



La révolution du génome

Jean Rosa

J. Rosa : Professeur émérite à l'université Paris XII, Académie des sciences, 23, quai de Conti, 75006 Paris, France.

RÉFÉRENCES

1. Cohen F, Desanti J, Guyot R, Vassails G. *Science bourgeoise et science prolétarienne*. Paris : Les Éditions de la Nouvelle Critique, 1950.
2. Avery OT, McLeod CM, McCarty M. Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *J Exp Med* 1944 ; 79 : 137-58.
3. Kaplan JC, Delpech M. *Biologie moléculaire et médecine*. Paris : Flammarion-Médecine Sciences, 1993.
4. Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 1975 ; 98 : 503-17.
5. Maniatis T, Hardison R, Lacy E, et al. The isolation of structural genes from libraries of eucaryotiv DNA. *Cell* 1978 ; 15 : 687-701.
6. Kan YW, Dozy AM. Polymorphism of DNA sequence adjacent to human β globin structural gene : relationship to sickle mutation. *Proc Natl Acad Sci USA* 1978 ; 75 : 5631-7.
7. Saiki R, Scharf S, Faloona F, et al. Enzymatic amplification of β globin genomic sequences and restriction sites analysis for diagnosis of sickle cell anemia. *Science* 1985 ; 230 : 1350-4.

Dans un ou deux ans, la séquence du génome humain sera totalement établie et l'on sera entré de plein pied dans l'ère « post-génomique ».

Comment en est-on arrivé là ? Où en est-on actuellement ? Quelles suites peut-on, à l'heure actuelle, imaginer ? Étrange aventure que celle qu'a vécue ma génération. Jusqu'aux années 1930 (on ne peut pas se borner à remonter seulement 20 ans en arrière pour bien saisir la cinétique de cette révolution), le génome était un concept abstrait puisqu'on n'en connaissait aucun support physique. Vint Morgan et ses chromosomes et a *contrario*, dans les années qui suivirent la guerre, l'incroyable controverse déclenchée en URSS par les Mitchouriniens et autres Lissenkistes [1] pour lesquels les chromosomes n'étaient que « des artéfacts » nés de l'imagination de la science bourgeoise pour détruire la vérité vraie qui était la prééminence absolue de la transmission des caractères acquis (il faut que les jeunes générations aient connaissance de ces abominables péchés contre l'esprit). En 1944, Avery [2], suivi d'autres, montre qu'un facteur chimique isolé d'un pneumocoque virulent peut rendre pathogène un individu d'une variété non virulente ainsi que sa descendance. Le facteur est appelé acide désoxyribonucléique en raison de sa

composition élémentaire et l'aventure génomique commence.

A la même époque, l'utilisation des molécules marquées apparaît et la biochimie devient reine mais est cependant incapable d'élucider la structure des gigantesques polymères que sont les acides désoxyribonucléiques. Ce sont un biochimiste (marginal) – Crick – et Watson, un botaniste (!) qui, en 1953, génialement inspirés, délaissant les techniques biochimiques utilisent les spectres de diffraction aux rayons X et font apparaître la Pierre de Rosette de la biologie : la double-hélice [3].

Époque révolutionnaire et passionnante qui comme toute période révolutionnaire comporte des événements contradictoires.

D'un côté, un très bon mais un peu naïf biologiste croit avoir reproduit l'expérience d'Avery sur des canards. Mais leur descendance, si elle est bien « transformée », ne l'est pas, contrairement à ce qu'il croyait, du fait de l'injection d'un ADN, bien incapable de le faire (il faudra 40 ans pour arriver à ce résultat chez les eucaryotes supérieurs) mais en raison d'ébats nocturnes et clandestins de volatils irrespectueux de la recherche scientifique.

Par ailleurs, de vagues microbiologistes – J. Monod et F. Jacob – nichés dans le grenier occupé par A. Lwoff à l'Institut Pasteur ouvrent la boîte de Pandore du fonctionnement du

génomique d'*E. coli* avec la découverte du « Messenger » et de « l'Opéron ». Il s'en suit 20 ans d'une cuisine macromoléculaire qui s'autobaptise « Biologie Moléculaire » et à laquelle participent naturellement des biochimistes mais aussi des virologues, des microbiologistes, des généticiens et même des physiciens qui lâchent la physique trop impliquée selon eux dans des débats concernant l'utilisation militaire de leurs travaux.

