

La génétique : une brève histoire en pleine expansion

Jean-Yves LE GALL *

En 1820 par lettres patentes, Louis XVIII regroupait dans notre compagnie l'ancienne Société Royale de Médecine et l'Académie de Chirurgie, mettant ainsi fin à un certain antagonisme historique entre médecins et chirurgiens et ouvrant la voie à un abord rationnel de la pathologie qu'elle soit interne ou externe comme on le disait à l'époque. Notons au passage que jusqu'au 1^{er} janvier 2015 l'Académie de Médecine est restée sous le seul régime juridique et administratif de ces lettres patentes ; mais depuis cette date nous sommes régis par un nouveau statut, semblable à celui de l'Institut de France.

Le début du XIX^e siècle marque l'amorce d'une science médicale dont l'évolution, aujourd'hui très accélérée, n'a jamais été régulièrement continue, mais au contraire a procédé par paliers successifs, plus ou moins intriqués dans le temps. Depuis 1820 trois grandes phases peuvent ainsi être distinguées :

La première phase est celle de la méthode anatomo-clinique, initiée par Laennec et l'invention du stéthoscope ; elle a conduit, en particulier avec Trousseau, Bretonneau, Charcot et bien d'autres, à l'établissement de la séméiologie, à la description des maladies et à leur classification, c'est-à-dire à leur nosologie.

La deuxième phase débute en 1865 avec la publication de *l'Introduction à l'étude de la médecine expérimentale* de Claude Bernard, fondateur de la médecine scientifique. Élève de François Magendie dont le portrait orne le bureau du président de notre Académie, Claude Bernard fut en effet l'initiateur de la physiologie ; il suffit de citer à son actif : la fonction glycogénique du foie, le diabète, les propriétés du suc pancréatique, la pharmacologie du curare, la définition du milieu intérieur et de l'homéostasie, la fonction respiratoire du sang, le système vaso-sympathique.

Une troisième phase est celle de l'ère pastoriennne. Rappelons que Louis Pasteur était chimiste de formation, raison pour laquelle il fût admis dans cette compagnie en qualité de membre libre, et qu'au début de sa carrière il fût l'initiateur de la stéréochimie avec la description des isomères dextrogyres et lévogyres, avant d'être le fondateur de la microbiologie moderne, conduisant à la description des maladies infectieuses, à la sérothérapie, au développement des vaccinations, aux règles d'antiseptie et d'aseptie, avant l'apparition des sulfamides et des antibiotiques.

La quatrième phase, que nous vivons aujourd'hui, au-delà de multiples progrès médicaux, est celle de la génétique moléculaire au sens large du terme et de ses outils

* Président de l'Académie nationale de médecine

techniques. Permettez-moi d'en rappeler rapidement son histoire, histoire récente puisqu'il est habituel de considérer son émergence aux environs de 1870.

En tant que discipline scientifique reconnue, la génétique elle-même n'est apparue qu'au début du xx^e siècle, ses prémices remontant au xix^e avec en 1859 la publication par Darwin de sa théorie sur l'évolution et son texte *l'Origine des espèces*, et en 1865 celle des lois de Mendel. Soulignons que, dans son monastère de Brunn, le moine Grégor Mendel visait à élucider les lois gouvernant l'origine et la formation des hybrides végétaux. Travaillant sur des variétés de pois, en raison de leurs caractères phénotypiques clairs et distincts, tels que la couleur et la forme, il démontre que les facteurs héréditaires existent par paires et qu'ils se comportent de manière indépendante, se réunissant ou se séparant au hasard des générations. Il introduit surtout de façon précise les concepts majeurs de dominance et de récessivité.

Les travaux de Mendel sont donc restés inconnus jusqu'au début des années 1900, à une époque où l'agriculture devait s'améliorer et s'adapter aux conséquences du développement industriel, de l'urbanisation et de sa population grandissante. Il fallait pouvoir « manipuler » les variétés horticoles et agricoles existantes et en créer de nouvelles plus productives, éventuellement plus résistantes aux maladies, au vent, à la sécheresse. Mes goûts personnels pour le jardinage et le travail de la terre me conduisent, bien entendu à me féliciter du rôle de la biologie végétale dans l'essor de la génétique.

