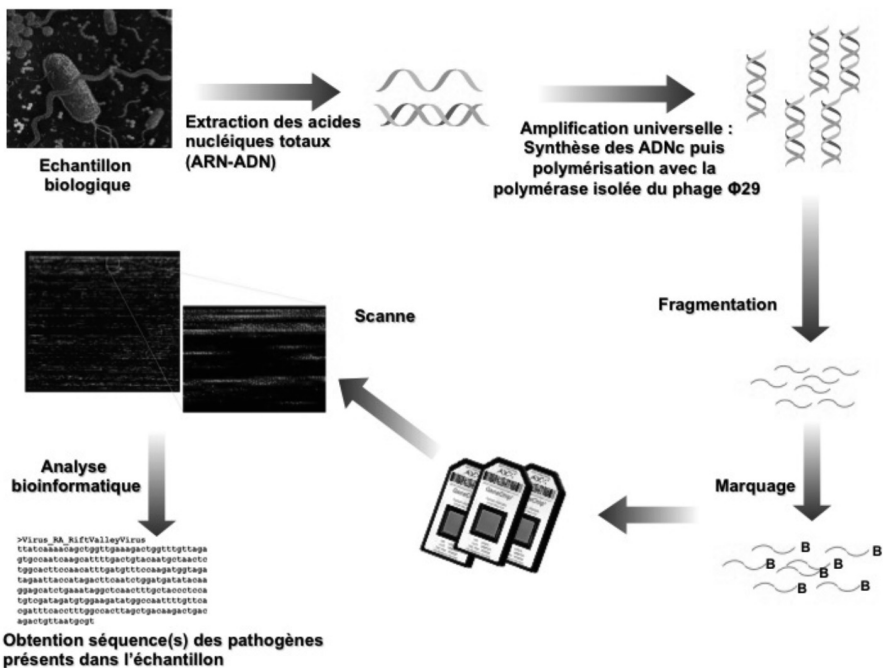


FIG. 3. — Approche globale de l'identification d'un pathogène à partir d'un échantillon biologique par hybridation sur une puce à ADN de reséquençage. L'identification d'un agent microbiologique par les puces à ADN commence par l'extraction de tous les acides nucléiques à partir d'un échantillon biologique où ils sont amplifiés de façon universelle et non spécifique puis fragmentés de façon enzymatique. Ceux-ci sont marqués à la biotine en position terminale, hybridés sur les puces à ADN qui sont lavées. Les réactions d'hybridations positives sont révélées par scan. L'analyse des données ainsi obtenues permet d'identifier le ou les pathogènes présents dans l'échantillon.

