

LA RÉVOLUTION EST EN MARCHÉ

Le séquençage du premier génome humain avait coûté 3 milliards de dollars et duré treize ans. Aujourd'hui, les séquenceurs disponibles offrent le même résultat en une journée pour 1 000 dollars. Et l'on nous promet déjà le génome à 100 dollars d'ici à 2020. De jeunes entreprises américaines de génomique visent par ailleurs à dépasser l'usage des puces à ADN, préconisé par les leaders actuels du marché des tests génétiques, pour rendre accessible au grand public le séquençage complet de leur génome. En 2018, deux entrepreneurs du secteur, Razib Khan et David Mittelman, soutenaient que l'ère des puces de SNP des années 2010 serait supplantée par la révolution du séquençage complet des génomes qui s'opérerait au grand jour dans les années 2020. Selon eux, le génome complet de 60 millions d'Américains sera séquençé d'ici à 2025. Si ces informations plus exhaustives contribueront à augmenter la précision des résultats des tests génétiques, elles augmenteront surtout la valeur commerciale des bases de données privées des entreprises qui, d'ores et déjà, font saliver les laboratoires pharmaceutiques.

avec les effets collectifs de cette information biologique qui a la particularité d'être à la fois très stable dans le temps et partagée entre personnes d'une même famille et d'une même population. L'ADN est ainsi une donnée fort utile pour qui cherche à trier, sélectionner, discriminer, assigner à des individus une identité figée, sur la base d'informations biologiques pouvant être présentées comme «scientifiques». C'est en outre une information bien plus fragile qu'elle n'en a l'air. Les liens entre risque de cancer du sein ou de l'ovaire et mutations dans les gènes *BRCA* en sont une bonne illustration.

Alors même que le test *BRCA* est proposé en routine clinique depuis la fin des années 1990, l'opération reste délicate. Avancer des prédictions de cancers fondées sur la seule analyse génétique, alors que le déterminisme de ces maladies est souvent plus complexe, requiert des précautions. Le faire pour des mutations qui n'ont pas toutes les mêmes conséquences sur le fonctionnement cellulaire complique encore la situation. En France, dès 1991, les professionnels de ce domaine, conscients de ces difficultés, se sont organisés en un réseau national. Ce Groupe génétique et cancer a ainsi permis de proposer les tests *BRCA* sur tout le territoire, tout en évaluant les pratiques. L'enjeu était de faire remonter toutes les difficultés et les solutions expérimentées de façon à gérer au mieux, collectivement, les nombreuses incertitudes associées à ces tests. De fait, si la mise en place de bases de données internationales dont la taille ne cesse de croître permet d'affiner les prédictions, elle ne réduit pas totalement l'incertitude. Le risque associé à certaines mutations reste toujours impossible à préciser et, pour d'autres, les estimations de risque sont évaluées à la baisse.

Aujourd'hui, toutefois, la spécificité de l'information génétique semble s'amoindrir du fait de l'explosion du marché des tests génétiques.

L'information génétique n'en serait plus qu'une parmi d'autres dans l'océan de données personnalisées sur notre biologie et nos modes de vie. Le droit à disposer de son information génétique individuelle serait du même ordre que celui de savoir ce que Google connaît de nous. Et même mieux, il nous permettrait de devenir encore plus acteurs que nous sommes déjà censés l'être de notre santé, du bonheur de nos enfants (en leur évitant d'être atteints de certaines maladies), de notre quête d'identité...

Dans ce récit lisse, largement porté par des discours de promesses que sous-tendent des intérêts économiques qu'il faut avoir en tête (*voir l'encadré page 30*), l'information génétique a une valeur donnée, fiable, scientifiquement validée. Peu de place est laissée aux difficultés, aux ajustements, aux incertitudes évoquées. Ou alors, il est considéré que les individus s'y adapteront, et les systèmes avec.

QUELLE LOI DE BIOÉTHIQUE ?

Les questions soulevées par les tests génétiques en libre accès se posent évidemment dans tous les pays. Mais tous n'entretiennent pas la même histoire avec la génétique. Ils n'ont pas fait les mêmes choix en matière d'organisation de leur système de santé. Ils n'ont pas tous la même confiance dans le marché pour organiser la vie sociale. Ils n'évaluent pas de la même façon la balance entre libertés individuelles et bien-être collectif. Dans un pays comme la France, où l'accès aux soins est pris en charge par une solidarité nationale, sous fortes tensions financières, on ne peut réfléchir à l'accès aux tests génétiques individuels sans penser aussi à ses conséquences sur le système de santé dans son ensemble, et notamment à celles de la multiplication d'actes médicaux liés à des tests génétiques réalisés en dehors de tout contexte clinique pertinent.

La généralisation des tests génétiques préconceptionnels *via* le diagnostic prénatal et l'augmentation du recours à la fécondation *in vitro*, qui en est souvent un corollaire, impliquent de confier encore un peu plus nos projets reproductifs au contrôle des médecins. Un choix qui n'a rien d'automatique. Cela pourrait renforcer encore la pression sociale qui s'exerce déjà sur les couples en les poussant à utiliser tous les moyens technologiques disponibles pour faire naître des enfants «en bonne santé», pour optimiser encore la sélection de vies considérées comme «valant le coup d'être vécues».

Ces interrogations, qui sont loin d'épuiser le sujet, sont importantes au moment où le Parlement français examine le nouveau projet de loi de bioéthique, dont un volet concerne les examens génétiques et la médecine génomique. D'autant que les progrès en termes de séquençage sont loin de se tarir. ■

BIBLIOGRAPHIE

P. DUCOURNAU, *S'entreprendre avec ses gènes: Enquête sur l'auto-généralisation*, PUR, 2018.

M. MATHIEU ET P. MALZAC, *Tests génétiques: illusion ou prédiction?*, Le Muscadier, 2017.

C. BOURGAIN ET P. DARLU, *ADN superstar ou superflic? Les citoyens face à une molécule envahissante*, Le Seuil, 2013.

L'ESSENTIEL

- La biologie enseigne que les gènes s'expriment dès lors que des protéines les activent.
- Or l'expression des gènes est souvent aléatoire: pour un même gène, elle varie d'une cellule à l'autre et au cours du temps.
- La première cause de l'expression aléatoire est la quantité de facteurs de transcription disponibles pour «lire» un gène.
- Une autre cause est spatiale: l'ADN est organisé de telle façon que l'activation d'un même gène est plus ou moins probable.

L'AUTEUR



THOMAS HEAMS est maître de conférences à AgroParisTech à Paris et chercheur à l'Inra dans l'unité Génétique animale et biologie intégrative (Gabi), à Jouy-en-Josas.

La fin du programme génétique

Le génome n'est pas le chef d'orchestre exclusif du fonctionnement d'une cellule et de l'expression de ses gènes. Le hasard y a aussi sa part.

D

ans toutes les universités du monde, les professeurs de biologie enseignent le même grand principe où le déterminisme est roi: les cellules obéissent à un programme inscrit dans le génome, l'ensemble des gènes localisés dans l'ADN du noyau cellulaire. Ce programme génétique est le chorégraphe d'un ballet

bien réglé, où chaque cellule et chaque molécule contribuent au bon fonctionnement de l'organisme entier. En particulier, chaque gène s'exprime, c'est-à-dire transmet un message, par l'intermédiaire de protéines qui se fixent sur l'ADN. Si l'une d'elles est présente en quantité suffisante, le gène qu'elle contrôle est activé ou réprimé. Cette description strictement déterministe ne reflète qu'une partie de la réalité.

