

Premier lever d'une carte physique du génome humain

Établir la cartographie du génome humain représente l'un des grands défis de la génétique moderne. Les équipes du CEPH et du Génethon ont relevé ce défi et sont en passe de lever la toute première carte physique de l'ensemble des chromosomes humains. Pourquoi autant d'émoi autour de résultat décrits dans deux publications, l'une dans *Cell* [1], l'autre dans *Nature* [2]. Il convient tout d'abord de replacer ces travaux dans le contexte du programme Génome Humain et de ses retombées attendues dans le domaine médical. Ce grand projet fut proposé par des scientifiques américains en 1986 et prit corps à l'échelon mondial à la fin des années 1980. Son objectif, aujourd'hui consensuel, est d'établir la cartographie et la séquence complète des génomes de l'homme et de quelques autres organismes d'ici l'an 2005. Opérationnellement, cartographier un génome, c'est positionner le plus précisément possible le long des chromosomes, des segments d'ADN issus de ce génome quels que soient leurs rôles ou la façon d'identifier ces fragments et quelles que soient les méthodes utilisées pour les ordonner les uns par rapport aux autres. Quand les fragments correspondent à des *loci* polymorphes, il est possible d'établir leur ordre respectif en comptant, par des études familiales, le nombre de recombinaisons méiotiques survenues entre chaque *locus*. Moins il y en a et plus les *loci* sont proches, on dit aussi « génétiquement liés ». C'est de cette définition que provient le terme de « carte génétique » pour désigner (improprement) la mise en ordre le long du génome de *loci* polymorphes par analyse de liaison génétique. Il ne s'agit donc pas de la carte des gènes mais de la pose de balises entre lesquelles il sera possible, toujours par des analyses familiales, de positionner d'autres *loci* polymorphes et notamment ceux des gènes intervenant en pathologie humaine.

Il existe d'autres façons de positionner des balises (polymorphes ou non-

polymorphes) le long du génome. Il s'agit des méthodes dites physiques utilisant l'hybridation *in situ* de chromosomes métaphasiques ou des collections de lignées porteuses de cassures ou de délétions chromosomiques bien caractérisées. Mais, d'une façon générale, lever une carte physique d'un génome, c'est établir le long de chaque chromosome un continuum de fragments clonés se chevauchant les uns les autres. Cette carte physique est une étape indispensable du programme Génome Humain. Elle constitue en effet un pré-découpage des chromosomes en fragments d'une taille compatible avec les opérations de séquençage. Mais, dans l'immédiat, cette carte représente également un gain de temps considérable dans la chasse aux gènes mutés responsables des maladies héréditaires. En effet, ces mutations sont découvertes par des méthodes dites positionnelles, c'est-à-dire qu'on les cherche en balayant le génome humain plutôt qu'en analysant l'altération à l'échelon des protéines par des tests fonctionnels. Ce balayage s'effectue en plusieurs étapes. On positionne tout d'abord grossièrement les gènes mutés entre deux balises de la carte génétique. En pratique, cela s'obtient par l'analyse de familles de malades en comptant le nombre de recombinaisons méiotiques entre chacune des balises et le gène muté (il existe aujourd'hui près de 3 000 balises polymorphes le long du génome humain). On repère ainsi les deux balises qui encadrent le gène muté et qui définissent un intervalle de 2 à 10 millions de paires de bases, le génome humain comptant 3 500 millions de paires de bases. Il faut ensuite isoler l'ADN correspondant à cet intervalle : c'est la phase du clonage. Puis enfin, et c'est la dernière étape, une analyse fine des fragments clonés de l'intervalle permet de dresser un inventaire des gènes candidats et de comparer leur séquence chez des sujets atteints et non-atteints. Dans cette longue succession d'opérations, il faut encore un à deux ans pour la phase

de localisation, deux à cinq ans pour la phase de clonage et un à deux ans pour la dernière phase. Il est clair que l'existence de la carte physique raccourcit à quelques semaines la phase de clonage. Il suffit en effet que les deux balises encadrant le gène soient localisées sur le continuum de fragments chevauchants, et tous les fragments d'ADN clonés sont immédiatement accessibles dans l'ordre d'une balise à l'autre. Ainsi, si l'on considère qu'il existe 3 000 gènes impliqués dans les maladies héréditaires et que l'on gagne au minimum trois ans par maladie et par équipe, le gain de travail s'élève à 9 000 années-équipe, soit en pratique à plusieurs décennies.

Cette carte physique était attendue vers les années 1995-1997. La stratégie envisagée reposait sur une répartition de la tâche chromosome par chromosome dans plusieurs centres, pour la plupart américains. Des développements technologiques particuliers et une mise de fonds considérable de l'Association Française contre les Myopathies nous permirent d'aborder une autre stratégie, celle d'une approche globale où l'on cartographie tous les chromosomes en une seule opération. Pour cela, nous avons amélioré à l'extrême des méthodologies existantes ayant déjà été utilisées sur des génomes plus petits comme celui de la levure (15 millions de paires de bases) ou *C. Elegans* (100 millions de paires de bases).

Nous avons analysé individuellement 22 000 fragments clonés dans la levure sous la forme de chromosome artificiel (YAC = *yeast artificial chromosome*) d'environ un million de paires de bases. Pour chaque YAC fut déterminée la liste des tailles des fragments de restriction contenant la séquence répétée humaine L1 (en moyenne 15 fragments par YAC digéré par trois enzymes de restriction). Pour obtenir ces « empreintes » près de 3 000 *Southern blots* ont été produits automatiquement en quelques mois grâce à un instrument, le Multiblotter, développé avec la Société Bertin.

