



par Bertrand JORDAN

*Les HGM se suivent...
et ne se ressemblent pas*

Nos lecteurs avaient peut-être remarqué la disparition des « Chroniques Génomiques ». Disparition toute temporaire, et due uniquement au manque de disponibilité de l'auteur. Ces chroniques reprennent, et tous les efforts seront faits pour les assurer sur un rythme bimestriel, car la manière reste abondante et variée.

Le dernier *Human Gene Mapping Workshop*, HGM 11, avait eu lieu à Londres en août 1991; à la mi-novembre 1993, HGM 93 se tenait à Kobe (Japon). HGM 93 juste après HGM 11: ce saut soudain dans les numéros d'ordre n'est pas le fait d'une machine à voyager dans le temps. Il manifeste l'évolution d'une formule qui, après avoir eu son heure de gloire, était devenue obsolète.

Les HGM, « ateliers » biennaux depuis 1973, avaient pour but de faire le point sur l'avancement de la carte des gènes humains, chromosome par chromosome, et d'établir un document résumant les conclusions qui résultent d'un consensus - éventuellement provisoire. Ils comportaient quelques conférences, et de nombreux *posters* présentant les récents travaux des participants; mais l'essentiel se passait dans de petites salles où les chercheurs se réunissaient pour débattre des derniers résultats obtenus sur chaque chromosome. Chacun de ces mini-conclaves rassemblait quelques dizaines de cartographes du génome, sous la houlette d'un ou deux présidents, experts du chromosome en question et chargés de veiller au bon déroulement des échanges ainsi que d'en rédiger les conclusions. On a bataillé ferme, dans ces assemblées, pour savoir s'il fallait enregistrer la localisation du gène de la myopathie de Duchenne (avant qu'il ne soit cloné) entre les repères DXS41 et DXS84, bien que les données fussent un peu préliminaires et qu'une autre équipe indiquât une position légèrement différente. On

y parlait aussi de nomenclature, afin d'éviter de répertorier deux fois le même gène sous des noms différents, ou au contraire de donner la même appellation à des entités distinctes; on se demandait si les sites fragiles ou les points de translocation devaient être catalogués au même titre que les gènes... Après trois ou quatre jours de ce régime, le responsable de l'atelier suivant était élu par l'assemblée; les présidents des comités travaillaient très tard le soir à la rédaction d'un document final de plusieurs centaines de pages qui, au prix de prodiges d'organisation, était reproduit à des centaines d'exemplaires durant la nuit et remis à chaque congressiste avant son départ.

Ce système convivial et sympathique ne devait pas résister à l'avalanche de résultats qui apparut dès la deuxième moitié de la décennie 1980, pour s'accélérer encore avec le démarrage effectif des programmes d'étude du génome humain. La création de bases de données susceptibles d'emmagasiner toutes ces informations s'avérait cruciale: la *Human Gene Mapping Library* de Yale, Genatlas à Paris en 1987, *Genome Data Base* de Baltimore à partir de 1989 s'y attelaient avec des succès divers. Mais l'enregistrement de l'ensemble des données au cours d'un atelier de quelques jours devenait acrobatique, on atteignait la limite des possibilités matérielles et surtout financières. Rassembler mille personnes, c'est presque banal s'il suffit de leur assurer gîte et couvert et de leur faire écouter, trois jours durant, quelques dizaines

Bertrand R. Jordan

Directeur de recherche au Cnrs, responsable du groupe « Structure du Génome et fonctions immunitaires ». CIML, Inserm/Cnrs, case 906, 13288 Marseille Cedex 9, France.

