

## L'ontophylogénèse – un nouveau paradigme biologique

**Baptiste Morizot** : Bien on va commencer. D'abord bonjour à tous, merci d'être venu si nombreux. Je parle au nom du laboratoire junior EHVI dont on est plusieurs membres ici. On est très honoré d'accueillir Jean-Jacques Kupiec dans le cadre du séminaire qu'on organise. Je vais me passer de faire une présentation ; normalement, si vous êtes là, c'est que vous avez reçu le mail sur lequel on vous proposait une bibliographie et une courte présentation. D'ailleurs Jean-Jacques Kupiec sera toujours plus intéressant que ce que je pourrai en dire ici. Je voudrais simplement noter qu'on est ravi de l'accueillir dans un espace au nom si prestigieux : l'amphithéâtre Schrödinger sur le site Jacques Monod. C'est d'autant plus intéressant que la théorie de l'ontophylogénèse se fonde sur une critique d'abord de la thèse de Schrödinger dans *Qu'est-ce que la vie ?*, et dans un second temps sur la théorie de la stéréospécificité défendue en partie par Jacques Monod. C'est donc un hommage à l'esprit critique de la science et à sa capacité à produire des conflits sains et vigoureux que l'on se trouve ici. Je voudrais enfin conclure sur le fait que la séance sera enregistrée et sera disponible au plus tôt sur le site internet du laboratoire junior EHVI – enquête sur l'homme vivant – référencé sur Google. On peut y trouver des informations sur l'actualité du laboratoire, et cette conférence. Jean-Jacques, merci encore.

**Julie Henry** : Je vais faire passer une feuille pour ceux qui veulent nous donner leurs adresses email afin qu'on vous rappelle si besoin l'adresse du site internet, mais aussi pour ceux qui voudraient être inscrits à la liste de diffusion du laboratoire junior, si cela vous intéresse d'être prévenus des différents événements du laboratoire et de tout ce qui tourne autour de sa problématique – enquête sur l'homme vivant.

**Jean-Jacques Kupiec** : Merci pour l'invitation. Pendant le séminaire, je vais essayer de développer deux idées principales : la première, c'est que le vivant est intrinsèquement probabiliste, et que de ce fait la notion de programme génétique est caduque. Vous savez peut-être, ou devez savoir, que depuis une dizaine d'années il est maintenant démontré que l'expression des gènes est un phénomène stochastique, à l'issue d'une série de travaux qui se sont accumulés – ça a commencé un peu avant, à partir des années 90. Mais l'expression stochastique des gènes y est interprétée comme du « bruit ». On appelle ça, dans la littérature anglo-saxonne, « noise of gene expression ». Donc on considère que l'expression des gènes intègre une dimension stochastique, mais ça ne remet pas en cause la théorie du programme génétique. On considère que c'est en quelque sorte une marge d'ajustement, une marge de fluctuation aléatoire dans l'expression du programme génétique. Donc ce que je vais essayer de montrer, c'est qu'en fait cette dimension stochastique implique un réaménagement de très grande ampleur en biologie, et que ce réaménagement mène obligatoirement à la dissolution du concept même de programme génétique.

La deuxième idée que je vais essayer de développer, c'est cette idée que l'ontogenèse et la phylogénèse ne forment qu'un seul processus. Vous savez également qu'en biologie on considère qu'il y a deux phénomènes qu'on considère comme totalement distincts : d'un côté l'évolution des espèces (phylogénèse) et d'un autre côté la genèse des organismes individuels (l'ontogenèse ou l'embryogenèse). Et donc on considère qu'il s'agit de deux phénomènes distincts qui sont donc expliqués par deux théories distinctes et clairement définies : d'un côté la théorie de la sélection naturelle (pour la période moderne) pour l'évolution des espèces, et puis l'embryogenèse qui est expliquée comme étant l'expression, le développement, l'actualisation du programme génétique par

la biologie moléculaire. En fait, cette manière de séparer pose toute une série de problèmes récurrents, notamment en ce qui concerne l'évolution des espèces. L'an dernier, c'était l'année Darwin, on a entendu cette expression tout au long de l'année, mais en réalité c'est une expression qui est un abus de langage car à proprement parler, les espèces n'évoluent pas. Pourquoi est-ce que les espèces n'évoluent pas ? Parce que l'espèce est une abstraction. Dans la nature, il n'y a pas une espèce « cheval » ou une espèce « lapin » ou un prototype d'espèce « cheval » ou un prototype de l'espèce « lapin » qui évolue. Ce qui existe réellement, ce sont des individus, des populations d'organismes individuels – et des populations d'organismes individuels qui se multiplient eux mêmes par la génération : c'est-à-dire que ce qui existe réellement, ce sont des lignées généalogiques qui voient la multiplication ininterrompue d'individus. Et c'est à partir de ces populations qu'on reconnaît et qu'on abstrait des espèces. Donc ce qui évolue réellement, en fait, ce sont ces individus, à travers la génération. Ce que je viens de dire, c'est un résumé extrêmement synthétique de l'analyse qu'a faite Darwin dans *L'Origine des espèces*. Si vous l'avez lu ou si vous le lisez vous verrez qu'il procède à une déconstruction du concept réaliste de l'espèce, pour adopter un point de vue nominaliste, et il finit par proposer une définition de l'espèce qui est une définition généalogique : pour Darwin, textuellement, une espèce est un ensemble d'individus qui ont un ancêtre commun. On passe donc d'une définition typologique de l'espèce à une définition généalogique, et ce qui est implicite dans cette manière de voir, c'est que ce qui est réel, ce n'est pas l'espèce (qui est une abstraction) mais c'est la lignée généalogique. Cette lignée généalogique qui a commencé à partir du moment où le vivant est apparu.

Alors maintenant, en ce qui concerne l'ontogénèse. On est persuadé qu'il y a un processus, qui s'appelle l'embryogénèse, qui a pour but de créer des organismes individuels. Mais cette manière de voir, elle est tout aussi arbitraire. Il est évident que nous avons tendance à faire une projection anthropocentrique, et on est persuadé que ce processus a pour but de nous créer. Mais si l'on prend un peu de distance, ici aussi, ce à quoi on a affaire, c'est une succession d'êtres qui se reproduisent à l'infini – une lignée généalogique qui voit se succéder des formes, des phases cycliques (des phases unicellulaires (cellules germinales) qui donnent des êtres multicellulaires qui vont redonner des cellules germinales, etc.). Et nous, évidemment, on fige un instant, qu'on appelle l'individu, sur un processus continu et ininterrompu. On pourrait très bien inverser le point de vue, et dire que ce qui est réel, après tout, pourquoi ce ne serait pas l'ovogénèse ? Et d'une certaine manière, c'est le point de vue de la génétique, dans ses interprétations extrêmes. Pour Weissman, ce qui était réel, c'était la lignée germinale. Et pour quelqu'un comme Dawkins – on peut considérer que c'est l'ultra-généticien par excellence – ce qui est réel, ce sont les gènes, et nous ne sommes que des avatars destinés à propager les gènes. En vérité, les deux points de vue sont aussi vrais et aussi faux l'un que l'autre – ce ne sont que des points de vue. La seule chose dont l'on peut dire que c'est ce qui est réel (si l'on essaie de s'abstraire de toute projection), c'est qu'on a affaire à un phénomène cyclique ininterrompu, à nouveau, qui voit la succession, la reproduction, de ce qu'on appelle des organismes individuels, et qui voit se succéder des phases cellulaires, multicellulaires, cellulaires, multicellulaires. Donc, si on procède à cette analyse, on s'aperçoit que dans les deux cas – dans l'ontogénèse comme dans la phylogénèse – on bute en quelque sorte sur une réalité unique, qui est la lignée généalogique. Donc ce point de vue sur l'ontogénèse, à nouveau, à la même époque que Darwin, un biologiste l'a exprimée très clairement et très explicitement, c'est Claude Bernard dans les *Leçons sur les phénomènes de la vie*.

