

**Chroniques génomiques**



par Bertrand JORDAN

« Feu sur le quartier général » :  
le projet Génome en balance ?

**Un programme de nouveau critiqué aux États-Unis  
dans un climat de relative pénurie.**

**Le malentendu du séquençage...**

**qui, d'ailleurs, n'est peut-être pas si idiot que cela !**

**Banques d'ADNc et normalisation, une voie prometteuse.**

**La cartographie fait à juste titre l'unanimité.**

**Mais tout n'est pas joué, et il faudra parler de l'éthique...**

Nous avons évoqué, le mois dernier, la mise en place imminente d'un programme « Génome » en France. Paradoxalement, c'est au moment où notre pays semble se décider à investir sérieusement dans ce domaine et où le programme CEE sort enfin des limbes que s'élèvent aux États-Unis — leader incontestable — des critiques très sévères laissant craindre un ralentissement des recherches en question [L. Roberts, « Tough times ahead for the Genome project », *Science* 1990 ; 248 : 1600-1]. On se souvient que la décision de lancer le projet aux États-Unis avait été prise fin 1988, après trois ou quatre années de discussions assez vives pendant lesquelles de nombreux scientifiques avaient critiqué trois aspects du programme : son objectif final (est-il vraiment utile de séquencer tout le génome, constitué à 95 ou 98 % de séquences intergénomiques sans fonction connue ?), son ampleur financière (qui risquait d'avoir pour conséquence la réduction des dépenses dans des secteurs moins à la mode mais scientifiquement plus importants) et son aspect semi-industriel faisant entrer la biologie dans l'ère de la *big science* massive et hiérarchisée, ne laissant plus de place à l'individu et à son impré-

visible inventivité. Mais, après de nombreuses discussions, il semblait que les scientifiques se soient majoritairement ralliés au projet.

En fait la controverse a rebondi, ou du moins a repris de l'intensité, ces derniers temps. Les problèmes financiers du NIH (*National Institute of Health*), et donc ceux des laboratoires de recherche qui dépendent de ses contrats, ont certainement joué un rôle important dans le regain de ces critiques. En effet, le NIH (dont le budget augmente, mais lentement) a du mal à satisfaire les très nombreuses demandes de contrat qu'il reçoit, et les laboratoires totalement dépendants de cette aide sont en difficulté, tout particulièrement les nouveaux laboratoires que démarrent de jeunes chercheurs après leur « post-doc ». Toujours est-il que l'on en est depuis le début de l'année au stade des pétitions contre le projet Génome (envoyées au directeur du NIH et à certains parlementaires), pétitions qui traduisent l'irritation de nombreux biologistes devant ce qu'ils considèrent comme un programme trop onéreux, peu intéressant scientifiquement et risquant de mettre en danger la « vraie recherche » [voir en particulier B.D. Davis *et al.*, « The human

genome and other initiatives », *Science* 1990 ; 249 : 342-3]. Au moment où la discussion porte sur le budget 1991 du NIH, qui comporte une ligne budgétaire de 108 millions de dollars pour le génome (contre 60 en 1990), ces critiques rencontrent un certain écho... d'autant plus qu'elles mettent explicitement en cause un des éléments clés du programme : les *Genome Research Centers*, destinés à concentrer un certain nombre de moyens pour s'attaquer à des tâches comme la cartographie physique complète d'un chromosome. Le financement prévu pour ces centres se chiffre en dizaines de millions de dollars par an, et une réduction significative du budget les mettrait certainement en danger.

