

Recherche en biologie : état des lieux et perspectives

Pierre Tambourin

Directeur du génopole d'Évry, coordinateur du réseau des génopoles de France, président du conseil scientifique de l'Inra

Je présenterai essentiellement trois aspects de l'évolution actuelle en biologie, que celle-ci concerne la médecine ou d'autres domaines :

- je résumerai d'abord brièvement les développements actuels de la biologie dite à grande échelle,
- j'insisterai plus longuement sur les conséquences sur l'organisation et la vie quotidienne de la recherche en biologie sur cet aspect,
- dans un troisième temps, même si cela ne constitue pas l'objectif de ce colloque, je soulignerai à nouveau l'importance et la nécessité de la présence, à côté des laboratoires de recherche publique, dont la mission relative à ces questions reste évidemment claire, d'un ensemble d'activités appelées les biotechnologies qui, pour un pays, une nation ou un continent, simplement pour les futures conséquences du travail que nous effectuons dans nos laboratoires, représente un élément-clé.

Abordons d'abord la biologie à grande échelle.

Si vous vous référez à des publications pourtant récentes, puisque datées de 1993, vous pouvez lire que la date à laquelle nous aurions dû terminer le séquençage du génome de l'homme était attendue, selon les auteurs, entre 2005 – pour les plus optimistes – et jusqu'à 2015 dans la majorité des cas. Or vous n'ignorez pas que l'année 2002 a été marquée par la publication, non pas certes de la séquence complète du génome de l'homme, qui reste incomplètement défini, mais, comme M. Kahn en prend souvent l'image, d'un certain nombre de pages, pas toujours classées dans le bon ordre et dont certaines sont encore blanches. Globalement, nous connaissons actuellement 90 à 95 % de ces quelques milliers de pages, ce qui n'est pas négligeable. Le travail est cependant loin d'être terminé, car les 5 % restants ne sont certainement pas les plus faciles à définir.

Pourquoi cette accélération extraordinaire, qui fait remettre en cause tout un ensemble de stratégies ? Comment peut-on passer de 2015 à 2001, comment peut-on s'attendre à une décryptage complet en 2003, 2004 au plus tard ? Il y a deux raisons essentielles à cela.

- Tout d'abord, la biologie a connu au cours de la dernière décennie une évolution considérable – le terme de révolution est peut-être un peu excessif – des méthodes et des modes de pensée quant à l'approche de ces questions, que l'on a résumée sous le vocable de biologie à grande échelle.
- Par ailleurs, ne l'oublions pas, une compétition s'est développée, certes un peu plus tardivement, mais elle a constitué un réel accélérateur. Des sociétés de biotechnologies se sont lancées dans l'entreprise et se sont fait fort de mener à bien le travail bien plus rapidement que les laboratoires de recherche académiques qui, jusque-là, étaient naturellement les lieux de réalisation de ce type de recherches. L'entrée en lice de sociétés privées comporte un risque pour l'ensemble du dispositif, non seulement pour la recherche publique, mais aussi pour l'ensemble des entreprises, de voir une société « s'approprier » certaines portions du génome. Même si la réalité de ce risque était discutable, il a conduit certains chercheurs – Gene Watson, par exemple, mais d'autres aussi dans d'autres pays – à demander une augmentation massive des moyens alloués au secteur public dans la recherche de la séquence du génome. Cette revendication comportait bien sûr une petite part d'opportunisme, comme toujours, mais reposait tout de même sur des analyses assez approfondies. C'est un élément qui fera dorénavant partie en permanence de ce paysage nouveau, qui sera présent dans la démarche, que celle-ci soit académique ou non.

Le développement de la biologie à grande échelle n'est rien d'autre que l'apparition dans certains laboratoires – pas tous, bien sûr (c'est alors que se pose le problème de l'organisation, sur lequel je reviendrai) – de moyens qui n'ont en eux-mêmes rien d'extraordinaire, d'appareils que beaucoup d'ingénieurs connaissent : les automates. La biologie à grande échelle repose sur la mise en œuvre de techniques automatisées grâce à des robots et sur la mise en parallèle de séquenceurs pour analyser la carte du génome. C'est cet ensemble de moyens coûteux qui a permis l'accélération du processus de séquençage.

