

# Un mécanisme de hasard-sélection pourrait expliquer la différenciation de plusieurs types cellulaires

Jean-Jacques Kupiec

*Selon la théorie dite de hasard-sélection, la différenciation cellulaire s'effectue en deux phases à chaque stade du développement. Durant la première phase, l'expression des gènes est instable. Cette instabilité, provoquée par la diffusion des facteurs de régulation de l'expression des gènes, produit différents types cellulaires. Durant la seconde phase, des interactions cellulaires activent des mécanismes qui limitent les possibilités*

*de diffusion des facteurs de régulation et stabilisent ainsi l'expression des gènes. Cependant, cette stabilisation ne peut se produire que si une combinaison adéquate de types cellulaires est présente. Ce mécanisme de sélection ordonne l'expression des gènes et dirige l'embryon vers le stade adulte. Un des mécanismes de stabilisation pourrait être la phosphorylation de ces facteurs de régulation.*

La différenciation cellulaire est habituellement expliquée par des modèles déterministes (ou instructionnistes). Dans le cadre de ces modèles, les cellules reçoivent une information spécifique (ou instruction) qui provoque une différenciation bien déterminée (figure 1A). Cette information est supposée être véhiculée par des interactions cellulaires, ou bien par des facteurs de différenciation diffusibles (inducteurs spécifiques ou morphogènes). Cependant, dans ces modèles, l'origine de l'hétérogénéité cellulaire n'est pas expliquée. En effet ils sous-entendent que les cellules sont déjà différentes au départ du processus puisque certaines d'entre elles synthétisent un inducteur spécifique transmis aux autres cellules (figure 1A). Ils nécessitent donc une hypothèse supplémentaire. On postule, par exemple, l'existence de mitoses asymétriques ou bien de

gradients morphogénétiques pré-existants dans l'œuf. Il existe une deuxième catégorie de modèles qui expliquent la différenciation cellulaire par la réalisation d'événements aléatoires internes aux cellules, sans l'intervention d'un signal extérieur (figure 1B). Grâce à ce cadre conceptuel, il est possible de comprendre, sans hypothèse supplémentaire, comment deux cellules peuvent devenir différentes même si elles sont identiques au départ du processus. Par exemple, selon que l'événement aléatoire a ou b se produit (figure 1B), la cellule se différencie en type A ou B, respectivement. En accord avec un tel modèle probabiliste, il a été montré que la cinétique de différenciation de nombreuses lignées cellulaires est mieux décrite lorsqu'on assigne aux cellules une probabilité de se différencier par unité de temps ou par mitose [1-3]. Plus récemment une composante aléatoire a égale-

ment été impliquée dans la différenciation des myoblastes, des cellules de la crypte intestinale, des hépatocytes, des lymphocytes et des cellules de la crête neurale [4-10]. Cependant, la nature des événements moléculaires qui sont à la base de ce phénomène est encore inconnue. Je propose une théorie fondée sur le concept de hasard-sélection [11, 12], dans laquelle les interactions cellulaires sélectionnent la voie de différenciation parmi un ensemble d'états issus d'événements d'activation aléatoire des gènes. Dans ce cadre théorique deux aspects doivent être discutés: l'activation des gènes et la sélection cellulaire.

## Modèle d'expression aléatoire des gènes

Ce modèle est fondé sur la propriété qu'ont les molécules qui interagissent avec l'ADN d'atteindre leur séquence

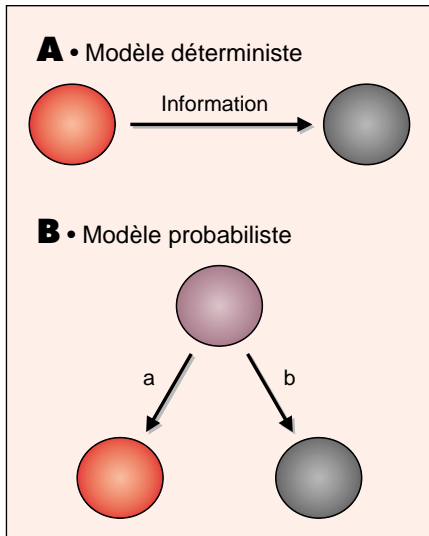


Figure 1. **Modèle déterministe et modèle probabiliste.** **A.** Dans le modèle déterministe les deux cellules sont déjà différentes parce que l'une d'entre elles fabrique et transmet une information à l'autre cellule. **B.** Dans le modèle probabiliste, selon que c'est l'événement *a* ou l'événement *b* qui se produit la cellule se différencie en A ou B, respectivement. Dans une population cellulaire la proportion des cellules A et B dépend des probabilités des événements *a* et *b*.

d'affinité par diffusion le long de l'ADN [13-15], y compris dans le cas des eucaryotes [16], et sur l'hypothèse qu'il y a une compétition entre les séquences affines de l'ADN pour les régulateurs de la transcription. Plusieurs auteurs ont proposé une hypothèse similaire de compétition. C'est le cas de l'activation des gènes de la globine et celui de la génération des sites d'hypersensibilité à la DNase I dans la chromatine [17-19]. L'exemple le plus simple du modèle est décrit en figure 2A. Une seule molécule de régulateur sert à activer deux gènes (ou ensembles de gènes) *a* et *b*. Lorsque le régulateur interagit avec le promoteur de *a*, ce gène est activé; lorsqu'il se dissocie de ce promoteur, il se déplace par diffusion dans le nucléoplasme. A cause des lois de la diffusion, il peut alors atteindre n'importe laquelle de ses séquences d'affinité avec une probabilité inversement proportionnelle à la distance

qui le sépare de ces séquences. Ainsi, après dissociation de *a*, le régulateur aura la plus forte probabilité de revenir en *a*, mais il pourrait également, avec une probabilité plus faible, atteindre le gène *b* et l'activer.