Il va sortir de cette période fondatrice et fabuleuse des techniques clés pour le futur décryptage du génome. Il en est ainsi de la découverte et de l'utilisation des enzymes de restriction qui vont permettre de fragmenter de façon spécifique les acides nucléiques; la fabuleuse électrophorèse de Southern, l'une des publications les plus citées au monde [4], qui sépare et objective les fragments de restriction; la transcriptase inverse qui transforme le fugace ARN messenger en sa stable copie, l'« ADNc ». Les microbiologistes élucident la physiologie des plasmides, qui vont être les premiers vecteurs capables d'importer des morceaux d'ADN dans des bactéries, ouvrant la porte au clonage. C'est le prélude à toute une série de recherches, permettant l'introduction dans les cellules, d'abord procaryotiques puis eucaryotiques, de fragments d'ADN de plus en plus grands ce qui va permettre des études fonctionnelles des gènes: recherches de séquences régulatrices, une marche à grands pas sur les chromosomes et plus tard transgénèse et essais de thérapie génique. Enfin, le séquençage devient lui aussi possible grâce à Gilbert d'un côté de l'Atlantique et à Sanger de l'autre.

Le bricolage génétique devenu possible inquiète les chercheurs qui se fixent un moratoire de un an pour envisager si toutes les précautions peuvent être prises pour éviter des drames biologiques, ce qui est tout à l'honneur de la communauté scientifique mais retarde d'autant le passage aux eucaryotes et naturellement à l'homme.

Asilomar ayant permis de rassurer la communauté scientifique, les recherches reprennent et l'on commence vers les années 1975 à s'attaquer aux eucaryotes.

Première surprise quand on s'attaque aux eucaryotes, contrairement au

génome procaryotique, le leur est discontinu, les parties codantes ne représentent qu'une très faible partie de leur génome et de plus les gènes eux-mêmes sont fractionnés par des séquences non codantes, les introns dont on ne saisit pas initialement bien le rôle [3].

Puis le génome de l'hémoglobine sert de prototype: les gènes des chaînes connues sont localisés chromosomiquement, clonés, et l'ADN de la chaîne β de lapin inaugure le séquençage des gènes eucaryotes [5]. Au passage, on découvre les gènes des chaînes embryonnaires, inconnus jusqu'alors, ainsi que des pseudo-gènes. L'application à l'homme devenait possible et dans ce qui suivit le génome humain normal et/ou pathologique devint leader dans la révolution génomique avec une dialectique très fournie entre recherches de base et pathologie génétique et moléculaire. Ce furent d'abord les besoins en diagnostic prénatal qui tirèrent les recherches. Le polymorphisme de restriction du gène β de l'hémoglobine découvert par Y.W. Kan [6] permit, dans les débuts des années 1980 un diagnostic prénatal ultraprécoce de la drépanocytose avant de servir de marqueur anthropologique inégalé et, étendu à d'autres gènes, être l'un des outils-clés de la marche sur le génome et de l'établissement des cartes génétiques. Ce sont également les besoins nés du diagnostic prénatal qui amenèrent à la naissance d'une autre fantastique innovation, la PCR (*polymerase chain reaction*) [7], qui permet la reproduction à l'infini, *in vitro*, de la moindre séquence nucléique et va devenir un outil universel de recherche en biologie.

Toutes ces possibilités ouvertes par la biologie moléculaire « interpellent » alors les chercheurs en génétique médicale et les pionniers de la pathologie moléculaire. Les maladies monogéniques deviennent des cibles privilégiées. Tout d'abord celles dont on connaît la protéine normale ou déficitaire, et dans ce domaine on découvre avec une certaine surprise que, moléculairement parlant, ces maladies sont en fait des syndromes pouvant être causés par divers événements génétiques. On découvre aussi, par exemple à propos du gène de la myoglobine [8] puis de celui de l'insuline, qu'ils sont flanqués par des

séquences répétitives de longueur variable selon les individus. Ces microsatellites vont servir d'outils à la médecine légale en étant le support des « empreintes génétiques » et aussi constituer un nouveau type de marqueur pour l'établissement des séquences génomiques.

