

# Chroniques génomiques

## Notre ADN est partout !

Bertrand Jordan



### Analyser l'ADN dans l'environnement

L'extraordinaire amélioration des techniques de détection et d'analyse de l'ADN [1] (→) fait qu'il est aujourd'hui possible de récupérer l'ADN présent dans le sol d'une caverne ou même le sable d'une plage et d'en déduire les espèces qui ont été présentes dans cet environnement [2]. En pratique, on agite l'échantillon de terre ou de sable avec une solution *ad hoc*, on récupère le surnageant et on le fait passer sur un filtre qui va retenir les cellules et les fragments d'ADN. Pour analyser l'eau ou l'air, on en fait passer un certain volume sur les mêmes filtres [3]. Ceux-ci sont alors traités pour en extraire l'ADN qui sera ensuite analysé. Dans un premier temps, on a recherché dans cet ADN des séquences spécifiques d'une espèce par des techniques comme la PCR (*polymerase chain reaction*) quantitative, ce qui donne une bonne sensibilité mais, bien sûr, limite l'étude aux espèces définies *a priori* et correspondant aux amorces de PCR employées. Plus récemment, la baisse toujours rapide des coûts de séquençage [1] a rendu possible et relativement abordable une approche globale dans laquelle on séquence l'ensemble de l'ADN récupéré (*shotgun sequencing*) et l'on identifie ensuite les fragments correspondant à une espèce donnée par voie informatique, en les alignant sur la séquence génomique de cette espèce. On obtient ainsi une vue non biaisée des différents ADN présents dans l'échantillon, et on peut quantifier la proportion de chacun d'eux en comparant le nombre de lectures (*reads*)<sup>1</sup> spécifiques d'une espèce par rapport au total des (courtes) séquences obtenues.

(→) Voir la Chronique  
génomique de B. Jordan,  
*m/s* n° 5, mai 2023,  
page 469

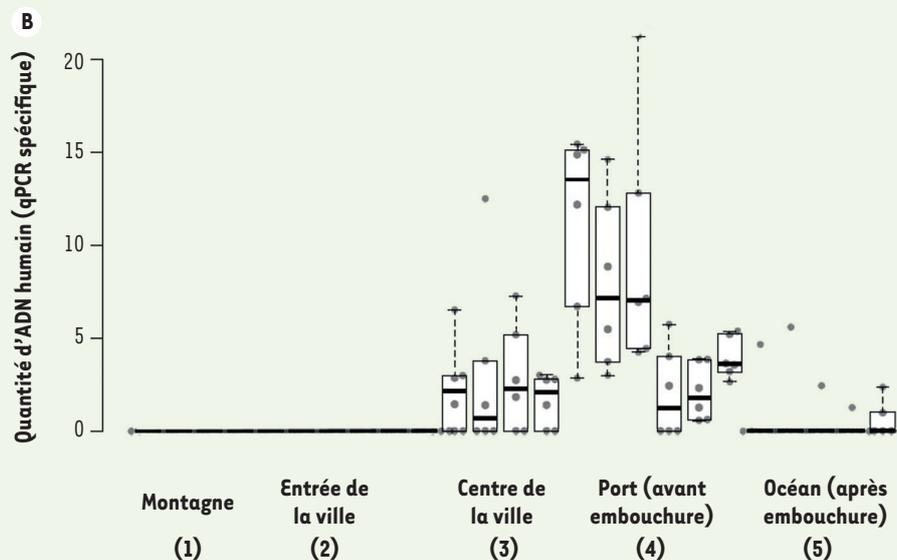


Biologiste, généticien et immunologiste, Président d'Aprogène (Association pour la promotion de la Génomique), 13007 Marseille, France.  
[brjordan@orange.fr](mailto:brjordan@orange.fr)