Ces critiques ne peuvent pas être écartées sans discussion. Il est vrai que les buts scientifiques du séquençage du génome ne sont pas évidents, que cette recherche ne fait pas particulièrement appel à l'imagination (sauf, nuance importante, du point de vue de la technologie) et qu'elle coûte très cher : le budget du CEPH (Centre d'étude du polymorphisme humain, principal représentant de ce type de recherche en France) permettrait de faire vivre

plusieurs laboratoires « classiques ». Mais la controverse repose à mon avis, et en partie du moins, sur un malentendu. L'aspect le plus critiquable du programme, et celui sur lequel porte en effet la quasi-totalité des critiques, est le séquençage exhaustif du génome humain. C'est en effet un travail très onéreux, très répétitif, demandant une organisation quasi militaire et dont les retombées scientifiques prévisibles sont — peut-être — un peu légères face à l'investissement nécessaire avec les techniques actuelles (1 dollar la base en étant optimiste, soit 3 milliards de dollars pour l'ensemble). Mais ce n'est pas un objectif du programme Génome aux États-Unis dans sa phase actuelle, et le séquençage massif n'est envisagé que si et quand le coût aura baissé d'un ordre de grandeur ; l'essentiel du travail financé actuellement et à moyen terme porte sur la cartographie génétique et physique, et accessoirement sur l'amélioration des techniques de séquençage. Il y a donc, à mon avis, un gros malentendu et, parfois, peut-être un peu de mauvaise foi dans ces critiques.

Il ne faut, d'ailleurs, sans doute pas enterrer le séquençage massif aussi sommairement que je l'ai fait ci-dessus : un fait expérimental récent montre que des surprises restent possibles. L'exemple en question est fourni par le programme européen de séquençage de la levure dont la mise en place avait été entourée d'un certain scepticisme : il s'agissait — du moins dans un premier temps — de séquençer le chromosome III de la levure *Saccharomyces cerevisiae*, dans une collaboration impliquant plusieurs dizaines de laboratoires. Vu de l'extérieur, le projet ne paraissait pas impressionnant : le chromosome III mesure à peine 340 kb, tout juste le double de la taille du virus d'Épstein-Barr séquencé par un seul laboratoire, celui de B. Barrell à Cambridge (G-B) il y a plusieurs années [R. Baer *et al.*, « DNA sequence and expression of the B95-8 Epstein-Barr virus genome », *Nature* 1984 ; 310 : 207-11] ; l'argent distribué aux nombreuses équipes, divisé par le nombre de nucléotides à séquençer, donnait un prix élevé de cinq dollars la base et pouvait faire penser qu'il

s'agissait plus d'une subvention à fonds perdus que d'un projet précis ; et, surtout, la part d'inconnu semblait très faible dans ce chromosome déjà très étudié et dans lequel plusieurs dizaines de gènes avaient déjà été repérés : peut-être un séquençage systématique dévoilerait-il quelques gènes non encore décelés ou quelques régions intergéniques intéressantes ? En fait, le séquençage a révélé plus d'une centaine de gènes nouveaux, indiquant ainsi une densité génique insoupçonnée : un gène toutes les deux kilobases en moyenne. De plus, les données les plus récentes indiquent qu'il s'agit bien de gènes fonctionnels (et non de simples cadres de lecture ouverts), puisqu'il s'avère qu'ils sont pratiquement tous transcrits ; et certains de ces « nouveaux » gènes semblent particulièrement intéressants... On voit donc que même le séquençage « bête et méchant » de l'ADN génomique peut fournir des nouveautés conceptuelles non négligeables.

Certes il s'agit, dans cet exemple, de la levure, dont le petit génome (13 600 kb en tout) doit porter une très importante somme d'informations, d'où une densité génique sans doute particulièrement élevée. Une entreprise analogue menée sur l'ADN humain serait certainement moins fructueuse dans la mesure où l'on sait que les séquences codantes des gènes n'occupent qu'une petite partie de la séquence, quelques pour cent, le reste étant constitué par des introns, des séquences répétées et des séquences intergéniques sans fonction connue. D'où l'idée périodiquement agitée de procéder chez l'homme à un séquençage massif mais sélectif sur des clones d'ADNc (représentant par définition l'ARN messager, et donc les séquences codantes) : c'est un serpent de mer qui revient régulièrement dans les discussions sur les programmes Génome. Il est séduisant, en effet, de cibler l'effort sur les 1 ou 2 % utiles du génome ; mais cette approche se heurte à deux problèmes sérieux : d'une part, le fait que dans chaque tissu ou chaque lignée cellulaire beaucoup de gènes (la majorité) sont silencieux, non exprimés, non transcrits et donc absents des bibliothèques d'ADNc correspondantes ; d'autre part et surtout,