Dans ce modèle, l'activation des gènes est un processus probabiliste dépendant de deux paramètres: les distances séparant les gènes et les constantes de dissociation entre les régulateurs et leurs séquences d'affinité dans l'ADN. Plus les gènes sont proches et plus grandes sont les constantes de dissociation, plus grande sera la probabilité de déplacement d'un gène à l'autre.

La figure 2B décrit un autre exemple du modèle, légèrement plus compliqué. Une molécule régulatrice peut activer cinq gènes. A partir de *a*, elle a une probabilité égale de passer en *b* ou en *d*; et à partir de ces gènes la probabilité la plus forte est qu'elle aille respectivement en *c* ou *e*. La position relative des gènes définit différentes séquences de déplacement du régulateur. Deux de ces séquences ont la probabilité la plus forte de se réaliser ( $a > b > c$  et  $a > d > e$ ). Cependant, toutes les autres séquences peuvent se réaliser, mais avec des probabilités plus faibles. Dans le cadre de ce modèle, ce sont les interactions cellulaires qui contrôlent ce processus en empêchant que

se produisent les séquences qui ne correspondent pas à des voies de différenciation adéquates. Dans une population de cellules qui se multiplie, les probabilités de réalisation des séquences de déplacement déterminent trois paramètres qui sont essentiels: (1) les différentes voies de différenciation; (2) la chronologie de l'activation des gènes; (3) la proportion de cellules qui s'engagent dans chaque voie de différenciation. En effet, la probabilité de déplacement détermine le temps nécessaire pour le déplacement d'un gène à l'autre. Il est évident que plus grande est la probabilité de déplacement vers un gène, plus tôt ce gène sera exprimé, et inversement. Il en découle que la taille des populations cellulaires de l'embryon dépend également des probabilités de déplacement des régulateurs. Ce temps nécessaire à l'activation d'un gène est une variable statistique. C'est-à-dire que lorsque ce phénomène se répète plusieurs fois, il est statistiquement reproductible, décrit par une moyenne et une variance. Les probabilités de déplacement dépendant de la position des gènes sur le chromosome et des constantes de dissociation des régulateurs, elles sont génétiquement transmissibles, assurant la reproductibilité du processus de différenciation à chaque génération.

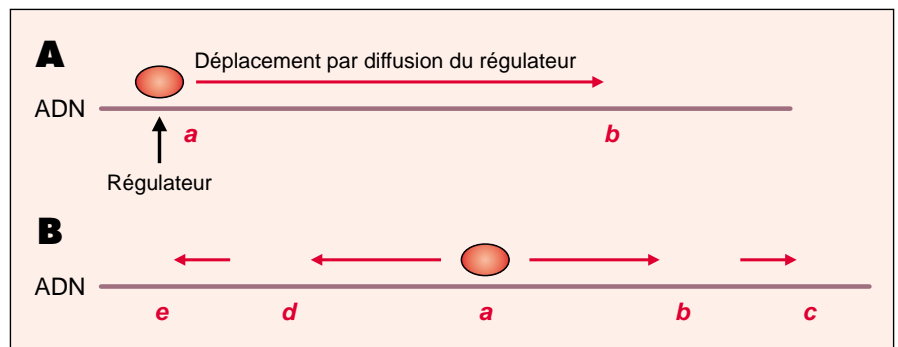


Figure 2. **Modèle d'expression aléatoire des gènes.** Le régulateur se déplace de manière aléatoire par diffusion le long de l'ADN. **A.** *a* et *b* sont deux gènes (ou ensembles de gènes). La probabilité de déplacement de *a* vers *b* est fonction de la constante de dissociation entre le régulateur et les séquences d'affinité dans les régions promotrices de *a* et *b*. **B.** Dans cet exemple, le régulateur sur le gène *a* est équidistant de *b* et *d*, et de *c* et *e*. Les gènes *b* et *d* sont plus près, respectivement, de *c* et *e* que de *a*. On suppose également que les constantes de dissociation sont égales pour tous les promoteurs. Cette organisation définit deux séquences d'activation  $a > b > c$  et  $a > d > e$  qui ont la probabilité de réalisation la plus grande.

Les exemples décrits sur la *figure 2* sont évidemment très simples. Cependant, les principes du modèle resteraient les mêmes dans le cas de systèmes plus complexes, impliquant un plus grand nombre de molécules régulatrices et de gènes. On peut les formuler de la manière suivante: (1) Premier principe: il y a moins de régulateurs transcriptionnels que de séquences d'affinité dans l'ADN et différentes distributions de ces régulateurs sur les séquences d'affinité sont possibles. Chacune de ces distributions correspond à un état d'organisation de la chromatine et donc à l'activation ou la répression de différents ensembles de gènes; (2) Deuxième principe: après dissociation de l'ADN et diffusion, les régulateurs sont redistribués sur les séquences d'affinité, changeant l'état d'organisation de la chromatine et donc d'expression génétique.