Des expériences récentes ont confirmé des intuitions théoriques vieilles de plusieurs décennies: l'expression des gènes (*voir les Repères, page 6*) est souvent aléatoire: elle n'est pas la même, à un instant donné, dans des cellules dotées des mêmes gènes, ni au cours du temps dans une cellule particulière. En d'autres termes, loin d'être rigoureusement programmée, l'expression des gènes obéit à des lois de probabilité. Et cela change tout.

Avec une expression déterministe, l'existence de gènes impose un ensemble de >



Des variations aléatoires à chaque division cellulaire dans les profils épigénétiques des cellules expliqueraient en partie pourquoi deux vrais jumeaux n'ont pas exactement les mêmes caractéristiques physiques.

> caractéristiques cellulaires spécifiques, un phénotype. À l'inverse, avec une conception probabiliste, à un groupe de gènes ne correspond pas un phénotype unique; des aléas modulent l'expression de ces gènes, de sorte que certaines cellules n'acquiescent pas ce phénotype, tandis qu'il apparaît par hasard dans d'autres cellules.

En réalité, cette opposition n'est qu'apparente. D'une part, un événement plus ou moins probable n'échappe pas aux règles de la causalité. Le jeu de dés (*az-zahr* en arabe), par exemple, obéit aux probabilités, bien que les dés suivent les lois de la mécanique classique. D'autre part, et c'est ce qui nous intéresse ici, le déterminisme peut être envisagé à l'aune des probabilités: en effet, tout événement déterminé a une probabilité de s'accomplir égale à 0 (il ne se produit pas), ou égale à 1 (il se produit); quand on étudie la probabilité qu'apparaisse une combinaison d'un très grand nombre d'événements déterminés, ayant tous une probabilité égale à 0 ou à 1, on conçoit qu'elle puisse prendre toutes les valeurs entre 0 et 1, et devenir déterministe seulement dans ce dernier cas.

IDENTIQUES, MAIS DIFFÉRENTS

Les premiers indices de la variabilité aléatoire de l'expression des gènes datent de la fin des années 1950. Aaron Novick et Milton Weiner, de l'université de Chicago, travaillaient sur le modèle cellulaire de l'opéron lactose, dont l'équipe de Jacques Monod, à l'institut Pasteur, avait montré l'intérêt pour étudier l'expression des gènes: en présence de lactose,

des bactéries *Escherichia coli* fabriquent l'enzyme nécessaire à l'utilisation de ce sucre.

Ils montrèrent que des bacilles génétiquement identiques et placés dans les mêmes milieux de culture ont un comportement différent: les uns sont capables de métaboliser le lactose (ils fabriquent l'enzyme); les autres ne le peuvent pas, car ils produisent l'enzyme en trop petite quantité, voire pas du tout. Ces différences proviennent de variations de l'expression des gènes, mais aussi de fluctuations génétiquement déterminées des concentrations d'autres composants cellulaires.

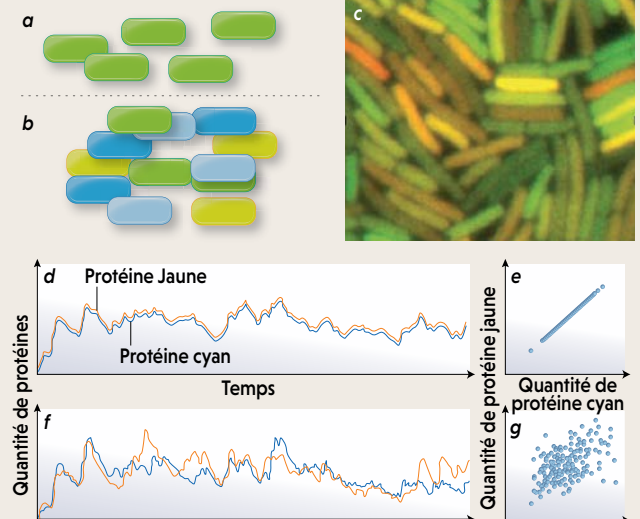
Après ces prémices, l'histoire de l'expression aléatoire des gènes s'est écrite en pointillé, faute de technique efficace. En 1976, John Spudich et Daniel Koshland, de l'université de Berkeley, décrivent chez des bactéries une «individualité non génétique», caractérisée par des cellules ayant des comportements différents en milieu homogène, qu'ils relient au hasard. Dans les années 1990, l'utilisation astucieuse, mais sporadique, de techniques de biologie moléculaire ou d'imagerie permet d'en préciser la nature.

Ainsi, Jamel Chelly, dans l'équipe d'Axel Kahn et Jean-Claude Kaplan, à l'institut Cochin de génétique moléculaire, décrivent en 1988 un phénomène d'expression «illégitime», où n'importe quelle cellule exprime faiblement n'importe quel gène. Mark Wijgerde et Peter Fraser, à Rotterdam, en 1995, envisagent une activation aléatoire des gènes d'une cellule à l'autre, et le groupe d'Edna Hardeman, en Australie, détecte une expression génique stochastique chez la souris, en 1998.

VARIATIONS SUR UN MÊME THÈME

Des colibacilles expriment deux gènes codant deux protéines fluorescentes, l'une bleu cyan, l'autre jaune. On s'attend à ce que les bactéries produisent la même quantité des deux protéines, et qu'elles deviennent vertes (a). Or les unes deviennent jaunes, d'autres bleues, et d'autres encore prennent des couleurs intermédiaires (b); sur la photographie (c), les couleurs des bactéries résultent de l'utilisation d'un filtre. Si l'expression des gènes des deux protéines n'était pas

aléatoire, l'abondance de la protéine jaune varierait au cours du temps de la même façon que celle de la protéine bleue (d). Autrement dit, ces quantités varieraient linéairement (e). En revanche, l'expression aléatoire des gènes se traduit par des variations indépendantes des quantités des deux protéines (f). La variation de la protéine jaune en fonction de la bleue se traduit alors par une distribution aléatoire des points de fluorescence: chaque point correspond à une bactérie (g).



Ces travaux expérimentaux se déroulent alors que Jean-Jacques Kupiec, à l'hôpital Saint-Louis, puis à l'institut Cochin, développe une théorie probabiliste de la différenciation cellulaire depuis le début des années 1980. Mais ce sont des techniques utilisant la protéine fluorescente verte, ou GFP, qui vont faire la part des choses, en prouvant que des gènes s'expriment de façon stochastique.

La nature stochastique de l'expression des gènes est confirmée, des bactéries jusqu'aux humains

En 2002, l'équipe de Peter Swain et Michael Elowitz, de l'université Rockefeller, à New York, intègre dans le génome d'*Escherichia coli* deux gènes codant deux variants de la GFP: l'un, une protéine à fluorescence bleu cyan, l'autre, une protéine fluorescente jaune. La construction génétique est identique dans les deux cas. D'après la théorie déterministe, chaque bactérie doit exprimer ces deux gènes dans des proportions quasi semblables; la fluorescence résultante devrait donc être similaire dans toutes les bactéries, qui auront peu ou prou la même «couleur»: un mélange à parts égales de bleu et de jaune, soit du vert. De petites différences ne représenteraient qu'un «bruit» extrinsèque: les bactéries pourraient être plus ou moins vertes.