## PRINCIPE D'ANALYSE DES DONNÉES DE RESÉQUENÇAGE

L'analyse des données générées par les puces à ADN de reséquençage ou de la détection de la variation est une étape clé de la maîtrise de cette technologie. Au fur et à mesure de l'utilisation et des applications dans divers domaines, de nombreuses améliorations ont été apportées au processus analytique principalement au niveau de la détermination de l'exactitude dans les variations détectées par une approche statistique nommée ABACUS (*Adaptative Background genotype Calling Scheme*). Cette méthode assigne à chaque base un score de qualité (QS pour *Quality Score*) [12]. Si la valeur du QS est au-dessus du seuil défini, la base est validée. Dans le cas contraire, la base est indéterminée et un N est assigné à cette position. Afin de gérer les quantités importantes de données générées par ces études des méthodes d'analyse automatique ont été développées. Elles permettent non seulement de réduire le temps de calcul, de standardiser les identifications, mais aussi d'améliorer l'analyse des faux positifs dans les zones à fort taux de mutations ce qui peut notamment affecter la qualité des études portant sur l'évolution des pathogènes. En effet, un algorithme spécifique a permis d'améliorer l'analyse des bases ambiguës, en particulier dans les zones à fort taux de mutations ou ayant un rôle dans la virulence, la résistance aux antiviraux ou l'échappement vaccinal. Celui-ci s'est révélé être efficace pour la surveillance des virus grippaux en améliorant nettement la valeur du taux d'appel global avec un haut niveau d'exactitude [13]. Afin de détecter un mélange de plusieurs pathogènes spécifiques, il est important d'améliorer la reconnaissance de chaque séquence nucléotidique obtenue. Une première approche consiste à rechercher une homologie entre la ou les séquence(s) obtenue(s) et celles présentes dans les banques telles que GenBank, soit par des alignements deux à deux ou multiples, soit en utilisant BLAST [14]. En fonction de la taille de la séquence obtenue et du nombre de bases indéterminées, une assignation taxonomique peut ainsi être attribuée. Cependant, la détection de pathogènes fortement variables peut conduire à obtenir des séquences plus courtes. Ces dernières années, de nombreux algorithmes ont été développés pour permettre d'identifier correctement une liste de pathogènes définie mais aussi de réaliser un filtrage des séquences obtenues, ce qui a permis d'utiliser les séquences partielles, en nombre plus important, après déplétion des zones comprenant des bases indéterminées [15-16]. L'ensemble de ces séquences filtrées peut être utilisé pour construire une séquence consensus qui permet d'améliorer la longueur et l'exactitude de la séquence d'un virus qui n'aurait pas été déposée sur la puce [16]. Ces algorithmes fonctionnent aussi bien pour l'identification de séquences virales que pour le typage de plusieurs souches bactériennes de références de *Clostridium botulinum* [17]. En effet, parmi toutes les séquences retenues après filtrage, la quasi-totalité des séquences où un résultat de BLAST a pu être obtenu ont permis une assignation taxonomique cohérente de la souche, ou du moins jusqu'au niveau du toxinotype. De plus, l'analyse détaillée de toutes les séquences non retenues a indiqué qu'aucune séquence informative n'avait été exclue par cette approche de filtrage [17].

## APPLICATIONS POUR MIEUX COMPRENDRE UNE ÉMERGENCE MICROBIENNE

Historiquement, la première puce à ADN de reséquençage produite spécifiquement pour la détection d'un agent viral fut développée par Kozal et *al.* et utilisée pour la recherche de mutations de résistance de souches de VIH-1 (Virus de l'Immunodéficience Humaine de type 1) aux antirétroviraux [18]. Cependant, l'évolution très rapide du génome du VIH ne rendit pas cette approche concluante sur le long terme, malgré des résultats prometteurs. L'identification de plusieurs espèces de *Mycobacterium* ainsi que la détection de la résistance à la rifampicine ont pu être réalisées en se basant également sur la détection de variations ponctuelles spécifiques [19-20]. D'autres puces furent également mise au point aussi bien pour le séquençage d'une région d'environ de 29kb du génome de *Bacillus anthracis* que pour la caractérisation de *Neisseria meningitidis* en se basant sur le reséquençage de plusieurs régions variables du génome du pathogène [21-22]. En parallèle, deux principales plateformes de puce à ADN de reséquençage, RPM et PathogenID ont été développées ces dernières années pour l'identification et la caractérisation moléculaire de plusieurs pathogènes, bactériens et viraux, impliqués dans différentes types de pathologies. Elles diffèrent par la taille du « spot » d'où un nombre de bases lues différentes, ainsi que par la stratégie de sélection des sondes. Dans le cadre de cette revue, trois exemples d'applications de cette technologie seront présentés. Ces résultats se basent sur l'utilisation de la deuxième version d'une puce à ADN développée pour la détection d'agents viraux et bactériens. Très récemment, une troisième génération a été conçue pour l'identification et la caractérisation d'un grand nombre d'agents viraux d'importance aussi bien médicales que vétérinaires. Ces travaux illustreront l'apport que cet outil peut fournir dans la compréhension d'une émergence bactérienne ou virale.

