

*médecine/sciences
a présenté dans ses
premiers numéros
une série de brèves
mises à jour
conceptuelles
sur la structure
et l'expression
du génome eucaryote.
La génétique moléculaire
n'est cependant pas
seulement une nouvelle
discipline destinée
à approfondir
nos connaissances
fondamentales des
processus biologiques.
Elle est aussi un
ensemble de techniques
qui ont commencé à
modifier l'exploration
de certaines maladies
et dont on peut prévoir
que la place augmentera
encore considérablement
dans l'arsenal
préventif, diagnostique
et, ultérieurement
thérapeutique
des médecins.
Nous commençons donc
la présentation des
méthodes de base de
génétique moléculaire
qui s'appliquent dès
aujourd'hui
à l'exploration
de multiples
problèmes pathologiques.*

L'hybridation moléculaire

Le « Southern blot »

Principe

C'est en 1975 qu'E.M. Southern décrit la méthode consistant à détecter très spécifiquement des fragments d'ADN transférés sur filtre par leur hybridation à des séquences identiques radioactives. Le schéma montre le principe de cette technique appliquée à la détection d'une séquence particulière d'ADN génomique total. L'ADN purifié de cellules quelconques, par exemple de leucocytes circulants, est coupé en des sites très spécifiques par des enzymes de restriction qui reconnaissent des séquences nucléotidiques bien particulières, dispersées au hasard sur l'ADN. Une très grande variété de fragments est ainsi produite, dont l'un (ou un tout petit nombre) correspond à une partie ou à la totalité du gène étudié. La longueur totale de l'ADN humain est de l'ordre de $2,5 \times 10^6$ kilobases (kb) si bien qu'un fragment de 10 kb portant la séquence étudiée ne représentera que $1/250\,000$ de l'ADN. Les fragments sont séparés selon leur taille par électrophorèse dans un gel d'agarose, puis dénaturés par un traitement alcalin : le gel est immergé dans une solution de soude qui sépare les deux brins d'ADN de chaque fragment. Les fragments, maintenant simples brins, sont transférés par capillarité sur un filtre de nitrocellulose ou de nylon. Le filtre est alors incubé dans des conditions bien déterminées avec des petits fragments d'ADN (ou parfois

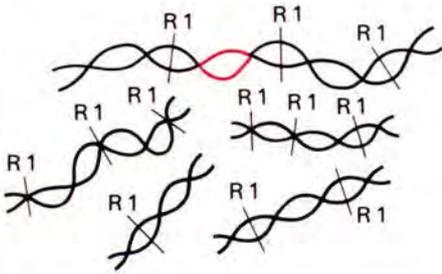
d'ARN), eux aussi dénaturés, de séquence identique à celle d'une partie du fragment recherché; ces fragments radioactifs sont appelés « sonde ». Les bases complémentaires forment alors des hybrides stables qui persistent après lavages répétés du filtre pour éliminer toute trace de sonde radioactive non spécifiquement liée à une séquence complémentaire.

Le ou les fragments correspondant au gène ou aux séquences intergéniques étudiés apparaît alors en autoradiographie sous la forme d'une bande radioactive dont la position dépend de la taille du fragment qui migre d'autant plus vite qu'il est plus petit.

Utilisation en médecine

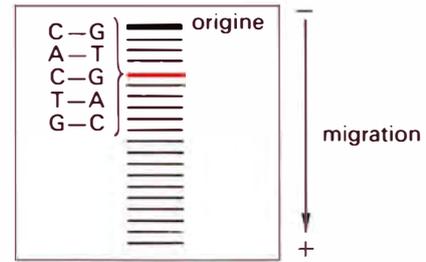
La technique du « Southern blot » est d'ores et déjà d'application très large. Elle permet d'étudier la présence et la structure de gènes particuliers impliqués dans des maladies congénitales, de rechercher des séquences virales (par exemple du virus de l'hépatite B), de mettre en évidence le remaniement et l'amplification d'oncogènes dans des cancers, de caractériser des lignées dérivées de lymphocytes B ou T par le remaniement des gènes d'immunoglobuline dans le premier cas et du récepteur T dans le deuxième (voir la Nouvelle sur ce sujet dans le prochain n° de m/s). Plus précisément que les techniques classiques de cytogénétique, cette méthode permet de détecter et d'explorer des