Le résultat déjoua ces prédictions: la couleur de fluorescence variait d'une bactérie à l'autre, allant du jaune à l'orange et au vert, avec plusieurs couleurs intermédiaires. Cela signifiait que l'expression des gènes de fluorescence dans une bactérie donnée était aléatoire, indépendamment de la concentration des acteurs de la transcription (voir l'encadré page ci-contre).

La même année, l'équipe de Robert Singer, du collège de médecine Albert-Einstein, à New York, a étendu ce type d'analyse à plusieurs

gènes dans des cellules humaines. Des sondes d'ADN fluorescentes se fixaient sur les molécules d'ARN codées par 11 gènes. À chaque gène correspondait ainsi un «code à barres» coloré spécifique, composé par les sondes fluorescentes. Or, dans une population de cellules identiques, chaque cellule présentait une combinaison de codes à barres spécifique qui changeait au cours du temps; et toutes les cellules avaient des comportements différents.

Depuis, de nombreuses équipes ont confirmé la nature stochastique de l'expression des gènes, des bactéries jusqu'aux cellules humaines. Mais pourquoi des gènes s'expriment-ils ainsi? On peut identifier deux types de raisons. D'abord, des causes moléculaires, dues aux interactions des molécules sur les sites de la transcription (enzymes, facteurs de transcription, autres protéines, ADN). Ensuite, des causes topologiques, mettant en jeu la facilité avec laquelle les facteurs de transcription accèdent aux gènes selon la forme que prend la molécule d'ADN.

LA DYNAMIQUE DE L'EXPRESSION

Pour comprendre, rappelons qu'un gène est une séquence d'ADN, une suite de nucléotides (les lettres A, C, T et G), qui contient un message codé. Le code génétique établit une correspondance entre des séries de trois nucléotides et les acides aminés qui composent les protéines. L'expression d'un gène comporte plusieurs étapes qui aboutissent à la synthèse d'une protéine (voir *les Repères*, page 6). Elle commence par l'activation du promoteur, une petite séquence située au début du gène. Puis, une enzyme (l'ARN polymérase II) se fixe sur le promoteur et se déplace le long de l'un des brins de l'ADN: ce faisant, elle transcrit la suite de nucléotides composant le gène en un ARN dit messager; celui-ci sort du noyau et rejoint des ribosomes qui traduisent le message en protéines.

Ce mécanisme d'expression est régulé par des mécanismes dits épigénétiques, qui ne dépendent pas directement de la séquence d'ADN initiale: de petites molécules se fixent sur l'ADN ou sur les histones, ces protéines autour desquelles l'ADN s'enroule dans les cellules eucaryotes, dotées d'un noyau (voir *L'hérédité sans gène*, par M. Skinner, page 44). Par ailleurs, de petites molécules d'ARN, les micro-ARN, bloquent la traduction de certains ARN messagers. Or, à chacune de ces étapes, les structures moléculaires en jeu interagissent de façon dynamique, c'est-à-dire que leurs interactions évoluent constamment. C'est ce qu'a montré l'utilisation de sondes fluorescentes.

Par exemple, le groupe de Gordon Hager et Jim McNally, de l'Institut américain du cancer, a ainsi mis en évidence un ballet aléatoire de certains facteurs de transcription autour du promoteur de leurs gènes cibles: chaque molécule ne ➤

➤ reste liée au promoteur que durant quelques secondes ou dizaines de secondes avant d'être remplacée par une autre molécule. De même, on sait désormais que l'association des histones avec l'ADN est beaucoup plus dynamique que ce que l'on imaginait: le temps moyen de cette association varie de quelques secondes à quelques minutes, ce qui modifie sans cesse les zones de l'ADN qui peuvent ou non être transcrites.

On sait aussi, notamment grâce à l'équipe de Peter Fraser, à l'institut Babraham de Cambridge, que l'expression des gènes n'est pas homogène dans le noyau cellulaire; elle se déroule dans quelque 2500 régions particulières, nommées «usines à transcription». Ces régions sont riches en ARN polymérase. Plusieurs gènes partagent la même usine à transcription, sorte de tête de lecture sur laquelle chacun vient tour à tour «s'activer».

Faute de connaître les mécanismes tridimensionnels plausibles qui contrôleraient précisément ces mouvements, on doit privilégier l'hypothèse la plus simple: celle de changements aléatoires de la conformation de l'ADN qui exposeraient les gènes aux usines à transcription. Ce mode dynamique coïncide avec le fait que les ARN et les protéines correspondant à un gène sont produits par bouffées, parfois sous l'impulsion d'une seule molécule.

UNE EXPRESSION EN FONCTION D'UNE CLOCHE

Par ailleurs, les molécules qui régulent l'expression des gènes – facteurs de transcription et répresseurs – sont souvent en petite quantité dans la cellule. Ainsi, 80% des protéines d'une bactérie sont présentes en moins de 100 exemplaires; certaines, impliquées dans la division cellulaire ou dans la réparation de l'ADN, n'ont que quelques dizaines de copies.

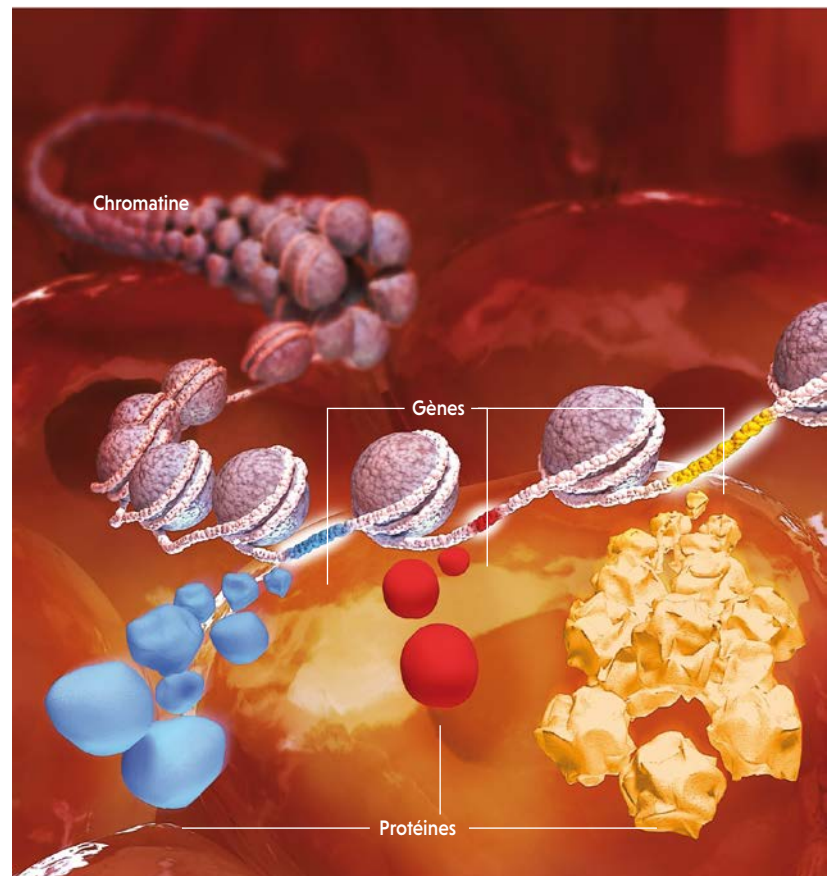
Prenons le cas d'une protéine que chaque cellule d'une population contient en 50 exemplaires. En théorie, lors de la division cellulaire, les cellules filles reçoivent, en moyenne, 25 copies chacune. Mais en réalité, cette répartition suit la loi de Gauss; elle est en forme de cloche centrée sur une moyenne. On peut donc s'attendre à ce qu'autour de cette moyenne, surviennent des situations très diverses. Dans environ 6% des cas, la répartition est même très inégale, certaines cellules filles recevant moins de 19 copies, tandis que d'autres en récupèrent plus de 30 (ces estimations sont issues d'une répartition gaussienne standard). De sorte que certaines cellules ont 50% de protéines de plus que les autres. Au final, la transcription d'un gène fonctionne sur un mode instable: son expression n'a lieu que si les molécules de régulation sont suffisamment nombreuses à proximité de son promoteur. Elle peut se produire dans une cellule (et pas dans ses

voisines), puis diminuer avec la dégradation des facteurs de transcription.

