



HAL
open science

Actualités du séquençage d'ADN

Bertrand Jordan

► **To cite this version:**

Bertrand Jordan. Actualités du séquençage d'ADN. Médecine/Sciences, 2019, 35 (1), pp.81-83. <10.1051/med-sci/2018306>. <hal-02970143>

HAL Id: hal-02970143

<https://hal.science/hal-02970143v1>

Submitted on 17 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Chroniques génomiques

Actualités du séquençage d'ADN

Bertrand Jordan



Une absorption qui fait du bruit : *Illumina* et *PacBio*

L'annonce du prochain rachat de *Pacific Biosciences* (*PacBio*) par le géant *Illumina* a pris par surprise le petit monde du séquençage d'ADN et suscité de nombreux commentaires [1]. La nouvelle est d'importance car *PacBio* était un des seuls concurrents sérieux d'*Illumina* ; cette absorption, si elle est confirmée (elle pourrait tomber sous le coup des lois antitrust), mettrait en effet *Illumina* dans une position quasi-hégémonique et aurait de multiples conséquences. Rappelons que l'entreprise est actuellement dominante sur le marché du séquençage d'ADN, au point que l'on oublie parfois qu'elle a encore quelques concurrents. Son point fort est le séquençage à très haut débit et à bas coût, soit moins de 1 000 dollars pour un génome humain lu à une redondance de 30 fois, pour ses machines les plus productives. Son point faible est la taille des lectures effectuées, environ 150 nucléotides (lectures courtes), ce qui est gênant tout particulièrement pour le séquençage *de novo*. La technologie mise au point par *PacBio* [2]¹ permet au contraire des lectures beaucoup plus longues, jusqu'à une dizaine de kilobases, moyennant un coût qui diminue au fil des années mais est encore cinq fois supérieur à celui d'*Illumina*. Les deux approches sont donc complémentaires, et dans les grands centres de séquençage, on trouve souvent une batterie de machines *Illumina* et un ou deux appareils *PacBio*. L'incorporation de *PacBio* au sein d'*Illumina* permettrait d'intégrer les deux approches et sans doute d'accélérer l'amélioration de l'approche *PacBio* grâce aux capacités technologiques et financières d'*Illumina*. Le rachat se ferait moyennant une somme (1,2 milliards de dollars) qui représente presque le double de la capitalisation boursière actuelle de *PacBio* [1].



UMR 7268 ADÉS, Aix-Marseille, Université/EFS/ CNRS ; CoReBio PACA, case 901, Parc scientifique de Luminy, 13288 Marseille Cedex 09, France.
brjordan@orange.fr

Un peu de recul

En fait, le monde du séquençage d'ADN ne se limite pas à ces deux entreprises, et un rapide tour d'horizon s'impose pour avoir une vision correcte des forces en présence. L'entreprise dominante est donc *Illumina*, spécialiste des machines très productives avec des lectures courtes, avec toute une gamme d'appareils dont le plus puissant, le NovaSeq 6000, peut fournir jusqu'à 6 térabases² en deux jours, un chiffre d'affaires de près de trois milliards de dollars (avec un bénéfice net de 26 % de ce chiffre) et des milliers d'appareils installés à travers le monde. Sur le plan technique, il s'agit de séquençage par synthèse, après amplification par PCR (*polymerase chain reaction*). *PacBio*, avec un chiffre d'affaires d'un peu moins de cent millions de dollars (et un déficit presque équivalent), ne joue clairement pas dans la même cour malgré l'intérêt de sa technologie, qui repose sur la lecture de molécules uniques sans étape d'amplification grâce à des systèmes optiques très sophistiqués [2]. Une troisième entreprise, *Oxford Nanopore Technologies* (ONT), exploite, elle, un tout autre principe de séquençage : les molécules d'ADN sont lues alors qu'elles passent au travers de nanopores disposés sur une membrane [3] (→).

(→) Voir la Chronique génomique de B. Jordan, *m/s* n° 8-9, août-septembre 2017, page 801

Le taux d'erreurs est élevé (plus de 5 %), mais des lectures très longues sont possibles (plus de mille kilo-

¹ On trouvera une description succincte mais claire des différentes techniques dans la récente revue de Heather et Chain [2].

² Soit plusieurs dizaines de génomes humains (6 gigabases chacun) lus avec une redondance de 30 fois (une térabase = 1 000 gigabases).

bases !). Le système est ultra-compact (de la taille d'une grosse clé USB) et bon marché (moins de 1 000 euros, mais en principe à usage unique). Le chiffre d'affaires de cette entreprise britannique (non cotée en bourse) est de l'ordre de 20 millions de dollars : c'est actuellement une *niche player*, une entreprise s'adressant à un marché de niche, qui a néanmoins une forte présence médiatique grâce à son principe original et à son adéquation pour des analyses sur le terrain, comme cela a été le cas lors de l'épidémie d'Ébola en Afrique de l'Ouest [4].