### La présence fréquente d'ADN humain

L'article qui fait l'objet de cette chronique [4] émane de groupes qui s'intéressent à la biologie de la tortue de mer et, à cet effet, effectuaient des analyses cherchant à caractériser la présence dans divers environnements d'ADN provenant de ces animaux et de leurs virus. Procédant par *shotgun*, ils ont pu analyser les séquences obtenues en tentant de les aligner avec la séquence du génome humain, et ont eu la surprise de constater que notre ADN est présent dans une grande variété d'échantillons, par exemple dans des traces de pas sur le sable. Il ne s'agit pas d'une contamination expérimentale puisque cet ADN est absent de prélèvements effectués dans des zones interdites au public ; par ailleurs, l'analyse de l'eau d'une rivière (la rivière Avoca en Irlande), depuis sa source jusqu'à l'embouchure, montre que l'ADN humain n'apparaît qu'au moment où elle traverse la petite ville d'Arklow et persiste ensuite (tout en diminuant) jusqu'à l'embouchure (*Figure 1*). C'est donc bien du véritable ADN humain qui est détecté dans ces analyses. Les auteurs montrent aussi l'obtention d'ADN humain à partir de l'air d'une pièce si celle-ci est occupée par des personnes. Bien sûr, il n'est présent qu'à l'état de traces dans ces différents prélèvements, et les séquences d'origine humaine sont très minoritaires. Le séquençage d'un échantillon avec les dernières machines (*Illumina Novaseq 6000*) fournit plusieurs milliards de *reads* par session ; les séquences humaines (identifiées par voie bioinformatique) constituent au maximum 0,1 ou 0,2 % de cet ensemble pour les échantillons positifs. C'est très faible, mais néanmoins significatif compte tenu de la qualité des données et du résultat des contrôles négatifs.

<sup>1</sup> Il s'agit là de séquençage de type Illumina, qui produit en une session un très grand nombre (plusieurs milliards) de courtes séquences (100 à 200 nucléotides).



**Figure 1. ADN humain dans l'eau de la rivière Avoca (Irlande), de la source à l'embouchure. A. Vue satellite de la rivière Avoca** (source : OpenStreetMap) **avec indications des zones de prélèvement.** Les chiffres entre parenthèses correspondent aux zones indiquées dans la Figure 1B. **B. Contenu en ADN humain des différents prélèvements.** Pour cette étude très détaillée, l'ADN humain a été caractérisé par PCR spécifique, approche plus économique que le séquençage en *shotgun*. Chaque mesure individuelle est représentée par un point, la barre indique la médiane. Malgré la dispersion des points, on voit clairement que l'ADN humain n'apparaît dans l'eau qu'à la traversée de la ville (extrait partiel et remanié de la figure supplémentaire 1 de [4]).

## Un ADN révélateur

On pourrait penser que l'ADN humain récupéré à partir de tels prélèvements est très fragmenté et que, du coup, son analyse fournit peu d'informations sur son origine. Ce n'est pas le cas. Les auteurs ont répété leurs expériences en utilisant le séquençage par nanopores (*Oxford Gene Technology*) [1] qui permet de lire de longues séquences si l'ADN analysé est en bon état. Ils ont obtenu de très longues séquences, généralement quelques kilobases et jusqu'à presque 150 kilobases, montrant ainsi qu'une partie au moins de cet ADN « environnemental » est en bon état et suggérant que son analyse peut fournir des informations sur les individus dont il provient. *A priori*, l'ADN récupéré va provenir de plusieurs personnes (sauf peut-être pour les traces de pas), ce qui complique un peu l'interprétation. Néanmoins, les auteurs montrent que l'on peut, à partir de ces résultats, retrouver des délétions fréquentes dans la population européenne, et même évaluer la proportion de différents haplotypes caractéristiques au sein des échantillons – donc avoir une idée des groupes d'ascendance (« groupes ethniques ») dans la population qu'ils représentent. Différentes mutations mitochondriales sont également détectables. On peut regretter ici que les auteurs n'aient pas poussé plus loin leurs analyses sur l'ADN humain récupéré dans des empreintes de pas, qui provient vraisemblablement d'une seule personne et aurait permis des conclusions plus nettes<sup>2</sup>. Quoi qu'il en soit, on voit que l'activité ou la simple présence humaine laisse une indéniable trace ADN, et que celle-ci est susceptible de fournir des informations précises sur ceux qui en sont à l'origine.