TIRÉS A PART

B.R. Jordan.

d'orateurs. Mais, pour *Human Gene Mapping*, il fallait de surcroît prévoir la réunion de plus de vingt comités spécialisés dans autant de salles ; il fallait surtout que chacun d'eux ait à sa disposition, en permanence, tous les moyens informatiques nécessaires pour collationner, vérifier et enregistrer dans GDB les résultats obtenus au cours des deux années précédentes. HGM 11, dernier de la série, devait ainsi coûter très cher : plus d'un million de livres sterling, soit dix millions de nos francs, pour la seule informatique - sans compter les voyages et les séjours des conférenciers, présidents de comités et congressistes. Aussi fut-il décidé de tenir à l'avenir des colloques spécialisés portant chacun sur un seul chromosome, réunissant beaucoup moins de participants, et plus faciles à mettre sur pied. Décision qui devait provoquer quelques remous : les chercheurs étaient fort attachés à ce rassemblement général, occasion de prendre connaissance de l'ensemble des travaux et de revoir les collègues. Et puis tout le monde n'est pas spécialisé dans l'étude d'un seul chromosome ; de nombreux thèmes de recherche sont transversaux : leurs protagonistes doivent-ils alors assister à une quinzaine de colloques dans l'année pour se tenir au courant ? Confrontés à une grogne inattendue, les états-majors effectuèrent un repli stratégique, et l'on termina sur la notion, somme toute raisonnable de réunions par chromosome pour l'enregistrement des données, doublées d'un grand congrès biennal permettant à tous de se rencontrer et de faire le point - sans introduction de données sur les systèmes informatiques, donc sans les complications et les frais encourus à Londres.

C'est au Japon, qui depuis plusieurs années souhaitait en être le théâtre, qu'a eu lieu cette HGM « nouvelle formule » précédée d'une séance de coordination (CCM, *Chromosome Coordinating Committee*) rassemblant les présidents des réunions par chromosome. Moins couru que les précédents (environ cinq cents inscrits), hésitant un peu entre l'atelier et le symposium classique, ce colloque ne manquait pourtant pas d'intérêt à

bien des égards. Comme on pouvait s'y attendre dans ce pays où règne le souci du détail, l'organisation matérielle était impeccable, l'équipement audiovisuel parfaitement adapté et le personnel nombreux et efficace, quoique pas toujours très anglophone. La participation japonaise était naturellement massive (une bonne moitié des assistants), ce qui donnait l'occasion de rencontrer de jeunes chercheurs locaux rarement présents dans les congrès internationaux. Le programme Génome, qui m'avait semblé prendre un réel départ au Japon il y a deux ans [1], confirme son décollage. Le montant des financements en 1993 est de 3,6 milliards de Yen, soit environ 180 millions de nos francs, et une augmentation de l'ordre de 20 % est prévue pour 1994. Rappelons que notre ministère de la Recherche attribue une soixantaine de MF au GREG (plus vingt-huit au CEPH), et qu'aux États-Unis la somme totale (NIH plus DOE) tend vers les deux cent millions de dollars. La liste des projets financés par le Monbusho (ministère de l'Éducation) donne une idée des lignes de force : une vingtaine portent sur la cartographie physique et le clonage positionnel, une dizaine sur l'ADNc, quinze sur des mises au point technologiques... et trente-deux sur la bio-informatique appliquée au génome. Des structures spécialisées sont en bonne voie de réalisation : mentionnons l'Institut de séquençage de Kazusa (à Chiba, près de Tokyo), financé par le gouvernement local et dont la construction devrait s'achever au printemps 1994, ou le *Human Genome Center* de l'Université de Tokyo (cartographie physique, bases de données, développements d'outils d'analyse de séquences). Le rôle de Kenichi Matsubara, vice-président de Hugo, responsable du programme du ministère de l'Éducation et organisateur de HGM 93 semble bien établi, ce qui est de bon augure car il s'agit d'un scientifique particulièrement compétent, étranger aux coteries et d'une intégrité sans faille.