### Une émergence virale, le virus influenza A(H1N1)pdm comme paradigme

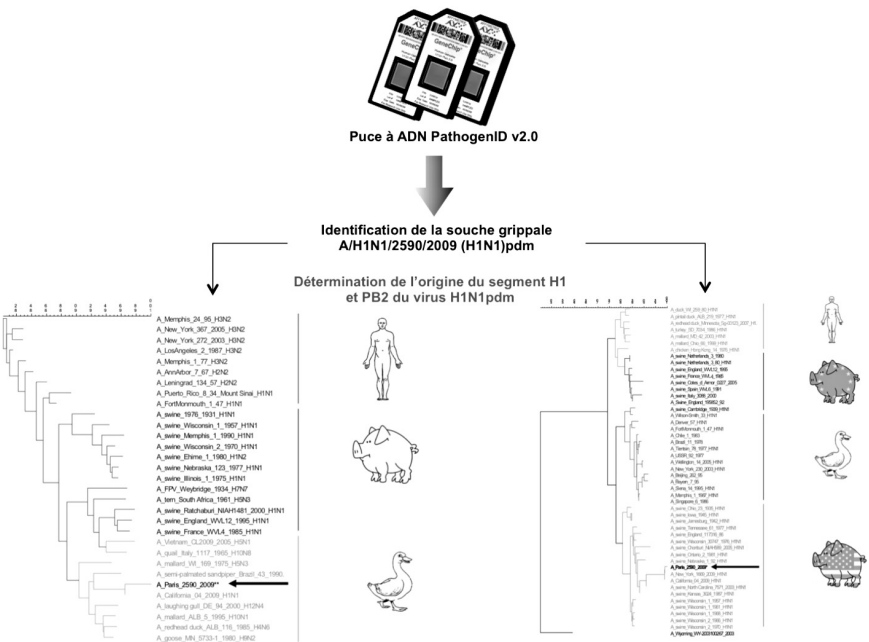
Dans ce premier travail, nous avons montré comment cet outil technologique pouvait donner accès rapidement au diagnostic d'un virus émergent pour un laboratoire focalisé sur ces problèmes de santé publique. Les virus grippaux de type A (VGA) ont été utilisés comme paradigme. La classe des Oiseaux constitue la clé de voûte de l'écologie des VGA car elle abrite tous les VGA formés de la combinaison d'un type moléculaire d'hémagglutinine (H) parmi les 16 possibles, et d'un type moléculaire de neuraminidase (N) parmi les 9 possibles. De plus, les VGA dont le génome est constitué de 8 segments d'ARN simple brin, qui dérivent tous de virus aviaires, circulent chez de nombreuses espèces animales (cheval, porc) et chez l'homme. Cet outil permet le reséquençage simultané de plusieurs portions des différents segments du génome des VGA. En effet, 5 séquences du complexe de polymérisation (PB2) et une du gène de la matrice (M) ont été déposées sur le support et correspondent à un variant donné représentatif d'un groupe bien défini.



En revanche, les séquences sélectionnées pour le séquençage des gènes de l'hémagglutinine (H) et de la neuraminidase (N) sont représentatives des principaux groupes de H (H1, H2, H3, H5, H7 et H9) et de N (N1, N2 et N3) et correspondent à un consensus couvrant une longue période de temps et une large variété de réservoirs. Cette stratégie a permis de détecter et d'identifier des souches de référence de VGA de H3N2 humaines ancienne (1975) ou récente (2003), ainsi que des souches porcines et équines [23]. Les données des deux segments PB2 et M sont compatibles avec l'hypothèse du réassortiment entre les souches équines H7N7 et aviaires H3N8 survenu à la transition des années 1950 et 1960. Après les différentes étapes de validation de cette puce, celle-ci a rendu possible l'identification complexe de plusieurs VGA isolés à partir de prélèvements cloacaux venant de canards sauvages. Cette approche décrit très rapidement le génotype des VGA. De plus, elle doit permettre de confirmer ou d'infirmer la présence d'un VGA H5N1. En effet, l'analyse des séquences obtenues avec différents virus H1N1 pandémiques et saisonniers provenant aussi bien de cas humains que d'un surnageant de culture ont permis de déterminer : 1) la période de circulation de ces virus (2008-2009) et 2) l'origine des différents segments du virus H1N1 pandémique. Ces résultats indiquaient une origine aviaire du gène PB2 et une origine porcine pour les gènes H1 et N1, appartenant aux lignages nord-américain et eurasiens, respectivement. Ces origines géographiques et des hôtes réservoirs pour ces segments H1, N1 et PB2 ont été déterminées sur la base des analyses phylogénétiques (Fig. 4) [24].