Au-delà de cet aspect purement quantitatif (accumuler des données plus rapidement), quelque chose se dessine de l'ordre du qualitatif, que je vais essayer de vous expliquer. Il ne s'agit plus seulement d'aller plus vite, d'acquérir des parts de marché pour les entreprises ou bien, pour certains pays qui se veulent très présents dans ce domaine, d'être les premiers. Il y a également l'idée que, grâce à ces technologies relativement nouvelles – pour les biologistes –, pourrait apparaître une vision plus ou moins nouvelle – en tout cas très différente, j'essaierai de l'illustrer – du monde du vivant : une vision nouvelle de la cellule, de l'organe, de l'organisme, mais aussi, plus largement, une nouvelle vision des écosystèmes, que l'on peut aujourd'hui aborder dans toute leur complexité.

C'est ce terme de complexité, dont on parle beaucoup depuis longtemps en biologie, qui devient un enjeu de connaissance. Les chercheurs ambitionnent désormais

de donner du fonctionnement de la cellule (par exemple) une représentation globale probablement beaucoup plus proche de la réalité que celle que nous avons montrée jusqu'à maintenant et qui se bornait à isoler un gène puis à essayer de comprendre sur quel autre gène celui-ci pouvait agir et ensuite, de cascade en cascade, à essayer de remonter le fil.

Les sociologues m'ont un jour donné une image de ce que nous étions en train de vivre : c'est un changement d'échelle comparable à la différence existant entre comprendre l'individu et comprendre le groupe d'individus dans sa complexité et ses interactions, voire des ensembles de groupes d'individus. C'est peut-être cette analogie qu'il faut mettre en place, mais comme toujours avec beaucoup de prudence.

C'est ainsi que se sont créés dans le monde des laboratoires de séquençage dotés de moyens très importants. Un élément peut vous donner une idée de la vitesse à laquelle cela évolue : quand, en 1996, la France s'est lancée dans l'entreprise de séquençage du génome – Elisabeth Dufourcq puis François Daubert étaient alors les ministres de la recherche scientifique –, on avait calibré le centre de séquençage français pour qu'il soit le troisième au monde, après le Sanger Centre en Angleterre et le laboratoire du NIH aux États-Unis. Aujourd'hui, ce laboratoire, qui a pourtant disposé de moyens importants, est tombé au huitième ou neuvième rang mondial, selon les critères retenus, simplement parce que d'autres pays se sont beaucoup investis.

Que savons-nous aujourd'hui et quel est l'apport de la connaissance du génome ?

À ce jour, un peu plus d'une cinquantaine de microorganismes présentant un intérêt soit médical (dans le cadre de l'étude des maladies infectieuses), soit agro-alimentaire, soit même purement académique (en tant que modèles d'étude), ont été entièrement séquencés.

Dans le monde des eucaryotes – vous pouvez trouver sans difficulté toutes ces données en vous connectant à des sites internet tels que celui du NCBI (National Center for Biotechnology Information) ou bien d'autres –, voici l'état des connaissances acquises :

– six génomes ont été totalement séquencés, ce sont essentiellement des espèces dites académiques (la drosophile par exemple), liées à la recherche, ce qui n'est pas anormal,

– on termine actuellement le séquençage d'une petite vingtaine de génomes, parmi lesquels ceux d'un certain nombre de plantes d'intérêt industriel et académique, d'organismes d'intérêt médical ou académique, dont l'énumération serait fastidieuse.

Du point de vue qualitatif, l'approche par ces méthodes de séquençage massif – c'est là que je souhaitais en venir et je voudrais le faire pour chacune des technologies que j'évoquerai – permet de connaître les génomes plus rapidement, évidemment. Mais elle permet également, par exemple, d'aborder d'une manière relativement

satisfaisante des populations complexes de microorganismes que l'on ne pourra probablement pas avant longtemps cultiver, isoler ni caractériser. Je fais référence aux écosystèmes complexes en termes de microorganismes que l'on trouve dans les sols, à ceux présents dans les espaces aqueux, mais également à la flore intestinale de l'homme, laquelle constitue un monde de complexité.