Cette coopération, génétique et agriculture, s'est développé aux États-Unis, en Grande-Bretagne, en Scandinavie et en Russie impériale. Alors que notre pays avait joué un rôle primordial et déterminant lors des précédentes phases du développement médical et scientifique, il est resté largement absent dans les premières étapes de la génétique ; une grande proportion de biologistes français était en effet attachée, et ceci jusque dans les années 40, à l'influence du « lamarkisme », c'est-à-dire schématiquement à la notion d'une transmission aux descendants de certaines modifications acquises par les êtres vivants au cours de leur existence, donc en définitive à une certaine hérédité acquise .

Cette méfiance vis-à-vis de la génétique mendélienne peut s'expliquer, au moins en partie, par des raisons politiques et culturelles ; ce sont sans doute ce même type de raison qui explique encore aujourd'hui le rejet obscurantiste des OGM. Le plus remarquable exemple des interactions entre système socio-politique et génétique fût à partir des années 30 en Union Soviétique, l'influence du biologiste Lyssenko. Celui-ci fit de la génétique, représentée par les travaux de Mendel et ceux de l'américain Thomas Morgan, qui avait démontré chez la drosophile l'existence de mutations spontanées, une science bourgeoise et capitaliste. Lyssenko était entre autres défenseur du processus de « vernalisation » : les graines de blé précocement germées puis exposées au froid, deviennent prétendument résistantes au gel et peuvent ainsi être semées plus tôt que les graines non traitées. Pour la petite histoire et à titre personnel je garde le souvenir d'avoir eu en PCB un éminent professeur de biologie végétale, membre officiel du parti communiste, qui vantait aux étudiants les mérites des théories de Lyssenko.

Avant d'arriver à la période contemporaine, quelques dates me semblent intéressantes à rappeler :

1869 : découverte de l'acide désoxyribonucléique par Miescher.

1879 : découverte des chromosomes par Walter et Remming.

1902 : description de l'alcaptonurie par Garrod (déficit en homogentisate oxydase).
Il s'agit d'une date majeure car apparait le concept d'erreur innée du métabolisme.

1944 : démonstration que l'ADN est le support biochimique des caractères héréditairement transmis.

1953 : est également une date majeure, celle de la publication par Watson et Crick de la structure de l'ADN en double hélice.

1958 : Jérôme Lejeune (élève de Raymond Turpin) décrit la trisomie 21.

Les années qui suivirent virent l'élucidation du mécanisme de réplication de l'ADN, c'est-à-dire la façon dont une cellule transmet l'intégralité de son patrimoine génétique à ses deux cellules filles ; celui également du mécanisme de biosynthèse des protéines, au centre du dogme central de la biologie moléculaire (ADN-ARN- Protéine), dont les artisans furent des chercheurs européens au premier plan desquels François Jacob, Jacques Monod et François Gros. Une autre avancée majeure fut dans les années 1963/64 le déchiffrement du code génétique (Nirenberg, Khorana et Ochoa), c'est-à-dire la correspondance entre le langage de l'ADN formé des seules quatre lettres ATGC désignant les quatre désoxyribonucléotides constitutifs et le langage des polypeptides constitués de vingt acides aminés différents : chaque acide aminé est programmé au niveau de l'ADN par la succession de trois bases ou codon. Ainsi en 1965 les éléments de base de la nouvelle biologie moléculaire étaient établis : nature du message héréditaire et mode de traduction de ce message en protéine.