1 De l'ADN de haut poids moléculaire est coupé avec l'enzyme R1

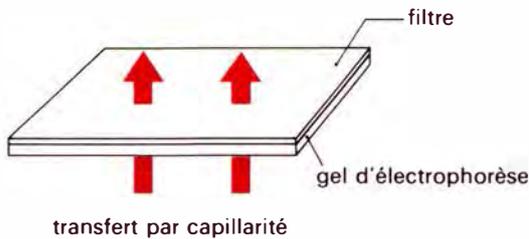


2 Séparation des fragments selon leur taille

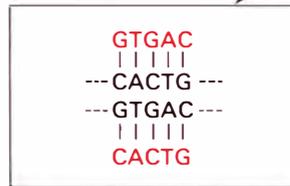
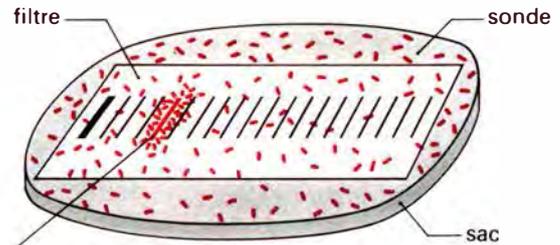
électrophorèse



3 Les brins des fragments sont séparés par dénaturation alcaline, puis transférés sur un filtre.



4 Le filtre est hybridé avec des petits fragments d'ADN (ou d'ARN) radioactifs identiques à la séquence recherchée



5 Après lavages, la « sonde » radioactive est éliminée, sauf au niveau du brin d'ADN avec lequel elle a pu s'hybrider : une bande radioactive correspondant au fragment étudié est ainsi détectée en autoradiographie

remaniements chromosomiques, des délétions de petite taille, etc. Enfin, par l'utilisation des polymorphismes d'enzymes de restriction dont nous reparlerons dans le *Lexique* du prochain numéro, le « Southern blot » constitue une méthode de choix pour prévoir la susceptibilité à des affections dont la lésion génétique précise n'est souvent pas connue et fournit donc un outil précieux pour un nombre croissant de diagnostics prénataux. Il s'agit dans l'ensemble d'une technique simple qui peut être mise en œuvre avec peu de moyens et qui peut s'appliquer à l'ADN purifié à partir de 10 ml de sang total ou de 10 mg de tissu placentaire prélevé par biopsie.

A. K.

Principe du « Southern blot » d'ADN génomique. L'encart situé après la quatrième étape du schéma montre comment la sonde radioactive (ici un fragment d'ADN double brin dénaturé) s'hybride spécifiquement sur le filtre avec le fragment d'ADN dénaturé comportant la même séquence.

Electrophorèse de l'ADN en champ pulsé

La technique d'électrophorèse en gel d'agarose ou de polyacrylamide permet de séparer des molécules d'ADN suivant leur taille à condition qu'elle n'excède pas 50 000 paires de bases (50 kilobases ou kb). Les gels d'agarose ou d'acrylamide sont constitués d'un réseau désorganisé de longues fibres baigné dans un milieu aqueux. On place les molécules d'ADN qu'on désire séparer en un point du gel et on applique un champ électrique. Sous son action, les molécules d'ADN chargées négativement se déplacent à travers le réseau. On constate expérimentalement que plus elles sont petites, plus elles se déplacent rapidement. Ainsi peut-on les séparer. Cette séparation est due aux fibres du réseau, car, en leur absence, les molécules d'ADN se déplacent toutes avec la même vitesse quelle que soit leur taille. Cependant l'électrophorèse perd tout pouvoir résolutif pour des molécules d'ADN de taille supérieure à 50 kb ; par exemple, l'ADN du bactériophage lambda (50 kb) et l'ADN des chromosomes de la levure (de 300 à 2 000 kb) migrent à peu près à la même vitesse (*figure 1*). En d'autres termes, au-delà de 50 kb, la mobilité électrophorétique de l'ADN ne dépend plus de sa longueur. En 1984, deux chercheurs de l'université de Columbia à New York, Schwartz et Cantor [1], proposent une méthode d'électrophorèse sur gel d'agarose dont le pouvoir de séparation peut aller au moins jusqu'à 9 000 kb

[9]. Comment cela est-il possible ? Commençons par quelques ordres de grandeur. Les distances moyennes qui séparent des fibres voisines dans ces gels sont de l'ordre de 150 à 400 nanomètres (nm) ($1 \text{ nm} = 10^{-9} \text{ m}$). La flexibilité des molécules d'ADN peut être appréciée en mesurant la distance minimale devant séparer deux points A et B quelconques de cette molécule pour que son orientation au point A n'ait pas d'influence sur son orientation au point B. Cette distance, appelée longueur de persistance, est de l'ordre de 150 nm. Entre deux fibres du gel, l'ADN peut donc tout juste faire une boucle.

