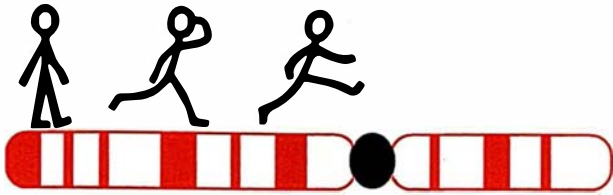


Chroniques génomiques



par Bertrand JORDAN

USA :
*un programme Génome
solidement installé*

RÉFÉRENCES

1. Jordan BR. Les yeux plus gros que le ventre ? *médecine/sciences* 1990 ; 6 : 576-9.
2. Petersen MB, Slaugenhaupt SA, et al. A genetic linkage map of 27 markers on human chromosome 21. *Genomics* 1991 ; 9 : 407-19.
3. Jordan BR. La montée en puissance des YAC. *médecine/sciences* 1990 ; 6 : 470-2.
4. Anand R, Riley JH, Smith JC, Markham AF. A 3.5 genome equivalent multi access YAC library : construction, characterisation, screening and storage. *Nucleic Acids Res* 1990 ; 18 : 1951-6.
5. Albertsen HM, Abderrahim H, Cann HM, Dausset J, Le Paslier D, Cohen D. Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. *Proc Natl Acad Sci USA* 1990 ; 87 : 4256-60.
6. Larin Z, Monaco AP, Lehrach H. Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. *Proc Natl Acad Sci USA* 1991 ; 88 : 4123-27.
7. Jordan BR. Des vecteurs de clonage à la pelle. *médecine/sciences* 1991 ; 7 : 503-4.
8. Adams P, et al. Complementary DNA sequencing : expressed sequence tags and human genome project. *Science* 1991 ; 252 : 1651-6.
9. Jordan BR. Génome et informatique : condamnés à s'entendre ? *médecine/sciences* 1991 ; 7 : 726-8.
10. Aldhous P. Human genome databases at the crossroads. *Nature* 1991 ; 352 : 94.
11. Jordan BR. Programme génome : et la France ? *médecine/sciences* 1990 ; 6 : 807-9.
12. Jordan BR. Génome au Japon : au-delà des mythes. *médecine/sciences* 1991 ; 7 : 851-3.

Plantons le décor

Le programme Génome aux USA est bien vivant ; il est maintenant enraciné dans le tissu scientifique du pays et l'on n'imagine pas qu'il puisse être abandonné, ou même sérieusement ralenti, ce qui n'était pas le cas il y a un ou deux ans. Sa direction bicéphale est assurée par le DOE (*department of energy*) et le NIH (*national institutes of health*), le premier étant à peu près l'équivalent de notre CEA, l'autre de l'INSERM ; la *National Science Foundation*, qui correspond au CNRS, est peu engagée dans la recherche biologique. Les laboratoires qui effectuent les travaux vont de l'équipe universitaire classique, très focalisée sur des problèmes biologiques précis (l'isolement du gène impliqué dans une maladie, en général) à l'« usine à cartographier » qui se consacre à l'établissement de cartes physiques ou génétiques, utilise beaucoup de machines et fonctionne selon un *planning* de type industriel. Mais situons d'abord le terrain des uns et des autres. Le DOE dispose d'un budget (annuel) de l'ordre de 60 millions de dollars qui alimente trois centres principaux : Lawrence Livermore (Anthony Carrano, chromosome 19) ; Los Alamos (Bob Moysis, chromosome 16) et Lawrence Berkeley (ancien responsable Charles Cantor, chromosome 21), ainsi que des équipes plus restreintes localisées dans ses autres laboratoires : Argonne, Brookhaven, Oak Ridge. Enfin, un quart des crédits alimente des groupes extérieurs comme celui de George