Ces deux principes définissent un modèle d'auto-assemblage probabiliste de la chromatine. A la différence des modèles d'assemblage déterministes, les interactions des molécules chromatiniennes ne sont pas fixes et statiques. Cela permet les transitions entre les différents états d'organisation de la chromatine.

Le répresseur de la transcription MeCP2 fournit un exemple de régulateur qui pourrait participer à ce mécanisme. En effet, il a pour cible le dinucléotide méthyl-CG. Il existe  $4.10^7$  copies de cette cible pour  $10^6$  molécules de MeCP2 dans un noyau de cellule de mammifère et il est très peu probable que ces molécules soient distribuées de manière identique dans toutes les cellules (*m/s n° 10, vol. 13, p. 1210*) [20].

Ce mécanisme suppose également une instabilité de la chromatine fondée sur des événements de dissociation entre les régulateurs et l'ADN. Les quelques données expérimentales disponibles semblent compatibles avec le modèle. En effet, on sait que dans le cas du gène de la globine il y a chez l'embryon un changement d'expression allélique toutes les 15 à 30 minutes (*m/s n° 1, vol. 12, p. 105*) [21] et que dans le cas de la protéine TBP de la levure, la durée de vie du complexe TBP-ADN est de 23 minutes [22].

Étant donné la variabilité des systèmes biologiques, de nombreuses variantes pourraient exister. Le déplacement des régulateurs pourrait se faire par diffusion dans un espace à une dimension, le long de l'ADN, comme cela a été montré pour plusieurs molécules [13-16], ou bien par une diffusion classique dans un espace à trois dimensions. La probabilité de déplacement d'un régulateur d'un gène à un autre pourrait s'exprimer directement en fonction du temps ou bien en fonction du nombre de mitoses. Il pourrait également exister d'autres variations concernant la nature du régulateur ou son mode d'action. Il pourrait également s'agir d'un complexe de plusieurs facteurs. Dans tous ces cas la logique du modèle resterait la même.

On peut, également, reposer le problème de la répllication de la structure de la chromatine. Ce problème est difficile à résoudre, y compris dans le cadre des modèles déterministes [23, 24]. Le modèle d'expression aléatoire n'apporte pas, lui non plus, de solution définitive. Cependant, les facteurs de régulation détachés de l'ADN au moment de la répllication pourraient se réassocier aux molécules filles selon les mêmes principes que ceux du modèle général de la *figure 1*.

### Colinéarité

Dans le cadre du modèle d'expression aléatoire, il existe une corrélation entre la position des gènes et leur séquence temporelle d'expression. Cette corrélation a effectivement été démontrée pour les gènes de la globine et les homéogènes, et appelée colinéarité [18, 25]. On peut également faire une autre prédiction simple: si l'on inverse la position des gènes on doit également inverser leur ordre d'expression. Cette prédiction est également vérifiée pour les gènes de la globine et les homéogènes [18, 26].

Il a, cependant, été possible de reconstituer une expression tissulaire et chronologique apparemment normale pour plusieurs homéogènes isolés de leur locus. Au contraire des résultats précédemment cités, ces

résultats semblent suggérer que ces homéogènes n'ont pas besoin d'être positionnés dans le locus pour que leur expression soit correctement réglée. Il y a là un paradoxe, parce qu'il semble inconcevable que la position des homéogènes ait ainsi pu être conservée par l'évolution sans qu'elle ait une signification fonctionnelle. Ce problème a déjà été analysé par Robb Krumlauf (Londres, GB) [25]. Les expériences de transgénèse ont en fait été réalisées dans un contexte génétique sauvage, dans lequel les protéines Hox endogènes sont normalement synthétisées. Dans la mesure où il a été montré que certains homéogènes exercent des effets autorégulateurs, les résultats obtenus pourraient ne refléter que cette propriété.

Il est important de remarquer que, d'un point de vue purement mécanique, la différence est minime entre un modèle d'expression aléatoire et un modèle classique. Elle ne tient qu'à la proportion relative des régulateurs vis-à-vis des séquences d'affinité de l'ADN. Si cette proportion est inférieure à 1 (conditions non saturantes), on est dans le cadre du modèle aléatoire, et si elle est égale (ou supérieure) à 1 (conditions saturantes), on est dans le cadre d'un modèle classique. Il est possible que les deux modèles co-existent. C'est-à-dire que certains gènes fonctionnent selon le modèle d'expression aléatoire, alors que d'autres gènes fonctionneraient selon un modèle classique, indépendamment de leur position chromosomique. Les gènes dépendant du modèle d'expression aléatoire engendreraient une diversité d'expression génétique nécessaire à la différenciation, alors que les gènes fonctionnant selon un modèle classique assureraient l'expression coordonnée de nombreux gènes.