Un autre *challenge* va ensuite, se présenter, celui d'identifier la lésion génétique d'une maladie dont on ne connaît pas la lésion moléculaire phénotypique. La génétique inverse, dénommée maintenant génétique positionnelle, va aboutir à ce résultat, en isolant et en caractérisant en première intention le gène lésé. Cela fut inauguré à propos du gène responsable d'une des dystrophies musculaires [9]. La séquence protéique déduite permet ensuite d'isoler une protéine-clé de la contraction musculaire, protéine totalement inconnue jusqu'alors: la dystrophine. Cette nouvelle stratégie extrêmement ingénieuse mais très lourde, inaugure les incontournables collaborations entre généticiens, cliniciens, biologistes moléculaires, cytogénéticiens, informaticiens et associations de malades. Elle permet ensuite l'isolement des gènes d'autres maladies monogéniques dont, entre autres, celui de la mucoviscidose.

Si les recherches en génétique moléculaire ne permirent pas d'avancer aussi rapidement dans le domaine des maladies polygéniques, elles ont fourni par ailleurs des résultats très importants par exemple dans le cas du syndrome du X fragile où J.L. Mandel démontra l'implication dans la maladie de séquences répétées en nombre anormal et sujettes à des méthylations s'accroissant de génération en génération ce qui entraîne l'aggravation jusqu'alors inexpliquée des signes cliniques [10].

Ce nouveau concept de physiopathologie moléculaire permet ensuite d'élucider rapidement les lésions moléculaires de nombre d'autres maladies génétiquement déterminées dont la fameuse chorée de Huntington dont le gène était localisé depuis des années sans qu'ait été identifiée la lésion génétique.

Les avancées technologiques et conceptuelles dérivées des recherches en génétique furent naturellement mises à profit pour les recherches sur les affections malignes. Découvertes

des oncogènes et des anti-oncogènes, de chimiokines, de gènes impliqués dans l'apoptose se succèdent. Il en résulte d'importantes avancées des connaissances sur les mécanismes physiopathologiques à l'origine de nombreuses tumeurs, solides ou leucémiques, sur certaines méthodes diagnostiques ou thérapeutiques à leur propos sans que l'on n'ait cependant réussi à transformer encore radicalement le statut général de ces maladies. Parallèlement, les recherches de base se poursuivaient, fournissant de nouveaux outils d'un intérêt immense telles que la mutagenèse dirigée, la recombinaison homologue et plus récemment les « puces ».

La mutagenèse dirigée est un processus qui implique la modification *in vitro* d'une ou plusieurs bases d'un ADNc qui est ensuite recombiné dans des cellules hôtes. Cette technique l'emporte de très loin sur la recherche fastidieuse et aléatoire de mutants naturels. Elle permet de synthétiser une protéine mutée dont on peut étudier les éventuelles perturbations qu'elle induit dans des cellules productrices ou *in vitro*, les corrélations structures / fonctions. Elle est naturellement aussi utilisée pour la création de modèles animaux.

Les possibilités de transgénèses chez l'animal supérieur qui se sont rapidement développées ont créé une situation quasi révolutionnaire en recherche biologique et médicale en permettant, associée à la recombinaison homologue également rendue possible, de détruire ou de remplacer un gène d'intérêt chez l'animal.

L'industrie ne s'y est pas trompée qui fait un effort gigantesque de reconversion dans ce domaine tant en ce qui concerne la production de molécules recombinées génétiquement que pour ses tests pharmacologiques. Ces possibilités viennent d'ailleurs de connaître un nouvel et considérable incrément en raison du développement de la technique des « puces ». Sur le modèle des plaques utilisées pour les radio- ou enzymo-immuno-essais, on greffe des rangées de sondes nucléiques. En y adjoignant les principes de la chimie combinatoire, on arrive à des dispositifs entièrement automatisés capables de pratiquer des milliers de tests par jour (ou de synthétiser de multiples oligonucléotides) [11].