### Investigation virologique de lésions maculo-papulaires

Depuis que l'éradication de la variole a été réalisée en 1977, la principale orthopoxvirose est causée par le virus de la variole du singe (Monkeypox virus) qui appartient à la famille des *Poxviridae*, sous-famille des *Chordopoxvirinae*, genre *Orthopoxvirus* (OPXV). Nous avons tout d'abord montré la possibilité de détection et de différenciation du MonkeyPox par rapport à la variole, par la détection de 3 mutations spécifiques du virus MonkeyPox directement à partir d'un prélèvement cutané d'une microépidémie au Soudan en 2005 [25]. De plus, la mise en évidence de la présence de séquences de bactéries avait permis dans la même expérience de confirmer l'origine humaine du prélèvement avec la détection de séquence d'ADN mitochondrial humain ainsi que la présence d'un *Staphylococcus aureus* multirésistant sur-infectant probablement la croûte cutanée de l'enfant analysé. Plus récemment en juin 2010, deux enfants Pygmées vivant à la bordure d'une forêt profonde en République Centrafricaine avaient présenté des éruptions cutanées associées à des pustules et des croûtes cutanées après avoir mangé un rongeur sauvage. Cette deuxième génération de cette puce à ADN a été appliquée dans l'investigation de cette microépidémie due au virus MonkeyPox. Les séquences obtenues du gène de l'hémagglutinine et de l'ADN polymérase-ADN dépendante ont permis de confirmer l'identification du virus suspecté et de déterminer également son génotype appartenant au clade Zaire [26].



\*Séquence générée par notre puce à ADN de reséquençage.

FIG. 4. — Identification du virus de la grippe A(H1N1)pdm avec la puce à ADN de reséquençage PathogenID v2. Après une mise en culture du virus de la grippe A(H1N1) pandémique à partir d'un cas humain, les acides nucléiques totaux ont été extraits, amplifiés puis hybridés sur cette puce à ADN. Les séquences ainsi générées après analyse pour les fragments des gènes PB2 et de H1 ont été analysées dans deux arbres phylogénétiques basés sur la méthode de *Neighbor-Joining* avec les séquences disponibles dans GenBank. L'analyse des séquences a permis de déterminer une origine aviaire du gène PB2 et une origine porcine appartenant au lignage nord-américain pour le gène H1 [24].

### Identification et typage moléculaire des *Clostridium* productrices de neurotoxines

Le botulisme est une maladie grave causée par une toxine neuroparalytique botulique qui se caractérise par une paralysie flasque aigüe descendante. Ces toxines connues comme parmi les plus puissantes sont produites par la bactérie *Clostridium botulinum* et qui peuvent être subdivisées en sept toxinotypes qui se basent sur leurs propriétés antigéniques. La diversité génétique rencontrée au niveau des différents gènes composant les clusters nécessite le développement d'outils de techniques moléculaires pour la détection et le typage simultanés des différents toxinotypes de *C. botulinum*. L'association de méthodes d'amplifications universelles couplées à des technologies de séquençage ciblant simultanément plusieurs marqueurs moléculaires pourrait permettre de dépasser certaines limitations observées. Cette puce à ADN de reséquençage PathogenID possède de nombreux gènes ciblant aussi bien les gènes de ménage pour la détection et la caractérisation des espèces de bactéries