C'est d'ailleurs son équipe qui présentait certains des résultats les plus

intéressants du congrès, du moins à mon sens. Il s'agit en fait de la poursuite du projet portant sur l'ADN complémentaire dont j'avais vu les débuts en 1991 [1] et dont les premiers résultats ont paru l'an dernier [2, 3]. L'objectif est d'établir une carte transcriptionnelle des deux cents cellules de base du corps humain, une *body map* selon la terminologie choisie par les auteurs. Le moyen : pour chaque cellule, construire une banque d'ADNc calibrée de façon à représenter fidèlement la population des ARN messagers, et séquencer deux ou trois cents nucléotides à l'extrémité 3' de mille clones pris au hasard. La première publication du groupe [2] avait montré toutes les informations que pouvait fournir cet exercice *a priori* un peu répétitif. Il avait en effet défini une brassée de gènes « nouveaux » (jusqu'à présent inconnus), et, surtout, donné une évaluation du taux d'expression de centaines de gènes connus. La détermination de la multiplicité des clones correspondant à des séquences déjà répertoriées donnait pour la première fois une vision d'ensemble de l'activité transcriptionnelle d'une cellule - avec quelques surprises, comme par exemple le fait que le gène le plus exprimé soit celui du facteur d'élongation α , retrouvé vingt-deux fois parmi les mille clones... Le but annoncé d'établir la même liste pour les cent quatre-vingt-dix-neuf autres types cellulaires pouvait paraître irréaliste. Cela ne semble pas être le cas, puisqu'à Kobe le groupe présentait l'analyse des données obtenues sur vingt types cellulaires : il est devenu le centre d'une nébuleuse incluant une dizaine de laboratoires (et un industriel, Hitachi), qui appliquent cette méthode d'analyse et mettent en commun les résultats obtenus pour les différentes cellules.

Leur confrontation, grâce à une base de données qui facilite les recoupements, se révèle très féconde. On peut, par exemple, effectuer par voie informatique une sorte de « Northern blot synthétique » dessiné à partir de la fréquence à laquelle se retrouve une séquence dans chacune des vingt cellules étudiées ; ou, à l'aide des premières

assignations chromosomiques effectuées sur deux ou trois centaines de clones, constater que ce sont les gènes à expression forte et ubiquitaire qui ont le plus souvent deux ou trois localisations chromosomiques (correspondant sans doute à des pseudogènes), au contraire des gènes spécifiques d'un tissu.

Du côté des Occidentaux, peu de nouveautés (et beaucoup d'absents, comme le programme français Genexpress) ; seul l'informaticien de l'*Institute for Genomic Research*, le nouveau laboratoire de Craig Venter, exposait son *Expressed Gene Anatomy Database* visant à regrouper et comparer les séquences partielles obtenues par différentes équipes. Chose intéressante, l'orateur estimait le nombre total de séquences d'EST actuellement déterminées à environ 130 000, chiffre supérieur au nombre total des gènes humains (50 000 à 100 000). Même si l'on tient compte des inévitables redondances (intra- et interlaboratoires), ce chiffre indique que la moitié sans doute des gènes humains a fait l'objet d'un étiquetage - chiffre imposant bien que beaucoup de ces séquences ne soient pas encore dans les banques de données. D'une façon générale, les progrès de ces programmes montrent que l'exploration par séquençage partiel présente, dans l'état actuel des techniques et lorsqu'il est mis en œuvre de manière rigoureuse, un excellent rapport qualité/prix.

Sur le front de la technologie, le bilan japonais est plus nuancé. L'usine à séquencer de Tsukuba, cet ensemble de robots destiné à automatiser presque complètement le séquençage de l'ADN [3, 4], devait entrer en fonction courant 1992 : elle n'a, en réalité, pas vraiment démarré. Un *poster* décrivait la séquence du chromosome VI de la Levure, 280 kilobases, effectuée à Tsukuba... sans passer par l'usine située dans le même laboratoire, et à l'aide de séquenceurs Applied Biosystems ; un peu plus loin, un autre *poster* montrait, lui, quelques photos de cette installation et indiquait, avec une humilité rare, « *the performance is very poor* ». Cette tentative, qui m'avait fait forte impres-