Aujourd'hui, grâce à ces méthodes de séquençage massif, on peut commencer à aborder cette population dans son ensemble et à recueillir des indications sur les constituants de cette population. Il est d'ores et déjà possible, par exemple, de suivre l'évolution de ces populations, certes de manière indirecte.

Deux aspects, quantitatif et qualitatif, sont donc à considérer. Certains problèmes que l'on ne pouvait pas aborder auparavant peuvent commencer aujourd'hui à être envisagés.

Parallèlement à ce séquençage massif, mais également parce que ces technologies de séquençage massif le permettent, une autre ambition s'est développée, en même temps que la création du centre de séquençage en France ou très peu de temps après : l'idée de rechercher comment on peut repérer, au sein de nos génomes respectifs, les régions, les zones responsables de certains caractères génétiquement transmis au sein du génome et qui jouent un rôle par exemple – mais pas uniquement – dans l'apparition d'un certain nombre de pathologies non habituellement classées parmi les maladies génétiques.

Axel Kahn l'a très bien dit : nous sommes très marqués, dans notre pensée, par les maladies génétiques rares (elles ne sont d'ailleurs pas toutes rares, car certaines d'entre elles sévissent dans des pays où l'on n'a pas beaucoup l'habitude de travailler, c'est le cas en particulier de certaines maladies génétiques dont la pénétrance est si élevée que la présence d'une seule mutation se traduira presque inmanquablement par l'apparition de la maladie).

Il existe de nombreuses autres pathologies face auxquelles nous sommes inégaux. Ceci est vrai même pour des maladies réputées non génétiques, comme les maladies infectieuses. En dehors peut-être des accidents de la route (... et encore !), on peut dire, même si cela peut en choquer certains, que toutes les pathologies ont un certain contenu génétique. Bien sûr, la part jouée par le génome peut être plus ou moins importante. On touche ici aux grandes pathologies de l'homme, celles auxquelles nous pensons tous : le diabète, les obésités – et pas simplement les obésités héréditaires, bien sûr –, l'inégalité face à la consommation.

Je cite souvent l'exemple suivant : supposons que l'un de vos gènes vous prédispose au cancer du poumon lié au tabagisme ; si vous ne fumez pas, si vous fumez peu ou si vous vivez dans un environnement où l'exposition à ce type d'agents agressifs est relativement faible, l'expression de ce gène se verra moins, bien évidemment. C'est cette interaction entre, d'une part, un gène et son environnement génétique dans le génome – puisque les autres gènes jouent un rôle – et, d'autre

part, l'environnement au sens classique du terme, que nous nous proposons de débroussailler en repérant, au sein des génomes, les régions susceptibles de jouer un rôle dans le développement des maladies.

Toutes les grandes pathologies : cardiovasculaires, métaboliques, systémiques, neurologiques, etc. sont actuellement étudiées dans le monde du point de vue de la génétique. Ici encore, deux aspects sont à considérer, celui du diagnostic et celui du traitement.