Ainsi, des cellules ayant les mêmes gènes produisent des quantités différentes de facteurs de transcription et expriment des gènes distincts. Pour un gène donné, le promoteur de la transcription passe d'un état actif à un état inactif et le rythme de transition d'un état à l'autre détermine à quelle vitesse fluctue l'expression du gène: plus la période séparant deux activations successives du promoteur est courte, plus la variation de l'expression du gène est rapide. Or on a montré que la structure du promoteur influe sur ce rythme. Ainsi, en 2006, l'équipe de James Collins, à l'université de Boston, a découvert chez la levure que les promoteurs qui contiennent un motif de nucléotides, la boîte TATA (riche en thymine et en adénine), réagissent plus vite que les promoteurs qui en sont dépourvus.

Ce sont donc les variations naturelles des interactions moléculaires sur les sites de la transcription qui rendent l'expression génique aléatoire. Cependant, la structure même de l'ADN dans la cellule y contribue également. Cette longue molécule occupe un volume dans

L'expression aléatoire des gènes se traduit par la synthèse de protéines différentes ou dans des proportions différentes dans les cellules génétiquement identiques d'un même tissu. Par exemple, une protéine (ici en rouge) peut être synthétisée en petite quantité tandis que d'autres (en bleu et en jaune) sont produites en plus fortes concentrations. En revanche, la même protéine sera plus abondante dans une cellule voisine, ou absente dans une troisième, tandis que les deux autres protéines seront moins présentes.



HASARD ET ÉPIGÉNÉTIQUE

Pour les Grecs, la Terre était le centre du monde. Quand Ptolémée, à l'aube du I^{er} millénaire, se rendit compte que cette hypothèse n'expliquait pas correctement la trajectoire des planètes, il proposa non pas d'en changer, mais d'ajouter des hypothèses supplémentaires, les « épicycles ». Il fallut attendre Copernic pour que l'on abandonne l'hypothèse initiale et que l'on adopte le système héliocentrique. Aujourd'hui, en biologie, la révolution copernicienne correspond à l'avènement de l'épigénétique : depuis quelques années, on rassemble sous cette appellation plusieurs phénomènes moléculaires disparates qui provoquent des différences héréditaires et réversibles de l'expression des gènes sans changer leur séquence (voir *L'hérédité sans gène*, par M. Skinner, page 44). Ces phénomènes remettent en cause la vision classique, déterministe, du fonctionnement des gènes, puisqu'il ne suffit plus de connaître la séquence précise d'un génome pour espérer expliquer le fonctionnement de l'organisme qui le porte. Le phénomène d'épigénèse le mieux connu est celui de la méthylation des gènes, l'ajout d'un groupe méthyle. Ces petites modifications chimiques de certaines bases de l'ADN peuvent empêcher ou moduler l'expression du gène qui les porte. La méthylation et la déméthylation sont des phénomènes dynamiques qui durent quelques dizaines de secondes au plus. De même, les histones, protéines autour desquelles s'enroule l'ADN, peuvent être méthylées, acétylées (par ajout d'un groupe acétyle) ou phosphorylées (groupe phosphate), et ce statut, lui aussi

dynamique, a un impact sur l'accessibilité des facteurs de transcription à l'ADN. On sait que les profils épigénétiques d'une cellule (par exemple l'ensemble des groupes méthyles portés par son génome) subissent des variations aléatoires à chaque division cellulaire. Ces différences expliqueraient en partie pourquoi deux vrais jumeaux n'ont pas exactement les mêmes caractéristiques physiques. D'autres phénomènes épigénétiques sont de natures ou d'échelles différentes. On peut mentionner les micro-ARN. Ces très nombreux petits ARN qui ne sont pas traduits en protéines peuvent perturber l'expression des autres gènes : quand ils s'apparient avec un ARN messager, ils provoquent sa dégradation ! Enfin, à une plus large échelle, les chromosomes occupent dans le noyau, même quand ils sont décondensés, des « territoires chromosomiques », dont l'organisation interne comporte des zones où l'expression des gènes est facilitée et d'autres où elle est réprimée. Il existe en outre des enzymes « remodeleuses » de la chromatine, qui modifient sa structure et rendent possible ou non l'expression des gènes. Il semble enfin que la position d'un gène dans le noyau, plutôt vers le centre ou plutôt vers la périphérie, ait des conséquences sur son expression. Ces mécanismes épigénétiques sont-ils la preuve qu'un ordre déterministe est à l'œuvre, dans la mesure où ils permettraient de contraindre la dimension aléatoire de l'expression des gènes ? Ou au contraire, sont-ils des indices que le hasard règne en maître, que toute volonté de codifier et de prévoir le

fonctionnement cellulaire est illusoire tant l'écheveau des interactions internes est complexe ?

L'existence d'un « code histone », qui se superposerait au code génétique, irait dans le sens d'un déterminisme caché. Ce code, encore hypothétique et contesté, résulterait des modifications chimiques des histones, dont les combinaisons correspondraient à différents états de transcription des gènes. Il reste à comprendre si les fluctuations aléatoires des modifications chimiques autorisent une stabilité suffisante du code, nécessaire à son « interprétation ». Le raisonnement est le même avec les micro-ARN dits interférents. On considère qu'ils représentent un nouveau mode de régulation, emboîté dans ceux déjà décrits. Mais si l'on se place dans le cadre probabiliste, on s'interroge : de tels éléments de régulation peuvent-ils vraiment s'emboîter de la sorte ? À tout le moins, on comprend qu'ils contribuent à la variabilité intercellulaire. Quant aux territoires chromosomiques, alors que l'on a longtemps pensé qu'ils se disposaient les uns par rapport aux autres de façon non aléatoire, ils fluctuent en fait selon les cellules ou les générations cellulaires. De sorte que si leur position a un impact sur l'expression des gènes qu'ils contiennent, cette dernière pourrait en être modifiée. On voit mal quel code complexe pourrait, en coordination avec les autres « codes », dicter ces variations topologiques. Elles sont plus facilement explicables par l'intervention du hasard, éventuellement suivie d'une stabilisation des états adaptés à la fonction de la cellule ou à sa survie.

le noyau, car elle se replie sur elle-même. Chez les eucaryotes, nous l'avons dit, elle s'enroule autour de protéines, les histones, formant un long filament, la chromatine. Elle ressemble alors à un long collier de perles enroulé de façon compacte, les perles correspondant aux assemblages d'ADN et de molécules d'histones, nommés nucléosomes.

