



HAL
open science

Le génome, une “ donnée personnelle ” d’un nouveau genre : nouvelles questions posées par la généralisation du séquençage de l’ADN

Hervé Seitz

► To cite this version:

Hervé Seitz. Le génome, une “ donnée personnelle ” d’un nouveau genre : nouvelles questions posées par la généralisation du séquençage de l’ADN. *Journal de médecine légale, droit médical, victimologie, dommage corporel*, 2022, Dossier spécial “ Les données génétiques ”, 65 (1), pp.17-23. <hal-03257146>

HAL Id: hal-03257146

<https://hal.science/hal-03257146v1>

Submitted on 10 Jun 2021

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

Le génome, une « donnée personnelle » d'un nouveau genre : nouvelles questions posées par la généralisation du séquençage de l'ADN

Hervé Seitz, directeur de recherche au CNRS

Institut de Génétique Humaine, UMR 9002 CNRS et université de Montpellier, 141, rue de la Cardonille, 34396 Montpellier CEDEX 5, France; email : herve.seitz@igh.cnrs.fr

Diaporama donné lors du colloque « Santé et droit de choisir » de l'école de droit de l'université de Clermont-Auvergne le 28 novembre 2019 : disponible à <https://www.igh.cnrs.fr/images/microsite/herve-seitz/files/Sequencage.pdf>

1 Introduction : le génome, une information héréditaire

De manière universelle dans le monde vivant, les caractéristiques d'un individu sont déterminées à la fois par une information innée, héritée de son ou ses parent(s), et de ses interactions avec l'environnement. L'information innée, transmise héréditairement, est appelée « information génétique », et, matériellement, elle est constituée d'une séquence de groupes d'atomes le long des molécules d'ADN des cellules constitutrices de l'organisme¹. On appelle « génome » d'un individu la totalité de la séquence de son ADN.

Le lecteur trouvera une présentation détaillée de la nature matérielle du génome, et de ses propriétés, dans l'introduction de l'article « La technologie CRISPR/Cas9 », publié dans le n°15 de la Revue du Centre Michel de l'Hospital (Seitz, 2018). Seules seront rappelées ici quelques propriétés de l'information génétique :

- La presque-totalité des cellules d'un organisme pluricellulaire contiennent la même séquence génomique : il y a donc un sens à parler du « génome d'un individu », alors que la plupart de ses cellules possède leur propre ADN.
- Le génome d'une espèce donnée est typiquement segmenté sous forme de « chromosomes », chacun portant une fraction de l'information génétique totale.
- Il existe deux modes de reproduction :
 1. La reproduction asexuée, dans laquelle un parent (unique) bourgeonne un nouvel individu à partir d'une partie de son corps. Ce nouvel individu est génétiquement identique à son parent (les cellules de l'enfant sont issues de divisions des cellules du parent, qui ont dupliqué fidèlement leur information génétique au cours de la division), on dit que l'enfant est un « clone » du parent. Ce mode de reproduction est inexistant chez l'Homme, mais il est fréquent chez d'autres espèces (exemples : la division bactérienne, le bouturage des plantes, ...).
 2. La reproduction sexuée, dans laquelle deux parents (de types sexuels compatibles) mettent en commun une partie de leurs génomes : le génome de l'enfant est un assemblage original de la moitié du génome de chacun de ses parents, il n'est donc le clone d'aucun de ses deux parents, mais a hérité de la moitié de l'information génétique de chacun (voir Figure 1).
- On appelle un « gène » une portion du génome, qui contrôle une caractéristique moléculaire précise (classiquement, un gène donné permet l'expression d'une protéine ou d'un ARN donné, qui pourra jouer une fonction ou une autre dans la cellule). Chaque chromosome porte de nombreux gènes. Les gènes contrôlent donc certaines caractéristiques des individus. Dans la reproduction asexuée, l'enfant hérite des gènes de son parent unique (il partagera donc de très nombreuses caractéristiques avec lui). Dans la reproduction sexuée, l'enfant hérite d'une copie de chaque gène par son père, et d'une autre copie par sa mère ; les caractéristiques de l'enfant seront donc similaires à celles de ses parents, sans être tout à fait identiques à celles de l'un ou l'autre.
- Matériellement, l'information génétique est constituée d'une séquence de groupes d'atomes le long de la molécule d'ADN des chromosomes. Des techniques de biologie moléculaire permettent de lire cette séquence à partir d'ADN purifié.
- Certaines caractéristiques semblent héréditaires, mais sans être contrôlées par la génétique (elles peuvent être dues à l'éducation, aux habitudes alimentaires, ..., qui ont tendance à se transmettre dans les familles). Il existe donc d'autres formes d'héritabilité que la génétique, et il n'est pas toujours simple de définir la contribution du génome, et celle des formes héréditaires du mode de vie, qui seraient responsables de la transmission de caractéristiques entre les générations.