- Le premier aspect est le repérage, le dépistage, utilisé en médecine préventive. Si on sait qu'un sujet donné quelqu'un est exposé à tel ou tel problème, le dépistage peut conduire à lui recommander de modifier son mode de vie ou son alimentation, par exemple, de façon à empêcher la maladie de se développer. Le médecin doit cependant rester conscient des limites de son action. Je cite souvent l'exemple de l'épilepsie, qui notre touche en France 400 000 personnes de manière plus ou moins grave. L'épilepsie est une maladie du système nerveux central, organe auquel on n'a pas facilement accès, si ce n'est par des moyens externes, non invasifs. De plus, jusqu'à ces dernières années, nous ne disposions pas pour l'épilepsie de modèles animaux pertinents (n'en déplaise à certains, le fait de disposer de modèles animaux permet d'avancer très vite dans la connaissance de certaines pathologies). Ces obstacles expliquent que l'on n'ait que peu progressé dans la connaissance de la physiopathologie de cette maladie. La découverte de gènes de prédisposition à certaines formes d'épilepsie peut fournir l'occasion de repérer un fil conducteur qui nous permette de comprendre un peu mieux le(s) mécanisme(s) à l'origine de ces affections.
- Le second aspect est proprement thérapeutique. L'identification d'un gène qui joue un rôle dans telle ou telle pathologie – M. Kahn en a évoqué quelques-uns – peut permettre de concevoir de nouvelles cibles dont on espère, dans une vision évidemment très discutable, pour ne pas dire parfois caricaturale, qu'elles seront plus spécifiques, ce qui signifie que les médicaments seront plus efficaces et moins toxiques. Ces perspectives sont sous-tendues par une vision largement idéalisée, voire totalement fautive de la situation future. Il n'en est pas moins vrai que dans l'ensemble des approches possibles des maladies, la génétique peut ouvrir la voie à des cibles et des molécules nouvelles, auxquelles nous n'aurions pas pensé ou même que nous n'aurions pas pu trouver par les méthodes classiques.

Les génomes de deux individus donnés pris au hasard diffèrent l'un de l'autre une fois tous les mille gènes, ce qui fait au total plusieurs millions de différences possibles. Le problème consiste à rechercher, parmi ces différences, celles qui sont pertinentes et celles qui ne le sont pas.

Pour certains aspects des techniques de séquençage et de génotypage, les méthodes à grande échelle sont un complément indispensable : il est impossible, par les moyens conventionnels, de scanner rapidement un génome ou comparer plusieurs génomes. Pour tenter d'approcher tous les niveaux d'organisation de la cellule et du

vivant, les nouvelles méthodes de la biologie ambitionnent de fournir ni plus ni moins qu'une image instantanée du génome en fonctionnement : il s'agit par exemple de photographier de tous les ARN messagers présents dans l'ensemble d'une cellule à un moment donné. Ce type d'étude est la transcriptomique. Tout en progressant vite et bien, elle reste encore loin de satisfaire toute la curiosité du chercheur. Par exemple, nous sommes encore incapables, aujourd'hui, d'observer l'un des éléments probablement très importants du fonctionnement des eucaryotes supérieurs : l'épissage alternatif de l'ARN messager, c'est-à-dire le fait que l'on puisse, à partir d'un seul gène, engendrer plusieurs ARN messagers, donc plusieurs protéines pouvant exercer des fonctions très différentes, légèrement différentes, voire antagonistes. L'étude de ces événements est un domaine nouveau, qui pourrait conduire à l'apparition d'un néologisme « épisomique ». Il s'agirait par exemple d'embrasser tous les événements d'épissage alternatif d'une manière globale dans une cellule et d'en étudier, à un niveau supérieur, la mécanique interne.

Puis, en aval, le biologiste a l'ambition d'aborder de manière globale l'ensemble des protéines présentes à un instant donné dans une molécule, une cellule ou un organe, à travers ce que l'on appelle la protéomique à grande échelle, domaine que nous commençons juste à aborder. Tout récemment, par exemple, des travaux réalisés dans une sous-fraction de cellule végétale ont permis de comparer le bilan des protéines effectué par nos méthodes actuelles dans cette sous-fraction au bilan global. On s'est aperçu que 5 à 10 % seulement des protéines présentes avaient été identifiées par les méthodes classiques. On ne voit qu'une très faible partie de ce qui existe réellement, ce qui n'a rien d'étonnant pour les biologistes qui savent que la complexité du monde des protéines est supérieure d'un ou de deux ordres de grandeur à celle du monde de l'ADN, molécule universelle qui, comme chacun le sait, se laisse aisément manipuler (lorsque l'on a travaillé à la paillasse, on sait que l'on fait un peu ce qu'on veut avec l'ADN). Les protéines représentent un monde de difficultés ; il en existe même certaines classes que l'on ne sait pas encore correctement isoler.

