



HAL
open science

Réussir à l'école, une prédisposition génétique ?

Catherine Bourgain

► **To cite this version:**

Catherine Bourgain. Réussir à l'école, une prédisposition génétique?. Carnets rouges, 2024, 29, pp.21-24. hal-04397970

HAL Id: hal-04397970

<https://hal.science/hal-04397970v1>

Submitted on 18 Jan 2024

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.

Réussir à l'école, une prédisposition génétique ?

A lire une certaine presse grand public, la réussite scolaire serait avant tout une question de prédisposition génétique. Cet article conteste vivement ces affirmations, en revenant sur les études scientifiques qui les soutiennent. Il montre que ces études reposent sur l'hypothèse selon laquelle les comportements humains résulteraient d'effets innés et d'effets acquis, séparables et mesurables de façon indépendante. Pourtant, cette hypothèse, invalidée par de nombreuses études expérimentales menées sur les animaux en laboratoire, a encore moins de validité s'agissant de comportements humains aussi complexes que la réussite scolaire.

Un cadre conceptuel historiquement situé et idéologiquement marqué

Réfléchir la question des comportements humains en terme d'influences distinctes de facteurs innés et de facteurs acquis n'a rien d'une évidence. C'est un cadre de pensée particulier, qui a été construit à la faveur d'un intense travail scientifique, dont le contexte de production est important à préciser, pour en saisir les partis pris.

A la fin du 19ème siècle, une question travaille des statisticiens principalement britanniques, proches de sociétés savantes eugénistes. Il s'agit pour eux d'évaluer dans quelle mesure les savoirs qu'ils contribuent à élaborer pour la sélection des plantes et des animaux d'élevage pourraient être mis au service de politiques publiques permettant de renforcer dans la population la proportion d'individus naturellement doués et de réduire celle des individus naturellement « déviants » (dangereux, idiots, avec une propension au crime ou à l'alcoolisme...). Ils en sont persuadés : ces caractères (entendu ici comme tout type de trait physique, psychologique ou comportemental, mesurable chez un individu) résultent très largement d'effets « de nature », fixés dès la naissance, « innés ». De façon plus précise, Francis Galton, figure de proue de ce mouvement, avance que ces effets de nature peuvent être séparés des conditions sociales, naturelles et culturelles dans lesquelles les individus vivent. Cette idée de séparation est fondamentale. Elle s'inscrit d'une part, en rupture avec les théories alors dominantes, selon lesquelles, si des effets de nature existaient bien, ils étaient inextricables des effets de l'environnement et en conséquence, non mesurables. D'autre part, en permettant la quantification des effets de nature, la séparation permet d'en proposer une objectivation scientifique. Puisqu'ils sont mesurables, c'est qu'ils existent. Et plus ils sont importants, plus les interventions publiques qui les prennent pour cibles peuvent être présentées comme légitimes car plus efficaces que les interventions ciblant les conditions de vie des individus.

« Tous les modèles statistiques sont faux, mais certains sont utiles »

La force de ce courant scientifique va ensuite résider dans sa capacité à développer des outils statistiques qui incorporent cette conception de la distinction entre effets innés et effets d'environnement

pour pouvoir les quantifier. Ainsi, en 1918, Ronald Fisher, autre membre éminent du courant, présente un modèle mathématique selon lequel tout caractère peut être décomposé en une somme de petits effets innés indépendants les uns des autres et d'un effet de l'environnement, également indépendant. Appliqué sur des données de taille humaine, ce modèle permet à Fisher de proposer des mesures des effets innés sur la taille, à partir desquelles, il conclut « qu'il y a peu, si ce n'est aucune indication de causes non génétiques^[1] Les données collectées par la suite ont modifié la compréhension de l'impact sur la taille du contexte environnemental et culturel. Entre 1896 et 1996, les femmes de Corée du Sud ont grandi en moyenne de 20,5 cm et les hommes d'Iran de 16,5 cm, évolutions non explicables par des changements génétiques. ». Tous les modèles statistiques sont faux parce qu'ils constituent une simplification du réel, mais certains sont utiles, disait le statisticien George Box. Celui de Fisher est utile, en ce qu'il conforte les vues eugénistes qu'il partage avec ses pairs : si les effets de nature sont aussi importants, alors l'objectif eugéniste d'améliorer la « qualité génétique » des populations humaines devient réaliste.