Certains gènes sont inaccessibles, car situés au cœur de la « bobine ». De même, dans une cellule, par nature très encombrée par des centaines de protéines, de petites molécules et d'organites, le trajet d'un facteur de transcription jusqu'au promoteur de son gène cible est pour le moins aléatoire : s'il parvient à proximité de la molécule d'ADN, il faut encore qu'il pénètre dans les replis de la chromatine et qu'il trouve son chemin parmi les nucléosomes. Rien ne permet de penser que cette

progression se produise à l'identique d'une cellule à une autre, même si finalement la transcription du gène cible a bien lieu.

L'ALÉA CONTRÔLÉ

Un troisième type de mécanisme en jeu dans l'expression aléatoire des gènes est son... contrôle. Cette notion, qui sous-entend que l'on réduit l'incertitude, semble contredire celle de phénomène stochastique, par nature imprévisible. En fait, elle traduit, là encore, le fait que l'expression des gènes se déroule dans le contexte de la cellule et donc sous certaines contraintes.

Le mécanisme de contrôle le mieux connu est le rétrocontrôle négatif, ou autorégulation négative : des protéines synthétisées à la suite de l'expression d'un gène inhibent une étape située en amont de leur production, ce qui tend à réduire leur concentration. Si la cellule les

>

- > dégrade, leur production est à nouveau favorisée, car le rétrocontrôle diminue.

Un exemple bien étudié est celui des gènes et des protéines qui déterminent le fonctionnement de l'horloge circadienne. Deux protéines, *Per* et *Cry*, synthétisées sous l'action des protéines *Clock* et *Bmal1*, inhibent en retour ces deux molécules, ce qui stoppe l'expression des gènes *per* et *cry*. La concentration des deux protéines devient alors si faible qu'elles ne bloquent plus *Clock* et *Bmal1*, ce qui relance le cycle (les gènes *per* et *cry* pouvant de nouveau s'exprimer).

Résumons-nous. L'expression aléatoire des gènes est un phénomène répandu dans le monde vivant, explicable par des mécanismes moléculaires et structuraux, et contrôlé. Cette conclusion semble contre-intuitive. Le hasard serait-il contrôlé – ou déterministe – quand il s'agit de l'expression des gènes ?

La cellule constitue un milieu organisé et contraignant où les interactions moléculaires mettent en jeu des forces physicochimiques. Certes, les molécules en jeu sont si nombreuses qu'il semble vain d'espérer connaître leurs interactions. Mais cela ne signifie pas que l'expression aléatoire des gènes n'a pas de causes, qu'elle n'est pas déterminée.

Inversement, un processus déterministe a une probabilité maximale de se produire. En génétique, un tel processus se résume à la prédiction suivante : si une cellule a les gènes *a*, *b*... *n*, dans des conditions *c*, elle se différenciera inéluctablement en adoptant le phénotype *p* ; les conditions initiales imposent le comportement de la cellule. Ainsi, on peut prédire qu'une cellule souche du sang se différenciera en globule rouge si une série particulière de gènes est transcrite, alors que l'expression d'une autre combinaison de gènes produira un globule blanc. La prédiction considère donc les événements avec une probabilité égale à 0 (rien ne se passe comme prévu) ou à 1 (l'événement prédit se produit).

Nous entrevoyons alors une grammaire commune. La grille de lecture probabiliste affirme qu'une cellule va exprimer le phénotype *p*, non avec une probabilité de 0 ou de 1, mais avec une certaine probabilité située entre ces valeurs. Dans un tissu composé de nombreuses cellules, compte tenu des effets de moyenne appliqués aux variations entre cellules, la probabilité que les phénomènes aléatoires donnent un résultat prédictible et reproductible se rapproche alors de 1.

Le fonctionnement déterministe de la cellule représente donc un cas particulier de l'expression aléatoire des gènes. Selon cette conception, l'expression des gènes serait plus ou moins aléatoire, ou plus ou moins déterminée. Certains gènes seraient transcrits de façon quasi déterministe, grâce à des mécanismes acquis par sélection naturelle : augmentation

du nombre de copies, position sur le génome... D'autres gènes le seraient plus ou moins aléatoirement.

Le hasard joue sans doute un rôle notable quand le nombre de molécules produites est faible. En effet, la variation aléatoire de l'expression d'un gène – une mesure de l'intervention du hasard – est plus importante quand le nombre de molécules du facteur de transcription qui l'active est petit : si 100 molécules d'ARN sont formées dans une cellule, en produire cinq de plus correspond à une variation de 5%. En revanche, si dix molécules d'ARN seulement sont transcrites en moyenne, une transcription supplémentaire de cinq molécules représente une variation de 50%.

De nombreux travaux laissent penser que le génome n'est pas tant le support d'un programme qu'une « boîte à outils ». Face aux contraintes de l'environnement, la cellule y pioche pour « inventer » de nouvelles solutions. L'expression aléatoire des gènes confère alors un avantage : les cellules dont l'expression des gènes est stochastique survivent mieux aux changements de l'environnement ; certaines d'entre elles trouveront une solution adaptée au bout d'un certain nombre de cycles d'expression aléatoire, alors que les cellules qui expriment toutes les mêmes gènes risquent d'être condamnées. Ce comportement exploratoire suivi d'une stabilisation est une forme de sélection naturelle cellulaire.

UN PHÉNOMÈNE FONDAMENTAL

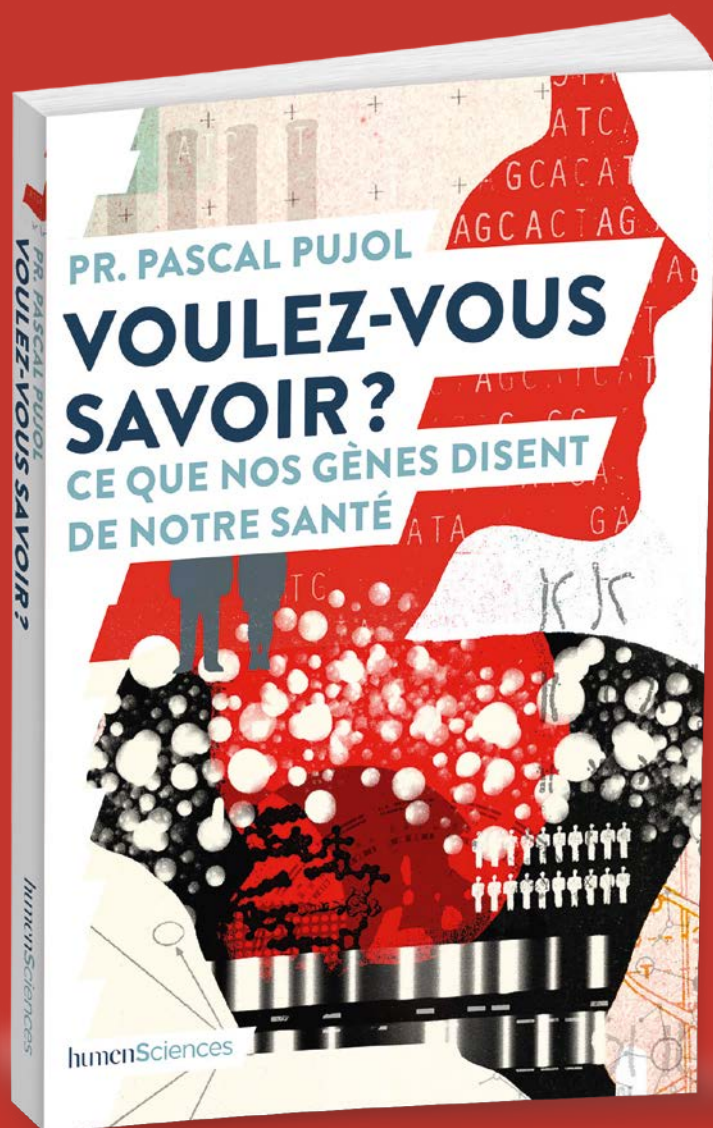
Cependant, il serait absurde de considérer que toute variation aléatoire de l'expression des gènes est avantageuse. Si les produits de certains gènes ainsi activés se révèlent nuisibles, les cellules qui les synthétisent risquent d'être éliminées par la sélection naturelle. Effectivement, plusieurs résultats révèlent que les processus évolutifs ont minimisé l'expression aléatoire des gènes qui limitent la capacité de l'individu à s'adapter. Par exemple, des gènes toxiques pour la levure sont transcrits de façon moins aléatoire que les autres gènes. Certains gènes sont donc plus soumis que d'autres aux fluctuations stochastiques de leur transcription.