Malgré ces progrès spectaculaires, la biologie moléculaire se heurtait néanmoins à un obstacle technique majeur : il était impossible d'analyser un gène de cellule eucaryote ; cette analyse impliquait en effet préalablement l'isolement de ce gène, c'est-à-dire d'un fragment de quelques centaines ou de quelques milliers de bases, à partir d'un génome entier, formé dans le cas du génome humain, par la succession de plus de 3 milliards de nucléotides. Ce que l'on peut qualifier de révolution, celle du génie génétique, leva ces obstacles au début des années 70, avec la mise au point d'un certain nombre de méthodologies et d'outils technologiques, dont je citerai les plus importants :

- endonucléases de restriction, enzymes des bactéries, découpant l'ADN en un nombre précis de fragments et ceci de façon reproductible (et non plus au hasard comme le font les nucléases du suc pancréatique) ;
- transcriptase inverse, enzyme des virus oncogènes à ARN capable de transcrire une séquence d'ARN en séquence d'ADN qualifiée alors de complémentaire ;
- technique de Southern, visualisant après séparation électrophorétique des frag-

ments d'un ADN coupés par une endonucléase une séquence génique particulière par hybridation avec une sonde spécifique complémentaire ;

— technique de séquençage de l'ADN par la méthode de Sanger et Coulson en 1975, puis par la méthode de Maxam et Gilbert en 1977.

Par ailleurs en 1972 le laboratoire de Paul Berg, à l'université de Stanford, était parvenu à recombiner des fragments d'ADN appartenant à deux virus différents (bactériophage lambda et virus de singe SV40), et à les amplifier dans la bactérie *Escherichia Coli* ; pour la première fois un être vivant provenant de deux ADN différents avait ainsi été construit au laboratoire. Cette recombinaison *in vitro*, volontiers qualifiée de « manipulation génétique », suscita de multiples craintes, et un débat sociétal et médiatique important, reflet peut-être de l'ambiance idéologique post 68. Ce débat fini par s'éteindre en quelques années ; il en demeure peut-être quelques échos dans le mouvement anti OGM actuel. Il apparut en effet rapidement qu'aucune des craintes évoquées n'était fondée ; tous les secteurs de la biologie ont alors fait appel aux techniques du génie génétique, conduisant à une transformation de la biologie médicale et à l'apparition d'une nouvelle discipline : la génétique moléculaire médicale, dont je rappellerai les premières dates :

1976 : premier diagnostic prénatal de drépanocytose, en utilisant les polymorphismes de restriction associés au gène de la globine ;

1977 : début du clonage des gènes humains : bêta globine et hormone lactogène placentaire, puis insuline, somatostatine, etc. ;

1978 : première banque génomique humaine en phages, c'est-à-dire une bibliothèque de clones représentant a priori l'ensemble d'un génome.

La découverte des polymorphismes de restriction, c'est-à-dire des variations individuelle des génomes portant sur un nucléotide (maintenant qualifiées de SNP pour « Single Nucléotide Polymorphism ») et entraînant l'apparition ou la disparition d'un site de restriction d'une endonucléase donnée, ainsi qu'en 1980 le clonage de la première sonde anonyme polymorphe, initièrent les travaux de cartographie du génome. La multiplication des sondes anonymes polymorphes et leur utilisation dans les études familiales conduisit ainsi à construire chromosome par chromosome des cartes de liaison génétique. Il s'agissait bien sûr d'un énorme travail qui n'a pu être réalisé que grâce à une coopération internationale (symposium « Human Gene Mapping » devenu ensuite « HUGO » se réunissant tous les deux ans) et au développement de programmes informatiques nouveaux (LINKAGE, M-LINK...).

Ce travail de cartographie génétique était également à la base de la stratégie de génétique dite inverse (inverse car il s'agissait pour identifier la cause d'une maladie d'abord de localiser et de caractériser le gène et seulement ensuite la protéine). Cette stratégie, également qualifiée de clonage positionnel, a connu ses premiers succès en 1987 avec l'identification de gènes et de protéines jusqu'alors inconnus : gène RB du rétinoblastome et gène DMD (dystrophine) de la myopathie de Duchenne, et en

1989 le gène CFTR de la mucoviscidose. Le rythme d'identification des gènes responsables de maladies génétiques s'est ensuite accéléré en parallèle avec les progrès technologiques ; leur nombre est aujourd'hui de plusieurs milliers.

