

Chroniques génomiques



par Bertrand JORDAN

Où l'auteur explicite ses intentions.
S'agit-il de séquence ou plutôt de carte ?
Le point sur les orientations actuelles.
Sujets futurs et appel aux lecteurs.

Pendant plusieurs années, on a beaucoup parlé du projet de séquençage du génome humain, depuis le Workshop de Santa Cruz (USA) en mai 1985 [1] jusqu'au rapport au Congrès américain de l'Office for Technological Assessment en avril 1988 [2]. Le fait nouveau est que ce programme a maintenant effectivement commencé, avec, aux USA, un financement de l'ordre de 100 millions de dollars en 1989, et que l'on en voit déjà les retombées scientifiques et technologiques [3], parfois dans des domaines *a priori* éloignés de la génétique humaine comme la biologie végétale [4]. Que l'on soit enthousiaste ou au contraire réservé sur ce programme, il est indéniable qu'il existe, et que les travaux menés dans ce cadre ont et auront une incidence directe ou indirecte sur beaucoup de domaines de la biologie. Il est donc important de faire circuler l'information sur l'état et les développements de ce projet, et c'est ce que je vais tenter de faire dans une chronique que je souhaite régulière, parfois ponctuelle, parfois polémique mais surtout, je l'espère, utile à la communauté scientifique et tout particulièrement au lectorat de la revue *médecine/sciences*, qui est directement concerné par ces développements.

Je commencerai aujourd'hui par donner une idée du contenu actuel des travaux réalisés dans ce cadre des programmes « Génome », travaux qui ne sont pas actuellement centrés sur le séquençage, contrairement à ce que l'on pourrait penser. Le projet a été en effet présenté et discuté au départ sous l'angle de la séquence : la détermination de la séquence totale d'un génome humain, soit environ trois milliards de bases, en était le but annoncé, et l'on discutait docement pour savoir s'il valait mieux « séquencer » un homme ou une femme (un homme : afin d'avoir l'information du chromosome Y), s'il fallait compter sur un prix de revient de 0,1 ou plutôt de 1 dollar US par base séquencée, et s'il fallait mettre en place des centres spéciaux dédiés à ce séquençage massif. Il apparaît clairement aujourd'hui que ces discussions étaient prématurées : l'essentiel du travail financé par les différents programmes « Génome » en cours rentre dans deux grandes catégories dont je vais donner une idée ci-dessous.

La plus importante est celle de la cartographie physique détaillée du génome : construction de cartes à l'aide d'analyses par électrophorèse en champs pulsés, clonage de grands segments d'ADN sous forme de chromosomes artificiels dans la levure (YAC), alignement de centaines ou milliers de cosmides (*contigs*) à l'image de ce qui a été réalisé il y a déjà deux ans par A. Kohara pour le génome d'*Escherichia coli* (cinq millions de bases seulement...). L'investissement dans ce type de travail correspond à une évidence : avant d'en arriver à l'analyse hyperdétaillée que constitue la séquence, il faut jalonner le territoire, y placer des repères

pour l'analyser de façon moins fine mais néanmoins exhaustive. Un séquençage à grande échelle prématuré aboutirait à un fouillis de séquences dont on ne saurait ni où elles se situent exactement dans le génome, ni comment elles se raccordent entre elles, et ne servirait pas à grand-chose. La phase de cartographie physique détaillée est donc une phase de structuration indispensable permettant d'organiser l'étape suivante (celle du séquençage) et si on le désire de l'effectuer de façon sélective. Une estimation très grossière indique que plus de la moitié des crédits investis dans le programme « Génome » va à ce type d'étude, une bonne partie du reste étant d'ailleurs consacrée aux études génétiques (RFLP, étude de *linkage*), qui se situent, elles, encore en amont de la carte physique.

La seconde catégorie de recherches impulsée par les divers programmes « Génome » porte bien sur le séquençage proprement dit, mais sous l'aspect de l'amélioration des techniques plus que sur l'accumulation de données. (Cet aspect technologique sur lequel je reviendrai dans une prochaine chronique est d'ailleurs une constante dans tout le programme et un des aspects les plus facilement « transférables » à d'autres secteurs.) L'amélioration des méthodes de séquençage est recherchée soit par l'automatisation pure et simple des méthodes manuelles existantes (voie suivie par les Japonais par exemple), soit au contraire — et de façon plus intéressante me semble-t-il — par le développement de technologies nouvelles comme le « séquençage multiplex » de Georges Church (Harvard), futuristes comme l'électrophorèse capillaire couplée à la spectrométrie

1. Sinsheimer RL. The Santa Cruz workshop, may 1985. *Genomics* 1989 ; 6 : 954-6.
2. Mapping our genes. The Genome projects : how big, how fast ? OTA report BA 373. Washington DC : US Government Printing Office, avril 1988.
3. Jordan BR. Au rendez-vous des cartographes. *médecine/sciences* 1989 ; 5 : 500-3.
4. Ezzell C, Swinbanks D. Plant researchers eager for genome programme. *Nature* 1989 ; 340 : 491.

de masse (Genomyx, San Francisco) ou même carrément exotiques comme la microscopie à effet tunnel, qui pourrait permettre de « voir » directement les différentes bases d'une molécule d'ADN et donc d'en lire directement la séquence (*voir le numéro de Nature du 9 novembre 1989 et m/s p. 76 de ce numéro*).

Parallèlement, les informaticiens (ou plutôt les biologistes branchés sur l'informatique) affûtent leurs programmes afin de les rendre capables d'interpréter les données de séquence et en particulier d'y reconnaître des gènes. Le problème est loin d'être résolu, il est d'ailleurs loin d'être trivial. Nous aurons sans doute aussi l'occasion d'y revenir ; l'interprétation informatique des séquences est un domaine difficile mais très important, intéressant et peut être un « créneau » pour la recherche européenne, qui est plutôt bien placée dans ce domaine.

Cette première chronique restera brève puisque c'est la loi du genre. Je parlerai à l'avenir d'autres aspects des programmes « Génome » : l'évolution technologique et l'automatisation des opérations de génie génétique, qui méritent bien un chapitre, de même qu'une « visite guidée » des organismes intervenant dans ce domaine (NIH, DOE, HHMI, HUGO...) pour expliciter leur nature, leurs objectifs et leurs moyens. Et si comme je l'espère cette rubrique intéresse des lecteurs de *m/s* je les invite à manifester cet intérêt en me faisant part des thèmes qu'ils souhaiteraient me voir traiter ; je le ferai, bien sûr, dans la limite de mes compétences ! ■

Bertrand Jordan

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