

Chroniques génomiques

Notre génome est-il prédictif ?

Bertrand Jordan



CoReBio PACA, case 901,
parc scientifique de Luminy,
13288 Marseille Cedex 9,
France.
brjordan@club-internet.fr

Une évidence ?

Oui, bien sûr, notre génome est prédictif puisqu'il porte l'ensemble de notre information génétique et que quasiment toutes les maladies, à des degrés divers, sont influencées par les allèles présents dans notre ADN. La perspective d'une « médecine personnalisée » n'est-elle pas la justification des efforts faits pour caractériser nos gènes et définir les phénotypes associés à leurs différentes versions ? Les nouvelles techniques de séquençage ne vont-elles pas aboutir à une lecture systématique de l'ADN des nouveau-nés, de manière à connaître leurs vulnérabilités génétiques et à mettre en place des mesures de prévention ? Sans oublier qu'il s'agit là du fonds de commerce de *Navigenics*, *23andMe* et de nombreuses autres officines du même genre. Certes, nos connaissances restent aujourd'hui très imparfaites, le passage du génotype au phénotype comporte encore de nombreux mystères, et les études d'association génome entier (GWAS, *genome-wide association studies*) ne rendent compte que d'une faible partie de l'héritabilité des affections étudiées [1, 2]. Mais cela va s'améliorer au fil des années, et l'on découvrira bientôt tout l'intérêt médical et clinique de l'avalanche de séquences qui déferle actuellement.

Pourtant, un très intéressant article paru récemment dans *Science Translational Medicine* [3] remet en cause cette affirmation, à l'aide d'une approche originale indépendante des limites de nos connaissances actuelles. La méthode employée revient, dans les grandes lignes, à étudier la ressemblance entre les

destinées médicales de jumeaux homozygotes (donc portant exactement les mêmes jeux d'allèles¹), et à en déduire dans quelle mesure l'état de santé d'une personne est gouverné par sa constitution génétique. Il est alors possible d'évaluer les prédictions que l'on attend d'une connaissance parfaite de son génome. L'avantage de ce procédé est qu'il n'est pas influencé par les imperfections actuelles de la génomique clinique : il donne, en quelque sorte, la limite supérieure de ce que pourrait nous indiquer la séquence du génome d'un individu... si nous avions une connaissance complète de tous les variants de gènes qu'il contient et une compréhension parfaite de leurs corrélats physiologiques. De manière un peu surprenante, le résultat, pour une bonne vingtaine d'affections fréquentes, suggère que cela ne nous apprendrait pas grand-chose de réellement utile.

L'analyse des « génotypes »

Les auteurs ignorent, bien entendu, la séquence d'ADN des jumeaux étudiés. Mais ils savent que, au sein de chaque couple, ces séquences sont identiques, et que le risque génétique pour une affection donnée est donc le même pour chaque jumeau. Ils peuvent, du coup, définir une distribution de « génotypes » (pour l'affection en cause) au sein de la population étudiée,

¹ En toute rigueur, « presque exactement », puisqu'on sait qu'il peut exister quelques petites différences [4]. Je dirai « exactement » par la suite pour ne pas alourdir le texte.

chaque génotype correspondant à une valeur du risque génétique. Ils ajustent alors la distribution de ces génotypes (fréquence *versus* risque génétique) afin qu'elle rende compte au mieux des chiffres de la concordance et de la discordance entre ces vrais jumeaux. Une fois cette distribution obtenue, on peut évaluer l'information qu'apporterait – dans l'idéal – la séquence de l'ADN, et donc examiner dans quelle mesure elle aurait une réelle utilité clinique. Cette méthode se heurte cependant à un sérieux problème, le fait que la distribution des génotypes n'est pas unique : les concordances et discordances observées peuvent résulter d'un faible nombre d'individus très vulnérables, ou au contraire d'une forte proportion de personnes à risque modéré. On doit donc envisager plusieurs cas de figure, qui aboutissent à une gamme de valeurs pour la « prédictibilité » du risque, et non un chiffre unique. De plus, il est délicat d'évaluer quel niveau de risque doit être considéré comme cliniquement significatif (devant entraîner des mesures de prévention ou de prophylaxie). Comme l'indiquent les auteurs, les conséquences pratiques d'un risque indiqué par l'analyse dépendent de l'affection en cause, de sa fréquence et de sa gravité. Pour le cancer du sein, un risque de 90 % pousserait sans doute la patiente à envisager une mastectomie prophylactique ; une valeur de 15 %, en revanche, n'aurait guère d'impact étant donné que le risque dans la population générale est déjà de 12 %². Dans l'article considéré, les auteurs ont pris comme seuil un risque après analyse de 10 % pour les maladies peu fréquentes, et un doublement du risque *a priori* pour les autres : au-dessus de ces limites, ils considèrent que l'information obtenue a une valeur clinique.