Donc ce concept d'ontophylogénèse, que j'ai introduit, cherche à capter cette réalité unique sous-jacente à ce qu'on sépare habituellement et qu'on appelle l'ontogénèse et la phylogénèse. Et si on va un peu plus loin et qu'on reconnaît qu'on a affaire à un phénomène unique, a priori il n'y a

aucune raison pour qu'on ne puisse pas avoir une théorie unique, qu'il n'y ait pas un mécanisme unique qui permette d'expliquer ce phénomène unique. A priori, on doit pouvoir procéder à une réduction et ne pas avoir deux théories, mais une seule, qui permette d'expliquer ce phénomène unique. Et c'est ce que j'appelle le « darwinisme cellulaire » et que j'essaierai d'introduire également. Et dans la troisième partie de l'exposé, on verra comment – à partir de ce cadre théorique général, cette réinterprétation générale des faits biologiques – on peut revenir vers des phénomènes plus concrets, notamment la différenciation cellulaire, déboucher sur des interprétations plus précises de ces faits biologiques, et également sur un programme de recherche expérimental.

## Première partie

Pour commencer et pour comprendre donc comment se posent ces questions aujourd'hui, on va revenir sur les débuts de la biologie moléculaire. Comme Baptiste l'a indiqué, on se trouve dans l'amphithéâtre Schrödinger. Donc, il se trouve que Schrödinger a joué un rôle important en biologie moléculaire. C'est un physicien, et il est surtout connu pour ses travaux de physique quantique. En fait la biologie moléculaire – vous devez le savoir également – au départ a été créée par des physiciens, notamment un physicien qui s'appelle Maxime Delbrück qui avait créé le « groupe du phage » dans les années 30 – 40. Et comme beaucoup de physiciens, à la même époque, Schrödinger s'intéressait beaucoup à la biologie, mais d'un point de vue théorique, et il a écrit un livre qui s'appelle *Qu'est-ce que la vie ? (What is life ?)* publié en 1944, dans lequel il a popularisé tous les concepts de la biologie moléculaire naissante, notamment beaucoup des concepts et des travaux qui avaient été introduits par Max Delbrück. Et donc dans la première partie de ce livre, Schrödinger pose la question : d'où vient l'ordre dans les systèmes naturels ? Et par là même, il pose la question du rôle de l'aléatoire dans la genèse de l'ordre dans les systèmes naturels, et il est amené à tracer une séparation très nette entre la physique et la biologie : donc, selon Schrödinger, en physique opère ce qu'il appelle un principe « d'ordre à partir du désordre ». Donc ce à quoi il fait allusion, c'est la physique statistique. En physique, au niveau microscopique, le comportement des particules (atomes, molécules, particules subatomiques) est stochastique, il est aléatoire, soumis à l'agitation thermique. Mais si vous considérez le même système au niveau macroscopique, du fait du nombre immense de particules qui est impliqué dans tout système physique, opère la loi des grands nombres, et donc il y a une réduction de la variance globale du système, et donc on observe en quelque sorte la moyenne des comportements des particules que l'on peut parfaitement décrire par des lois déterministes, parce que la variance du système devenant négligeable, on peut tout à fait le décrire ainsi. L'exemple canonique est la diffusion : vous mettez une goutte de vinaigre dans un verre d'eau, chaque molécule de vinaigre va bouger selon une marche aléatoire et vous ne pouvez la décrire que de manière statistique ; mais si vous considérez la goutte dans son ensemble et que vous regardez la cinétique de dilution du vinaigre dans le verre d'eau, vous pouvez la décrire par les lois de Fick, qui sont des lois déterministes. Donc en physique, principe d'ordre à partir du désordre, pour Schrödinger. Et ça veut dire également que l'ordre est subjectif, finalement, puisqu'il dépend du niveau d'observation et il dépend de l'observateur : selon que vous vous placez à un niveau microscopique ou macroscopique, vous avez affaire à un phénomène désordonné ou organisé, probabiliste ou déterministe. Ce n'est donc qu'une approximation, finalement.

Alors donc la question que pose Schrödinger au début de *Qu'est-ce que la vie ?* : est-ce que le même principe d'ordre à partir du désordre peut opérer en biologie ? Donc là je vais aller un petit peu vite : sa réponse, et à sa suite la réponse de tous les biologistes moléculaires, est : non. Les arguments qu'il développe sont un petit peu datés aujourd'hui, mais en gros ils consistent à dire qu'il

n'y a pas assez d'atomes dans un gène pour que la loi des grands nombres opère, puisqu'il suppose que le gène gouverne les phénomènes biologiques ; et donc que si la biologie était soumise au même principe, du fait que le nombre de particules est trop petit, la loi des grands nombres ne pourrait pas opérer, donc les phénomènes biologiques auraient une très grande variabilité – ce qui est contradictoire avec la très grande précision et la reproductibilité qui les caractérisent. Donc doit opérer ce qu'il appelle un principe « d'ordre à partir de l'ordre » : c'est-à-dire que contrairement à la physique, au niveau microscopique, les molécules biologiques ne sont pas soumises uniquement à l'agitation thermique, au hasard brownien, mais elles sont contraintes par un principe qui vient les ordonner directement au niveau microscopique et les gouverner pour qu'elles se comportent de manière adéquate pour construire l'organisme. À l'époque, on avait pas encore découvert la structure de l'ADN, il y avait encore des débats pour savoir ce qu'il y avait dans les chromosomes. Donc ce à quoi Schrödinger fait allusion, avec ce principe « d'ordre à partir de l'ordre », c'est l'hypothèse émise à la même époque par Maxime Delbrück qu'il y avait des chromatines, des protéines, de l'ADN – du matériau génétique – et c'était donc l'hypothèse du « cristal aperiodique » : il y aurait un code dans les chromosomes, et ce principe de l'ordre à partir de l'ordre est ce que l'on appelle aujourd'hui « l'information génétique ». Et c'est ce qui différencie les systèmes biologiques des systèmes physiques : les molécules des systèmes biologiques ne sont pas livrées aux lois de la physique statistique pure, mais elles sont gouvernées par l'information génétique qui vient directement leur donner l'information, les instruire pour leur dire ce qu'elles doivent faire pour construire l'organisme et se comporter de manière adéquate.