Il faut aussi noter qu'une colinéarité simple et linéaire comme pour la globine ou les homéogènes n'est pas forcément une loi absolue pour tous les gènes. En effet, la structure secondaire (ou tertiaire) et les mouvements de diffusion aléatoire de la chromatine pourraient permettre des translocations des régulateurs dans des régions distales et non proxi-

males [27], en particulier si certains régulateurs peuvent diffuser dans un espace à trois dimensions. Il est même possible que, dans certains cas, la probabilité de déplacement des régulateurs d'un gène à l'autre dépende principalement de la constante de dissociation entre le régulateur et les séquences d'ADN, et non de la position dans le génome. On n'a pas observé, jusqu'à présent, de manifestation phénotypique dans les cas de translocations chromosomiques équilibrées qui ont lieu sans perte de matériel génétique. Ce fait semble également contredire le modèle d'expression aléatoire. En fait, il est possible que les gènes fonctionnant selon le modèle d'expression aléatoire assurent les choix de différenciation fondamentaux des cellules mais soient quantitativement peu nombreux. La probabilité qu'une cassure chromosomique interrompe un locus correspondant à ces gènes serait alors faible. Il pourrait y avoir une autre explication aux translocations équilibrées sans manifestation phénotypique. D'après le modèle d'expression aléatoire, si l'on change l'ordre des gènes on va modi-

fier la probabilité de réalisation des séquences d'activation, probablement la diminuer. Cela se traduirait par une mortalité embryonnaire plus élevée, les individus survivants étant tout à fait « normaux ». Jusqu'à présent, la mortalité embryonnaire observée semble être due à des caryotypes déséquilibrés qui sont produits avec une fréquence élevée par les parents porteurs de translocations équilibrées. Cependant, la mortalité embryonnaire est souvent difficile à observer parce qu'elle se produit dans des phases trop précoces du développement. Ainsi, il pourrait y avoir une mortalité supplémentaire non détectée, directement causée par les translocations équilibrées. Finalement, la redondance fonctionnelle et génétique des génomes eucaryotes pourrait servir de mécanisme compensateur en dupliquant les séquences d'activation des gènes et en palliant ainsi la défaillance de l'une d'entre elles. Dans la perspective d'un modèle probabiliste, la redondance pourrait avoir comme effet global d'augmenter la probabilité de réalisation de la différenciation cellulaire.

## Sélection cellulaire

### • Interactions cellulaires stabilisatrices

Dans le cadre de ce modèle, la stabilisation de l'expression génétique est provoquée par les interactions cellulaires lorsqu'une combinaison correcte de phénotypes cellulaires est exprimée. Un exemple extrêmement simple est montré sur la *figure 3*. Un organisme est constitué par deux cellules A et B, déterminées par l'expression de deux ensembles de gènes, respectivement, *a* et *b*. Un des gènes de chaque ensemble *a* et *b*, code, respectivement, pour une protéine membranaire de type A et B. La stabilisation de l'expression génétique est provoquée par l'interaction entre une molécule membranaire de type A et une autre de type B. Les gènes *a* sont activés en premier (*figure 3*, stade 1; dans la cellule souche le régulateur se trouve dans la région située en 5' des gènes *a*, où se trouve située une première séquence d'affinité). Deux molécules de type A n'ont pas d'affinité et ne peuvent provoquer la stabilisation de l'expression génétique (stade 1).