Un échantillon important... mais imparfait

Les études de jumeaux utilisées pour ce travail proviennent naturellement de publications antérieures : il s'agit là de travaux de longue haleine, sur des échantillons difficiles à réunir, et il n'était pas imaginable de s'y lancer *de novo*. Du coup la liste de maladies analysées (Tableau 1) est assez hétéroclite et comporte des entités dont la définition est problématique : « Syndrome de fatigue chronique », « Prolapsus des organes pelviens », « Colopathie fonctionnelle (*irritable bowel syndrome*) ». La part génétique de beaucoup de ces affections ne paraît pas évidente au vu des valeurs de concordance et de discordance³. De plus les effectifs sont parfois très faibles (23 couples affectés pour le diabète de type 1, 10 pour la maladie d'Alzheimer), ce qui met en doute le caractère significatif des résultats. Notons néanmoins que ces chiffres correspondent à un effectif total de presque 16 000 couples de jumeaux : c'est considérable, et il sera difficile de faire mieux.

² Il s'agit là du risque sur la vie entière, plus élevé que la prévalence (à une date donnée) observée au sein d'un groupe de personnes (voir Tableau II).

³ Pour le cancer du sein, dans lequel l'influence héréditaire est modeste mais bien connue, on a une prévalence dans la cohorte de 3,5 %. Parmi les 547 couples de jumelles dans lesquels au moins une femme est atteinte (505 + 42, voir Tableau II), on attendrait donc, s'il n'y avait pas d'influence génétique, 19 couples concordants (3,5 % de 547). On en observe 42, ce qui montre bien que le génome joue un rôle dans le risque. Pour le syndrome de fatigue chronique masculine (chiffres dans [3]), la prévalence est de 12,7 %, on attend 40 couples concordants sur les 314 répondant au même critère, on en trouve 48 ce qui est peu significatif.

Cancer de la vessie
Cancer du sein
Cancer colorectal
Leucémie
Cancer de l'ovaire
Cancer du pancréas
Cancer de la prostate
Cancer de l'estomac
Thyroïdite auto-immune
Diabète de type 1
Lithiase vésiculaire
Diabète de type 2
Maladie d'Alzheimer
Démence
Maladie de Parkinson
Syndrome de fatigue chronique
Reflux gastro-œsophagien
Colopathie fonctionnelle
Maladie coronarienne
Décès en relation avec un accident vasculaire cérébral
Prolapsus des organes pelviens
Dystonie généralisée
Incontinence urinaire de stress

Tableau 1. Maladies étudiées chez des couples de jumeaux. Liste des maladies pour lesquelles les auteurs ont eu accès à des études de concordance chez des couples de jumeaux. L'effectif va de 15 668 couples (pour les cancers de la vessie et du côlon, les leucémie, les cancers du poumon, pancréatique et de l'estomac) à 284 couples pour la maladie thyroïdienne auto-immune (extrait du tableau 1 de [3]. Voir cet article pour les références des travaux utilisés).

Les résultats : une « prédictivité » limitée

Les auteurs ont donc calculé, pour chaque affection incluse dans l'étude, les distributions de « génotypes » compatibles avec les incidences, concordances et discordances observées dans l'échantillon de jumeaux analysé. Ils peuvent alors en déduire différents paramètres : la probabilité pour que le test soit positif chez les malades, qu'il soit positif dans la population générale, et surtout la valeur du risque relatif (par rapport à la population générale) chez les personnes pour lesquelles le test serait négatif. Ils présentent leurs résultats sous forme de gammes de valeurs (minimum/maximum), compte tenu de l'influence de la distribution choisie sur le résultat final. J'ai choisi d'extraire de leurs résultats ceux qui concernent trois affections : cancer du sein, maladie thyroïdienne auto-immune et diabète de type 2.