Church (Harvard, MA) ou même étrangers comme celui de Grant Sutherland (Adelaide, Australie) qui collabore avec le centre de Los Alamos pour la carte génétique et cytogénétique du chromosome 16. Le NIH, lui, finance (à l'intérieur d'un budget annuel de 108 millions de dollars) une dizaine de centres : Rick Myers (UCSF, San Francisco, CA : chromosome 4, en démarrage) ; David Schlessinger (Washington University, Saint-Louis, MO : un centre qui « tourne » déjà pour la carte physique par YAC des chromosomes X et 7) ; Glen Evans (Salk, San Diego, CA pour le 11) ; Ray White et Ray Gesteland (Université de l'Utah, Salt Lake City pour la carte génétique des 16, 17 et 5) ; Tom Caskey (Houston, TX pour des régions du X et du 17) ; Francis Collins (Université du Michigan, Ann Arbor, MI pour la recherche de gènes impliqués dans des maladies), et, tout récemment, Beverly Emanuel (Université de Pennsylvanie, PA pour la cartographie du chromosome 22). De plus Eric Lander (MIT, Boston, MA) est financé pour la cartographie du génome de la souris et David Botstein et Ron Davis (Stanford, CA) pour un projet exploratoire de séquençage de la levure. Chacun de ces laboratoires emploie au total trente à quarante personnes et reçoit un budget de l'ordre de 10 millions de dollars (salaires compris) pour quatre à cinq ans.

La sociologie des genome centers
Seuls certains de ces centres sont à l'heure actuelle vraiment opération-

nels ; les autres sont en cours d'installation, leurs contrats ayant été accordés fin 1990 ou début 1991. Mais on voit déjà les différences : à Saint-Louis, toute une organisation très structurée est mise en place pour effectuer systématiquement et de bout en bout la cartographie physique (par alignement de YAC en *contigs*) des chromosomes X et 7. David Schlessinger et ses collaborateurs ont la volonté de faire les choses à fond et de manière homogène sur l'ensemble de ces chromosomes, et la ferme intention de ne pas dériver sur l'étude d'une région particulière, aussi attractive puisse-t-elle être. A Houston, Tom Caskey, au contraire, a explicitement obtenu le contrat Génome pour renforcer les « services » (criblage de YAC, cytogénétique, séquence, informatique...) d'un ensemble d'équipes qui chacune poursuit avec vigueur et compétence l'isolement du gène impliqué dans une ou deux maladies.

Les différences entre DOE et NIH sont également marquées. Le DOE, plus « professionnel », plus technologique, avec un personnel plus stable et un climat de compétition un peu moins intense, fournit un cadre favorable au fonctionnement de *genome centers* voués à l'étude complète d'un chromosome : c'est le cas de Lawrence Livermore et de Los Alamos. Peu de post-docs et encore moins d'étudiants dans ces équipes, mais beaucoup de technologie, des programmes à long terme et une ambiance souvent moins tendue que dans les laboratoires « académiques » ; de bons résultats aussi, d'autant plus que les YAC sont arrivés à point nommé pour compléter une stratégie de cartographie physique engagée à l'époque des cosmides et qui commençait à atteindre ses limites. Les *genome centers* du NIH, quant à eux, ont, malgré leur diversité, le caractère commun d'être proches du monde universitaire, avec un flux de personnel important, des étudiants, des post-docs, du mouvement, de la vie... mais aussi du mal à assumer l'aspect systématique de l'entreprise, à maîtriser l'envie d'explorer les pistes ponctuelles mais prometteuses qui apparaissent au cours d'un travail systématique, et qui peuvent amener à le délaisser. C'est une contradiction

difficile à gérer mais importante : la raison d'être du programme Génome est justement de promouvoir une approche globale. Dans le système « académique » classique ce type de travail n'est sans doute pas suffisamment reconnu...

Quelques enseignements : les cartes sont réalisables...

La cartographie génétique a beaucoup progressé, et une carte comportant en moyenne un marqueur polymorphique tous les centimorgans constitue maintenant un objectif réaliste [1] et d'ailleurs presque atteint pour certaines régions [2]. Mais les progrès les plus sensibles concernent l'obtention de cartes physiques complètes, fondées sur du matériel cloné (cosmides et surtout YAC) et couvrant l'ensemble d'un chromosome. Celles-ci sont bien avancées pour le 7, le 11, le 16, l'X, le 19 ; elles seront sûrement obtenues rapidement pour les petits chromosomes comme le 21 ou le 22 ; puis, dans un deuxième temps, pour d'autres. La division du travail par chromosomes n'est pas artificielle dans ce cas, car il y a effectivement beaucoup d'avantages d'échelle à étudier l'ensemble d'un chromosome donné, et l'aboutissement est une collection de clones extrêmement utile pour toute étude portant sur une région particulière. Notons que la carte physique complète d'un chromosome moyen représente plusieurs années de travail pour un laboratoire d'une trentaine de personnes, et une dépense d'une dizaine ou une vingtaine de millions de dollars en l'état actuel des techniques. Dans toutes ces entreprises, les YAC ont pris aux cosmides une belle « part du marché ». La banque générale établie à Saint-Louis a maintenant été reproduite à une dizaine d'exemplaires et mise en œuvre dans chaque *genome center* ; l'effort des équipes porte sur l'amélioration des méthodes de criblage (criblage tout PCR ou sur *pools* de clones séparés en champs pulsés) et d'analyse (clonage d'extrémité par *vector PCR*, par recombinaison homologue ou par circularisation, obtention de sondes internes par Alu-PCR ou Line-PCR... [3]) plus que sur la construction de nouvelles librairies ;