le fait que ceux des gènes qui s'expriment dans une cellule donnée le font à des niveaux quantitativement très différents, de sorte que pour chaque molécule de l'ARN messager du gène A présent dans la cellule, il y aura, par exemple, mille molécules de celui du gène B... Les banques d'ADNc reflétant en principe fidèlement ces proportions, cela veut dire que dans un séquençage au hasard de ces clones, il faudra séquencer mille fois B pour avoir des chances de séquencer une fois A ! Le problème qui se pose est donc celui de la normalisation des banques d'ADNc, de l'égalisation de la représentation numérique des différentes séquences dans la banque. On peut envisager différentes façons de le faire et je donnerai simplement le principe d'une méthode proposée par Sherman Weissman (Yale, New Haven, États-Unis) qui n'a pas — à ma connaissance — encore été mise en œuvre complètement ni publiée, mais qui a le mérite de donner une idée de ce qu'il faudrait réaliser : une hybridation très longue et à très forte concentration d'ADN génomique avec le mélange des ARN messagers, suivie d'une élimination soigneuse des ARN non hybridés, puis de la dissociation de l'hybride, fournirait un jeu normalisé d'ARN messagers puisque chaque gène, présent à un exemplaire sur l'ADN, aura retenu un exemplaire de l'ARNm correspondant. Il « suffit » maintenant de construire la banque d'ADNc, qui sera normalisée elle aussi ! Certes la réalisation pratique de l'expérience pose quelques problèmes, mais guère plus sans doute que — en 1985 — l'acrobatie clonage PERT qui permit à Lee Kunkel d'isoler les premières sondes provenant d'une délétion causant la myopathie de Duchenne [L.M. Kunkel, A.P. Monaco, *et al.* « Specific cloning of DNA fragments absent from the DNA of a male patient with an X-chromosome deletion », *Proc Natl Acad Sci USA* 1985 ; 82 : 4778-82]. D'une façon ou d'une autre, je suis prêt à parier que des banques d'ADNc à peu près normalisées pourront être construites, et la question se posera alors réellement d'un séquençage massif de ces banques : compte tenu du fait que chaque séquence révélera un gène, encore inconnu dans 99 %

des cas (puisque l'on estime que 1 % des gènes humains sont clonés), un tel programme serait sans doute valable même au prix actuel de la base séquencée...

La séquence nous a entraîné un peu loin, peut-être ; mais je voulais à la fois traiter du malentendu (le programme Génome vu uniquement sous l'aspect « Séquence ») et montrer que les choses ne sont pas si simples et que l'on se posera peut-être bientôt la question d'un séquençage massif (de clones ADNc normalisés) : il y a d'ailleurs au moins un projet sérieux dans ce domaine en France. Pour en revenir à la mise en cause du programme Génome, il me semble que ses autres aspects (d'ailleurs peu critiqués) sont plus faciles à défendre : la carte génétique, après quelques problèmes (« Les yeux plus gros que le ventre », *Chronique Génomique*, juin-juillet 1990) semble bien repartie ; la carte physique, grâce en particulier à l'apport décisif des YAC (*yeast artificial chromosomes*) est en bonne voie. Et dans ces deux secteurs les innovations diffusent vers d'autres systèmes : les cartes génétiques de