Des usages de la statistique invalidés par la génétique expérimentale

La postérité de ce modèle statistique est remarquable. Plus de cent ans après son invention, il reste le cadre d'analyse avec lequel beaucoup de recherches en génétique humaine sont menées. Les perfectionnements dont il a été l'objet, s'ils sont nombreux, ne remettent pas en cause l'idée fondamentale selon laquelle il serait toujours possible de distinguer des effets mesurables et indépendants des gènes et de l'environnement. Les résultats accumulés en biologie expérimentale ont pourtant montré combien, même chez les animaux plus simples très étudiés en laboratoire comme le petit nématode *Caenorhabditis elegans* ou la mouche drosophile, ce cadre ne correspondait pas à la réalité du fonctionnement du vivant, fait d'interdépendances et d'interactions permanentes entre les gènes et les conditions de vie des organismes, dès les premières étapes de la vie. Mais le modèle résiste. Implémenté au sein de logiciels, relativement faciles à utiliser, il est devenu une boîte noire dont peu d'utilisateurs comprennent précisément les hypothèses et conditions de validité, et a été très largement invisibilisé. Lorsque des quantités importantes de données génétiques ont été rendues disponibles au début des années 2000, il s'est imposé comme un modèle pratique et efficace pour traiter ces données et produire des résultats jugés prometteurs. Ce faisant, le cadre de pensée voulant qu'il y aurait, pour chaque caractère humain, de l'inné et de l'acquis dont il serait possible de mesurer une proportion de causes respectives, s'est banalisé sans que ne soit plus vraiment relevée la contradiction pourtant flagrante entre cette simplification réductrice et la complexité croissante des descriptions scientifiques du vivant.

Le retour d'un discours public sur la génétique de la réussite scolaire

C'est donc sans grande surprise que des prises de parole publiques fortes d'universitaires ont trouvé place dans l'espace médiatique pour défendre l'importance de l'inné sur les traits de caractère, les comportements et capacités sociales. Pour ce qui est de la réussite scolaire, la part de l'inné serait ainsi de 60%^[2] Entretien avec R. Plomin, dans l'Express, le 12 janvier 2023. Ces universitaires sont pour certains des acteurs en fin de carrière, comme c'est le cas de Robert Plomin, psychologue intéressé à la

génétique humaine, engagé dans des débats similaires dès les années 90, à partir de résultats obtenus sur des paires de jumeaux ou des enfants adoptés à la naissance. À l'époque, une réaction s'était organisée dans la communauté des généticiens pour contester ses propos, et pointer les biais méthodologiques de ses études. Plomin revient aujourd'hui occuper la scène médiatique aux côtés d'acteurs et actrices plus jeunes, comme la psychologue Paige Harden. Les études de jumeaux ont désormais cédé la place à des techniques marquées du sceau de la modernité génétique : analyses sur le génome entier et scores de risque polygéniques.

Des moyens colossaux pour des résultats insignifiants

Initiées au milieu des années 2000, ces études (appelées GWAS en anglais, pour Genome-wide association studies) consistent à collecter des données relatives à la psychologie, les diplômes universitaires, les caractéristiques biométriques ou la santé chez un très grand nombre de personnes auprès desquelles un échantillon biologique est également recueilli. L'ADN présent dans chaque échantillon est ensuite analysé sur des séquenceurs à très haut débit pour caractériser des millions de points sur la séquence d'ADN susceptibles de différer entre individus. Des corrélations statistiques peuvent alors être calculées entre les données collectées et chacune des variations génétiques identifiées, pour identifier celles qui seraient plus fréquemment présentes chez les personnes plus grandes, dont la valeur de QI est plus haute, ou le niveau d'études plus élevé.

Les chiffres des dernières publications sur le niveau d'études donnent le tournis. La dernière, publiée en 2022^[3] Okbay et al, Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. Nature Genetics. 54:437-449, 2022. En ligne : <https://www.nature.com/articles/s41588-022-01016-z>, inclut plus de 3 millions de personnes, chez lesquelles près de 10 millions de variations d'ADN sont analysées. Selon un principe très général en statistiques, plus la taille d'un échantillon est importante et plus il est possible de détecter des corrélations faibles. De fait, si ces études ne permettent pas d'identifier un gène de la réussite scolaire sur lequel fonder une catégorisation réussite /échec, elles ont détecté des combinaisons de variations sur l'ADN – près de 4000 variations pour celle de 2022 – dont la corrélation avec le niveau d'études est jugée statistiquement significative. Au total, ces variations génétiques expliqueraient entre 12 et 16% des différences de niveau d'études entre individus. Lorsqu'une correction est introduite pour tenir compte du fait que l'environnement est largement partagé dans les familles, ces estimations tombent à entre 5 et 2,5%.