Cependant, l'ambition reste la même : comment appréhender la complexité de l'ensemble du monde des protéines dans une cellule, un organe ou un organisme ? Pourquoi le fait-on et pourquoi essaie-t-on d'arriver à cela ? L'objectif n'est pas d'arriver à un niveau de performances supérieur, mais simplement d'approcher le métabolisme et la vie de la cellule dans toute sa complexité, toute son interactivité, dans tout un réseau qui interagit en permanence.

Ces nouvelles approches ont déjà des conséquences cliniques, même si elles sont encore très modestes. Je citerai un exemple de transcriptomique appliqué à la cancérologie (secteur qui est le mien et auquel je me réfère toujours). Les tumeurs sont classiquement caractérisées par leur aspect sous le microscope, sur une lame ; c'est l'anatomopathologiste ou le cytologiste qui définit le type de la tumeur et la sous-classe à laquelle elle appartient. Le traitement qui en découle et l'attitude du

médecin, souvent codifiés par des protocoles internationaux, découlent de ce diagnostic morphologique.

Or les cliniciens savent bien que pour deux patients atteints de tumeurs apparemment identiques ou presque identiques aux yeux du microscopiste, de l'anatomopathologiste, au moins en première approximation, le résultat du traitement peut se révéler très différent d'un cas à l'autre, comme il l'est d'une classe de tumeurs à l'autre.

Cette hétérogénéité a conduit à se demander si l'on pouvait, grâce à ces méthodes, séparer en plusieurs sous-groupes ces tumeurs apparemment homogènes. On y est effectivement parvenu. Le point le plus remarquable, et qui vient confirmer la pertinence de la classification « transcriptomique », est que le pronostic et les résultats du traitement diffèrent nettement d'un sous-groupe à l'autre. Par exemple, dans un sous-groupe de lymphomes identifié par cette technique, le taux de réponse au traitement atteint 80 à 85 % des cas, tandis que dans l'autre les résultats sont bien plus médiocres. Ces résultats comportent deux conséquences :

- il faut redéfinir la classification des maladies,
- il faut affiner l'approche thérapeutique et définir des stratégies nouvelles dans les formes qui ne répondent pas au traitement.

Voilà un rapide tableau des développements actuels qui, selon moi, doivent aboutir inévitablement à une révision de l'organisation de la recherche en biologie que l'on a connue. Pour paraphraser un aphorisme célèbre, je dirais que désormais la biologie est devenue une affaire trop sérieuse pour être laissée entre les mains des seuls biologistes. J'aimerais insister sur le rôle crucial que doit jouer le CNRS pour mettre en place une vraie politique permettant de faire travailler ensemble des biologistes, des informaticiens, des mathématiciens... Ce n'est pas facile car elles n'utilisent pas le même langage et n'ont pas le même mode de pensée ; ce n'est pas simple à réaliser, même s'il est simple de l'énoncer. De ce point de vue, on a vu apparaître aux États-Unis des initiatives intéressantes. C'est d'ailleurs l'une des bases conceptuelles des génopoles, qui visent en effet à rapprocher ces cultures différentes ; je ne sais pas encore si nous réussissons mieux qu'ailleurs, mais cela constitue en tout cas l'une des missions.

Enfin, on a souvent oublié, en France, que les entreprises de biotechnologies, issues de nos laboratoires, développent de la recherche fondamentale et participent à ce grand mouvement d'avancée de la science, de même que par ailleurs les industriels. Il faut maintenant pallier cette négligence qui a duré une génération. Un certain nombre des mesures qui ont été prises me semblent aller dans le bon sens, même si celles-ci me paraissent être encore trop ponctuelles, ciblées d'une manière très précise, alors que ces entreprises de biotechnologies nécessitent tout un environnement, à différents moments de leur vie, leur permettant de s'épanouir et d'affronter la compétition mondiale de manière correcte.