Une molécule d'ADN de 50 kb possède une longueur totale de 18 000 nm et, dans le gel, en l'absence de champ électrique, elle apparaît comme une longue chaîne serpentant au hasard entre les fibres du gel : une sorte de pelote statistique dans un réseau enchevêtré (*figure 2*). La distance moyenne entre deux fibres étant comparable à la longueur de persistance, la chaîne est complètement contrainte et, seules les deux extrémités ont suffisamment de degrés de liberté pour avancer. Une des deux extrémités détermine le chemin et le reste de la molécule coulisse entre les fibres comme un serpent dans un jeu de quilles. En l'absence de champ électrique, le sens du déplacement est indifféremment déterminé par l'une ou l'autre des deux extrémités et peut s'effectuer dans toutes les orientations possibles. Ce mou-

vement de la chaîne appelé mouvement de reptation [2] est un coulisement selon un chemin parcouru au hasard.

Sous l'influence d'un champ électrique, la chaîne chargée est soumise à une force électrostatique. Elle se déplace toujours par reptation mais, cette fois-ci, l'une des deux extrémités au lieu de prendre une direction au hasard va se déplacer préférentiellement dans celle imposée par le champ électrique [3-5]. On devine l'effet produit par une telle dynamique : l'enveloppe de la molécule d'ADN qui était sphérique en l'absence de champ (pelote statistique) va s'allonger dans le sens du champ électrique sous son influence ; en d'autres termes, le serpent se déplie puisque sa tête se déplace dans une direction déterminée (*figure 2*). Cette déformation est un effet du gel car, en son absence, il faudrait appliquer des champs 100 fois plus forts pour l'observer. L'expression de la mobilité électrophorétique de ces longues molécules d'ADN est la somme de deux termes $m_0 + m_E$ qui se comprennent comme les deux premiers termes d'un développement selon les puissances du champ électrique : le premier terme m_0 qui ne dépend pas du champ électrique (mobilité à champ électrique faible) correspond au déplacement de l'ADN non déformé. Un calcul [7, 8] montre qu'il varie comme l'inverse de la longueur de la molécule. Le second terme m_E dépend du carré du champ électrique et rend compte