Cette méthode est naturellement aussi utilisée en recherche de base où par exemple elle permet de cribler tous les messagers produits par une cellule ou un organe à un instant donné fournissant un outil d'une puissance vraiment extraordinaire par exemple en biologie du développement ou de la différenciation pour ne pas parler de l'exploration physiopathologique.

Pour en revenir à l'industrie pharmaceutique, l'apparition de ces nouvelles possibilités dues aux progrès en génomique l'amènent à de profondes mutations stratégiques. Il en est ainsi d'une diminution considérable de ses besoins quantitatifs en expérimentation animale cependant qu'au contraire se développent ses besoins en modèles animaux.

Elle est également très attentive aux progrès des connaissances génomiques en ce qui concerne la pharmacogénétique qui commence à indiquer des comportements génétiquement déterminés vis-à-vis de certains produits.

Pendant que ces divers développements étaient appliqués aux maladies, à la connaissance de leur pathogénie, de leur diagnostic et de leur traitements, les recherches fondamentales sur les cartes génétiques d'abord [12], puis sur la séquence du génome progressaient elles aussi à une vitesse vertigineuse. Une première avancée notable fut obtenue dans les années 1990 grâce à des perfectionnements des moyens techniques déjà existants, à une semi-industrialisation des *Southern blots*, à l'apparition de séquenceurs d'ADN de plus en plus rapides et performants. Parallèlement, on assista au développement d'une informatique très puissante dédiée à l'interprétation des séquences obtenues et à leur stockage dans de gigantesques banques de données. Il en résulte un besoin aigu en bio-informaticiens spécialisés dans ces problèmes [13]. Les cartes génétiques établies, se posait le problème de décider quelle stratégie de séquençage choisir. Certains étaient partisans de séquencer seulement les gènes en les repérant grâce à leurs « étiquettes », début de leur séquences codantes obtenues par action de la reverse-transcriptase sur leurs ARNm. D'autres plaidaient pour que l'on se lance directement sur la séquence totale du génome humain, confortés par l'obtention d'une célé-

rité inattendue de la séquence totale de *Saccharomyces cerevisiae*. Pendant un certain temps, il ne fut pas clair de discerner quelles seraient la ou les stratégies qui devaient l'emporter et de ce fait nombre de laboratoires se lancèrent à la recherche et à l'identification d'étiquettes. Cette démarche était confortée par le fait qu'elle permettait de réaliser une véritable carte fonctionnelle du génome. Elle avait aussi l'avantage de mettre à jour des messagers correspondant à des protéines inconnues. L'adhésion à cette solution était aussi nourrie par l'immensité du problème tel qu'apparaissait il y a encore quelques années la détermination de la séquence du génome entier. Pour cette dernière, la stratégie du « Consortium Saccharomyces » fut d'abord envisagée, consistant à répartir entre de nombreux groupes le séquençage de différentes parties du génome. A l'usage, il apparut rapidement que cette solution n'était pas vraiment adaptée au séquençage d'un génome de la taille de celui de l'homme. C'est finalement le concept du grand « combinat » qui s'imposa : constitution de quelques très grandes structures disposant des moyens financiers nécessaires pour acquérir un nombre impressionnant de séquenceurs rapides asservis par une informatique ad-hoc et utilisés par des personnels parfaitement spécialisés. Ces quelques combinats situés aux États-Unis, à Cambridge (Grande-Bretagne), au Japon et au Centre français de Grand Séquençage se sont répartis les différents chromosomes et l'on espère la séquence totale du génome humain pour les toutes prochaines années.