À cette époque, l'explosion des progrès concernait également les techniques de génie génétique : nouveaux vecteurs pour le clonage de grands fragments d'ADN comme les BACs (Bacterial Artificial Chromosome) et les YACs (Yeast Artificial Chromosome) qui serviront de supports pour l'établissement d'une carte non plus génétique mais physique du génome ; marquage par FISH c'est-à-dire par fluorescence des gènes sur les chromosomes, ; inactivation des gènes et création d'animaux dits « Known Out » ; mutagenèse dirigée ; inactivation génique par interférence d'ARN (RNAi) ; synthèse d'oligosondes nucléotidiques ; puces ADN, micro et macro-arrays, etc.

Parmi toutes ces innovations, l'une d'entre elles a eu une importance exceptionnelle : la technique de PCR (« Polymerase Chain Reaction »), décrite en 1988 par Kary Mullis (prix Nobel de chimie en 1993) grâce à la découverte et à l'utilisation d'une endonucléase thermostable (75-80 degrés) isolée des bactéries vivant dans les sources chaudes du parc de Yellowstones. Cette technique permet de multiplier le nombre d'exemplaires d'un petit fragment d'ADN, à condition de pouvoir hybrider par des oligosondes les séquences situées au voisinage de la zone à amplifier. D'une extrême sensibilité (amplification de plusieurs millions de fois), elle est en particulier à la base des techniques de diagnostic prénatal et préimplantatoire ; elle est également responsable du succès de ce qu'on appelle les « empreintes génétiques » dont les utilisations médico-légales sont multiples. Rappelons qu'il s'agit de l'amplification des microsattellites, c'est-à-dire de structure génomiques correspondant à la répétition un nombre variable de fois d'un même motif formé de 3 à 5 nucléotides. Avec d'autres structures répétées plus grandes comme les VNTR (« Variable Number of Tandem Repeat ») et les CNVR (« Copy Number Variants »), ces microsattellites ont fait apparaître deux concepts supplémentaires : celui de polyaléélisme sur un même locus et celui de polymorphisme de taille, ce qui revient à dire en définitive que nous n'avons pas tous exactement la même taille de génome.

Dès le début de l'identification des gènes des maladies héréditaires monogéniques, il était donc apparu que le clonage positionnel demandait une connaissance aussi fine que possible de la cartographie du génome. En France deux organismes ont joué un rôle déterminant : le CEPH (centre d'études du polymorphisme humain) de Jean Dausset et l'AFM, et les efforts engagés ont conduit à la description des premières cartes génétiques et physiques complètes du génome dans les années 93-95.

Mais en même temps se dessinait la perspective du séquençage systématique des génomes ; en 1995 premier génome séquencé, celui d'un procaryote « *Hemophilus Influenzae* » ; en 1996 séquence du premier génome eucaryote, celui de la levure « *Saccharomyces Cerevisiae* » ; en 1998 premier génome métazoaire : celui du ver « *Caenorhabditis Elegans* ». Notons également qu'en 1997 avait lieu la publication du premier clonage reproductif d'un mammifère (la brebis Dolly) obtenu par transfert d'un noyau de cellule somatique dans un ovocyte énucléé.

Organisé dans le cadre d'une coopération internationale de vingt laboratoires dont douze aux États-Unis et un en France le projet de séquençage du génome humain ou HGP a aboutit en 4-5 ans, et pour un coût relativement modeste de 2,7 milliards de dollars, à la publication d'une première ébauche en 2001, suivie d'une version finie en 2004-2008. Sur le plan technique la méthode utilisée a toujours reposée sur la réaction de Sanger aux didésoxynucléotides, mais bien entendu a entraîné des efforts importants d'automatisation, préludes à l'évolution actuelle des séquenceurs. Une surprise importante fût que l'analyse de la séquence génomique obtenue par des programmes informatique performants ne fit apparaitre, contrairement aux estimations qui avaient été faites, qu'un nombre de gènes inférieur à 25 000, c'est-à-dire à peine supérieur à celui d'espèces animales très inférieures dans la classification, et très inférieur à celui de beaucoup d'espèces végétales.