du genre *Clostridium* que des séquences ciblant les gènes codant le complexe de la neurotoxine botulique des toxinotypes A à G. De plus, des séquences supplémentaires ont été déposées sur la puce pour la détection et la caractérisation d'autres toxines spécifiques, produites par d'autres espèces de *Clostridium*. La capacité de cette puce à détecter les souches et sous-types de *C. botulinum* produisant des neurotoxines a été évaluée en utilisant de l'ADN génomique extrait à partir de onze souches caractérisées et représentatives des principaux toxinotypes et sous-types [17]. Afin de démontrer la capacité de l'approche de reséquençage à détecter les séquences cibles de *C. botulinum* dans le contexte d'une flambée de cas d'origine alimentaire, des extraits d'ADN obtenus à partir de salades naturellement contaminées ont été analysés [27]. Le reséquençage a été réalisé sur des extraits d'ADN obtenus à partir d'une culture d'enrichissement de 48h à partir d'un échantillon de salade. À partir de cet extrait, 15 séquences prototypes de *C. botulinum* ont été retenues et identifiées après filtrage. En outre, des séquences ciblant des gènes domestiques et des gènes de toxines spécifiques de *C. perfringens* ont été également récupérées. Toutes ces séquences ont permis d'obtenir au moins un résultat de BLAST et ceux-ci ont tous désigné l'espèce *C. botulinum* comme identité taxonomique appartenant au toxinotype de sous-type A2 [17]. Lorsque la souche de références de *C. botulinum* de type G a été testée, l'analyse des informations obtenues avec les séquences des gènes de résistance déposés sur le support a permis de mettre en évidence la présence d'un nouveau variant d'une ligase D-Alanine:D-Serine de type VanG. L'obtention de la séquence entière du gène de cette ligase et des séquences environnantes a pu être déterminée après l'obtention du génome de cette souche par séquençage à haut-débit. Les analyses bioinformatiques ont confirmées la présence de ce nouveau variant ainsi que la présence de ce gène au sein d'un groupe apparenté aux opérons de type *van* responsables de la résistance à la vancomycine.

## CONCLUSIONS

Depuis le développement de la première puce à ADN sur membrane, les améliorations technologiques successives ont permis de disposer d'une large gamme de puces à ADN différentes. Cette diversité technologique offre aujourd'hui la possibilité de cibler une application spécifique en fonction des besoins recherchés aussi bien au niveau de la précision du résultat attendu que du nombre de paramètres biologiques à étudier. Comme toutes les technologies, cette approche de puce à ADN de reséquençage possède des avantages et inconvénients. Au nombre des avantages, elle permet le regroupement, en une seule étape, d'une centaine voire plus d'expériences de détections individuelles et la recherche simultanée de nombreux pathogènes tout en cartographiant plusieurs régions du génome d'un virus ou d'une bactérie. Ceci permet aussi d'optimiser la logistique en limitant le nombre de réactifs nécessaires. De plus, quand le principal pathogène suspecté n'est pas détecté, elle permet de réaliser un diagnostic différentiel notamment lorsque la symptomatologie n'est pas