Cependant, bénéficiant des avancées suscitées par l'opération « Génome Humain », d'autres recherches progressent avec une vitesse dont la croissance est pratiquement exponentielle et ont déjà abouti à produire les séquences de nombreux génomes dont celui de la drosophile, crucial pour des développements rapides de génétique « expérimentale » et de biologie du développement. On attend dans ce secteur des possibilités totalement inconnues jusqu'alors d'identification des gènes contrôlant le développement. La définition récente du génome du BK (et de celui du BCG) ainsi que celui du tréponème pâle (cependant que ceux des plasmodiums

sont attendus pour prochainement) ouvre dès lors un espace vertigineux sur des possibilités thérapeutiques anti-infectieuses : identification de la séquence des épitopes antigéniques, leur obtention par génie génétique et potentiellement demain vaccination par injection d'ADN [14]. Inutile d'insister sur les avantages que l'on attend de la séquence génomique de la souris et en agro-alimentaire de celle de céréales comme le riz.

Bénéficiant naturellement également des retombées des opérations génome les possibilités de thérapie génique [15]. Dans ce domaine les premiers essais, contrairement à ce que pensent beaucoup, ne remontent pas à hier mais ont eu lieu il y a une quinzaine d'années à propos de β -thalassémies dans des conditions qui violaient totalement les règles éthiques les plus élémentaires et valurent à leurs auteurs de très sévères sanctions. Une dizaine d'années s'écoulèrent au cours desquelles les connaissances sur la régulation des gènes, la domestication de nouveaux vecteurs tels que des adénovirus, les progrès en matière de biologie cellulaire avec en particulier l'isolement de cellules souches et de chimiokines induisant différenciation ou division, la réussite de transgénèses avec création de modèles animaux rendirent possibles des essais rationnels de thérapie génique. Un nouvel et fascinant chapitre de recherche thérapeutique est ainsi ouvert suscitant espoirs, convoitises et débats qui interpellent beaucoup les grands comités d'éthique et aussi... les pouvoirs publics et l'industrie.

En essayant de faire du prospectif on peut raisonnablement imaginer que l'origine de toutes les maladies monogéniques sera établie avant 10 ans et avec un peu de chances cette connaissance s'étendra à un grand nombre de maladies polygéniques. En ce qui concerne les cancers il est très vraisemblable que l'on aura pu identifier la quasi-totalité des facteurs génétiques impliqués dans leur survenue et leur développement. On peut espérer une véritable révolution en psychiatrie si, comme certaines prémisses semblent l'indiquer, nombre de syndromes maniaco-dépressifs, schizophréniques ou autistiques ressortent de composantes génétiques. On peut aussi imaginer

voir surgir des maladies dont une composante génétique est actuellement totalement insoupçonnée.

Pour parvenir à cet avenir radieux il y aura beaucoup de grain à moudre pour les acteurs du postgénomique [16]. Cela concernera d'abord l'identification et la caractérisation fonctionnelle des dizaines de milliers de protéines, retombées directes de la séquence génomique : quels exemplaires à travailler en priorité ? Par quelles méthodes ? La spectrométrie de masse permettra-t-elle comme le pensent certains des identifications quasi industrielles [17] ? Où ce travail sera-t-il effectué ? Dans les laboratoires de recherche ? Dans de grands centres privés à l'image de ce qui est en train de se produire pour les séquences d'ADN ? On peut prévoir sans beaucoup de risques des besoins massifs en reconversions conceptuelles, opérationnelles et aussi éducationnelles [18]. Mais que ces recherches seront passionnantes avec ces logiciels qui pourront établir la structure dans l'espace et les interactions des macromolécules d'intérêt, aidés qu'ils seront par les développements de la RMN, de la microscopie confocale et par tant d'autres techniques révolutionnaires de physico-chimie telles que, par exemple, les biocapteurs plasmoniques. Passionnantes avec ces animaux « génomisés » qui seront de fantastiques nouveaux modèles pour les neurosciences, la biologie du développement... et la pathologie expérimentale ! au grand dam des défenseurs de la gent animale. Ces défenseurs pour lesquels les fantastiques progrès de la genèse médicamenteuse (retombée directe de la génomique) constitueront une source accrue de fureur. On a vu plus haut apparaître la pharmaceutique nouvelle et les fabuleuses possibilités dont elle dispose avec les puces associées à la chimie combinatoire. Qui aurait pu prévoir, il y a seulement quelques années la révolution copernicienne induite dans ce secteur ? criblages « taylorisés » à l'aide de ces puces d'un nombre sans cesse croissant d'effecteurs ou d'inhibiteurs de récepteurs, de médiateurs à finalisation thérapeutique, produits ensuite par génie génétique, testés sur des cellules transformées de façon ad-hoc et sur des modèles animaux humani-

sés et reproduisant la maladie à traiter, et, enfin, optimisés en fonction des données pharmacogénétiques de chaque patient qui disposerait à la limite de médicaments personnalisés !