L'évolution contemporaine des techniques de séquençage illustre le caractère imprévisible du progrès technologique : à partir des années 2005 est apparue une nouvelle génération de séquenceurs dits à haut débit, dont les performances sont en amélioration constante et ceci de façon tellement rapide qu'elles mettent en défaut la fameuse loi de Moore ; cette loi établie pour le développement des microprocesseurs prévoit un doublement de leur puissance et une division des coûts par deux tous les dix-huit mois ; dans le cas des séquenceurs nous sommes bien au de là de cette vitesse d'évolution.

Les méthodes de séquençage classiques des ADN de grande taille sont basées sur une stratégie de shotgun, c'est-à-dire de leur coupure au hasard en petits fragments, suivie de leur clonage en vecteur phagique avant leur séquençage. La nouvelle génération d'appareils s'affranchit de cette contrainte grâce à l'ultraparallélisation des réactions de séquençage. Pour reconstituer la séquence originelle l'assemblage du nombre considérable de séquences obtenues nécessite un traitement bioinformatique extrêmement performant ; des programmes informatiques sont également nécessaires à l'interprétation fonctionnelle de la séquence assemblée, ils sont encore aujourd'hui loin d'être parfaits. Enfin le stockage des données est une difficulté majeure : un génome de 3 milliard de nucléotides transcrit en caractères typographiques mêmes petits occuperait l'équivalent de 200 volumes d'annuaire téléphonique.

En termes de diagnostic médical, l'analyse du génome peut se faire suivant plusieurs possibilités : séquençage d'un, de quelques ou d'un groupe de quelques dizaines ou centaines de gènes potentiellement en cause dans un type de pathologie ; séquençage de l'exome c'est-à-dire de l'ensemble des régions codantes mais ne représentant pour autant qu'environ 2 % du génome ; séquençage du génome total. Cette analyse est aujourd'hui accessible en quelques heures pour un prix de quelques centaines ou quelques milliers d'euros ; ces chiffres sont à rapprocher de ceux concernant le premier séquençage d'un génome (plusieurs années, 2,5 milliards de dollars) pour percevoir les progrès phénoménaux réalisés en 10 ans.

Cette évolution technologique a et aura des conséquences majeures sur la pratique médicale. Trois d'entre elles méritent déjà d'être soulignées :

- La capacité de rechercher simultanément les mutations dans un grand nombre de gènes est en définitive un bénéfice considérable en termes de temps et de coût pour le diagnostic des maladies mendéliennes. Il est probable que ces possibilités soient étendues aux maladies multifactorielles lorsqu'aura été identifiée la totalité des variants génétiques de prédisposition.
- La possibilité d'analyser l'ADN fœtal circulant dans le sang maternel dès la cinquième ou sixième semaine d'aménorrhée et donc de faire de façon non invasive, sans ponction de liquide amniotique et donc sans prendre le risque de provoquer un avortement, le diagnostic d'aneuploidie comme la trisomie 21. Il est probable que ce séquençage de l'ADN fœtal circulant sera également rapidement utilisé pour le dépistage des maladies mendéliennes.
- La caractérisation des spectres mutationnels tumoraux dans les affections cancéreuses héréditaires ou non, susceptibles de conduire à des traitements reposant sur un ou sur une combinaison de médicaments moléculairement ciblés. Le premier exemple de ce type de traitement a été l'Imatinib inhibant spécifiquement l'activité tyrosine/kinase de la protéine de fusion bcr/abl dans la leucémie myéloïde chronique.

