

> 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.

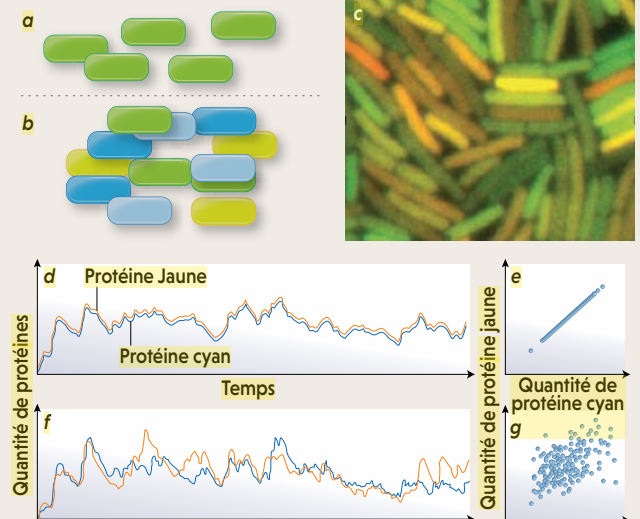
Après ces prémices, l'histoire de l'expression aléatoire des gènes s'est écrite en pointillé, faute de technique efficace. En 1976, John Spudich et Daniel Koshland, de l'université de Berkeley, décrivent chez des bactéries une «individualité non génétique», caractérisée par des cellules ayant des comportements 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 ➤