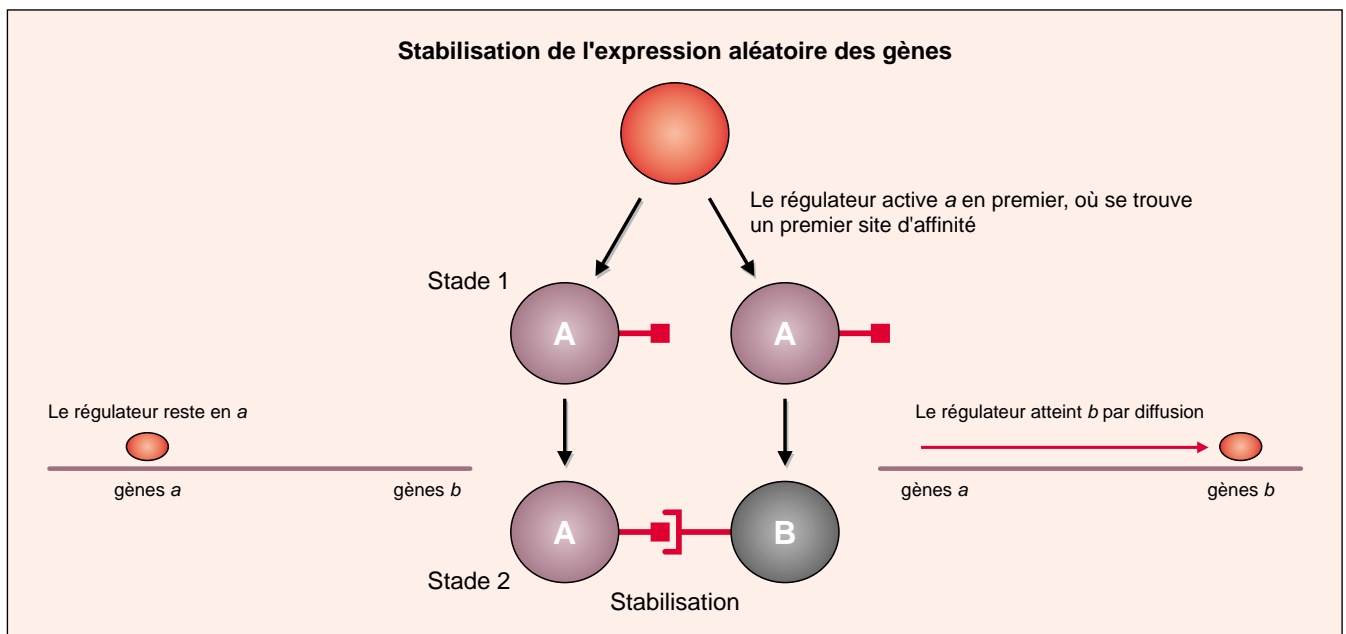


Figure 3. **Stabilisation de l'expression aléatoire des gènes.** Seule l'interaction de A et B provoque la stabilisation du régulateur. De ce fait, cet organisme est dirigé vers le stade 2.

Pour des raisons identiques, la stabilisation ne peut se produire si les deux cellules produisent des molécules de type B. Cette stabilisation ne peut se produire que lorsque le déplacement du régulateur aux gènes *b* a eu lieu dans une cellule, alors qu'il est resté au locus *a* dans l'autre cellule (stade 2). Ce système de stabilisation équivaut à une sélection de la cellule A par la cellule B et réciproquement, permettant ainsi une coordination de l'apparition de ces phénotypes cellulaires.

Une série de résultats obtenus dans différents systèmes expérimentaux est en accord avec ce modèle. L'instabilité de l'expression génétique a été décrite pour les gènes de la globine. Dans des embryons de souris transgéniques, les allèles  $\alpha$  et  $\beta$  de la globine humaine s'expriment alternativement (*m/s n°1*, vol. 12, p. 105). Pendant le développement de ces embryons, il y a une stabilisation progressive de l'expression de l'allèle  $\beta$  [21]. L'importance des interactions membranaires est bien connue mais elle est le plus souvent interprétée dans un cadre déterministe. En fait, on sait que la différenciation des cellules AC et VU chez *C. elegans* implique un événement aléatoire et des interactions cellulaires dépendantes des gènes *lin12/notch* [28, 29]. Ce phénomène de spécification latérale pourrait correspondre à un mécanisme de stabilisation. Finalement, l'effet des interactions membranaires pourrait être remplacé par des facteurs diffusibles pénétrant à l'intérieur des cellules. Cette variante du modèle permettrait une stabilisation de populations cellulaires entières. Cela a été montré chez la drosophile chez laquelle la protéine *wingless* est exprimée dans des cellules adjacentes à celles exprimant *engrailed*. Dans ces cellules, *wingless* n'est pas nécessaire à l'induction de *engrailed* mais à la stabilisation de son expression [30-32].

En culture, on obtient une population homogène quand on induit la différenciation de nombreuses lignées. Cette expérience peut sembler en contradiction avec mon modèle car il postule un mécanisme stabilisateur impliquant deux types cellulaires. Cependant, cette expérience contredit également les

modèles instructionnistes parce que la stimulation des cellules est généralement provoquée par un agent qui n'a rien d'un inducteur spécifique (le 5-bromouridine par exemple) et qui, malgré ce manque de spécificité, est capable de provoquer des différenciations vers des types cellulaires tout à fait déterminés en fonction de la lignée utilisée. En fait, on peut comprendre que certains agents puissent provoquer une déstabilisation des lignées cellulaires qui pourront alors se différencier selon des voies qui leurs sont propres, en fonction des conditions dans lesquelles elles sont placées. Les interactions entre différents types cellulaires ne sont probablement nécessaires, quel que soit le modèle (instructionniste ou sélectif), que pour les événements impliquant un choix de différenciation entre plusieurs lignées. Ainsi, il n'est pas étonnant que les cellules aboutissent à des populations homogènes différenciées si elles sont déjà engagées dans une voie de différenciation au moment de la stimulation. Il est également possible que le milieu de culture exerce un effet sélectif, dirigeant préférentiellement les cellules dans une voie unique. Il faut d'ailleurs remarquer que la cinétique de ces phénomènes est souvent décrite par des modèles probabilistes [1].

• **La phosphorylation des protéines pourrait être l'un des mécanismes de stabilisation de l'expression des gènes**

Les interactions cellulaires activent des voies de transmission de signal, aboutissant à la phosphorylation ou à la déphosphorylation de facteurs de transcription [33]. Cependant, du fait du caractère ubiquitaire des kinases, comment ce phénomène peut aboutir à une expression spécifique des gènes n'est pas clair dans le cadre d'un modèle déterministe. Au contraire, une explication simple peut être donnée dans le cadre du modèle d'expression aléatoire. Dans ce modèle, les constantes de dissociation entre les régulateurs et l'ADN déterminent les probabilités de déplacement. On sait que la phosphorylation ou la déphosphorylation des facteurs de transcription peut changer complètement leur affinité pour

l'ADN [34-39]. Par cet effet, les probabilités de déplacement des régulateurs transcriptionnels seront également modifiées, pouvant conduire à la stabilisation de l'expression génétique. Si la constante de dissociation entre un régulateur et l'ADN devient très faible, ce régulateur ne se déplacera plus sur l'ADN et l'expression des gènes sera figée. Un tel mécanisme pourrait expliquer l'omniprésence des protéine-kinases et des phosphatases dans les mécanismes de régulation de l'expression génétique. Cela pourrait permettre, de plus, de rompre le cloisonnement actuel entre métabolisme et régulation des gènes. En effet, outre les interactions du type de celles décrites dans la *figure 3*, le niveau d'activité métabolique d'une cellule pourrait modifier l'état d'activation des kinases ou des phosphatases, comme c'est le cas pour l'induction génique par le glucose dans le foie [40], et donc sur la stabilisation ou la déstabilisation de l'expression génétique. Une telle régulation métabolique pourrait correspondre à l'origine de la différenciation cellulaire.