Donc la question qui s'est posée et qui se pose aussitôt qu'on a admis un principe tel que celui de l'ordre par l'ordre c'est de savoir comment l'information codée dans les chromosomes (qui est quelque chose de virtuel) va pouvoir être transformé en une structure physique réelle au cours d'un processus physique réel qui est l'ontogénèse. Autrement dit : comment la représentation virtuelle de l'organisme qui se trouve dans les chromosomes peut être transformé en un processus physique réel et à trois dimensions ? Donc, Schrödinger avait un peu abordé cette question, mais je passe un peu. La réponse qui a été adoptée massivement par les biologistes moléculaires, notamment à partir des années 1960, c'est le principe de l'auto-assemblage stéréospécifique. Ce principe est le suivant : les protéines ont une structure tridimensionnelle, et en fonction de leur structure tridimensionnelle, les protéines ne se reconnaissent pas au hasard – en fait elles ne se reconnaissent que de manière univoque, puisque chaque protéine ne peut reconnaître qu'un seul partenaire (ou un nombre très très petit de partenaires) de telle sorte que toute possibilité combinatoire dans les interactions soit exclue, et que soit donc exclue par là même toute possibilité de hasard ou d'effet aléatoire. En fait, on est dans le système du puzzle, en quelque sorte – c'est ce que j'ai figuré sur le schéma – pour qu'on comprenne bien le principe : vous mettez un ensemble de pièces d'un puzzle, elles sont contraintes par leur structure à interagir d'une certaine manière – ici, ces quatre pièces ne pourront interagir que de manière à former la structure en bas à droite. Et tout cela dans le cas des protéines : comme la structure tridimensionnelle des protéines dépend de leur structure linéaire d'acides aminés qui est la traduction directe de la séquence en nucléotides qui est contenue dans les gènes, ce mécanisme d'auto-assemblage stéréospécifique est bien la traduction physique de l'information génétique contenue dans les gènes. Ce principe a été proposé par Caspar et Klug au début des années soixante dans des papiers séminaux historiques, d'abord pour expliquer dans un premier temps la genèse des particules virales. Ensuite, ça a été très vite étendue à toutes les morphogénèses de structures cellulaires. Et à la même époque, le même principe d'interaction stéréospécifique a été utilisé pour expliquer les phénomènes de régulations, notamment par Monod et Jacob – à la même époque. Le modèle de Monod-Jacob date même de 1961, il est antérieur d'une année. Ils ont expliqué la régulation du gène du lactose dans la bactérie *Escherichia coli* par

l'existence d'un régulateur. La régulation des gènes a été vue comme liée à la présence de molécules qui interagissent avec les régions promotrices de manière stéréospécifique : le répresseur de l'opéron lactose venant se coller sur les gènes, en amont de la région promotrice des gènes du lactose, pour le réprimer. Ces événements ont été vus comme des interrupteurs cybernétique « on/off » – on dit qu'un gène est activé ou réprimé en fonction d'une interaction spécifique – et à partir de là, on a conçu par extrapolation toute la régulation des génomes comme des cascades de régulations, des successions de signaux « on/off », des gènes pouvant synthétiser des molécules qui vont aller activer d'autres gènes ou réprimer d'autres gènes ; et on a conçu toute la régulation du génome comme un circuit cybernétique fait d'une succession de switch « on/off ». Et cette cascade, c'est ce que l'on appelle le « programme génétique ». Donc le concept de programme génétique repose également sur cette notion d'interaction stéréospécifique. Pareil pour la transduction du signal : donc, pour les non-biologistes, les cellules échangent des signaux à travers la membrane, ça va être véhiculé jusqu'au noyau, et ce que l'on appelle la transduction du signal a été également conçu comme des cascades d'interactions, de réactions entre molécules, sur le mode de la stéréospécificité.

Donc cette vision, cette théorie a une conséquence : c'est que sous-jacentes à tous les processus biologiques, il y a des cascades d'interactions stéréospécifiques – et c'est ce que l'on appelle les « réseaux de protéines ». Et donc on doit être capables de dresser des cartes sous-jacentes : si on dresse la carte et que l'on reconstitue les chaînes d'interactions sous-jacentes aux phénomènes biologiques, on doit pouvoir décortiquer au niveau moléculaire et donc avoir l'explication causale des phénomènes qu'on étudie. Donc rapidement, cette même notion de stéréospécificité est à la base de ce que l'on appelle le « réductionnisme génétique ». C'est l'idée que les gènes codent pour des molécules – c'est la construction du phénotype à partir du génotype – les molécules sont stéréospécifiques, ce qui va permettre de construire des cellules, ces cellules vont échanger des signaux qui sont eux-mêmes spécifiques et vont permettre de transmettre de l'information ; cela permet donc aux cellules de se différencier et de s'organiser en tissus qui vont donner les organes, et ainsi on peut passer du génotype au phénotype par niveaux d'organisation – et chaque niveau est produit par les interactions stéréospécifiques du niveau inférieur. Et dans cette vision, l'organisme qui existe est bien le déploiement de l'information au départ qui est contenue dans les gènes. Là je paraphrase des passages de Monod-Jacob, *Le Hasard et la nécessité*, etc. et d'autres.

Cette théorie permet de comprendre en même temps le programme de recherches qui a été mis en place par la biologie moléculaire, puisqu'à partir du moment où vous avez une théorie de ce type là, il est logique que pour tout phénomène qu'on veut étudier – normal ou pathologique – on cherche à isoler un gène qui est en relation avec ce phénomène, ensuite isoler la protéine, et grâce aux techniques qu'on a, on essaie de trouver les partenaires moléculaires de cette protéine, pour reconstituer ainsi les chaînes d'interactions, et trouver les cartes d'interactions protéiques, ou les cartes d'interactions entre gènes, sous-jacentes aux phénomènes. Et c'est effectivement le programme de recherches qui est mis en œuvre massivement par la biologie moléculaire depuis quarante ans. Comme vous le savez, on a isolé un très très grand nombre de gènes, ce qui a permis d'éclairer tous les phénomènes, et qui a constitué des progrès énormes. En même temps, il y a quelque chose qui est apparu petit à petit, et maintenant de manière massive et claire, c'est que contrairement à ce qui a été prédit, la prédiction de stéréospécificité n'a pas été vérifiée. Au début, c'était par des études au cas par cas, on isolait ce qu'on pensait être dans un premier temps le facteur de transcription spécifique d'un gène spécifiant une lignée au cours de la différenciation, etc. ou une protéine impliquée dans une maladie (le cancer, par exemple).

Mais petit à petit, on s'est aperçu que ces mêmes molécules étaient capables d'interagir avec d'autres partenaires, avec un nombre de plus en plus grand, et donc maintenant on a cartographié des réseaux de gènes entiers correspondant à des organismes entiers (des bactéries – E. coli – la levure, la souris, la drosophile, l'homme) et donc on a les cartes totales des réseaux de toutes les interactions d'organismes entiers, et maintenant il apparaît clairement que contrairement à ce qui était prédit, la connectivité de ces réseaux est très grande. J'ai représenté ici l'image classique d'un réseau de protéines qu'on trouve dans la littérature. Ces réseaux ont une structure caractéristique : la connectivité moyenne est de l'ordre de 7 à 8 – ce qui est beaucoup – et il y a une région centrale dans ces réseaux qui est très très dense, avec plus de 10% des protéines qui peuvent interagir avec au moins 100 partenaires. Donc à la périphérie, la connectivité est plus faible, évidemment, mais avec une connectivité de l'ordre de 7 à 8, cela peut générer des possibilités combinatoires immenses. Cela a une conséquence, et notamment l'existence de cette région très dense en connexions, avec ces protéines qu'on appelle des « hub-protéines », des protéines centrales qui peuvent interagir avec plus de 100 partenaires : c'est que toutes les voies de régulation sont inter-connectées entre elles. Cela veut dire que si vous vous placez à la périphérie du réseau, où que vous entriez dans le réseau, vous allez arriver à des noeuds où il y a plusieurs options, vous allez en prendre une, et potentiellement vous pouvez sortir partout ailleurs dans le réseau. Voilà qui pose un problème, évidemment : c'est une négation directe du principe de l'ordre par l'ordre. C'est-à-dire que, quand une cellule reçoit un signal, elle a une réponse qui consiste à activer un nombre limité de voies de régulations – elle n'active pas toutes les voies de régulation possibles. Cela a donc une conséquence : contrairement à ce que l'on croyait initialement, la structure du réseau lui-même et sa cartographie ne peuvent pas expliquer le fonctionnement de la cellule. Il doit obligatoirement y avoir un ou des mécanismes supplémentaires qui viennent se greffer sur ce réseau et qui vont limiter à un moment donné l'utilisation du réseau – dans la mesure où, quand une cellule se différencie ou reçoit un signal, elle ne met pas en œuvre tout son réseau mais n'utilise qu'un sous-ensemble de son réseau. Donc, pour bien comprendre ce dont il est question, il s'agit d'un exemple local, plus restreint, maintenant qu'on a bien décortiqué certaines voies de transduction du signal. C'est ici un schéma pris dans la littérature, tiré d'un article proposé par Schwartz et Madani et publié dans la revue *Annual Review of Genetics* – une revue sur cette question de la spécificité des protéines – il s'agit donc d'une levure : on a trois signaux A,B,C qui aboutissent à trois réponses A', B', C'. Quand la cellule est activée par une phéromone sexuelle, ça déclenche la recombinaison génétique – c'est la réponse A', et non B' ou C'. Le signal B déclenche la synthèse de solute pour répondre à une trop grande pression osmotique, etc. Donc on a décortiqué ces trois voies, et comme c'est très souvent le cas, il y a un tronc commun entre ces trois voies : elles utilisent des éléments communs, elles convergent toutes vers un tronc commun constitué de cascades de réactions, ce sont des kinases en général. Et ensuite, on sait que ce tronc commun diverge pour donner les trois réponses. C'est une vue schématique. Or quand la cellule reçoit un signal donné, elle a une seule réponse ; d'où la question qui se pose, c'est : quand le signal arrive en bas du tronc commun, qu'est-ce qui lui dit d'aller vers la réponse spécifique, et pas d'aller activer les trois réponses ? C'est un exemple bien caractérisé, un exemple local parmi une multitude – un réseau de protéines, c'est ça multiplié et reproduit.