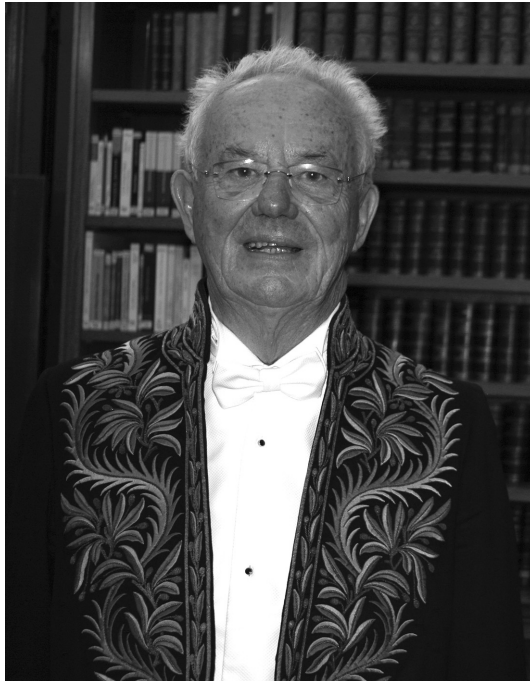
Le séquençage à haut débit non seulement en raison de ses performances et des développements multiples à en attendre, mais aussi du fait de sa capacité à mettre à jour de multiples informations génétiques concernant à la fois le sujet testé et ses apparentés, soulève tout une série de questions médicales, éthiques et réglementaires. Elles concernent notamment :

- le dépistage néonatal systématique des maladies génétiques qui dans notre pays est phénotypique et ne concerne que cinq affections.
- le dépistage éventuel des hétérozygotes pour les maladies mendéliennes
- l'analyse de l'ADN fœtal du sang maternel à des fins autres que médicales
- l'analyse de l'ADN tumoral qui nécessite également une analyse de l'ADN constitutionnel.
- la garantie de la confidentialité des énormes quantités d'informations génétiques ainsi recueillies.

L'histoire de la génétique médicale que je viens de parcourir, sans doute de façon trop rapide et trop schématique, montre la rapidité exponentielle d'évolution de nos connaissances scientifiques et médicales, des potentialités thérapeutiques qui en découlent, mais aussi des questions philosophiques qu'elle pose. L'exemple le plus récent de ces sauts évolutifs est celui de la caractérisation en 2011 du système CRISPR-cas9 ; il s'agit d'un système d'endonucléases bactériennes ayant pour objet de lutter contre les infections par les bactériophages ; transposé et utilisé en biologie animale ce système permet de modifier un génome de façon ponctuelle (remplacement d'un nucléotide par un autre) et ciblée. En envisageant de réparer directement le gène muté, cette technologie améliore considérablement les possibilités de la

thérapie génique somatique, qui malgré quelques succès remarquables comme ceux de notre confrère Alain Fischer, sont restés limitées ; mais elle offre également et théoriquement la possibilité d'une thérapie germinale (ce qui a été très rapidement fait chez l'animal). Cette intervention ciblée peut bien sûr être envisagée pour corriger toute anomalie génétique, mais aussi éventuellement pour améliorer le patrimoine génétique des individus et des espèces ; les applications en biologie végétale et en agriculture devraient être rapides et nombreuses. Nous sommes ainsi revenus, après des progrès scientifiques considérables, aux questions d'eugénisme qui, au début du xx^e siècle faisaient l'objet d'un large consensus dans le milieu des généticiens, essentiellement dans les pays anglo-saxons et d'Europe du Nord.

Arrivé au terme actuel de cette histoire, je me garderai de la conclure car l'évolution technologique, conceptuelle et médicale de la génétique se poursuit très rapidement ; si les circonstances m'amenaient à refaire un tel exposé, nul doute qu'il comporterait quelques pages supplémentaires. Ma conclusion sera d'une nature différente. L'évolution scientifique, en particulier de la génétique, a de multiples implications sociétales, éthiques, réglementaires, philosophiques ou religieuses. Il me paraît important que le corps médical n'y soit pas systématiquement impliqué et qu'il soit rappelé que le rôle de notre profession est simplement de diagnostiquer, de soigner et si possible de prévenir.



Jean-Yves LE GALL