Dans le même ordre d'idées, une équipe japonaise spécialisée dans l'étude du microbiome humain vient de démontrer [5] que les séquences obtenues par séquençage global du microbiome contiennent suffisamment d'ADN humain pour permettre l'identification du donneur de l'échantillon. Du coup, ils prévoient de retirer les séquences de microbiome déjà déposées dans les bases de données publiques pour en éliminer les séquences humaines avant de les redéposer, afin d'éviter la diffusion des données génétiques personnelles des donneurs.

## Surprises, opportunités et problèmes

La puissance des techniques d'analyse actuelles a visiblement surpris les auteurs de ce travail, qui ne s'attendaient sans doute pas à détecter de l'ADN humain (en bon état, qui plus est) partout où sont passés des

êtres humains, et même dans l'air qu'ils ont respiré ou le sable qu'ils ont foulé. Cela apparaît dans le terme qu'ils ont choisi pour désigner cet ADN : *Bycatch DNA*, qui renvoie au problème des captures accessoires (ou involontaires) dans la pêche océanique, ces espèces que l'on ne cherche pas à pêcher mais qui se retrouvent parfois dans les filets alors même qu'elles sont protégées ou sans valeur commerciale. Cette découverte ouvre évidemment des perspectives très intéressantes : on peut imaginer des études génétiques de populations (présentes ou passées) à partir de l'analyse de l'ADN dans l'environnement sans nécessité de prélèvements ; des applications en criminologie viennent immédiatement à l'esprit, tout comme la surveillance de l'environnement (détection des pollutions d'origine humaine). Mais la perspective d'accéder aux caractéristiques génétiques d'une population (peut-être même d'une personne particulière) en l'absence de tout consentement et même de toute information pose évidemment un sérieux problème de confidentialité (*privacy*) [6], d'autant que les techniques vont continuer à se perfectionner et deviendront plus précises et plus abordables. À l'heure actuelle, les études génétiques de terrain sur des espèces animales (comme celles des auteurs sur les tortues de mer) ne sont pas subordonnées à l'approbation d'un comité d'éthique humaine, mais cela ne devrait-il pas être le cas si l'on analyse de l'ADN humain, même par inadvertance ? Que devient le consentement éclairé du sujet (quasi impossible à obtenir) dans ces conditions ? Voilà des questions difficiles auxquelles il va néanmoins falloir répondre... ♦

## SUMMARY

### DNA everywhere

Advanced analysis of environmental DNA for diversity monitoring using deep sequencing reveals the presence of human DNA in many samples connected to human activity. Moreover, this DNA is in relatively good condition and can be used for genetic survey of populations and even individuals. This opens many interesting scientific opportunities but also raises serious privacy issues. ♦

## LIENS D'INTÉRÊT

L'auteur déclare n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

## RÉFÉRENCES

1. Jordan B. Séquençage d'ADN, la fin d'un quasi-monopole ? *Med Sci (Paris)* 2023 ; 39 : 469–73.
2. Farrell JA, Whitmore L, Duffy DJ. The promise and pitfalls of environmental DNA and RNA approaches for the monitoring of human and animal pathogens from aquatic sources. *Bioscience* 2021 ; 71 : 609–25. <https://doi.org/10.1093/biosci/biab027>
3. Hinlo R, Gleeson D, Lintermans M, Furlan E. Methods to maximise recovery of environmental DNA from water samples. *PLoS One* 2017 ; 12 : e0179251.
4. Whitmore L, McCauley M, Farrell JA, et al. Inadvertent human genomic bycatch and intentional capture raise beneficial applications and ethical concerns with environmental DNA. *Nat Ecol Evol* 2023 ; 7 : 873–8.
5. Tomofuji Y, Sonehara K, Kishikawa T, et al. Reconstruction of the personal information from human genome reads in gut metagenome sequencing data. *Nat Microbiol* 2023 ; 8 : 1079–94.
6. Vogel G. Human genomic 'bycatch' threatens privacy. *Science* 2023 ; 380 : 676–7.

<sup>2</sup> Ils s'en sont peut-être abstenus pour des raisons éthiques (voir la discussion à la fin de cette chronique).