Reste la question de fond : l'expression stochastique des gènes n'est-elle pas, plus simplement, un phénomène parasite, un bruit qui empêche le fonctionnement optimal d'organismes gouvernés par leur programme génétique ? Nous sommes nombreux à penser qu'on ne peut la réduire à cela. Elle constitue vraisemblablement aussi la base d'un phénomène biologique fondamental, utile et potentiellement nécessaire qui, en multipliant les combinaisons de gènes exprimés dans une cellule, ferait apparaître celle qui sera adaptée à l'environnement cellulaire, en une forme de sélection naturelle interne à l'organisme. ■

BIBLIOGRAPHIE

- T. HEAMS, *Expression stochastique des gènes et différenciation cellulaire. Le hasard au cœur de la cellule. Probabilités, déterminisme, génétique*, sous la direction de J.-J. Kupiec, O. Gandrillon, M. Morange et M. Silberstein, Syllepse, 2009.
- T. HEAMS, PH. HUNEMAN, G. LECOINTRE & M. SILBERSTEIN, *Les Mondes darwiniens. L'évolution de l'évolution*, Syllepse, 2009.
- A. RAJ ET A. VAN OUDENAARDEN, *Nature, nurture, or chance: stochastic gene expression and its consequences*, *Cell*, vol. 135, pp. 216-226, 2008.
- M. KERN ET AL., *Stochasticity in gene expression: from theories to phenotypes*, *Nat. Rev. Genet.*, vol. 6, pp. 451-464, 2005.
- M. B. ELOWITZ ET AL., *Stochastic gene expression in a single cell*, *Science*, vol. 297, pp. 1183-1186, 2002.

LA MÉDECINE VIT UNE RÉVOLUTION



« UN OUVRAGE DIDACTIQUE
QUI ÉCLAIRE LE DÉBAT. »
MIDI LIBRE

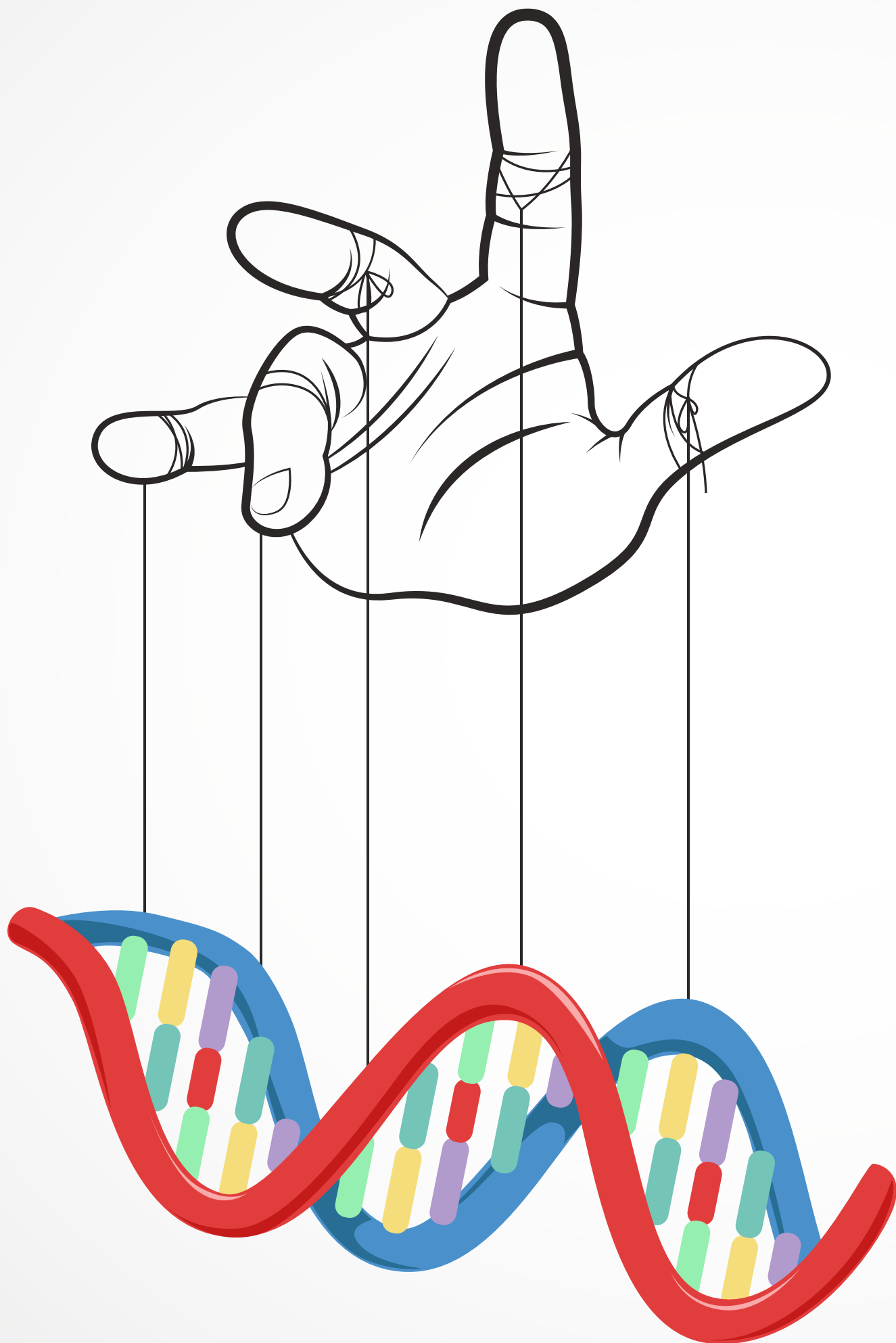
« UN PANORAMA
DES POTENTIALITÉS
DE CETTE NOUVELLE
MÉDECINE. »
LE MONDE

humenSciences

fait entrer la science dans votre vie



humensciences.com



UN GÉNOME SOUS INFLUENCE

Notre génome n'est pas enfermé dans une tour d'ivoire d'où il régenterait le fonctionnement de la cellule et donc de notre organisme. Bien au contraire, il est soumis à des aléas venus de l'extérieur. D'abord, ceux de l'environnement, notamment la pollution, qui imprime ses marques sur l'ADN et influe sur son fonctionnement : c'est l'épigénétique. Nous recevons aussi le renfort de nos microbiotes dont les gènes, trente fois plus nombreux que les nôtres, nous sont devenus au cours de l'évolution indispensables. Enfin, l'ADN n'a jamais été aussi vulnérable et malléable que depuis l'avènement de CRISPR-Cas9, cet outil rendant possible n'importe quelle manipulation génétique, avec une facilité déconcertante.