1. Il existe également une autre forme d'information transmise héréditairement, les « modifications épigénétiques », qui peuvent modifier les caractéristiques de l'organisme, à séquence de génome constant. Ce mode de contrôle des propriétés de l'organisme est clairement mis en évidence dans de nombreuses espèces vivantes, mais il n'a jamais été observé de manière convaincante, sous sa forme transmissible intergénérationnellement, chez les Vertébrés. Il n'en sera pas question ici.

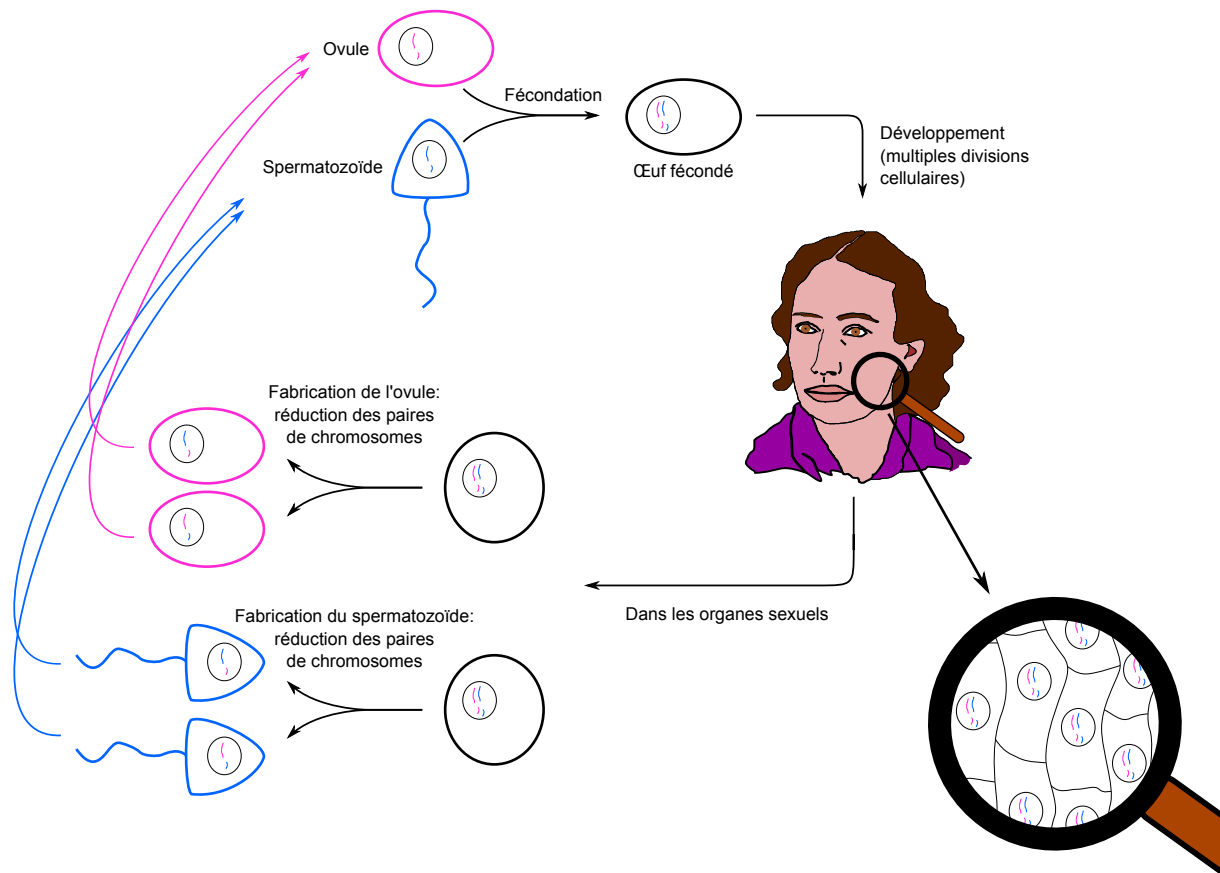


FIGURE 1 – **La reproduction sexuée (ici illustrée chez l'Homme)**. On appelle « gamète » (nom masculin) les cellules qui fusionneront pour donner un œuf fécondé (c'est à dire dans l'espèce humaine : le spermatozoïde et l'ovule). À la fécondation, chaque gamète apporte un lot de chromosomes (ici en bleu, les chromosomes d'origine paternelle, et en rose les chromosomes d'origine maternelle). L'œuf fécondé possède donc des paires de chromosomes. Cette cellule duplique ses chromosomes à l'identique avant de se diviser, puis chacune des cellules-filles en fait autant et se divise à son tour, et ainsi de suite ; au cours du développement, de nombreuses divisions cellulaires multiplie le nombre de cellules, et chaque cellule hérite d'un lot de chromosomes identiques à ceux de l'œuf fécondé initial (à de très rares exceptions près). Tout au long de sa vie adulte, l'individu possède donc, dans ses cellules, des paires de chromosomes : toute son information génétique est présente en deux copies (une copie d'origine paternelle, une copie d'origine maternelle). Dans ses organes sexuels, cet individu fabriquera de nouveaux gamètes (ce seront des spermatozoïdes