Les autoradiogrammes correspondants furent digitalisés puis analysés automatiquement par des algorithmes complexes de comparaison d'empreintes. Ces programmes permirent de comparer tous les clones deux à deux grâce à un superordinateur créé pour l'occasion à l'INRIA par la mise en réseau de tous les mini-ordinateurs de l'Institut. Ainsi, des calculs qui auraient nécessité plusieurs années furent obtenus en quelques jours. De ces comparaisons d'empreintes, nous avons pu déduire des chevauchements entre clones, ces 22 000 clones représentant cinq génomes haploïdes et donc en pratique cinq découpes différentes du génome humain. Ainsi près de 50 % de chaque chromosome furent reconstitués avec des *contigs* (ensemble de clones chevauchants) de cinq YAC en moyenne, 15 à 20 % du génome étant couverts par des contigs de plus de 3 millions de paires de bases. La validité de l'assemblage fut vérifiée en analysant 10 contigs au hasard parmi les 1 000 obtenus. Différentes simulations à partir des résultats obtenus sur ces 22 000 clones montrent que l'analyse de 24 000 clones supplémentaires permettra de lever une carte de 90 % du génome avec des contigs de plus de 5 millions de paires de bases. Ceux-ci seront obtenus dans les mois à venir. Les tronçons de ce premier assemblage seront ensuite positionnés le long du chromosome en utilisant quelques centaines de balises ordonnées de la carte génétique. Les solutions de continuité de cette carte physique primaire (environ 100 « trous ») seront traitées par l'ensemble de la Communauté Internationale. Les derniers trous correspondront aux segments d'ADN réfractaires au clonage.

Parallèlement à cette approche globale, et justement pour vérifier qu'il était possible d'obtenir 100 % d'un chromosome, la carte physique complète du bras long du chromosome 21 (43 millions de paires de bases, 1,2 % du génome) a été réalisée, le bras court, long de quelques millions de paires de bases, n'étant constitué que d'hétérochromatine. Cela fut fait en criblant une banque de YAC de 90 000 clones avec 200 balises du chromosome 21 dont près de la moitié n'étaient pas encore positionnées. Ces balises correspondaient à des sites

amplifiables par PCR (STS = *sequenced tagged sites*) représentant, soit des *loci* polymorphes, soit des gènes connus, soit enfin des sites anonymes. Nous avons obtenu effectivement un continuum de clones chevauchants s'étendant du centromère au télomère du bras long du chromosome 21.

Ce *contig* va accélérer l'identification des gènes de ce chromosome comme ceux responsables du syndrome de Down (trisomie 21), de la sclérose latérale amyotrophique, de l'épilepsie myoclonique. Il permettra également de clarifier le rôle du chromosome 21 dans la maladie d'Alzheimer.

L'ensemble de nos travaux démontre aussi qu'un changement d'échelle de la biologie s'opère peut-être plus tôt que prévu et là où on ne l'attendait pas. Les méthodes et les stratégies employées marquent définitivement la naissance d'une nouvelle branche de la recherche, plus technologique, dans laquelle la productique fait son entrée au service de l'homme malade. Des dizaines de nouvelles mutations pathologiques seront identifiées très rapidement, créant une véritable inflation par rapport aux efforts consentis pour développer des solutions thérapeutiques. Mais il faut espérer que cet afflux stimulera les investissements publics et privés.

Daniel Cohen

1. Bellané-Chantelot C, Lacroix B, Ougen P, *et al.* Mapping the whole human genome by fingerprinting yeast artificial chromosomes. *Cell* 1992 ; 70 : 1059-68.

2. Chumakov H, Rigault P, Guillou S, *et al.* Continuum of overlapping clones spanning the entire human chromosome 21 q. *Nature* 1992 ; 359 : 380-7.

■■■■ La protéine β -amyloïde (β -AP) déstabilise la régulation neuronale du calcium intracellulaire. La β -AP est le constituant majeur des plaques séniles observées dans le cerveau de patients atteints de la maladie d'Alzheimer (*voir m/s n° 3, vol. 7, p. 294*). On soupçonne depuis longtemps que cette protéine joue un rôle dans l'atteinte neuronale, et une activité neurotoxique a été observée sur des neurones en culture, notamment en présence d'acides aminés excitateurs [1] mais les mécanismes n'en sont pas connus. Mattson *et al.* (University of Kentucky, USA) apportent un élément de réponse important en démontrant que la β -AP modifie la régulation des concentrations de calcium intracellulaire dans les neurones [2]. La mise en culture de neurones de cortex humain fœtal en présence de β -AP provoque, après quelques jours, une augmentation de la concentration calcique intracellulaire à près de deux fois le contrôle. Toutes les stimulations qui, normalement, accroissent la concentration calcique intracellulaire sont, par ailleurs, nettement amplifiées (jusqu'à trois fois) en présence de β -AP. Cela est particulièrement net — et important pour les hypothèses concernant l'effet neurotoxique — en ce qui concerne les acides aminés excitateurs tel le glutamate. Or on sait que l'excitotoxicité associée au glutamate est due, en grande partie, à une amplification de l'influx de calcium conduisant à une surcharge calcique intracellulaire (*voir m/s n° 4, vol. 6, p. 394*). Une « hypothèse calcique » du mécanisme de la neurodégénérescence alzheimerienne émerge donc de ces travaux, selon laquelle l'accumulation de β -AP dans l'espace extracellulaire modifierait la régulation calcique. La protéine abaisserait ainsi le seuil de l'excitotoxicité des neurones qui seraient alors susceptibles de dégénérer en présence de concentrations pourtant physiologiques d'acides aminés excitateurs.

[1. Koh JY, *et al.* *Brain Res* 1990 ; 533 : 315-20.]

[2. Mattson MP, *et al.* *J Neurosci* 1992 ; 12 : 376-89.]