L'hérédité sans gène

L'expression des gènes varie au gré de modifications des chromosomes, indépendamment des séquences d'ADN. Ces marques dites épigénétiques, sensibles à l'environnement et aux polluants, se transmettent d'une génération à l'autre.

L'ESSENTIEL

● L'expression des gènes est régulée par des facteurs dits épigénétiques – des marques biochimiques sur les chromosomes qui laissent inchangée la séquence d'ADN.

● La plupart de ces marques sont réinitialisées peu après la fécondation.

● L'environnement (polluants, stress, alimentation...) introduit parfois des erreurs épigénétiques et modifie ainsi le comportement des cellules.

● Certaines de ces erreurs sont transmises aux descendants.

L'AUTEUR



MICHAEL SKINNER
est professeur de biologie
à l'université d'État de
Washington, aux États-Unis.

L

orsque mes enfants sont nés, il y a un peu plus de trente ans, je savais qu'à peu près la moitié de leur ADN était un héritage du mien. À l'époque, on pensait qu'il n'existait qu'une seule voie de transmission des informations biologiques héréditaires des parents aux enfants, du moins chez les humains et autres mammifères: le transfert de l'ADN du spermatozoïde et de l'ovule à l'embryon.

Bien sûr, on savait que l'ADN n'est pas tout. Si de nombreux traits y sont inscrits sous la forme de gènes codants – des séquences qui déterminent la forme et la fonction des protéines –, ce que nous absorbons compte aussi. Les conditions de vie, l'alimentation, la pollution, le stress... influent sur le fonctionnement des gènes. Par exemple, on invoque souvent les facteurs sociaux et environnementaux pour expliquer pourquoi de vrais jumeaux souffrent parfois de maladies différentes, malgré un patrimoine génétique très similaire.

Ce que l'on ignorait, c'est qu'on lègue à ses descendants bien plus qu'un code d'ADN. Nos enfants, mais aussi leurs enfants et petits-enfants, pourraient hériter d'une information dite épigénétique. Comme l'information génétique portée par l'ADN, l'information épigénétique est contenue dans les chromosomes et régule le fonctionnement des cellules. Mais il ne s'agit pas de

séquences d'ADN. Ses vecteurs sont divers, incluant de petites molécules biochimiques qui se fixent sur l'ADN et les protéines composant les chromosomes.

Cette information épigénétique est sensible à l'environnement. De fait, certains polluants, notamment des produits chimiques agricoles, le kérosène et même certains plastiques, sont en mesure d'entraîner des aberrations épigénétiques. Et ces anomalies peuvent se traduire en maladies et problèmes de fertilité, sans que la séquence d'ADN soit modifiée. Plus étonnant, de telles «épimutations» sont parfois transmises aux générations suivantes – avec des éventuels risques pour leur santé. Les implications en santé publique sont considérables. La forte progression de l'obésité, du diabète et d'autres maladies chez les enfants du baby-boom et les plus jeunes générations serait liée, en partie, à l'exposition de leurs parents et de leurs grands-parents à des polluants tels que la dioxine et le DDT.

LA MATIÈRE NOIRE DU GÉNOME

Les effets épigénétiques sont connus depuis quelque temps, mais l'étendue de leur rôle n'est apparue que récemment. Il y a plusieurs dizaines d'années, des biologistes ont remarqué que des groupes méthyle (CH_3) sont attachés à l'ADN en de nombreuses positions (*voir l'encadré page 48*). Chez les mammifères, dont l'homme, cette modification se rencontre souvent aux endroits où la lettre C du code génétique (la cytosine) précède la lettre G (la guanine) – une configuration que l'on retrouve environ 28 millions de fois dans l'ADN humain.

On a tout d'abord cru que la fonction principale de cette méthylation de l'ADN était de réprimer les transposons – des fragments d'ADN capables de se déplacer dans le génome, provoquant parfois des maladies. Aujourd'hui, nous savons que la méthylation contribue aussi à la régulation de l'activité des gènes et qu'elle est perturbée dans de nombreuses maladies, dont les cancers.

Des mutations d'ordre épigénétique, c'est-à-dire laissant inchangée la séquence d'ADN, peuvent-elles se transmettre d'une génération à l'autre, le long de l'arbre généalogique ?

© Milaski/Shutterstock.com

> Dans les années 1990, des biologistes ont mis en évidence d'autres types de marques épigénétiques et leur fonctionnement. Ils ont découvert que des groupes méthyle, acétyle (CO-CH_3) et d'autres modifications biochimiques marquent des groupes de protéines nommées histones. Dans un chromosome, la molécule d'ADN s'enroule autour de chaque groupe d'histones, formant comme un collier de perles. Les marques épigénétiques contrôlent le degré de resserrement des boucles d'ADN autour des «perles» et l'écart entre chaque perle, ce qui active ou inhibe des ensembles entiers de gènes. Par exemple, les gènes situés dans les endroits très compacts sont réduits au silence, car inaccessibles à la machinerie qui lit les gènes.

Depuis, on a identifié d'autres acteurs épigénétiques tels que la structure tridimensionnelle changeante de l'ADN ou des ARN dits «non codants». Les ARN sont des molécules produites par la cellule lors de la lecture du code ADN. Nombre d'entre eux sont des intermédiaires dans la fabrication des protéines. Mais parmi les ARN non codants (non traduits en protéines), certains modulent les marques épigénétiques de l'ADN et des histones.

Tous ces acteurs épigénétiques régulent l'activité des gènes de façon complexe, sans changement de la séquence d'ADN. L'interaction des gènes et de l'épigénome, l'ensemble des marques épigénétiques, est dynamique et reste encore mystérieuse. On sait que

d'ADN des altérations, la carte des marques épigénétiques varie au cours du développement de l'individu. Ces changements contribuent à déterminer la spécialisation des cellules, c'est-à-dire leur acquisition des caractéristiques de cellules de la peau ou de neurones, par exemple: de légères modulations de l'information épigénétique suffisent à changer la carte des gènes actifs dans tout l'organisme. De même, les substances chimiques dangereuses, les carences nutritives et autres types de stress pourraient entraîner l'apparition ou la disparition de marques épigénétiques, et modifier l'activité des gènes.

Aujourd'hui, le rôle crucial des marques épigénétiques dans le développement, le vieillissement et même le cancer ne fait aucun doute. Mais les biologistes débattent sur la portée de leurs effets: des épimutations anormales peuvent-elles se transmettre sur plusieurs générations de mammifères?

UN HÉRITAGE ACCIDENTEL

Mon premier contact avec cette question a été fortuit. Il y a quelques années, avec des collègues, nous étudions les effets, sur la reproduction des rats, de deux substances chimiques employées dans l'agriculture: le méthoxy-chlore, un pesticide, et la vinclozoline, un fongicide. Ces molécules sont des perturbateurs endocriniens: elles interfèrent avec les signaux hormonaux qui orchestrent la formation des organes de reproduction et leur fonctionnement. Nous les avons injectées à des rates au cours de leur deuxième semaine de gestation – au moment où les gonades de l'embryon se développent. Presque tous les descendants mâles nés de ces mères ont présenté des testicules anormaux, fabriquant des spermatozoïdes médiocres en qualité et en quantité.