s'il est mâle, des ovules s'il est femelle) : au cours d'une division cellulaire qui produit ces cellules particulières, les chromosomes ne sont pas dupliqués avant d'être répartis entre les deux cellules-filles, si bien que chaque spermatozoïde ou ovule ne reçoit qu'un seul lot de chaque chromosome, au lieu de paires entières. Lorsque cet individu se reproduira, ces gamètes pourront donc fusionner avec un gamète issu d'un individu de type sexuel compatible, et créer un nouvel assemblage génomique original dans le nouvel œuf fécondé.

2 L'organisation du génome

Dans les cellules de Mammifères, les chromosomes sont donc présents en deux copies (la copie d'origine paternelle et la copie d'origine maternelle). D'aspect extérieur, les deux chromosomes de la paire sont identiques (même taille, même contenu en gènes), à l'exception d'une paire : la paire de « chromosomes sexuels ». Chez les Mammifères femelles, les deux chromosomes de cette paire ont le

même aspect (ces deux chromosomes sont appelés les « chromosomes X »), mais chez les Mammifères mâles, cette paire est constituée d'un chromosome X et d'un chromosome beaucoup plus court, appelé « chromosome Y » (voir Figure 2).

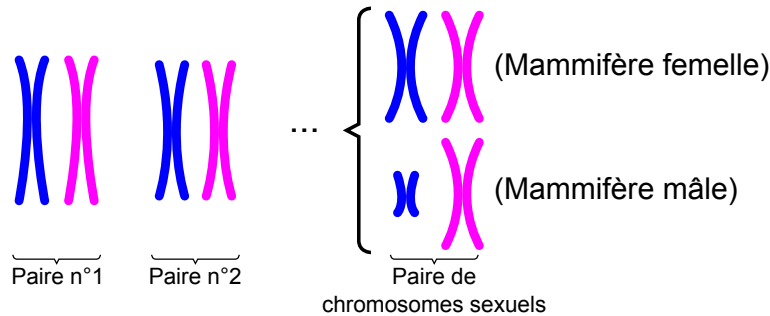


FIGURE 2 – **Les chromosomes sexuels de Mammifères.** Dans chaque paire, le chromosome d'origine paternelle est en bleu, le chromosome d'origine maternelle est en rose. Dans leur paire de chromosomes sexuels (à droite), les Mammifères femelles possèdent deux chromosomes X (l'un est hérité de leur père, l'autre de leur mère), alors que les Mammifères mâles possèdent un chromosome X et un chromosome Y (le X est toujours hérité de leur mère, le Y de leur père).