La question qui est posée est la suivante : cette cartographie ne permettant pas d'expliquer la régulation, comment la régulation biologique est possible ? On se retrouve en quelque sorte au point de départ : comment se fait-il que, quand une cellule est dans une situation donnée, elle est capable de répondre adéquatement ? Il faut donc qu'il y ait quelque chose qui limite la combinatoire moléculaire, une contrainte qui vienne s'exercer. Voilà pour fixer les idées. Si on reprend mon

schéma, on pensait au départ que les protéines étant stéréospécifiques, elles viennent s'emboîter de manière adéquate, puisque leur structure est la traduction de l'information, et elles vont ainsi donner naissance à une structure – elles ne peuvent pas faire autre chose. Or ce que l'on trouve c'est qu'il y a de grandes possibilités combinatoires, et si l'on prend cet exemple, elles vont pouvoir interagir de manières différentes et ainsi donner d'autres structures, d'autres types d'interaction. Donc la question qui se pose, c'est donc : dans un processus physiologique normal, comment on fait pour éviter les interférences des possibilités combinatoires, des autres interactions possibles ? Qu'est-ce qui permet de faire le tri ?

C'est une question qui est connue et qui a déjà été travaillée, elle a déjà fait l'objet de plusieurs revues dans la littérature, et il y a toute une série de mécanismes qui ont été montrés et proposés comme permettant de comprendre comment la cellule peut réduire la combinatoire moléculaire. Je les ai listé là, cette liste doit être à peu près exhaustive de tout ce qui a été proposé dans la littérature. Il y a toute une série de mécanismes :

Le premier mécanisme, c'est la compartimentation spatiale – une cellule n'est pas un verre d'eau, c'est une structure compartimentée, et les protéines ne se répartissent pas au hasard dans les différents compartiments (certaines cellules vont plutôt vers le noyau, d'autres vont plutôt à la membrane, etc.). Donc du fait que les cellules sont séparées physiquement, ça va permettre d'éliminer certaines interactions. Voilà pour la séparation spatiale.

Un autre mécanisme qui a été proposé est un mécanisme de séparation temporelle : les gènes ne s'expriment pas avec les mêmes dynamiques, pas forcément au même moment dans le cycle cellulaire, et de ce fait les protéines ne vont pas être présentes au même moment et ne pourront pas interagir. Cela peut donc permettre de réduire les combinatoires moléculaires.

Troisième mécanisme : la différenciation selon le type cellulaire a aussi été suggéré. C'est un peu analogue : les cellules, selon leur type cellulaire, n'expriment pas exactement les mêmes gènes, et donc on peut avoir la même voie de transduction, mais il suffit que les récepteurs dans le noyau soient différents pour qu'on aboutisse à l'activation de gènes différentes en fonction du type cellulaire en réutilisant les mêmes voies de transduction.

Ensuite, des mécanismes de micro-compartimentation : il y a certaines protéines qui peuvent fixer plusieurs protéines participant à la même voie, et qui donc vont concentrer localement toutes les protéines participant à une voie de transduction ou à une voie d'activation et donc favoriser leurs interactions, du coup.

Tous ceux-là sont des mécanismes de séparation, et il y a d'autres mécanismes encore qui ont été proposés, notamment l'idée que quand il s'agit de transduction d'un signal, ce n'est pas un signal mais des combinaisons de signal, ou des combinaisons de voies de signalisation, qui confèreraient la spécificité des réponses.

Ensuite des phénomènes d'inhibitions croisées : par exemple, il peut y avoir deux voies qui sont activées par le même signal, mais l'une des deux a un élément qui peut inhiber l'autre, et vice-versa, et aussitôt qu'il y a déséquilibre entre les deux voies, l'une des deux va s'éteindre.

Et aussi des mécanismes d'intensité du signal. Donc il y a toute une série de mécanismes qui ont été proposés et qui effectivement permettent de comprendre comment on peut réduire la combinatoire moléculaire. Sauf que ...

## Deuxième partie

Jusqu'à présent j'ai fait une synthèse de série de données – ce que vous pouvez trouver dans la littérature, je crois. Maintenant, je vais rentrer dans mon analyse personnelle. Ces mécanismes

permettent très bien de comprendre comment on peut réduire la combinatoire des molécules qui ne sont pas spécifiques, sauf qu'il y a un petit problème malgré tout, c'est que tous ces mécanismes ont tendance à déplacer l'explication de la stéréospécificité du niveau moléculaire sur le niveau cellulaire. C'est-à-dire que ces mécanismes présupposent l'existence d'une cellule qui est déjà structurée, organisée et différenciée – alors que c'est précisément ce qu'est censée expliquer la stéréospécificité des interactions moléculaires. On inverse en fait la cause et l'effet. Par exemple, j'ai cité le mécanisme le plus évident : le type cellulaire. Normalement, c'est l'existence de régulateurs spécifiques qui est supposée expliquer l'existence de types cellulaires, et non l'inverse ! Séparation temporelle : on dit que c'est la dynamique d'expression des gènes qui explique la stéréospécificité, et qui va réduire la combinatoire, mais normalement c'est la stéréospécificité des interactions qui doit expliquer la régulation de gènes ! Pareil pour tous les mécanismes : ils présupposent l'existence de cette structure déjà organisée. Le seul mécanisme pour lequel c'est un peu plus ambigu, c'est la compartimentation spatiale : parce qu'on sait que les protéines ont des régions qu'on appelle des « signaux d'adressage » qui vont les diriger préférentiellement vers un compartiment plutôt qu'un autre, sauf que pour diriger une protéine vers la membrane ou vers le noyau, ça suppose qu'on a déjà une structure existante. Donc globalement, ces mécanismes inversent la causalité initiale, c'est-à-dire en fait le fondement même de la biologie moléculaire. C'est ce que j'ai figuré ici : normalement, le point de départ de la biologie moléculaire consiste à dire que ce sont les interactions entre protéines (qui dépendent de leur structure, donc qui dépendent de l'information qui est codée dans les gènes) qui expliquent l'état macroscopique du système (le phénotype). Or maintenant, pour expliquer la stéréospécificité des interactions, on est en train de dire que c'est l'état macroscopique du système (c'est-à-dire l'existence d'une structure organisée et différenciée) qui explique la stéréospécificité de ce qu'il se passe au niveau moléculaire. Évidemment, c'est une contradiction massive des fondements même de la biologie moléculaire.