Nous ne pensions alors pas à un effet épigénétique, et encore moins à une transmission de ces anomalies par les mères exposées. Aussi n'avions-nous pas prévu de les élever et d'étudier leur descendance. Or un jour, par erreur, on accoupla des rats mâles et femelles issus de cette expérience (sans lien de parenté). Résultat? Plus de 90% des mâles de ces portées arboraient les mêmes anomalies des testicules que leur père, alors que leurs parents étaient des embryons de la taille d'une tête d'épingle lorsque leurs propres mères avaient été exposées aux substances chimiques.

Cela nous surprit. De nombreuses études toxicologiques avaient recherché si des substances répandues dans l'environnement, comme le méthoxy-chlore ou la vinclozoline, provoquaient des mutations de l'ADN. En vain. Ce que confirmait notre étude. De plus, la génétique classique ne pouvait expliquer l'apparition d'un nouveau trait chez 90% de rats issus de différentes familles.

**On lègue
à nos enfants,
petits-enfants...
bien plus qu'un
code génétique
fait d'ADN**

chaque fois qu'une cellule se divise, les marques épigénétiques de ses chromosomes sont copiées et se retrouvent dans ceux des cellules filles. Les événements épigénétiques survenus tôt dans la vie peuvent donc altérer le comportement des cellules à long terme.

De plus, alors que les cellules sont équipées d'une machinerie qui protège la séquence

Cependant, une autre explication était envisageable : l'exposition aux substances chimiques avait eu lieu alors que l'embryon contenait déjà des cellules germinales primordiales – des cellules qui se spécialiseront en spermatozoïdes ou en ovules. La substance chimique aurait pu influencer directement ces cellules, l'effet persister dans les spermatozoïdes ou les ovules qu'elles avaient donnés et, par conséquent, se retrouver dans les fruits de leur fécondation : les petits de la génération suivante, soit la troisième génération depuis les mères exposées. Si tel était le cas, alors la brève exposition aux substances chimiques aurait directement entraîné les anomalies des testicules des petits rats de la troisième génération, mais les générations suivantes devraient être normales.

LES RATS MAUDITS SUR PLUSIEURS GÉNÉRATIONS

Pour le vérifier, nous avons élevé une quatrième génération de rats, puis une cinquième, en accouplant à chaque fois des descendants des rats exposés à l'origine, afin d'éviter la dilution du trait étudié. À chaque fois, les rats souffraient des mêmes problèmes que leurs ascendants. Et tous ces changements résultaient d'une exposition brève (mais à des doses bien plus élevées) à des produits utilisés dans l'agriculture pendant des décennies.

Après plusieurs années d'expériences qui ont confirmé ces résultats, nous avons formulé l'hypothèse suivante : l'exposition aux

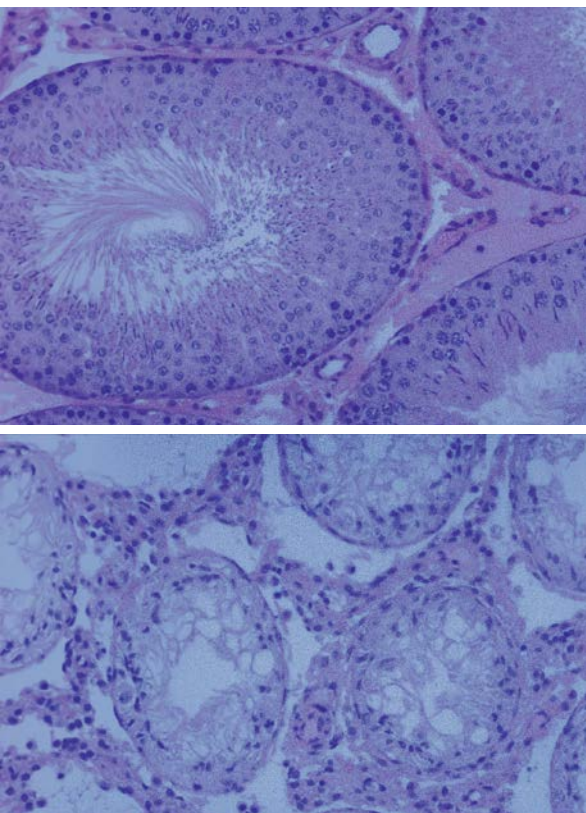
substances avait déclenché l'apparition d'épimutations pathologiques pour le développement des testicules. Au moment de la fécondation, ces épimutations passeraient du spermatozoïde à l'embryon, puis persisteraient dans ses cellules germinales primordiales, et ainsi de suite de génération en génération. En accord avec cette hypothèse, nos études préliminaires suggéraient des défauts de méthylation de l'ADN dans les spermatozoïdes des descendants des mâles exposés au fongicide.

Nos résultats déclenchèrent une tempête de réactions. Notamment, diverses équipes, issues de sociétés commercialisant la vinclozoline ou indépendantes, ont signalé qu'elles n'arrivaient pas à reproduire certains résultats. Ces incompatibilités pourraient être liées à des protocoles expérimentaux différents : administration par ingestion ou par injection, utilisation d'autres lignées de rats, mise en accouplement des mâles anormaux avec des femelles issues d'une lignée non exposée. Néanmoins, nous avons accumulé d'autres preuves suggérant que les épimutations pourraient persister sur plusieurs générations.

Notamment, nous avons mis en évidence des anomalies de méthylation de l'ADN dans les spermatozoïdes, les testicules et les ovaires de la quatrième génération issue des rats exposés au fongicide. L'activité des gènes dans leurs cellules germinales primordiales est aussi perturbée. Ces descendants ont également tendance à grossir et à être sujets au stress ; ils sélectionnent même leurs partenaires différemment.

Dans le même temps, la liste des agents polluants et stressants produisant des effets transgénérationnels s'est allongée. En 2012, nous avons montré que l'exposition de rates gravides à des polluants comme la dioxine, le kérosène, les insecticides ou un mélange de bisphénol A et de phtalates... déclenche des troubles qui persistent chez les rats de la quatrième génération : anomalies de la puberté, obésité, maladies des ovaires, des reins et de la prostate... Dans leurs spermatozoïdes, les profils de méthylation de l'ADN comptent des centaines d'anomalies spécifiques de l'agent auquel leurs ancêtres ont été exposés. De plus, les effets observés ne suivent pas les profils classiques de transmission génétique, ce qui suggère que des épimutations, et non des mutations génétiques, sont la cause de ces troubles.

Les doses utilisées pour ces études sont bien plus importantes que celles auxquelles l'homme ou l'animal sont exposés dans un environnement même fortement contaminé. Cependant, une étude de Jennifer Wolstenholme, de l'université de Virginie, suggère que des effets transgénérationnels sont aussi possibles à des concentrations plus réalistes. Du bisphénol A a été introduit dans l'alimentation de souris à des doses reproduisant des concentrations dans le sang équivalentes à celles ➤



Des arrière-petits bébés de rats exposés à un pesticide, la vinclozoline, se sont révélés peu fertiles, comme certains de leurs ascendants indirects. En témoigne la morphologie de leurs testicules en coupe histologique (*en bas*), anormale par rapport à celle d'un rat sain (*en haut*). Cette malformation serait due à un mécanisme épigénétique.