Les chromosomes de cette paire ont donc un comportement spécial : lorsqu'elle fabrique ses ovules, une femelle mammifère ne transmettra de cette paire que des chromosomes X. Mais lorsqu'il fabriquera ses spermatozoïdes, un mâle mammifère transmettra un chromosome X à la moitié de ses spermatozoïdes, et un chromosome Y à l'autre moitié. La fécondation d'un ovule (qui contient un chromosome X) avec un spermatozoïde (qui contient soit un chromosome X, soit un chromosome Y) produira donc, avec des probabilités pratiquement égales, soit un œuf fécondé contenant deux chromosomes X (cet œuf fécondé se développera donc en une femelle), soit un œuf fécondé contenant un chromosome X et un chromosome Y (qui se développera donc en un mâle).

3 La recombinaison et les mutations

La fabrication des gamètes est le cadre d'un événement particulier : les chromosomes d'une même paire échangent des fragments. Ce phénomène, appelé « recombinaison », aboutit donc à la production de chromosomes originaux, issus du mélange des chromosomes d'origine paternelle et maternelle (voir Figure 3). Un individu ne transmet donc pas à sa descendance des chromosomes absolument identiques à ceux qu'il a hérités d'un de ses parents : ce chromosome est panaché de fragments du chromosome de la même paire, hérité de l'autre parent.

La recombinaison génère donc des chromosomes originaux à chaque génération, et les chromosomes portés par un enfant se distinguent donc de ceux portés par ses parents.

Un autre type d'altération vient modifier la séquence portée par les chromosomes : les « mutations ». Ce sont des modifications de la séquence, qui peuvent être spontanées (par exemple au cours de la réplication de l'ADN avant une division cellulaire : en général la réplication est extrêmement fidèle, mais il lui arrive parfois de copier les séquences imparfaitement), ou induites (par exemple, suite à une exposition de l'ADN à des composés chimiques ou à des rayonnements). Une mutation peut être très courte (insertion, délétion ou substitution d'un seul nucléotide²), ou très longue (par exemple, la délétion simultanée de centaines de nucléotides).

Lorsqu'une mutation survient dans une cellule d'un organisme pluricellulaire, elle disparaîtra avec la mort de cet individu – sauf dans un cas particulier : si cette mutation survient dans un gamète, ou dans une cellule progénitrice de gamète, et que ce gamète est ensuite utilisé dans une fécondation

2. Les nucléotides sont les éléments constitutifs de la séquence de l'ADN ; par exemple, la séquence « ACCGTA » est constituée de 6 nucléotides : deux « A », deux « C », un « G » et un « T ».