On en est là aujourd'hui. Les protéines ne sont pas spécifiques, et c'est une contradiction directe du principe de l'ordre par l'ordre – ou son équivalent aujourd'hui – qui est à la base du déterminisme génétique et de la biologie moléculaire. Donc qu'est-ce qu'on fait dans une situation de ce type là ? C'est une situation, en fait, qui n'est pas inédite : c'est assez fréquent dans l'histoire des sciences, où l'on arrive à un point où les données expérimentales qui ont été produites par une théorie viennent à contredire la théorie elle-même. Il y a deux possibilités : on peut considérer que ce sont les données qui sont mauvaises ou insuffisantes, et qu'il n'y a pas besoin de changer la théorie ; il suffit donc de continuer à accumuler des données expérimentales. Ou bien, on peut considérer que c'est la théorie qui est mauvaise, car on a suffisamment de données, et que ça ne sert à rien de continuer à accumuler toujours le même type de données, et qu'il faut changer de théorie. Chacun est libre, c'est le débat scientifique. Moi, ce que je crois, c'est qu'en quelque sorte la coupe est pleine. Je suis peut-être le plus âgé dans la salle ; j'ai fait mes études juste après Monod-Jacob avec le paradigme de la stéréospécificité, et j'ai vu toute l'histoire de la stéréospécificité. Au début on croyait les avoir, et maintenant on a ces dernières données. Moi je crois que ça ne sert plus à rien maintenant de continuer à accumuler des données : clairement, il y a un vice dans la théorie et il faut changer de théorie. Qu'est-ce que ça veut dire, changer de théorie ? Ça veut dire trouver un nouveau cadre qui permettra à la fois d'expliquer les données expérimentales et puis d'aller plus loin, de proposer un nouveau programme de recherches fondé sur cette nouvelle théorie. C'est un peu ce qui m'a amené à proposer cette nouvelle théorie. En quoi consiste faire un travail théorique ? Ça consiste, dans un premier temps, à ne pas avoir peur de tirer les conséquences, même les plus extrêmes – même celles qui semblent les plus affolantes, en quelque sorte – des faits expérimentaux, de les pousser les plus loin, même si l'on a l'impression qu'on perd un peu de cohérence, parce qu'à terme, on sait que – peut-être, si l'on a de la chance – on pourra reconstruire

un nouveau cadre théorique qui sera plus cohérent que l'ancien, et qui permettra d'intégrer les données et d'aller plus loin ; parce qu'il permettra de générer un programme de recherche, c'est-à-dire d'imaginer des expériences nouvelles qu'on ne serait pas amené à faire dans la cadre de l'ancienne théorie.

Quelles sont les conséquences qu'on peut tirer des faits expérimentaux que je vous ai indiqués ?

La première, c'est que les réseaux de protéines ou de gènes n'existent pas. J'entends des murmures, oui, je sais, sachant que 95% des appels d'offre consistent à demander la caractérisation de tels réseaux, ça peut sembler un peu farfelu de dire ça. Mais ce que ça signifie, c'est que ça [schéma], ce n'est pas un réseau intelligible d'interactions – c'est un magma d'interactions. C'est pas une cartographie intelligible qui va vous donner l'explication, mais c'est un magma de réactions. Et donc, quand je dis « les réseaux de protéines ou de gènes n'existent pas », évidemment il y a dans la cellule des séquences de réactions qu'on est capables de décrire. Mais ces séquences – c'est-à-dire ce qu'on voit à un moment donné dans une cellule – c'est le *résultat* des processus cellulaires et non la cause ! C'est le résultat des processus cellulaires une fois que dans le magma de réactions, la cellule a sélectionné – « filtré », en quelque sorte – et n'active qu'un certain nombre de réactions qu'elle laisse se produire. Et donc effectivement, on pourra toujours trouver des séquences de réactions et tracer des cartes, mais ces cartes ne sont pas la cause des phénomènes, mais c'est leurs conséquences.

Deuxième conséquence qu'il faut tirer : les protéines ne sont pas spécifiques. Il y a des combinatoires potentielles immenses, et cela veut dire que les interactions entre molécules sont probabilistes. Il faut l'accepter, contrairement à l'idée de départ qu'il fallait éliminer la stochasticité parce que les protéines interagissaient avec un seul partenaire ou avec un petit nombre – non, elles ont des possibilités immenses ! Et en fonction de ce qui va être réalisé à un moment ou à un autre, elles vont pouvoir faire des choses différentes – et donc la nature même des phénomènes du vivant est stochastique. Là, j'ai mis : « les interactions moléculaires sont intrinsèquement aléatoires, et non pas bruitées ». Pourquoi ? Parce que, comme je l'ai déjà dit dans l'introduction, à l'heure actuelle il est couramment admis que l'expression des gènes est un phénomène aléatoire – mais on appelle ça du « bruit ». Alors ce qu'on entend par cela, l'explication qui est donnée, c'est que certains facteurs nucléaires transcriptionnels – on le sait – sont présents en petite quantité dans le noyau ; et donc, à cause des fluctuations thermiques (un facteur de transcription doit atteindre par diffusion le noyau, etc.) à tout moment, dans deux cellules différentes isolées vous n'aurez jamais exactement le même nombre de molécules de ce facteur de transcription. Donc vous aurez toujours des petites variations aléatoires. Et lorsque ces variations dues à la diffusion, donc à la fluctuation thermique, concernent des facteurs présents en petite quantité, ça peut avoir des conséquences importantes, parce que vous allez être très facilement au-dessus ou en-deçà du seuil d'action de ce facteur. Et donc comme ça vous allez pouvoir générer des différences de comportement cellulaire. Donc ça c'est ce que l'on appelle le bruit : c'est le fait qu'on ne peut pas moyenniser des fluctuations thermiques sur des facteurs qui sont présents en petite quantité. Mais cette manière de voir ne remet pas en cause l'idée qu'il y a une carte – qu'il y a un programme génétique – et que les molécules par ailleurs peuvent être stéréospécifiques, et que vous avez un réseau d'interactions. Donc c'est simplement des fluctuations dans la réalisation de ce programme. Là, ce dont je parle, c'est autre chose : je dis qu'il y a obligatoirement des phénomènes de compétition, du fait que les molécules ne sont pas spécifiques. Et en fait on sait aussi que ces phénomènes existent, c'est-à-dire qu'on sait que sur des promoteurs il y a de nombreux facteurs de transcription qui peuvent venir interagir et qu'il y a des phénomènes de compétition moléculaire. Et ça a été démontré par simulation, par une équipe de

Lyon, d'ailleurs, l'équipe d'Olivier Gandrillon. On peut avoir des effets stochastiques très forts qui ne dépendent pas du fait que le facteur de transcription est présent en petite quantité, c'est un phénomène différent. Et ça, par contre, pour les raisons que je vous ai expliquées, ça touche à la nature même : ce qui est mis en cause, c'est l'idée de programme génétique, c'est-à-dire l'idée qu'on a une carte sous-jacente fixe – un espèce de squelette fixe d'interactions microscopiques – qui explique ce qu'il se passe dans un être vivant.