évocatrice. La principale limite cependant réside dans la détection plus difficile de variants trop éloignés des séquences déposées sur la puce, et ce malgré le développement d'approches originales dans la sélection des sondes. La sensibilité de détection d'un agent pathogène est plus faible par rapport à une qPCR spécifique ou vis à vis des nouvelles technologies de séquençage depuis une matrice acellulaire. Cependant, elle peut apporter en une seule étape très rapidement une information sur l'origine de la souche et l'analyse des données générées reste plus simple que pour celles obtenues par séquençage à haut-débit [28-29]. Le coût financier d'une puce à haut-débit apparaît souvent comme une limite supplémentaire. Cependant, le coût de revient final peut diminuer si la demande augmente comme cela peut être le cas dans le cadre d'une utilisation en diagnostique clinique. À ce jour, une seule puce à ADN à haute densité de reséquençage (RPM-Flu v3.1), ciblée sur la détection des virus grippaux, a obtenu une autorisation d'utilisation par la FDA (*Food and Drug Agency*). D'un point de vue pratique, il est très peu probable qu'une puce à ADN ciblant une multitude d'agents puisse obtenir une autorisation de mise sur le marché pour le diagnostique, compte tenu de la difficulté à valider la totalité des cibles et des variants. Les nombreux travaux ont permis de mieux appréhender leurs limites et ainsi de pouvoir cibler plus précisément les contextes et les situations dans lesquelles elles peuvent être utilisées. Ces puces à ADN de reséquençage doivent plutôt être intégrées comme des outils d'expertise complémentaires à ceux existant déjà, afin d'aider les acteurs et les autorités de santé à orienter les décisions lorsqu'une situation d'urgence apparaît.

## REMERCIEMENTS

Ces différents travaux ont été soutenus financièrement par le *U.S. National Institute of Allergy and Infectious Diseases, National Institute of Health* (UC1 AI062613 — G.C. Kennedy), par l'Institut Pasteur à travers deux programmes de recherche transversaux (PTR DEVA n° 246 et CEVACAR n° 385), la fondation Total-Institut Pasteur et le prix Elisabeth Taub de l'Académie Nationale de Médecine. Je remercie le Dr J.C. Manuguerra pour la relecture de ce travail et la plateforme Génotypage des pathogènes et Santé Publique de l'Institut Pasteur pour l'utilisation de la station Affymetrix.

## BIBLIOGRAPHIE

- [1] FISCHER S.A., GRAHAM M.B., KUEHNERT M.J. *et al* — Transmission of lymphocytic choriomeningitis virus by organ transplantation. *N. Engl. J. Med.*, 2006, 25, 2235-2249.
- [2] MORSE S.S. — Factors in the emergence of infectious diseases. *Emerg. Infect. Dis.*, 1995, 1, 7-15.
- [3] PALACIOS G., DRUCE J., DU L. *et al* — A new arenavirus in a cluster of fatal transplant-associated diseases. *N. Engl. J. Med.*, 2008, 6, 991-998.
- [4] BERTHET N., LECLERCQ I., VALLAËYS T. *et al* — Les puces à ADN en virologie : concepts de base et applications pour la détection des virus. *Virologie*, 2013, 17, 61-80.