Enfin, et pour beaucoup, quel sera l'avenir en matière de thérapie génique ? Question qui intéresse non seulement les maladies génétiquement déterminées mais aussi les affections malignes ? Il est très difficile en fonction des connaissances actuelles de se livrer à un pronostic, mais ce dont on peut déjà être sûr est que les recherches développées à ce sujet amèneront fatalement à d'importantes innovations technologiques et à d'importants progrès en ce qui concerne les connaissances de base tant en biologie cellulaire qu'en génétique moléculaire et même en physiologie.

Qu'en sera-t-il de la perception au niveau de la société des retombées de la révolution génomique ? En ce qui concerne les effets pervers, nous en avons déjà un aperçu à partir des débats qui se déroulent dans la société civile : risques éthiques multiples, craintes pour l'environnement, possibilités d'accaparement militaire pour la création d'armes toujours plus dévastatrices, perturbations majeures dans certains secteurs de production, outre celui de la santé celui de l'agronomie et peut-être aussi celui d'autres secteurs industriels comme par exemple ceux du textile ou des matières plastiques.

Il est clair que tous les progrès techniques qui s'annoncent devront être accompagnés d'un développement accru des réflexions éthiques tant en ce qui concerne les problèmes de pratique médicale, de santé publique, de contrôles démographiques que ceux, de l'expérimentation animale, de la propriété scientifique et des rapports entre recherche publique et recherche privée.

Que souhaiter ? Avec une bonne dose d'angélisme, que notre société devienne enfin de plus en plus capable de faire des analyses en temps réel pour que toutes les perturbations provoquées par la révolution génomique soient gérées de la façon la moins traumatisante possible. Rude tâche pour les chercheurs, les médiateurs, les enseignants, et les politiques... entre autres [19] ■

RÉFÉRENCES

8. Jeffreys AJ, Wilson V, Thein SL. Hyper-variable « minisatellite » regions in human DNA. *Nature* 1985; 314: 67-73.
9. Kunkel LM, Monaco AP, Middlesworth W, et al. Specific cloning of fragments absent from DNA of a male patient with X chromosomal deletion. *Proc Natl Acad Sci USA* 1985; 82: 4778-82.
10. Mandel JL, Heitz D. Molecular genetics of the fragile-X syndrome: a novel type of unstable mutation. *Curr Opin Genet Dev* 1992; 2: 422-30.
11. Hélène C. Innovations scientifiques et technologiques dans la recherche de nouveaux médicaments. *CR Acad Sci Paris* 1999; 322: 531-40.
12. Cohen D, Chumakov I, Weissenbach J. Première génération de la carte physique du génome humain. *CR Acad Sci Paris* 1993; 316: 1484-8.
13. Durbin R. *Biological sequence analysis*. New York: Cambridge University Press, 1998.
14. Symposium sur la vaccinologie. *CR Acad Sci Paris* 1999; 323: 911-1013.
15. *La thérapie génique*. Rapport de l'Académie des Sciences. Paris: Tec et Doc, 1995.
16. *Développement et applications de la génomique*. Rapport de l'Académie des Sciences. Paris: Tec et Doc, 1999.
17. Sali A. 100 000 protein structures for the biologist. *Nat Struct Biol* 1998; 5: 1029-32.
18. Jordan B. *Génétique et génome: la fin de l'innocence*. Paris: Flammarion-Médecine Sciences, 1996.
19. Genome issue. *Science* 1998; 282: 581-832.

TIRÉS À PART

J. Rosa.