sion en son temps, semble donc se terminer peu glorieusement. Était-elle prématurée, ou s'est-elle fourvoyée en raison d'erreurs de parcours ? En tout cas, cet échec atteste que la persistance japonaise, cette capacité souvent payante à investir dans le long terme, peut aussi avoir des effets négatifs : difficulté à renoncer à un projet, à reconnaître qu'il est raté. Le programme d'ordinateurs de « cinquième génération », qui avait tant effrayé l'Occident lors de son lancement, en est un autre exemple : malgré son fiasco (il a produit des machines en principe très puissantes, mais en pratique inutilisables, dont certaines ont été fournies gratuitement à des laboratoires qui, après quelques essais, ont renoncé à s'en servir), il fut prolongé d'une ou deux années au-delà du terme prévu. Toujours à propos de séquençage, Shimadzu présentait un appareil proche des instruments occidentaux ; et surtout Hitachi, qui vend depuis trois ans au Japon un séquenceur analogue à l'« ALF » de Pharmacia, montrait les résultats obtenus avec un prototype utilisant l'électrophorèse sur capillaire. Technologie déjà largement discutée depuis deux ou trois ans, celle-ci promet une augmentation considérable du débit grâce à la vitesse de séparation, à la résolution qui doit permettre de lire en routine 1 000 bases (au lieu de 400 à 500 actuellement) et à la possibilité de coupler de nombreux capillaires, 96 dans le cas de la machine Hitachi, et d'effectuer ainsi autant de séquences en parallèle. La firme semble avoir résolu un des principaux problèmes techniques, celui de la détection des signaux de fluorescence avec une très haute sensibilité ; et il est instructif de voir qu'après s'être contentée, dans un premier temps, de la copie du système inventé par Wilhelm Ansorge, elle se trouve maintenant - pour autant que je puisse en juger - dans le peloton de tête de ceux qui mettent au point la prochaine génération de séquenceurs.

Quelques remarques pour conclure cette vision forcément très impressionniste d'un congrès dense et dont beaucoup de sessions se tenaient en

parallèle. Le front de la bio-informatique continue à bouger, et la saga des bases de données « officielles » n'est pas terminée : Peter Pearson n'est plus directeur de *Genome Data Base*, à qui l'on reproche toujours son manque de convivialité. Le nouveau responsable de l'informatique, Ken Fasman, promet de grandes améliorations, de nouveaux outils graphiques et la prise en compte des données préliminaires. Quant à Pearson, il est maintenant chargé de la restructuration d'OMIM (*On-line Mendelian Inheritance in Man*), la base de données sur les maladies génétiques tenue depuis trente ans par Victor McKusick, afin d'en faire une vraie base relationnelle : espérons qu'elle ne perdra pas dans le processus la relative facilité de consultation qui la caractérise aujourd'hui ! L'effort japonais déjà mentionné plus haut est considérable, et certaines des démonstrations présentées étaient tout à fait convaincantes. Elles témoignent, comme le système GNOME proposé par une fondation privée et qui offre une interface simple et uniforme pour toute recherche d'homologie de séquence, d'un louable souci de se mettre à l'écoute des biologistes et de développer les outils dont ils ressentent le besoin - plutôt que de vouloir à tout prix faire de la « vraie » recherche informatique en perdant un peu de vue ses applications. On aimerait que cette attitude soit plus répandue, singulièrement dans notre pays... Notons encore la brillante prestation de Yusuke Nakamura, dont la réputation n'est plus à faire dans le domaine de la cancérologie : résultats importants, appuyés sur des données impeccables. Il s'agissait cette fois d'une étude très complète sur les mutations dans le gène APC (impliqué dans la polyposé colique, et cloné par cette équipe), illustrant de façon parfaitement claire le mécanisme génétique en deux étapes de ce cancer : prédisposition héréditaire due à l'inactivation d'une copie du gène, puis inactivation de l'autre copie par une mutation somatique et apparition de tumeurs. La France n'était pas en reste grâce à l'exposé de Gilles Tho-