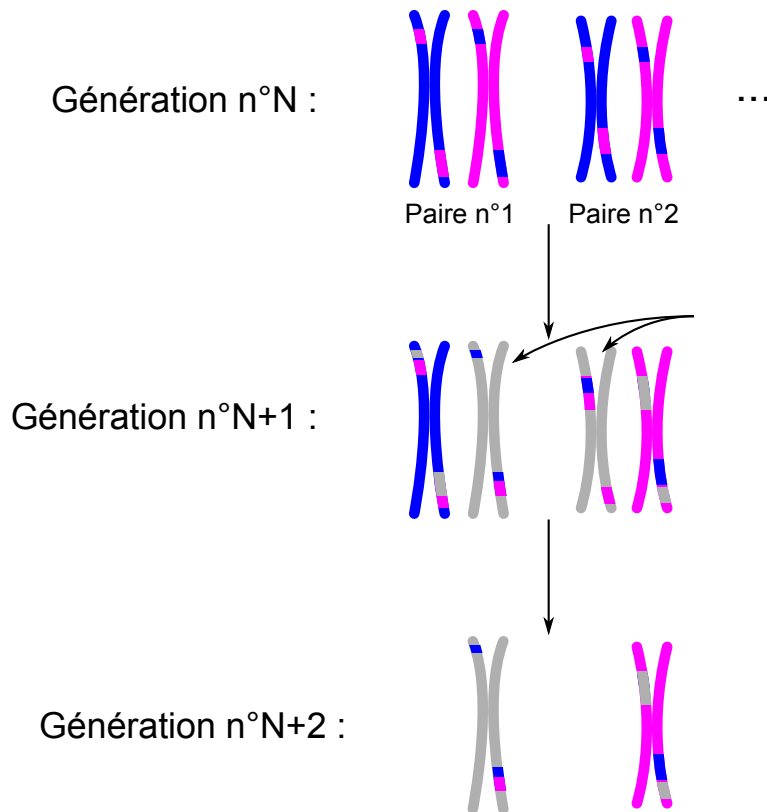


FIGURE 3 – **La recombinaison.** Lors de la production de ses gamètes, l'individu de la génération n°N (schéma du haut) recombine les chromosomes d'une même paire : les chromosomes qui seront transmis aux gamètes seront donc des mosaïques des chromosomes qu'il a hérités de ses parents. Chez son enfant (génération n°N+1, schéma du milieu), chacun de ces chromosomes constitue une paire avec un chromosome apporté par l'autre parent (représenté en gris). Quand cet individu fabriquera à son tour des gamètes, il recombinera ces chromosomes (des fragments de chromosomes gris seront échangés avec des fragments de chromosomes bleus-et-roses). Ce sont donc des chromosomes nouvellement recombines qu'il transmettra à sa propre descendance (génération n°N+2, schéma du bas).

qui donne naissance à un nouvel individu. Dans ce cas, la mutation sera héritée par cet individu, et elle pourra éventuellement se propager à long terme dans la population (si cet individu se reproduit lui-même, et la transmet à sa propre descendance).

Certaines régions du génome mutent à très haute fréquence dans les cellules progénitrices des gamètes, au point qu'en quelques générations d'individus, elles auront toutes les chances d'avoir muté – en d'autres termes : ces mutations ne seront partagées que par les membres d'une même famille. Certaines de ces régions mutent d'ailleurs tellement fréquemment qu'elles sont même différentes entre parents et enfants : elles peuvent servir de marqueurs témoignant de l'identité de chaque personne, puisqu'elles sont uniques à chaque individu.

Cette sorte d'« empreinte digitale génétique » est mise à profit dans plusieurs types d'applications biotechnologiques :

- L'identification d'une personne à partir d'un échantillon de son ADN (prélevé, par exemple, sur une scène de crime). Pour pouvoir relier la séquence d'ADN à l'identité de la personne, il faut donc au préalable disposer d'une base de données, dans laquelle on aura établi (sur la foi d'autres prélèvements, sur des personnes dont l'identité est indiscutablement connue) à quelle séquence correspond quelle personne.
- Dans les tests de parentalité (en utilisant des régions génomiques qui mutent suffisamment fréquemment pour distinguer les différentes lignées familiales, mais suffisamment rarement pour que l'enfant ait de grandes chances de porter la même séquence que le parent). Interroger

simultanément plusieurs régions du génome permet de limiter les risques d'erreur dues à ces deux approximations.

4 Apport pour la généalogie

Au fil de l'histoire humaine, les mutations s'accumulent donc dans le génome, elles sont héritées de parents en enfants, et s'enrichissent progressivement de nouvelles mutations qui finissent par distinguer nettement des lignées familiales issues d'un couple d'ancêtres communs. À chaque génération, au cours de la production des gamètes, des recombinaisons viennent assembler entre elles des mutations qui étaient initialement portées par des chromosomes différents (le chromosome d'origine paternelle et le chromosome d'origine maternelle; voir Figure 3).

Parmi tous les chromosomes du génome humain, deux chromosomes échappent cependant à la recombinaison : la plus grande partie du chromosome Y des hommes (puisque la paire de chromosomes sexuels des hommes est constituée d'un chromosome X et d'un chromosome Y, qui sont trop différents entre eux pour pouvoir recombiner – sauf sur de très courtes régions aux deux extrémités de ces chromosomes; chez les femmes en revanche, la paire de chromosomes sexuels est constituée de deux chromosomes X, qui peuvent recombiner); et un chromosome très particulier, le chromosome mitochondrial (qui ne se trouve pas dans le noyau des cellules comme les autres chromosomes, et qui est circulaire alors que les chromosomes du noyau sont linéaires).