Et enfin, troisième conséquence qu'on peut tirer de ces faits expérimentaux : c'est que l'ontogénèse et la phylogénèse ne font qu'un. J'en arrive à cette idée d'ontophylogénèse. En fait, si on admet – comme semblent le suggérer les données expérimentales – que la structure cellulaire restreint les combinatoires moléculaires, finalement on retombe aussi sur quelque chose qu'on sait. Je veux dire, la génération spontanée n'existe pas, il y a peut-être eu une origine, on n'en sait rien, mais la seule chose à laquelle on a accès aujourd'hui, c'est quand un être vivant se développe, il part jamais d'un ADN nu – il est toujours pris dans une structure, et on a toujours une structure cellulaire. Une cellule germinale, c'est une cellule : il y a toujours une structure au départ – on retombe là-dessus. Donc si l'on admet cette idée que c'est la structure cellulaire – comme l'indiquent tous ces mécanismes – qui permet de restreindre les interactions moléculaires au cours de l'ontogénèse, l'idée d'ontophylogénèse devient quasiment une déduction nécessaire.

C'est ce que je vais essayer d'expliquer maintenant, sur ce schéma. Ici, dans le panneau A, c'est la situation actuelle : vous savez, à l'heure actuelle, ce qu'on appelle la « théorie synthétique de l'évolution » (le cadre général dans lequel on baigne) est le fruit de la synthèse entre le darwinisme et la génétique, pour aller vite. Donc dans ce cadre théorique général, on a d'un côté l'évolution qui est le résultat de la sélection naturelle s'exerçant sur les phénotypes – les structures cellulaires ou multicellulaires – et puis d'un autre côté, on a l'ontogénèse qui est l'expression du programme génétique contenu dans l'ADN. Or ces deux phénomènes sont causalement disjoints : la sélection naturelle n'intervient pas dans l'ontogénèse, la seule chose qu'elle peut faire c'est – quand elle trie et qu'elle favorise une structure cellulaire plutôt qu'une autre – trier les mutations associées dans l'ADN. Mais elle n'intervient pas directement dans le processus lui-même. Maintenant dans le panneau B, on accepte cette idée, qu'on déduit des faits expérimentaux, que c'est la structure cellulaire qui trie les interactions moléculaires – ici on avait une flèche, la flèche classique qui va de l'ADN vers le phénotype. Ici, on a toujours la flèche montante, « bottom-top », mais on est obligé de mettre aussi ici une flèche qui va dans l'autre sens : c'est-à-dire que d'un côté l'ADN fabrique des protéines, ces protéines sont soumises à des combinatoires moléculaires, et cette combinatoire moléculaire est triée par la structure cellulaire – c'est ce que représentent ici ces deux flèches. Mais dans la mesure où la structure cellulaire est elle-même soumise et triée sous l'action de la sélection naturelle, le fait de rajouter cette petite flèche ici signifie qu'on rétablit une continuité causale qui part de la sélection naturelle jusqu'à ce qu'il se passe au niveau moléculaire pendant l'ontogénèse. Parce qu'on est bien d'accords : vous avez par exemple une cellule qui est d'une structure donnée plutôt que d'une autre, vous allez expliquer certaines interactions et pas d'autres, mais le fait que cette cellule ait à un moment donné de son histoire cette structure et pas une autre, c'est parce qu'elle a été triée, sélectionnée par la sélection naturelle. Donc, de fait, ce qui détermine de manière ultime les interactions qui sont triées au cours de l'ontogénèse, c'est la sélection naturelle qui agit via la structure cellulaire. Et les phénomènes ne sont plus disjoints, mais c'est un seul phénomène parce que vous avez une continuité causale.

C'est donc ce que j'ai écrit ici : l'ontogénèse et la phylogénèse, de ce fait, ne forment qu'un seul processus – et c'est ce que j'appelle « l'ontophylogénèse ». Et c'est une déduction quasi-

nécessaire à partir du moment où vous admettez que c'est la compartimentation spatiale, la séparation temporelle, etc. qui trient les interactions moléculaires. Donc on sort de cette analyse avec un cadre théorique pour l'ontogénèse qui est complètement chamboulé : il n'a plus rien à voir avec le cadre de départ. Le cadre de départ, c'était les interactions stéréospécifiques bien propres, bien nettes. Ici, vous admettez qu'il y a de la combinatoire, et puis – je reprends cet exemple – sur l'ensemble de la diversité potentielle, il y a un tri qui est opéré par un processus sélectif, et c'est ce qui permet de générer une structure. Donc on tombe dans un mécanisme qui est à la fois conceptuelle analogue à la théorie de Darwin – parce que c'est un mélange de hasard au niveau des interactions moléculaires trié par une sélection – mais ça va plus loin : ce n'est pas seulement analogue à la sélection naturelle, puisque c'est la sélection naturelle elle-même qui, *in fine*, est l'agent de ce processus – ce que j'ai expliqué dans la diapo précédente. C'est pour cela que j'ai introduit ce terme un peu barbare d'ontophylogénèse : pour indiquer que c'est vraiment un seul phénomène, car là maintenant on n'est plus capable de distinguer les deux. Les deux sont causalement intriqués, et donc on a un seul phénomène et on retombe sur cette idée que je vous ai expliquée dans l'introduction. Et en fait on retombe sur quelque chose qui est quasiment du sens commun, si l'on considère les remarques que j'ai faites en introduction, à savoir qu'il y a un processus unique sous-jacent à l'évolution des espèces et à l'ontogénèse : la reproduction.

Donc ça c'est une autre manière d'expliquer, et on arrive à cette idée que le développement n'est plus l'expression d'un programme génétique, mais c'est un mécanisme de hasard sous contrainte – le darwinisme cellulaire – d'interactions moléculaires aléatoires triées par la sélection naturelle. Pour réexpliquer tout cela d'une autre manière : la question qui se pose, qu'on pose souvent mais qui paraît assez contre-intuitive, c'est de savoir comment un mécanisme qui paraît aléatoire peut être compatible avec la très grande reproductibilité qui caractérise l'embryogénèse, l'ontogénèse et d'une manière générale les phénomènes du vivant. Donc je vais rentrer dans la troisième partie, et petit à petit on va retourner à questions plus précises et éventuellement déboucher sur un programme de recherches expérimentales. Pour comprendre cela – un « hasard sous contrainte » – imaginez une petite bille (la partie A du schéma) soumise à l'agitation brownienne, elle se déplace donc au hasard, et elle est libre de se déplacer comme elle veut. De ce fait, du fait qu'elle est libre de se déplacer, vous n'avez aucune prédictibilité, aucune reproductibilité autre que statistique si vous le faisiez un très grand nombre de fois, mais elle a un très grand nombre de possibilité à chaque fois et vous ne pouvez pas prédire ce qu'elle va faire. Maintenant (cas B) imaginez que vous mettez une contrainte sur son déplacement : par cette contrainte, vous l'obligez à se déplacer entre deux parois, deux murs ici. Que va-t-il se passer ? Elle va continuer à se déplacer au hasard, toujours, donc il y aura une part de hasard, mais ce hasard sera restreint puisqu'elle ne pourra pas sortir de cet espace que vous avez délimité. Donc on se trouve dans une situation en quelque sorte intermédiaire : vous avez restreint ses possibilités de déplacement aléatoire, il y a donc une forme de prédictibilité, mais ça reste encore très flou. Maintenant (cas C) imaginez que vous continuez à rapprocher les contraintes jusqu'au point où entre les deux parois il y a quasiment que l'espace correspondant au diamètre de la bille. Que va-t-il se passer ? La bille continue de se déplacer au hasard, mais elle va faire une marche aléatoire dans un espace à une dimension, c'est-à-dire selon une ligne droite. Et donc, bien que intrinsèquement, elle se comporte de manière aléatoire, à chaque fois que vous allez lâcher la bille ici, elle va tracer un segment de droite. Elle ne le tracera pas exactement de la même manière, parce qu'elle est toujours soumise à une marche aléatoire, mais elle tracera une droite parce qu'elle est contrainte par les deux murs. L'idée, c'est qu'à l'heure actuelle on est dans la même situation : la structure cellulaire qu'on connaît aujourd'hui, c'est le résultat de toute l'histoire de la vie. Tous les êtres vivants qui existent aujourd'hui, c'est des êtres qui ont survécu depuis que la vie est apparue à travers ce cycle ininterrompu de multiplications. Et