	Couples de jumeaux (total)	Affectés, concordants	Affectés, discordants	Prévalence (%)	% de tests positifs chez les patients	% de tests positifs (population générale)	Risque relatif si test négatif
Cancer du sein	8 437	42	505	3,5	2 - 33	0 - 8	0,72 - 0,90
Maladie thyroïdienne auto-immune	284	7	17	5,5	39 - 80	2 - 22	0,26 - 0,72
Diabète de type 2	4 307	29	113	2	32 - 46	0 - 3	0,56 - 0,69
Valeurs « idéales » (fortement prédictives)					Plus de 50	Similaire à la prévalence	Moins de 0,10

Tableau II. Valeurs obtenues pour trois affections. Résultats extraits des figures 1, 2 et 3 de [3]. Pour le cancer du sein, le calcul prévoit ainsi de 2 à 33 % de tests positifs chez les patientes, et de 0 à 8 % dans la population générale, l'éventail des prévisions étant dû aux différentes distributions possibles des « génotypes ». Si le test est négatif, le risque « après test » de la personne concernée est de 72 à 90 % du risque dans la population générale, il est donc peu différent du risque *a priori*.

La première affection a une composante génétique connue mais minoritaire, les deux suivantes (et surtout la deuxième) une composante génétique assez forte, comme l'indiquent les rapports entre concordants et discordants, respectivement 0,08, 0,29 et 0,20⁴.

On voit que, sauf pour la maladie thyroïdienne auto-immune (dont les chiffres sont sujets à caution compte tenu du faible effectif), la majorité des malades ne seraient pas identifiés par l'étude de leur ADN, et que, dans la plupart des cas, le risque après test reste supérieur à 50 % du risque *a priori*, c'est-à-dire qu'il n'est pas diminué de façon réellement significative. En d'autres termes, l'utilité clinique de telles analyses (pourtant simulée ici dans l'hypothèse d'une connaissance parfaite du génome ainsi que des corrélats cliniques de chaque variant) est douteuse pour la plupart des affections considérées. Ce résultat – certes affecté d'une grande marge d'incertitude comme en témoigne l'éventail très large des valeurs rapportées dans le *Tableau II* – met à mal une perception toujours très ancrée de l'ADN comme destin et porte un coup supplémentaire aux ambitions des entreprises qui prétendent révéler nos prédispositions génétiques. Petit

⁴ Pour une maladie monogénique, ce rapport est très élevé puisque les jumeaux sont presque toujours concordants. Pour une affection polygénique à forte influence génétique comme l'autisme, le rapport est de 1,38 selon les toutes dernières études (concordance 58 %) [5].

bémol selon les auteurs, si l'on se base sur les « meilleures » valeurs des intervalles indiqués, on devrait pouvoir repérer au moins un risque cliniquement significatif (pour l'une des affections envisagées) chez environ 90 % des personnes testées.

Les limites de l'étude

Cet article, dont le libellé (sinon le titre) est assez provocateur, n'a pas manqué de susciter des réactions. Trois commentaires de deux ou trois pages ont paru peu après dans la même revue [6-8], suivis d'une réponse des auteurs [9]. Les deux premiers [6, 7] discutent les méthodes statistiques utilisées et proposent des approches différentes qui améliorent nettement les résultats (c'est-à-dire la qualité de la prédiction à partir de l'ADN), et la réponse [9] explique d'une manière qui semble assez convaincante les défauts de ces approches alternatives. Le troisième commentaire [8], nettement plus polémique, critique la présentation médiatique de cette étude et affirme sa foi en l'avenir des méthodes de balayage du génome (GWAS), une fois complétées par la détection des variants rares grâce au séquençage. Or, comme le font très justement remarquer les auteurs [9], le choix de couples de jumeaux pour estimer la « prédictibilité » assure que l'on prend en compte l'ensemble du génome, que les variants soient rares, fréquents ou même de nature encore inconnue. Ce troisième commentaire me semble donc nul et non avvenu.

Cela dit, il faut noter que la plupart des affections envisagées ici ont une composante génétique peu marquée. Du coup, l'excellente idée d'utiliser la comparaison des histoires médicales de jumeaux monozygotes pour évaluer la prédictivité de l'ADN tombe un peu à

plat. Il aurait fallu disposer de grandes études de jumeaux portant sur des maladies dont la composante génétique est plus nette (hypertension, maladie bipolaire, sclérose en plaques, schizophrénie, maladie d'Alzheimer). J'imagine que les auteurs ont cherché assidûment... et ont utilisé tout ce qu'ils ont pu trouver. On peut aussi noter le faible effectif pour certains groupes, qui rend les conclusions assez aléatoires. Le résultat assez négatif de cette étude sur l'efficacité attendue d'une analyse « parfaite » du génome en tant que facteur prédictif de l'histoire médicale doit donc être modulé en tenant compte de ces limites. Elle n'en reste pas moins fort intéressante, de nature à dissiper quelques illusions et, peut-être, à avoir une influence sur le chiffre d'affaires de certaines entreprises. ♦

Is our genome predictive?