végétaux, d'animaux de ferme et même du chien (C. Holden, « Canine genome project », *Science* 1990 ; 248 : 1184) sont en cours d'établissement avec les méthodes à base de RFLPS et maintenant de microsattellites développées sur l'homme ; quant aux YAC, on ne compte plus les organismes dont les génomes sont étudiés à l'aide de ces nouveaux vecteurs. Et l'argument financier doit être relativisé en le mettant en parallèle avec d'autres coûts. Certes, en comparant le prix d'une recherche avec celui d'un avion de chasse (100 millions de dollars) ou du télescope spatial de Hubble dont un miroir est défectueux (2 milliards de dollars pour le télescope seul), tout paraît bon marché ; mais, plus raisonnablement, si l'on considère que le budget total (demandé) du NIH pour 1991 est de 7 milliards de dollars, les 1 à 2 % que représente le programme Génome (100 millions de dollars) ne paraissent pas exagérés. Pour le moment, la chambre des représentants n'a accordé au NIH que 66 millions de dollars pour ce chapitre (en laissant au directeur du NIH la lati-

tude d'y ajouter jusqu'à 18 millions) ; mais le Sénat a encore son mot à dire et peut réviser ce chiffre à la hausse. Il est sûr cependant que dans ces réticences entrent non seulement les arguments scientifiques discutés dans cette chronique, mais aussi, et d'une façon plus vague, la crainte que cette nouvelle avancée de la connaissance ne donne aux scientifiques, aux employeurs ou... aux assureurs de nouveaux pouvoirs exorbitants fondés sur la possibilité de connaître le patrimoine génétique de chacun. C'est une autre histoire, sur laquelle il y a beaucoup à dire... ce que nous ferons certainement un jour prochain.

Dernière nouvelle : le Sénat des États-Unis a rétabli l'intégralité des 108 millions de dollars demandés par le NIH pour le programme Génome : affaire à suivre... ■

**Bertrand Jordan**

Directeur de recherche au Cnrs, responsable du groupe génétique moléculaire humaine, CIML, Inserm/Cnrs, case 906, 13288 Marseille Cedex 9, France.

## FLASH

**Le programme français « Génome Humain » auquel je faisais allusion dans ma dernière Chronique Génomique (m/s n° 8, vol. 6, p. 807) a été officiellement annoncé par le Ministre de la recherche et de la technologie, Hubert Curien, le 17 octobre dernier. Rappelons que l'étude de ce programme avait été confiée à Philippe Lazar (Directeur de l'Inserm), et le projet établi par Philippe Kourilsky après consultation d'un certain nombre de scientifiques impliqués dans ce domaine.**

**Quelles sont les grandes lignes de ce « Programme National Génome Humain » ? Sur le plan scientifique, il privilégie le créneau de l'étude des régions codantes, c'est-à-dire du séquençage des ADNc (et non de l'ADN génomique « tout venant »). C'est une option raisonnable mais moins évidente qu'elle n'en a l'air : elle pose un certain nombre de problèmes méthodologiques qui sont justement évoqués dans la chronique de ce mois. Le séquençage exhaustif de petits génomes, ainsi que les développements informatiques, sont les deux autres points forts du programme, à côté d'un soutien à l'automatisation, à la distribution de matériel biologique et à la formation du personnel. Sur le plan organisationnel l'originalité du programme est d'être conduit par un groupement d'intérêt public (GIP), structure permettant d'associer organismes, associations et industriels. Ce GIP devrait disposer d'environ 50 MF en 1991 et 100 MF en 1992 : cela représente un investissement tout à fait notable. Le début de mise en place devrait se situer dans les semaines à venir, mais à l'heure actuelle le Président du GIP (dont la personnalité aura une grande importance) n'est pas désigné.**

**Au moment où le programme européen sur le Génome démarre enfin pour de bon (le premier appel d'offres vient de sortir), ces décisions sont très opportunes. J'espère pour ma part que les moyens annoncés sont réellement des moyens nouveaux, et que les problèmes de personnel auxquels j'ai déjà souvent fait allusion seront traités de façon adéquate dans le nouveau cadre. En tous cas, bonne chance au Programme National Génome Humain !**

**Bertrand Jordan**