Comme le chromosome Y et le chromosome mitochondrial ne recombinent pas, ils sont transmis fidèlement à la descendance, sans être modifiés par la recombinaison – les seules altérations qu'ils subissent, les mutations, sont quant à elles très rares. Ainsi, le chromosome Y d'un homme donné sera identique à celui de son père, lui-même identique à celui du grand-père paternel, et ainsi de suite.

Quant au chromosome mitochondrial, il est uniquement transmis aux enfants par leur mère (l'œuf fécondé l'a hérité de l'ovule, alors que celui du spermatozoïde est détruit à la fécondation). Ainsi, un enfant possède le même ADN mitochondrial que sa mère, qui elle-même possédait le même que la grand-mère maternelle, et ainsi de suite.

Finalement dans l'espèce humaine, et aux mutations (très rares) près, un chromosome Y de séquence donnée est propre à une lignée mâle, et un chromosome mitochondrial de séquence donnée est propre à une lignée femelle (voir Figure 4).

Ces propriétés ont été mises à profit pour vérifier l'attribution de restes humains à des personnages historiques. Nous allons ici détailler les conclusions de deux de ces études.

La première étude concerne l'attribution d'une tête momifiée à Henri IV. De premières études, basées sur des arguments historiques, morphologiques, mais également génétiques (mais à partir d'échantillons d'ADN de piètre qualité, extraits de cette relique, et d'un échantillon de sang séché attribué à Louis XVI, descendant d'Henri IV), avaient conclu que la tête était bien celle d'Henri IV, et le sang, bien celui de Louis XVI. Mais une meilleure extraction d'ADN de la tête momifiée, et sa comparaison avec des descendants actuels de la lignée mâle des Bourbons, a permis de ré-évaluer la question avec plus de précision (Larmuseau et al., 2014). Il apparaît que les chromosomes Y de la tête momifiée et du sang séché sont très différents entre eux, et très différents de celui des descendants Bourbon actuels, au point que les mutations spontanées du chromosome Y semblent trop peu fréquentes pour expliquer une telle différence en quelques siècles. Or d'après l'arbre généalogique officiel de la famille Bourbon, tous ces hommes devraient partager le même chromosome Y, hérité en lignée mâle à partir d'Henri IV.

Il reste possible que la lignée mâle ait été interrompue, et que le père biologique d'un de ces hommes ne soit pas son père légal (par exemple, à la suite d'un adultère). Mais l'ADN mitochondrial de la tête momifiée (qui aurait dû être celui de Jeanne d'Albret, mère d'Henri IV, si la tête était bien la sienne) n'est pas non plus celui de la lignée maternelle de Jeanne d'Albret. Il apparaît donc que la tête momifiée n'est pas celle d'Henri IV, et, incidemment, que le sang séché n'a pas non plus l'air d'être celui de Louis XVI.

La deuxième étude concerne le squelette retrouvé à Leicester en 2012, et qui pouvait être celui de Richard III, le dernier roi d'Angleterre de la dynastie Plantagenêt (plusieurs indices morphologiques le

suggéraient). Sa propre descendance s'était éteinte, mais la généalogie officielle permettait d'identifier des hommes actuellement vivants, et qui devaient porter le même chromosome Y que lui (ils descendaient, comme lui, en lignée mâle, du roi Édouard III, dont ils devaient porter le chromosome Y), ainsi qu'un homme et une femme actuellement vivants, qui devaient porter le même chromosome mitochondrial (ils descendaient, comme lui, en lignée femelle, de Cécile Neville : c'étaient des descendants d'Anne d'York, sœur de Richard III). Ces analyses ont donné des résultats contradictoires : le chromosome Y des descendants de la lignée mâle est très différent de celui prélevé sur le squelette de Leicester. En revanche, le chromosome mitochondrial des descendants de la lignée femelle est le même que celui du squelette³ : il apparaît donc que le squelette est bien celui de Richard III (qui a eu droit, en 2015, à une nouvelle inhumation, célébrée en grande pompe dans la cathédrale de Leicester) ; et ... qu'un adultère a dû avoir lieu dans sa lignée mâle, puisque les chromosomes Y trahissent une rupture de la transmission de père en fils biologique.