sur ce point l'Europe semble un peu en avance avec les banques établies par Rakesh Anand [4], le CEPH [5], et Tony Monaco [6]. Cela n'empêche pas d'autres équipes de travailler au développement de nouveaux vecteurs, en essayant, en particulier, d'assurer la propagation de grands segments d'ADN exogène dans des bactéries, plus faciles à manipuler que la levure [7].

... L'ADNc a le vent en poupe...

Il y a toujours eu de bonnes raisons de vouloir privilégier l'étude de l'ADNc par rapport au séquençage d'ADN génomique « tout venant » : accès direct aux gènes, séquence « utile », travail débouchant sur la fonction... mais il existait un sérieux obstacle technique : la représentation très inégale des différents ARNm dans toute banque d'ADNc. Cette difficulté semble maintenant résolue, puisqu'au moins deux laboratoires ont démontré que l'on pouvait « normaliser » ou « égaliser » ces banques (c'est-à-dire faire en sorte que les différentes séquences y soient à peu près également représentées). L'analyse systématique des ADNc, qui était déjà le point fort des programmes Génome britannique, français et japonais, devient très à la mode, et le DOE a lancé récemment un appel d'offres sur ce thème. Il s'avère d'ailleurs qu'un séquençage systématique peut fournir des informations très instructives même si la banque ADNc que l'on emploie n'est pas normalisée [8].

... L'informatique est très importante !

Un *genome center* typique comporte une équipe d'informatique de quatre à huit personnes, beaucoup plus que ce que l'on trouve au niveau de la plus grosse de nos unités INSERM ou CNRS ; la *Genome Data Base* (GDB) installée par Peter Pearson à Baltimore regroupe une trentaine d'informaticiens et de gestionnaires, avec un budget de plusieurs millions de dollars par an ; et la station de travail est en passe de remplacer le MAC haut de gamme ou le PC/AT dans les laboratoires et sur le bureau des chercheurs. J'ai exposé récemment les raisons de cette débauche

d'informatique [9] ; je répéterai simplement ici que l'amélioration des techniques et leur début d'automatisation aboutissent à multiplier les résultats, les clones, les données ; que certaines de ces méthodes font elles-mêmes de plus en plus appel à l'informatique pour la saisie directe des données et leur manipulation ultérieure ; que l'archivage de ces résultats et leur diffusion à la communauté scientifique, selon des modalités précises et après une procédure de validation à définir, revêtent une grande importance [10]. Il y a donc beaucoup de tâches à assumer, et les moyens débloqués, aux USA, sont à la hauteur — ce qui ne veut pas dire que tout aille pour le mieux : le fameux fossé entre informaticiens et biologistes est encore trop réel et la communication entre ces deux cultures très différentes ne se fait pas sans heurts. Mais dans l'ensemble les choses avancent, les carnets de laboratoire informatisés commencent à fonctionner, les bases de données privées, semi-privées et publiques se mettent en place ; l'accès à ces banques est assurément une des conditions d'un travail efficace en génétique humaine.

Et la France ? (bis) [11]

A la fin de ce bilan somme toute largement positif se pose inévitablement la question : qu'en est-il en France ? Cela d'autant plus que notre pays, qui occupe actuellement un honorable troisième rang mondial dans le domaine, vit une situation complexe. D'un côté des laboratoires « académiques » corsetés dans un cadre institutionnel (INSERM ou CNRS) très rigide, avec une gestion de personnel extrêmement lourde et une marge de manœuvre limitée ; de l'autre, des structures très « débridées » (sur le plan des moyens et de la gestion), le CEPH et le Genethon ; au milieu un « GIP génome » encore en devenir, tiraillé entre ses différents partenaires et qui a bien du mal à se constituer... Il y a beaucoup à dire, trop pour le faire ici : ce sera l'objet d'une prochaine chronique ! ■

Bertrand R. Jordan

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