LIENS D'INTÉRÊT

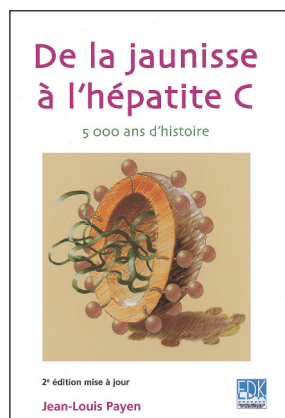
L'auteur déclare n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. Jordan B. Maladie de Crohn et GWAS, d'analyses en méta-analyses. *Med Sci (Paris)* 2011 ; 27 : 323-5.
2. Jordan B. Les tests génétiques grand public ont-ils une utilité clinique ? *Med Sci (Paris)* 2012 ; 28 : 325-8.
3. Roberts NJ, Vogelstein JT, Parmigiani G, et al. The predictive capacity of personal genome sequencing. *Sci Transl Med* 2012 ; 4 : 133ra58.
4. Bruder CE, Piotrowski A, Gijbbers AA, et al. Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. *Am J Hum Genet* 2008 ; 82 : 763-71.
5. Hallmayer J, Cleveland S, Torres A, et al. Genetic heritability and shared environmental factors among twin pairs with autism. *Arch Gen Psychiatry* 2011 ; 68 : 1095-102.
6. Begg CB, Pike MC. Comment on the predictive capacity of personal genome sequencing. *Sci Transl Med* 2012 ; 4 : 135le3.
7. Golan D, Rosset S. Comment on the predictive capacity of personal genome sequencing. *Sci Transl Med* 2012 ; 4 : 135le4.
8. Topol EJ. Comment on the predictive capacity of personal genome sequencing. *Sci Transl Med* 2012 ; 4 : 135le5.
9. Vogelstein B, Roberts NJ, Vogelstein JT, et al. Response to comments on the predictive capacity of personal genome sequencing. *Sci Transl Med* 2012 ; 4 : 135lr3.

TIRÉS À PART

B. Jordan



ISBN : 978-2-8425-4136-1 128 pages

La jaunisse est un symptôme facilement identifiable ; il paraissait bien naturel que l'homme, confronté à une modification de la couleur de ses yeux et de sa peau ait de tous temps recherché les causes de cette transformation.

Il n'est donc pas surprenant que le premier traité de médecine, écrit 3 000 ans avant J.C. par un médecin sumérien, décrive déjà la jaunisse. À chaque époque de l'histoire de la médecine, les praticiens, influencés par les concepts médicaux de leur temps, attribuèrent une ou plusieurs explications particulières à ce symptôme. Ainsi, du démon *Ahhâzu* des Sumériens à la sophistication des biotechnologies qui permirent la découverte du virus de l'hépatite C, le lecteur cheminera sur une période de 5 000 ans au travers des différents continents. Ici encore, l'histoire se révèle une formidable source de réflexion : le foie souvent impliqué dans l'apparition des jaunisses est-il le siège de l'âme ?

Les expérimentations humaines chez des volontaires ou chez des enfants handicapés mentaux étaient-elles justifiées pour permettre la découverte des virus des hépatites ?

Le formidable développement de la transfusion sanguine, des vaccinations, mais aussi de la toxicomanie explique-t-il les épidémies d'hépatites du XX^e siècle ?

Autant de questions qui sont abordées dans ce livre passionnant et accessible à tous.



BON DE COMMANDE

À retourner à EDK, 25, rue Daviel - 75013 Paris - Tél. : 01 58 10 19 05 - Fax : 01 43 29 32 62 - E-mail : edk@edk.fr

NOM : Prénom :

Adresse :

Code postal : Ville :

Pays :

Fonction :

Je souhaite recevoir l'ouvrage **De la jaunisse à l'hépatite C, 5 000 ans d'histoire** : 12 € + 3 € de port = **15 € TTC**

en exemplaire, soit un total de €

Par chèque, à l'ordre de **EDK**

Par carte bancaire : Visa Eurocard/Mastercard

Carte n° | | | | | | | | | | | | | | | | | | | | | |

Signature :

Date d'expiration : | | | | | |

N° de contrôle au dos de la carte : | | | | |