- [5] FODOR S.P., RAVA R.P., HUANG X.C. *et al* — Multiplexed biochemical assays with biological chips. *Nature*, 1993, 5, 555-556.
- [6] FODOR S.P., READ J.L., PIRRUNG M.C. *et al* — Light-directed, spatially addressable parallel chemical synthesis. *Science*, 1991, 15, 767-773.
- [7] PEASE A.C., SOLAS D., SULLIVAN E.J. *et al* — Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc Natl Acad Sci USA*, 1994, 24, 5022-5006.
- [8] BARONE A.D., BEECHER J.E., BURY P.A. *et al* — Photolithographic synthesis of high-density oligonucleotide probe arrays. *Nucleosides Nucleotides Nucleic Acids*, 2001, 20, 525-531.
- [9] MCGALL G.H., CHRISTIANS F.C. — High-density genechip oligonucleotide probe arrays. *Adv. Biochem. Eng. Biotechnol.*, 2002, 77, 21-42.
- [10] MCGALL G.H., FIDANZA J.A. — Photolithographic synthesis of high-density oligonucleotide arrays. *Methods Mol. Biol.*, 2001, 170, 71-101.
- [11] BERTHET N., REINHARDT A.K., LECLERCQ I. *et al* — Phi29 polymerase based random amplification of viral RNA as an alternative to random RT-PCR. *BMC Mol. Biol.*, 2008, 9, 77.
- [12] CUTLER D.J., ZWICK M.E., CARRASQUILLO M.M. *et al* — High-throughput variation detection and genotyping using microarrays. *Genome Res.*, 2001, 11, 1913-1925.
- [13] LEE C.W., KOH C.W., CHAN Y.S. *et al* — Large-scale evolutionary surveillance of the 2009 H1N1 influenza A virus using resequencing arrays. *Nucleic Acids Res.*, 2010, 38(9), e111.
- [14] ALTSCHUL S.F., GISH W., MILLER W. *et al* — Basic local alignment search tool. *J. Mol. Biol.*, 1990, 5, 403-410.
- [15] MALANOSKI A.P., LIN B., WANG Z. *et al* — Automated identification of multiple microorganisms from resequencing DNA microarrays. *Nucleic Acids Res.*, 2006, 34, 5300-5311.
- [16] DACHEUX L., BERTHET N., DISSARD G. *et al* — Application of broad-spectrum resequencing microarray for genotyping rhabdoviruses. *J. Virol.*, 2010, 84(18), 9557-9574.
- [17] VANHOMWEGEN J., BERTHET N., MAZUET C. *et al* — Application of high-density DNA resequencing microarray for detection and characterization of botulinum neurotoxin-producing clostridia. *PLoS One*, 2013, 8(6), e67510.
- [18] KOZAL M.J., SHAH N., SHEN N. *et al* — Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. *Nat. Med.*, 1996, 2, 753-9.
- [19] GINGERAS T.R., GHANDOUR G., WANG E. *et al* — Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic Mycobacterium DNA arrays. *Genome Res.*, 1998, 8, 435-48.
- [20] TROESCH A., NGUYEN H., MIYADA C.G. *et al* — Mycobacterium species identification and rifampin resistance testing with high-density DNA probe arrays. *J. Clin. Microbiol.*, 1999, 37, 49-55.
- [21] CORLESS C.E., KACZMARSKI E., BORROW R. *et al* — Molecular characterization of Neisseria meningitidis isolates using a resequencing DNA microarray. *J. Mol. Diagn.*, 2008, 10, 265-271.
- [22] ZWICK M.E., KILEY M.P., STEWART A.C. *et al* — Genotyping of Bacillus cereus strains by microarray-based resequencing. *PLoS One*, 2008, 3(7), e2513.
- [23] LECLERCQ I., BERTHET N., BATEJAT C. *et al* — Use of consensus sequences for the design of high density resequencing microarrays: the influenza virus paradigm. *BMC Genomics*, 2010, 11, 586.
- [24] BERTHET N., LECLERCQ I., DUBLINEAU A. *et al* — High-density resequencing DNA microarrays in public health emergencies. *Nat. Biotechnol.*, 2010, 28(1), 25-7.
- [25] BERTHET N., DICKINSON P., FILLIOL I. *et al* — Massively parallel pathogen identification using high-density microarrays. *Microb. Biotechnol.*, 2008, 1, 79-86.

- [26] BERTHET N., NAKOUNE E., WHIST E. *et al* — Maculopapular lesions in the Central African Republic. *Lancet*, 2011, 8, 1354.
- [27] ORIOT C. D'ARANDA E., CASTANIER M. *et al* — One collective case of type A foodborne botulism in Corsica. *Clin. Toxicol.*, 2011, 49, 752-754.
- [28] BERTHET N., PAULOUS S., COFFEY L.L. *et al* — Resequencing microarray method for molecular diagnosis of human arboviral diseases. *J. Clin. Virol.*, 2013, 56(3), 238-243.
- [29] CHEVAL J., SAUVAGE V., FRANGEUIL L. *et al* — Evaluation of high-throughput sequencing for identifying known and unknown viruses in biological samples. *J Clin Microbiol.*, 2011, 49, 3268-3275.