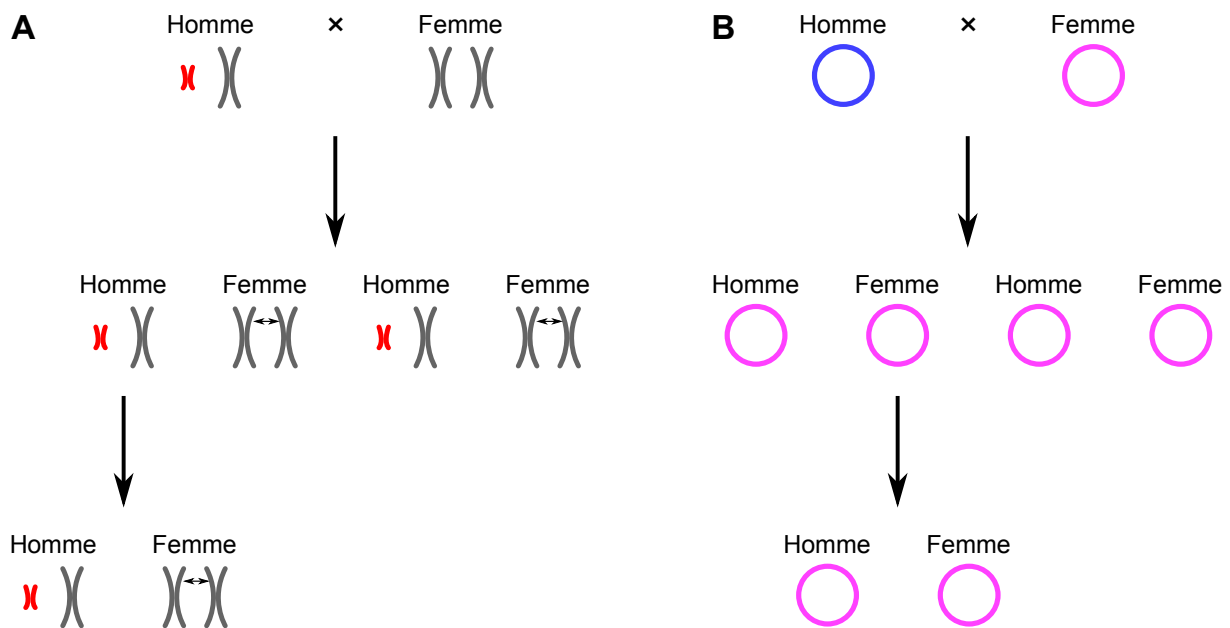


FIGURE 4 – Les chromosomes Y et mitochondrial, marqueurs des lignées mâle et femelle. **A.** Transmission du chromosome Y (en rouge) ; le chromosome X est en gris. Chaque homme hérite du chromosome Y de son père, qui n'aura pas recombiné (alors que chez les femmes, les deux chromosomes X recombinent au moment de la production des ovules : flèches horizontales noires). **B.** Transmission du chromosome mitochondrial (circulaire). Les descendants d'un couple donné héritent uniquement du chromosome mitochondrial de leur mère (ici en rose).

5 Génome et vie privée

La mise en évidence d'un adultère dans une ancienne famille royale anglaise peut faire sourire. Mais ce qui est vrai de la famille de Richard III peut l'être de la famille de M. et Mme. Tout le monde. Puisque l'analyse du génome des membres d'une famille peut révéler des détails embarrassants, et puisqu'on sème son génome facilement autour de soi (l'ADN peut être extrait de salive, de sang, même de cheveux), il existe une possibilité théorique d'emploi, à mauvais escient, de ce type de « donnée personnelle ».

3. Plus exactement : le chromosome mitochondrial de l'un des donneurs actuels, Michael Ibsen, est parfaitement identique à celui du squelette ; celui de l'autre donneuse, Wendy Duldig, présente une différence d'un seul nucléotide, sur les ≈ 16500 nucléotides de la séquence de ce chromosome. Il s'agit donc certainement d'une mutation, survenue au cours des 21 générations qui séparent Mme. Duldig de Richard III.