## Discussion

• **Modèle de hasard-sélection et reproductibilité de la différenciation cellulaire**

Le mécanisme du modèle de hasard-sélection s'exerce à deux niveaux. Comme tout processus faisant appel aux lois de la physico-chimie, il est probabiliste au niveau des événements moléculaires individuels, engendrant ainsi la diversité de types nécessaire à la différenciation. Cependant, au niveau des populations cellulaires entières, s'exerce un mécanisme de sélection-stabilisation qui n'a rien de probabiliste et qui optimise le résultat du processus probabiliste d'activation des gènes. On suppose souvent que des événements aléatoires ne peuvent pas être à la base d'un phénomène aussi reproductible que la différenciation cellulaire. En fait, un phénomène probabiliste est reproductible. Dans une population il se répète avec une variance, et est centré sur une moyenne. Si la variance est faible, il peut sembler parfaitement identique à chaque fois qu'on le

répète. C'est, par exemple, le cas de la décroissance de la radioactivité d'un produit marqué, qui sera, à notre niveau d'observation, la même à chaque fois qu'on la mesurera, alors que les événements moléculaires sous-jacents sont aléatoires. Il pourrait en être de même pour la différenciation cellulaire. D'autant plus que, dans le modèle de hasard-sélection, ainsi que nous l'avons vu, les événements aléatoires sont triés par les interactions cellulaires qui dirigent l'embryon vers le stade adulte.

On pense, habituellement, qu'une information est véhiculée par des gradients d'agents morphogènes et que les cellules se différencient, en fonction de leur position dans le gradient, selon la concentration de morphogène. Cependant, on ne sait pas comment la cellule peut répondre de manière multiple et spécifique à des petites variations de concentration de morphogène. En fait, on est confronté aux deux possibilités : soit l'interprétation instructionniste, soit l'interprétation par le modèle de hasard-sélection. Si le mode d'action du morphogène consiste en une stabilisation, comme c'est le cas de la protéine *wingless*, le modèle de hasard-sélection est tout à fait compatible avec l'existence de gradients.

#### • Modèle de hasard-sélection et réductionnisme

Les modèles déterministes sont liés à un point de vue réductionniste parce que l'ordre au niveau pluricellulaire découle directement (et pourrait être déduit) de l'ordre au niveau moléculaire. Dans ces modèles, les organismes et les cellules se forment par agencement des macromolécules à partir de phénomènes de reconnaissance stéréospécifique qui ne laissent aucune place au hasard [41]. Le modèle de hasard-sélection ne repose pas sur le concept de reconnaissance stéréospécifique. Au contraire, dans ce modèle, les interactions moléculaires sont soumises à des variations aléatoires causées par les lois de la diffusion, et l'ordre final est produit par des contraintes provenant des interactions qui se produisent au niveau cellulaire. Dans les modèles déterministes, les cellules sont stables et nécessitent une cause externe pour se

différencier (*figure 1A*); dans le modèle de hasard-sélection, les cellules sont spontanément instables et n'ont pas besoin d'une cause externe pour se différencier mais, au contraire, pour se stabiliser (*figure 1B*). Le comportement des cellules en culture soutient le modèle de hasard-sélection parce qu'elles peuvent se transformer ou se différencier spontanément, dans un milieu constant [42, 43]. Il est d'ailleurs très souvent nécessaire de cloner à nouveau régulièrement une lignée pour éviter sa dérive. Dans le modèle de hasard-sélection, les interactions cellulaires *in vivo* canalisent cette instabilité spontanée pour produire un organisme viable.

#### • Différenciation et évolution

Les comparaisons de séquences d'ADN montrent que les différences entre les protéines de deux espèces sont très faibles et suggèrent que l'évolution d'une espèce est plutôt due à l'évolution de ce qu'on appelle le « programme » de développement [44]. Cette notion est implicite dans d'autres théories [45, 46]. Dans le cadre du modèle de hasard-sélection, l'évolution de la différenciation peut être facilement comprise par le changement de l'ordre des gènes dans les chromosomes. En outre, toutes les séquences d'ADN situées entre les gènes, y compris les séquences non codantes, jouent un rôle essentiel en déterminant les probabilités de déplacement des régulateurs, donc le « programme » lui-même. Cela pourrait expliquer en partie la plasticité et l'excès d'ADN des génomes eucaryotes. Les éléments transposables, en changeant de position changent le « programme » de différenciation, procurant ainsi un potentiel adaptatif aux organismes.