donc ils ont emmagasiné leur structure. Pour survivre, ils ont obligatoirement emmagasiné dans leur structure toutes les contraintes successives de tous les environnements auxquels ils ont été confrontés. Et donc, que cette structure là est hautement contraignante, de manière analogue ici à ces deux murs (le cas C) de telle sorte qu'à chaque fois que vous avez une ontogenèse, c'est la même chose qui se reproduit. En même temps, cette structure réduit le hasard, mais ne l'élimine pas complètement. Ici, cette idée c'est que vous allez tracer une droite, mais à chaque fois de manière complètement différente. Donc ça ne le réduit pas complètement : cette marge de hasard est utile à l'organisme (et elle est contrôlée) car c'est ce qui lui sert à créer de la diversité, c'est ce qui permet d'expliquer pourquoi les systèmes vivants sont hautement plastiques et comment ils peuvent créer des lignées cellulaires différentes, s'adapter à tous types de situation auxquels ils sont confrontés, etc. Au cours de la différenciation cellulaire, il lui reste une marge d'aléatoire qui va lui permettre de créer quelques types cellulaires différents, mais que globalement la structure qui est celle de la cellule germinale – puisqu'on part toujours d'une cellule – qui est le résultat de toute l'histoire de la vie, par sa structure elle contraint le phénomène de telle manière qu'il soit reproductible.

Donc si on reprend le schéma ici, je vais aller vite, la relation entre génotype et phénotype n'est plus une relation unidirectionnelle qui va du génotype au phénotype ; mais c'est une relation bidirectionnelle, car la contrainte sélective issue du phénotype trie les événements moléculaires qui sont soumis à la variabilité. Dans ce schéma il n'y a pas d'origine unique du processus car il repose sur un mécanisme dual. C'est important, je pense, pour les philosophes de la biologie notamment. La biologie oscille, dans son histoire, entre le holisme et le réductionnisme. Et le holisme et le réductionnisme sont deux manières de voir qui assignent une origine, ce sont des théories de l'origine. Dans un cas, l'origine est la molécule, dans l'autre cas c'est le tour. Il y a toujours une origine, soit c'est l'ordre par le haut, soit l'ordre par le bas. Là, c'est ni l'un ni l'autre : c'est plutôt le facteur temps qu'il faut introduire. Il faut voir que c'est la flèche du temps qui est introduite et, à tous moments il y a des contraintes venant du haut et du bas qui interviennent et qui se résument dans l'organisme qui existe à un moment donné. Dans ce processus, il n'y a pas d'origine ontologique. Cela veut dire que c'est une description de l'état actuel, de ce qu'on voit, de ce à quoi on a accès, et on botte en touche pour ce qui est de la question de l'origine – y compris pour la question de l'origine de la vie. On admet aujourd'hui que tout processus d'ontogenèse part d'une cellule – c'est ce à quoi on a accès, peut-être qu'un jour on en saura plus, mais aujourd'hui on le prend comme point de départ, et c'est ce qu'on va décrire et expliquer. Donc on peut dire que c'est un changement d'ontologie, aussi.

### Troisième partie

Je vais maintenant aborder la dernière partie : maintenant qu'on a défini un nouveau cadre théorique, comment va-t-on va revenir vers quelque chose de plus expérimental ? La question qui est posée dans un premier temps est : comment peut-on revenir vers l'explication d'un phénomène plus concret (la différenciation cellulaire) ? Puis ensuite : comment est-ce qu'on peut l'aborder de manière expérimentale.

Donc classiquement la différenciation cellulaire est vue comme un mécanisme instructif : les cellules échangent des signaux, et selon les signaux qu'elles reçoivent, elles se différencient selon des voies de différenciation. Or, dans le cadre d'un modèle darwinien, on arrive à un modèle totalement différent : du fait que les molécules ne sont pas spécifiques et qu'il y a des possibilités combinatoires (notamment au niveau de l'expression des gènes à cause de ces phénomènes de

compétition) les cellules peuvent s'orienter spontanément en fonction des probabilités associées à ces événements (donc de fréquence dans des populations de cellules), c'est-à-dire qu'elles vont s'orienter selon des choix de lignages différentes sans avoir plus besoin de signaux. Ainsi, selon l'événement aléatoire  $a$  ou  $b$ , elles vont s'orienter vers un type cellulaire A ou B. Donc : hasard au niveau moléculaire. Et l'idée que c'est la structure cellulaire (donc le tout) qui contraint : se greffent là-dessus les interactions cellulaires qui sont sélectives ou stabilisatrices ; ce sont elles qui vont figer des combinaisons de phénotypes donnés. Donc dans ce modèle : expression aléatoire des gènes, engagement aléatoire dans des lignages cellulaires, et ensuite les types cellulaires interagissent entre eux avec stabilisation réciproque des phénotypes. Donc dans ce modèle on ne nie pas le fait que les cellules interagissent au cours du développement, simplement on ne leur assigne plus le rôle d'inducteur de la différenciation, mais de stabilisateur au contraire – stabilisateur de situation, de configuration qui ont été générées de manière aléatoire.

Rapidement, donc : aujourd'hui, quelles prédictions on peut faire à partir d'un modèle comme ça ? D'abord, j'ai pu dire que l'expression aléatoire des gènes était démontrée : ce modèle, je l'ai proposé la première fois en 1981, et à l'époque ça ne l'était pas du tout, donc je peux me permettre encore de dire (en tout cas dans mon cas) que c'était une prédiction du modèle qui a été vérifiée, pour des raisons historiques. Évidemment, expression aléatoire des gènes ; stochasticité dans la différenciation de la cellule ; et puis, les molécules ne sont pas spécifiques : pour qu'il puisse y avoir de l'aléatoire, il faut que les molécules ne soient pas spécifiques – je l'ai expliqué. Je voudrais juste dire deux mots sur l'expression stochastique des gènes et sur la variabilité cellulaire, parce que ce sont des données qui font aujourd'hui l'objet de plus en plus fréquemment de revues, mais ce n'est pas encore quelque chose qui est rentré dans la connaissance courante, ou massive. Donc aujourd'hui, le fait que l'expression des gènes est aléatoire, c'est démontré. Pourquoi est-ce démontré ? Parce que depuis le début des années 1990, on a des techniques qui permettent de discriminer ce qu'il se passe dans des cellules uniques, alors qu'avant, quand on voulait étudier l'expression des gènes, on était obligé de travailler sur des populations de cellules, on extrayait l'ARN par exemple de millions de cellules. Et donc quand on regardait l'ARN d'une population de cellules, on ne regardait pas les cellules une par une, mais on faisait la moyenne d'un grand nombre de cellules. Maintenant par des méthodes *in situ* notamment, en utilisant des marqueurs fluorescents, on peut visualiser ce qui se passe dans les cellules une par une. Et depuis qu'on a ces techniques, on sait qu'à un moment donné, dans des populations de cellules isogénique placées dans les mêmes conditions, on a toujours des variabilités de cellule à cellule. Voilà pour l'expression des gènes – ces diapos sont déjà datées, ce n'est pas exhaustif mais je ne vais pas rentrer dans le détail, voyez une bibliographie, il y a différentes revues qui ont été publiées récemment. Ça va assez loin, la variabilité transcriptionnelle, quand même : dès les années 1990 on avait des arguments expérimentaux très forts pour l'expression stochastique des gènes. C'était par exemple le cas de l'expression hétéro-alléliques : il y avait ainsi des données obtenues des techniques d'hybridation *in situ* où l'on visualise non seulement ce qu'il se passe dans des cellules différentes, mais dans les deux chromosomes correspondants, et on peut regarder ce qu'il se passe sur un chromosome donné ce qui se passe, et sur l'autre allèle sur l'autre chromosome. Et on avait alors montré que dans différents locus multi-alléliques, il y a toujours une fraction de cellules où ce n'est pas les mêmes allèles qui s'expriment à un moment donné dans la même cellule. Ça voulait dire clairement que ce n'était pas l'environnement en facteur de transcription qui déterminait son état transcriptionnel, et déjà dans ces papiers les auteurs disaient dans les discussions qu'ils ne voyaient pas comment faire autrement que d'admettre qu'il y avait une composante stochastique dans l'expression des gènes.