L'extraction et le séquençage d'ADN sont des techniques délicates, qui nécessitent un appareillage et une expertise particuliers : il ne faut donc pas s'attendre à ce que n'importe qui puisse séquencer votre génome à votre insu. En revanche, il existe des compagnies qui proposent, contre rémunération, de séquencer le génome de leurs clients (le client envoie un échantillon biologique, puis reçoit des estimations concernant les origines géographiques probables de ses ancêtres, et éventuellement, ses prédispositions à certaines maladies). L'usage des services de séquençage génomique est actuellement interdit en France hors procédure judiciaire ou médicale. La sécurité des données conservées par ces compagnies n'est garantie qu'à la hauteur des efforts qu'elles y consentent, et l'usage même qui en est fait par certaines de ces compagnies implique un stockage à long terme des données (certaines offrent de prévenir leurs anciens clients chaque fois que l'ADN d'un nouveau client suggère qu'ils pourraient être apparentés).

Outre son utilisation en généalogie, la séquence du génome peut souvent permettre d'identifier le donneur de l'échantillon d'ADN : dans certaines circonstances, la seule donnée de la séquence du génome, combinée à des informations publiquement accessibles, suffit à identifier avec confiance l'état-civil du donneur (Gymrek et al., 2013). Ce simple fait signifie qu'il est souvent illusoire d'anonymiser un génome, et les praticiens (scientifiques et cliniciens) amenés à manipuler des données de génomes humains individuels ne peuvent pas se contenter de masquer le nom du donneur dans les métadonnées des fichiers : la séquence elle-même peut permettre de le retrouver.

Enfin, la donnée du génome peut contenir des informations médicales sensibles, telles que des prédispositions à des maladies, que la personne ne souhaite probablement pas voir diffusées. Outre les régions génomiques déjà connues, prédisposant à des maladies identifiées, des informations médicales pourraient également se trouver dans des régions dont l'effet n'est pas encore connu, mais qui pourraient, dans le futur, s'avérer contrôler de nouvelles informations. Pour sécuriser le secret des prédispositions médicales, il ne suffit donc pas de masquer, dans un génome, les régions *actuellement* connues pour contrôler une caractéristique clinique.

6 Conclusion

Le public est informé des dangers du partage de ses données personnelles (nom, adresse, coordonnées téléphonique ou email, coordonnées bancaires, ...). La séquence du génome en est également une, qui présente de nombreuses particularités. Outre la richesse et la sensibilité des informations qu'il contient (preuves d'adultère, prédispositions médicales), sa simple donnée peut parfois suffire à identifier son propriétaire.

En y consacrant suffisamment d'efforts, il est possible de séquencer un génome, au moins partiellement, à partir d'échantillons biologiques qu'il est pratiquement impossible de ne pas disséminer sur son passage (cheveux, cellules de la peau, ...). Pour le moment, ces techniques ne sont accessibles qu'à des institutions scientifiques bien dotées, et à un personnel qualifié. Il n'est cependant pas trop tôt pour se poser la question des implications juridiques que portent les développements technologiques en matière de séquençage, et l'apparition du séquençage de loisir proposé au grand public.

REFERENCES

- Gymrek, M., McGuire, A. L., Golan, D., Halperin, E., and Erlich, Y. (2013). Identifying personal genomes by surname inference. *Science*, 339(6117):321–324.
- Larmuseau, M. H. D., Delorme, P., Germain, P., Vanderheyden, N., Gilissen, A., Van Geystelen, A., Cassiman, J.-J., and Decorte, R. (2014). Genetic genealogy reveals true Y haplogroup of House of Bourbon contradicting recent identification of the presumed remains of two French Kings. *Eur J Hum Genet*, 22(5):681–687. Texte intégral : <https://www.nature.com/articles/ejhg2013211>.
- Seitz, H. (2018). La technologie CRISPR/Cas9. *La Revue du Centre Michel de l'Hospital*, 15:28–38. Texte intégral : <https://en.calameo.com/read/00650526985add381001>.