La vision classique de la régulation de l'expression des gènes suppose l'existence de régulateurs spécifiques qui correspondraient, soit à des molécules individuelles, soit à des combinaisons de molécules. Or, quel que soit le cas, cela pose le problème de leur origine évolutive. Des régulateurs très spécifiques, comme le répresseur de l'opéron lactose, n'ont pu apparaître d'emblée au cours de l'évolution. Il a très certainement fallu l'amélioration progressive de

régulateurs non spécifiques. Les organismes primitifs ont dû subsister en l'absence de régulateurs spécifiques. Du fait que le modèle aléatoire de régulation des gènes fonctionne avec des régulateurs non spécifiques, il pourrait correspondre à un mode de régulation très ancien. Dans cette perspective l'origine de la différenciation cellulaire serait la conséquence de ce manque de spécificité des systèmes biologiques [47] ■

#### Remerciements

Je remercie Marc Alizon pour la relecture de ce manuscrit, Pierre Sonigo pour son soutien et de nombreuses discussions, Donny Strosberg pour avoir créé un environnement stimulant.

#### RÉFÉRENCES

1. Levenson R, Housman D. Commitment: how do cells make the decision to differentiate? *Cell* 1981; 25: 5-6.
2. Till JE, McCulloch EA, Siminovitch L. A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. *Proc Natl Acad Sci USA* 1964; 61: 29-36.
3. Bennett DC. Differentiation in mouse melanoma cells: initial reversibility and an on-off stochastic model. *Cell* 1983; 34: 445-53.
4. Baroffio A, Blot M. Statistical evidence for a random commitment of pluripotent cephalic neural crest cells. *J Cell Sci* 1992; 103: 581-7.
5. Davis CB, Killeen N, Crooks ME, Raulet D, Littman DR. Evidence for a stochastic mechanism in the differentiation of mature subsets of T lymphocytes. *Cell* 1993; 73: 207-8.
6. Le Douarin NM, Dupin E. Cell lineage analysis in neural crest ontogeny. *J Neurobiol* 1993; 24: 146-61.
7. Michaelson J. Cellular selection in the genesis of multicellular organization. *Lab Invest* 1993; 69: 136-51.
8. Paulus U, Loeffler M, Zeidler J, Owen G, Potten CS. The differentiation and lineage development of goblet cells in the murine small intestinal crypt: experimental and modelling studies. *J Cell Sci* 1993; 106: 473-83.
9. Chan SH, Benoist C, Mathis D. In favor of the selective model of positive selection. *Semin Immunol* 1994; 6: 241-8.
10. Lin Z, Lu MH, Scultheiss T, Choi J, Holtzer S, Dilullo C, Fischman DA, Holtzer H. Sequential appearance of muscle-specific proteins in myoblasts as a function of time after cell division: evidence for a conserved myoblast differentiation program in skeletal muscle. *Cell Motil Cytoskeleton* 1994; 29: 1-19.