Des échantillons non représentatifs et des mesures du niveau scolaire sorties de leur contexte

Ces éléments illustrent combien ces études sont de nature strictement statistique – elles ne reposent sur aucune démonstration expérimentale d'un lien entre variations génétiques et capacités intellectuelles. En conséquence, le modèle utilisé pour mesurer les corrélations, la composition des échantillons analysés, les mesures choisies pour définir le niveau d'étude, les autres données non génétiques qui sont intégrées au calcul sont des aspects fondamentaux pour en analyser la validité. En l'espèce, les trois millions de personnes incluses dans l'étude de 2022 sont toutes d'origine européenne et pour les deux tiers des clients d'une entreprise américaine vendant des services de tests génétiques en ligne. L'échantillon n'est donc pas construit pour viser une quelconque représentativité, mais au contraire

totale­ment biaisé. C'est un agrégat de données déjà disponibles dans lesquelles les participants avaient renseigné leur niveau d'étude. Pour les besoins de l'analyse statistique ces auto­déclarations ont été standardisées en utilisant la classification internationale de l'Unesco

(ISCED^[4]http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-1997-en_0.pdf, 1997). Celle-ci a été développée pour produire des données nationales agrégées à visée d'analyse et de comparaisons transnationales des politiques publiques et non pour interroger les déterminants individuels du niveau scolaire. Enfin, si l'ADN est décrit de façon très fine, les données non génétiques incorporées dans l'analyse sont frustrées, créant une situation de déséquilibre flagrant entre ces sources de variations. Mais, de façon bien plus fondamentale, le modèle statistique utilisé pour le calcul des corrélations et les estimations de variance, celui qui fonde la possibilité même de ces calculs, incorpore le cadre conceptuel proposé par Galton, Fisher et leurs collègues et donc le présumé d'une séparation fondamentale des effets de « nature » et de « culture ».

Une controverse scientifique sous influence

Certains défenseurs de ces approches revendiquent de plus en plus ouvertement la filiation de leurs travaux avec ceux de Galton. Hérauts de la science contre l'idéologie, ils s'enorgueillissent de trouver dans ces statisticiens eugénistes de glorieux précurseurs scientifiques, injustement déconsidérés par les errements d'un humanisme mièvre consécutif aux traumatismes des excès de deux guerres mondiales. D'autres feignent d'ignorer cette filiation conceptuelle, préférant insister sur les prouesses technologiques de la génétique alliées au raffinement des développements statistiques et informatiques mobilisés, et aux alliances internationales qui rendent possibles des études à si grande échelle. Puisqu'on arrive à mesurer ces corrélations, c'est qu'il doit bien y avoir derrière des mécanismes biologiques causaux. À défaut de pouvoir les identifier de façon expérimentale chez les humains, pourquoi ne pas utiliser ces corrélations pour faire des prédictions individuelles et soutenir le business des tests génétiques ? Pourquoi ne pas envisager de s'en servir pour organiser des politiques publiques ? Puisque la séquence d'un ADN est plus facile à caractériser que l'environnement social, familial et environnemental, n'est-ce pas là une approche d'optimisation pragmatique ?

Il est important de le réaffirmer clairement. Ni LA science, ni même LA génétique n'ont démontré que certaines personnes étaient génétiquement prédestinées dès la naissance pour réussir ou échouer à l'école. Les travaux de biologie expérimentale démontrent au contraire combien le vivant est complexe et les interdépendances permanentes entre génétique et environnement matériel et social. Les explications biologiques d'une propension individuelle innée à des comportements aussi socialement influencés n'existent pas. Seules demeurent des évaluations postulant cette propension.

Comment expliquer alors que des investissements importants soient mis au service de ces travaux, et qu'ils soient régulièrement publiés dans des revues scientifiques considérées comme prestigieuses ?

Intérêts professionnels de scientifiques construisant des carrières au sein d'institutions de recherche pilotées par des promesses de révolution médicale et de croissance économique par l'innovation ; intérêts financiers de firmes privées vendant des tests génétiques ou des médicaments associés ; visions morales et politiques de la responsabilité individuelle et de la place de la biologie dans les identités ; volonté de trouver des fondements scientifiques aux politiques de réduction de la dépense publique...

Ces raisons multiples et complexes sont un miroir des intérêts qui s'organisent dans et autour de la recherche biomédicale en général et de la génétique en particulier.

Catherine Bourgain

Directrice de recherche à l'Inserm, en génétique humaine et sociologie des sciences.

Directrice du Centre de Recherche Médecine, Sciences, Santé, Santé Mentale, Société (Cermes3)

Références[+]

Références

- ↑1 Les données collectées par la suite ont modifié la compréhension de l'impact sur la taille du contexte environnemental et culturel. Entre 1896 et 1996, les femmes de Corée du Sud ont grandi en moyenne de 20,5 cm et les hommes d'Iran de 16,5 cm, évolutions non explicables par des changements génétiques.
- ↑2 Entretien avec R. Plomin, dans l'Express, le 12 janvier 2023
- ↑3 Okbay et al, Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. Nature Genetics. 54:437-449, 2022. En ligne : <https://www.nature.com/articles/s41588-022-01016-z>
- ↑4 http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-1997-en_0.pdf