mas sur la neurofibromatose de type II et le sarcome d'Ewing, avec la mise en évidence du rôle dans différents cancers d'un produit de fusion (dû à une translocation) associant l'extrémité 5' de la protéine Ewing avec le site de fixation à l'ADN d'une autre protéine - d'où l'apparition d'un facteur de transcription anormal dont les effets sur la régulation cellulaire sont catastrophiques. Les maladies à déterminisme génétique complexe n'étaient pas oubliées; Mark Lathrop et William Cookson illustraient les voies d'approches et les difficultés de l'étude du déterminisme génétique de l'hypertension, du diabète ou de l'asthme, tandis que Francis Collins, nouveau directeur du programme Génome américain, montrait toute la complexité de l'« après-gène » dans le cas de la neurofibromatose de type I; deux ans après l'isolement du gène, on n'a encore qu'une idée très partielle des fonctions possibles de la protéine correspondante.

L'éthique avait aussi sa place et, contrairement à l'habitude, les interventions la concernant n'étaient pas placées tout à la fin du colloque, au moment où la salle est clairsemée et où chacun ne pense plus qu'à son avion. Tenu au début de la dernière journée devant un public nombreux et attentif, cette session fut marquée par les propos percu-

tants de Nancy Wexler, la blonde et infatigable égérie des recherches sur la chorée de Huntington. A partir du récent clonage d'embryons humains et de son exploitation médiatique, elle insistait avec force sur la nécessité d'une attitude à la fois rigoureuse et offensive des scientifiques pour limiter, par exemple, l'analyse d'ADN en vue de la recherche d'une maladie ou d'une prédisposition aux seuls adultes correctement informés et consentants, sur leur seule demande, à l'exclusion de tout tiers (assureur, employeur ou, même, membre de la famille). Quant au représentant français, Alain Pompidou, son exposé sur les problèmes de brevetage du génome évitait la rhétorique facile et maximaliste qui a parfois prévalu dans ce domaine, posait les bonnes questions et fut, me semble-t-il, bien reçu par l'auditoire.

Au total, cette première HGM nouveau style constitue un relatif succès. Elle fait certes un peu double emploi avec l'annuel réunion de *Cold Spring Harbor (CSH Genome Mapping and Sequencing Meeting)*, plus pointue, et les *HUGO meetings* dont le dernier se tint à Nice à l'automne 1992, plus ouverts mais aussi plus rigides sur le plan de leur organisation. Elle a perdu ce qui faisait la spécificité des HGM précédentes, l'accouchement d'un consensus sur les cartes et leur mise en mémoire,

mais elle pourrait dans l'avenir être un forum ouvert où se rencontrent tous les acteurs du programme Génome, dans un cadre moins numériquement restreint (et donc moins élitiste) que *Cold Spring Harbor* (limité à environ deux cents participants). Il faudrait pour cela accroître le rôle des *posters*: ils étaient fort bien traités au niveau de l'espace, une salle vaste et bien éclairée où ils sont restés exposés trois jours; mais aucune tranche horaire ne leur avait été réservée, ce qui leur a fait perdre beaucoup d'efficacité comme moyen de mise en contact. La prochaine édition, HGM 95, prévue en Europe, devrait être plus fréquentée et sera sans doute le véritable test de cette nouvelle formule ■

RÉFÉRENCES

1. Jordan BR. Génome au Japon: au-delà des mythes. *médecine/sciences* 1991; 7: 851-3.
2. Okubo K, Hori N, Matoba R, *et al.* Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nature Genet* 1992; 2: 173-9.
3. Jordan BR. Le festival des ADNc. *médecine/sciences* 1993; 9: 211-6.
4. Jordan BR. La robotique en biologie moléculaire: l'arlésienne? *Biofutur* 1992; 108: 22-5.