## RÉFÉRENCES

11. Kupiec JJ. A chance-selection model for cell differentiation. *Cell Death and Differentiation*, 1996; 3: 385-90.
12. Kupiec JJ. A darwinian theory for the origin of cellular differentiation. *Mol Gen Genet* 1997; 255: 201-8.
13. Berg OG, Winter RB, von Hippel PH. Diffusion-driven mechanisms of protein translocation on nucleic acids. *Biochemistry* 1981; 20: 6929-77.
14. Berg OG, von Hippel PH. Diffusion-controlled macromolecular interactions. *Annu Rev Biophys Chem* 1985; 14: 131-60.
15. Park CS, Wu FV, Wu CW. Molecular mechanisms of promoter selection in gene transcription. *J Biol Chem* 1982; 257: 6944-56.
16. Hannon R, Richards EG, Gould HJ. Facilitated diffusion of a DNA binding protein on chromatin. *EMBO J* 1986; 5: 3313-9.
17. Townes TM, Behringer RR. Human globin locus activation region (LAR): role in temporal control. *Trends Genet* 1990; 6: 219-23.
18. Hanscombe O, Whyatt D, Fraser P, Yanoutsos N, Greaves D, Dillon N, Grosveld F. Importance of globin gene order for correct developmental expression. *Genes Dev* 1991; 5: 1387-94.
19. Boyes J, Felsenfeld G. Tissue-specific factors additively increase the probability of the all-or-none formation of a hypersensitive site. *EMBO J* 1996; 15: 2496-507.
20. Nan X, Campoy FJ, Bird A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* 1997; 88: 471-81.
21. Wijgerde M, Grosveld F, Fraser P. Transcription complex stability and chromatin dynamics *in vivo*. *Nature* 1995; 377: 209-13.
22. Perz-Howard GM, Weil PA, Beechem JM. Yeast TATA binding protein interaction with DNA: fluorescence determination of oligomeric state, equilibrium binding, on-rate, and dissociation kinetics. *Biochemistry* 1995; 34: 8005-17.
23. Lucchini R, Sogo JM. Replication of transcriptionally active chromatin. *Nature* 1995; 374: 276-80.
24. Wolffe AP. Inheritance of chromatin states. *Dev Genet* 1994; 15: 463-70.
25. Krumlauf R. Hox genes in vertebrate development. *Cell* 1994; 78: 191-201.
26. van der Hoeven F, Zakany J, Duboule D. Gene transposition in the HoxD complex reveal a hierarchy of regulatory controls. *Cell* 1996; 85: 1025-35.
27. Dernburg AF, Broman KW, Fung JC, Marshall WF, Phillips J, Agard DA, Sedat JW. Perturbation of nuclear architecture by long-distance chromosome interactions. *Cell* 1996; 85: 745-59.
28. Félix M. Un ver, 959 cellules et 13000 gènes. *Med Sci* 1997; 13: 156-65.
29. Schweisguth F, Israel A. Signalisation intercellulaire par le récepteur Notch: conservation de la drosophile aux mammifères. *Med Sci* 1996; 155-63.
30. Heemskerk J, DiNardo S, Kostriken R, O'Farrell PH. Multiple modes of engrailed regulation in the progression towards cell fate determination. *Nature* 1991; 352: 404-10.
31. DiNardo S, Sher E, Heemskerk-Jongens J, Kassis J, O'Farrell P. Two-tiered regulation of spatially patterned engrailed gene expression during drosophila embryogenesis. *Nature* 1988; 332: 604-9.
32. Bellaïche Y, Perrimon N. La voie de signalisation Wingless chez la drosophile. *Med Sci* 1997; 13: 166-74.
33. Hill CS, Treisman R. Transcriptional regulation by external signals: mechanisms and specificity. *Cell* 1995; 80: 199-211.
34. Bourbon HM, Martin-Blanco E, Rosen D, Kornberg TB. Phosphorylation of the drosophila engrailed protein at a site outside its homeodomain enhances DNA binding. *J Biol Chem* 1995; 270: 11130-9.
35. Lefevre P, Gaub MP, Tahayato A, Rochette-Egly C, Formstecher P. Protein phosphatases 1 and 2A regulate the transcriptional and DNA binding activities of retinoic acid receptors. *J Biol Chem* 1995; 270: 10806-16.
36. Li CC, Dai RM, Chen E, Longo DL. Phosphorylation of NF-KB1-p50 is involved in NF-kappa B activation and stable DNA binding. *J Biol Chem* 1994; 269: 30089-92.
37. Rabault B, Ghysdael J. Calcium-induced phosphorylation of ETS1 inhibits its specific DNA binding activity. *J Biol Chem* 1994; 269: 28143-51.
38. Takenaka I, Morin F, Seizinger BR, Kley N. Regulation of the sequence-specific DNA binding function of p53 by protein kinase C and phosphatases. *J Biol Chem* 1995; 270: 5405-11.
39. Xu M, Sheppard KA, Peng CY, Yee AS, Piwnicka-Worms H. Cyclin A/CDK2 binds directly to E2F-1 and inhibits the DNA-binding activity of E2F-1/DP-1 by phosphorylation. *Mol Cell Biol* 1994; 14: 8420-31.
40. Doiron B. Le glucose induit l'expression génique par l'intermédiaire de la voie des pentoses-phosphates. *Med Sci* 1996; 12: 503-8.
41. Monod J. *Le hasard et la nécessité*. Paris: Éditions du Seuil, collection « Point Science », 80 et 118.
42. Chow M, Yao A, Rubin H. Cellular epigenetics: topochronology of progressive « spontaneous » transformation of cells under growth constraint. *Proc Natl Acad Sci USA* 1994; 91: 599-603.
43. Bohme K, Winterhalter KH, Bruckner P. Terminal differentiation of chondrocytes in culture is a spontaneous process and is arrested by transforming growth factor-beta 2 and basic fibroblast growth factor in synergy. *Exp Cell Res* 1995; 216: 191-8.
44. King MC, Wilson AC. Evolution at two levels in humans and chimpanzees. *Science* 1975; 188: 107-16.
45. Gould SJ. *Ontogeny and Phylogeny*. Cambridge, MA: The belknap press of harvard university press. 1977.
46. Darwin C. Embryology in: *The origin of species*. 1859; Chapter 13.
47. Peccoud J, Viville S. Aspects aléatoires de la dynamique de la différenciation cellulaire. *Med Sci* 1994; 10: 877-83.

### Jean-Jacques Kupiec

ICGM-CNRS. UPR415, 22, rue Méchain  
75014 Paris, France.

## Summary

### A chance-selection model could explain the differentiation of several cell lines

In this theory, cell differentiation is a two-step mechanism. In the first step, gene expression is unstable. It occurs stochastically and produces different cell types. In the second step gene expression is stabilized by means of cellular interactions. However, this stabilization cannot occur until the combination of cell phenotypes corresponding to the developmental stage is expressed. This selection mechanism prevents disorganizing consequences of stochasticity in gene expression and directs the embryo towards the adult stage. Instability and stochasticity in gene expression are caused by random displacement of regulators along DNA, whereas phosphorylation and/or dephosphorylation of transcriptional regulators triggered by signal transduction between cells are responsible for the stabilization of stochastic gene expression. This model is based on well documented phenomena: random diffusion of DNA binding molecules along DNA, phosphorylation or dephosphorylation of transcription factors by protein kinases or phosphatases and control of DNA binding of transcription factors through this latter process. The different explanatory powers of deterministic and stochastic models are discussed.

## TIRÉS À PART

J.J. Kupiec.