D'autres systèmes ont moins de présence médiatique, mais n'en sont pas moins solidement implantés : *Thermo Fisher*, qui a repris le système mis au point par *Ion Torrent* fondé sur le séquençage par synthèse, mais en utilisant une détection électronique et non optique [2]. Cela représente quelques centaines de machines principalement utilisées pour le séquençage clinique (séquençage de quelques dizaines ou centaines de gènes dans des échantillons tumoraux). Dans le même secteur, il faut mentionner *Qiagen*, qui commercialise un système de séquençage à visée clinique appelé *Gene Reader*, utilisant une technique proche de celle d'*Illumina* mais intégrée avec des kits de séquençage très complets. Il est difficile d'évaluer le nombre de systèmes installés, mais il est probablement de l'ordre de la centaine. Par ailleurs, certains acteurs ont disparu, comme *454* (premier acteur du secteur, racheté par *Roche* qui semble l'avoir laissé mourir), *Helicos* (pionnier du séquençage sur molécule unique, qui a fait faillite en 2012³) et *SOLiD*, utilisant le séquençage par hybridation, racheté par l'entreprise *Life Technologies* elle-même rachetée par *Thermo Fisher* mais dont les appareils ne sont plus commercialisés. Mentionnons enfin *Complete Genomics*, qui s'était positionné sur le marché du séquençage à façon avec ses propres machines (séquençage par hybridation après amplification linéaire par une technique originale baptisée *DNA nanoballs*). Cette firme annonçait une montée en puissance très impressionnante, mais n'a pas trouvé son marché et a été rachetée en 2013 par le conglomérat chinois *BGI* (anciennement *Beijing Genomics Institute*) pour 118 millions de dollars, une misère...

BGI, parlons-en, justement. On connaît cet institut de génomique géant, implanté principalement à Shenzhen (Chine) et intervenant dans de multiples domaines, depuis le séquençage à haut débit jusqu'aux porcs transgéniques en passant par la bioinformatique. Lors de sa création en 1999, il était équipé de machines occidentales (*Molecular Dynamics MegaBACE*, puis *Illumina* et *Ion Torrent*). Mais après l'achat de *Complete Genomics* en 2013, *BGI* a mis au point plusieurs modèles de séquenceurs fondés sur la technique des *DNA nanoballs*, notamment le *BGISEQ500*, et a tout récemment annoncé une machine haut de gamme, le *MGISEQ-77*, qui pourrait produire jusqu'à 6 térabases de séquences par jour, soit un débit équivalent à celui de la machine *Illumina NovaSeq 6000* mentionnée plus haut. Il convient d'être prudent, car une précédente tentative de *BGI*, le lancement en 2015 d'une plateforme haut de gamme appelée *Revology* (et censée concurrencer le système *Illumina HiSeq 10X*, haut de gamme de l'époque), s'était soldée par un flop retentissant avec l'abandon de sa commercialisation quelques mois plus tard. Mais les

appareils *BGISEQ 500* ont été comparés avec succès aux machines *Illumina HiSeq* et semblent présenter des performances équivalentes [5, 6]. Ils sont assez répandus en Extrême-Orient, avec près de 1 000 appareils déjà en place, principalement en Chine. Leur commercialisation en Europe est prévue, quelques machines y sont déjà installées, et le marché Nord-Américain n'est pas exclu à terme. S'agissant d'une méga-entreprise chinoise, il faut sans doute faire la part d'une certaine tendance à l'anticipation et au gonflement des performances annoncées, mais les équipes de l'Empire du Milieu ont montré leur capacité à progresser très rapidement, et *BGI* pourrait devenir un acteur important dans le marché mondial des systèmes de séquençage – il l'est déjà en Chine.

Un peu de prospective

L'absorption de *PacBio* par *Illumina* – si elle a bien lieu, ce qui semble très probable – devrait permettre l'intégration des deux approches (lectures courtes/lectures longues) qui est actuellement délicate en raison de la nécessité de combiner des informations provenant de deux systèmes très différents et fournissant les résultats dans des formats informatiques incompatibles. De plus, les moyens financiers et techniques d'*Illumina* devraient lui permettre d'accélérer l'évolution déjà en cours du système *PacBio* vers un débit plus important et un coût plus faible. Au total, *Illumina*, qui est déjà en position dominante, deviendrait quasiment imbattable avec une offre complète et bien intégrée. Cette acquisition, qui va coûter à *Illumina* un peu moins de deux fois son bénéfice annuel, est donc bien stratégique et susceptible d'assurer à la firme une domination durable du marché du séquençage – sauf surprise majeure. Dans ce paysage, les « petits » systèmes à visée clinique (*Thermo Fisher*, *Qiagen*) pourraient garder leur place, si du moins ils parviennent à évoluer assez vite pour garder leurs avantages (facilité d'emploi, adaptation poussée à des besoins spécifiques) tout en restant compétitifs au niveau des coûts.

Restent les inconnues. Quel est le devenir d'*Oxford Nanopore Technologies* ? L'entreprise fait depuis plusieurs années des efforts pour pénétrer sur le marché du séquençage de masse, avec des appareils qui regroupent plusieurs cellules de lecture (cinq pour le *GridION*, 48 pour le *PromethION*) de manière à atteindre (pour le *PromethION*) des débits comparables à ceux d'un système *Illumina* haut de gamme. Ces tentatives se heurtent à de nombreuses difficultés techniques et n'ont pas jusqu'ici donné de résultats très convaincants. On peut se demander si *ONT* ne ferait pas mieux de cultiver son marché de niche (appareils très bon marché, ultra-portables et permettant des lectures ultra-longues). À moins qu'un progrès

³ Mais dont la technologie est reprise par la firme chinoise *Direct Genomics* (<http://www.directgenomics.com/>) dont le fondateur n'est autre que He Jiankui, le « père » des jumelles génétiquement modifiées qui défraient (à juste titre) la chronique.