Et en ce qui concerne la variabilité cellulaire, c'est encore plus ancien. C'est toute une

littérature qui existe depuis très longtemps, mais du fait de l'hégémonie de la théorie du programme génétique, elle a été mise de côté. Dans les cas où l'on avait des marqueurs de différenciation cellulaire et qu'on pouvait regarder malgré tout cellule par cellule, dans des clones, dans différents modèles, dès 1964 avec les travaux de Jim Till, on pouvait regarder la cinétique de différenciation de cellules, avec un marqueur, dans une population de cellules, normalement vous faites votre cinétique, vous regardez combien de cellules vous avez à différents temps – je simplifie la logique de ce type d'expérimentation – vous faites des prédictions différentes selon que vous dites que vous avez un modèle déterministe ou stochastique. Dans un modèle stochastique, vous allez admettre que chaque cellule, à chaque mitose ou par unité de temps, elle a une certaine probabilité de se différencier, vous allez avoir une courbe exponentielle. Alors que si c'est un modèle déterministe, à un moment donné vous allez avoir une marche d'escalier, idéalement, plus ou moins avec du bruit. Et donc on est capable mathématiquement d'analyser ces courbes et de voir si c'est plutôt stochastique ou s'il y a sous-jacent un déterminisme. Et dès 1964 il y avait des gens qui avaient analysé des situations de ce type-là et ça s'est accumulé : ils disaient par exemple que la différenciation de la lignée hématopoïétique (ou d'autres) s'explique mieux par un modèle stochastique où l'on admet qu'il y a un événement stochastique à chaque mitose ou par unité de temps, etc. qui explique l'engagement de la cellule selon des voies.

Alors qu'est-ce qu'on peut faire expérimentalement ? Je conclurai là. Il y a deux types de choses qu'on peut faire.

Premier type d'expérimentation : si vous avez un modèle qui dit que la différenciation cellulaire est basée sur l'expression stochastique des gènes, clairement il faut étudier l'expression stochastique des gènes. C'est un travail que je fais en collaboration, je participe à ce programme avec Olivier Gandrillon qui est à Lyon avec l'Université Claude Bernard. Il faut faire une analyse quantitative : soit on dit que les variations de l'expression des gènes sont un phénomène accidentel qui n'a pas de signification biologique ; soit on dit que ça a une signification biologique, et dans ce cas là ça devient un paramètre physiologique à part entière. Or ce qui caractérise un paramètre physiologique, c'est que vous devez être capable de le quantifier, d'étudier ses modifications quantitatives au cours d'un phénomène physiologique et de corrélérer ces variations avec le phénomène que vous étudiez, et que la corrélation éclaire le phénomène – c'est cela, un paramètre physiologique. Donc clairement, c'est cela qu'il faut faire : analyser quantitativement. Qu'est-ce que ça veut dire ? Vous prenez un gène donné, par exemple, ainsi qu'une population de cellules, quantifier son niveau d'expression consiste à voir sa variance, avoir un coefficient de variation, de son niveau d'expression dans la population – c'est un exemple de quantification. Je donne là un exemple pour expliquer la logique du programme de recherche : si vous dites que c'est du bruit, que c'est accidentel, la variabilité d'expression des gènes, ça doit être plat – vous avez une cellule qui se différencie, c'est indépendant de la physiologie, c'est un phénomène marginal, et donc ça n'a aucune raison de varier. Par contre si vous dites que vous avez affaire à un phénomène, par exemple, sélectif, c'est autre chose. Qu'est-ce qu'un phénomène sélectif ? Vous partez d'une situation où vous avez une grande variabilité, une grande diversité, et vous sélectionnez et amplifiez un pattern d'expression, un profil d'expression. Alors vous faites une prédiction qui est aujourd'hui testable expérimentalement : vous devez avoir une réduction de la variabilité de l'expression des gènes. C'est ce que j'ai figuré ici : si vous partez d'une situation où les cellules, les pastilles de couleur, ces différents patterns d'expression – c'est l'expression aléatoire – eh bien dans cet aléa, ce magma aléatoire qu'est la différenciation cellulaire, il y a un phénomène sélectif qui va amplifier et stabiliser les rouges et les noirs. Ça va réduire votre variabilité parce que dans votre population, si vous faites une cinétique, au départ vous avez une grande variabilité – des jaunes, des gris, des bleus, etc. – alors qu'à la fin vous n'avez que du rouge et du noir – vous avez réduit, et vous pouvez

donc faire des prédictions. La prédiction, c'est donc que, si la base de la physiologie cellulaire est la stochasticité, alors la stochasticité doit devenir un paramètre que vous devez quantifier. La quantification doit vous permettre d'établir des corrélations avec les phénomènes biologiques – c'est le principe général et c'est ce qu'on est actuellement en train d'essayer de faire.

Le deuxième type d'expérimentation, c'est qu'on peut faire de la simulation – c'est ce que j'ai fait avec un physicien de l'Université Pierre et Marie Curie – au moins dans le principe. On a publié cela il y a déjà cinq ans. La grande question est : « est-ce qu'on peut créer de l'ordre et quelque chose qui ressemble à des tissus ? ». Je ne vais pas rentrer dans le détail, mais on peut facilement créer des modèles informatiques qui vont simuler le processus de hasard sélectif : on va créer des cellules virtuelles et on va les faire fabriquer des molécules, interagir, et se comporter selon des règles de hasard-sélection – le modèle darwinien de différenciation. Il y a pas mal de paramètres. On peut montrer qu'on peut créer des structures reproductibles, ici dans notre cas c'étaient des bicouches cellulaires. C'était un modèle rudimentaire, mais on a commencé par le cas le plus simple d'une différenciation rouge-vert. On part d'un petit nombre de cellules rouges et vertes qui ont un comportement aléatoire, elles sont soumises à des règles de hasard-sélection, et on laisse cette population se développer. Et on peut montrer qu'effectivement le modèle de hasard-sélection en tant que modèle est capable de générer des structures qui ressemblent à des structures cellulaires organisées. Donc par rapport à la question, ça ne prouve pas forcément que dans la nature, c'est comme cela, mais ça prouve que ça peut être comme ça, c'est-à-dire que c'est de toute façon un cadre théorique qui permet d'aborder la question.