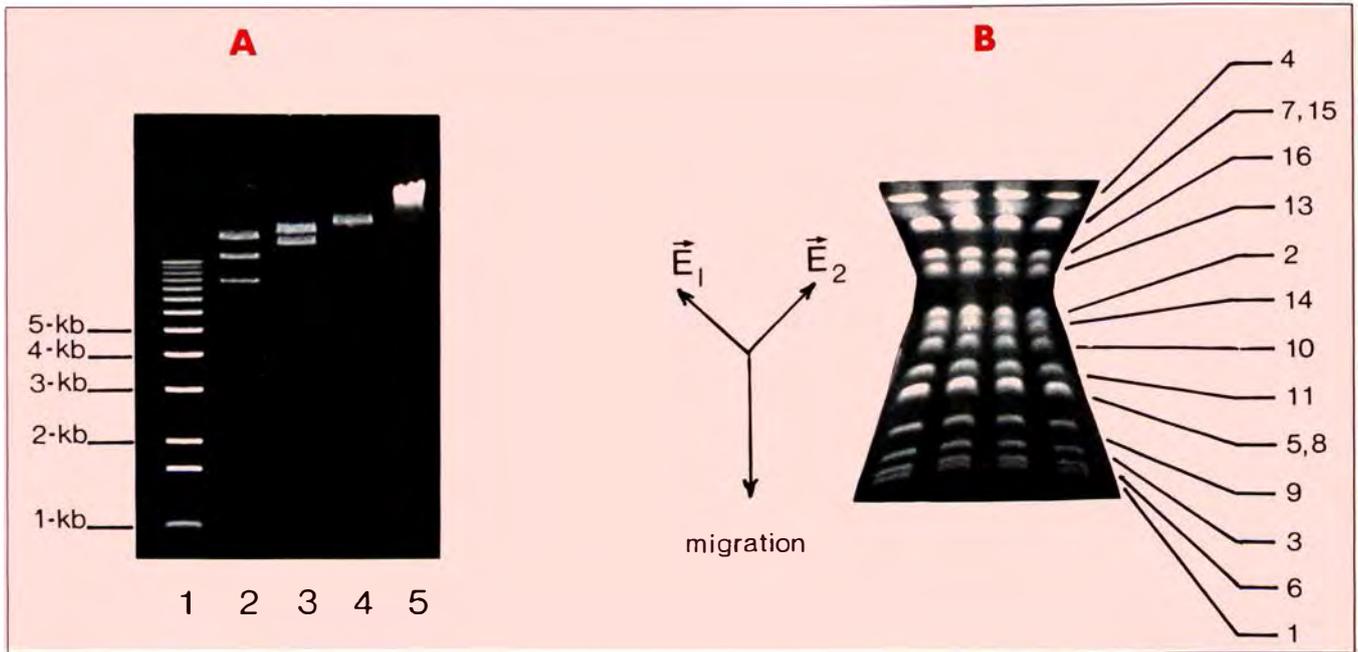


Figure 1. **A. Migration électrophorétique en champ électrique constant de molécules d'ADN de longueur variable.** La migration s'effectue de haut en bas. 1 = fragments d'ADN obtenus par polymérisation d'un fragment de 1 kb. Le fragment le plus long est de 12 kb ; 2 = trois fragments de restriction de l'ADN du bactériophage lambda de longueur 22 kb, 14 kb et 9 kb ; 3 = deux fragments de restriction de l'ADN du bactériophage de longueur 30 kb et 17 kb. 4 : L'ADN du bactériophage lambda, 48-kb. 5 : L'ADN total de la levure *Saccharomyces cerevisiae*. Bien que les tailles des ADN chromosomiques soient comprises entre 300 kb et 2 000 kb (voir B de cette même figure), leurs migrations électrophorétiques sont identiques et très voisines de celles de l'ADN du bactériophage lambda. **B. Migration électrophorétique en champs électriques alternatifs.** Cette méthode, analogue à celle en champs pulsés (voir texte), consiste à appliquer alternativement deux champs électriques suivant des directions orthogonales : E_1 et E_2 . La direction de migration globale se fera donc suivant la bissectrice de l'angle. Les numéros de chaque chromosome sont indiqués sur la droite. Deux chromosomes sont absents : le chromosome 17 qu'à ce jour personne n'a encore pu identifier et le chromosome 12. Ce dernier, le plus grand de tous, ne migre pas dans les conditions utilisées ici. Deux paires de chromosomes (5,8) et (7,15) ont des longueurs voisines et donnent naissance à deux doublets. On peut les résoudre en utilisant d'autres souches de levure, cet organisme présentant un polymorphisme des longueurs de ses ADN chromosomiques.

de la conformation allongée prise par la molécule soumise à son effet. Il est indépendant de la longueur de la molécule. Pour des molécules supérieures à 50 kb, et avec des champs électriques utilisables en pratique, le premier terme est petit devant le second et la mobilité est pratiquement indépendante de la longueur. Ceci explique la perte de résolution observée. Comment l'éviter ? L'idée de Schwartz et Cantor est fort simple. Au moment où l'on applique le champ électrique, l'enveloppe de la molécule d'ADN

commence à se déformer par son déplacement. Le temps mis par une molécule d'ADN pour s'allonger dans le sens du champ, qui est aussi le temps de passage de la forme à enveloppe sphérique à celle à enveloppe oblongue (figure 2) et qu'on appelle le temps caractéristique d'allongement, est à peu près égal au temps nécessaire à cette molécule pour se déplacer d'une distance égale à sa longueur : il dépend donc de sa longueur. Au moment où l'on applique le champ électrique, la mobilité est m_0 . Elle augmente

progressivement pour atteindre au bout du temps caractéristique d'allongement la valeur $m_0 + m_E$. Considérons maintenant deux molécules d'ADN de longueur différente et appliquons le champ électrique pendant un temps compris entre les deux temps caractéristiques d'allongement. Seule la plus petite aura la possibilité de s'allonger suffisamment pour atteindre la mobilité maximale. Pour pouvoir exploiter cette différence transitoire de mobilité, il suffit alors d'arrêter le champ pendant un temps suffisant

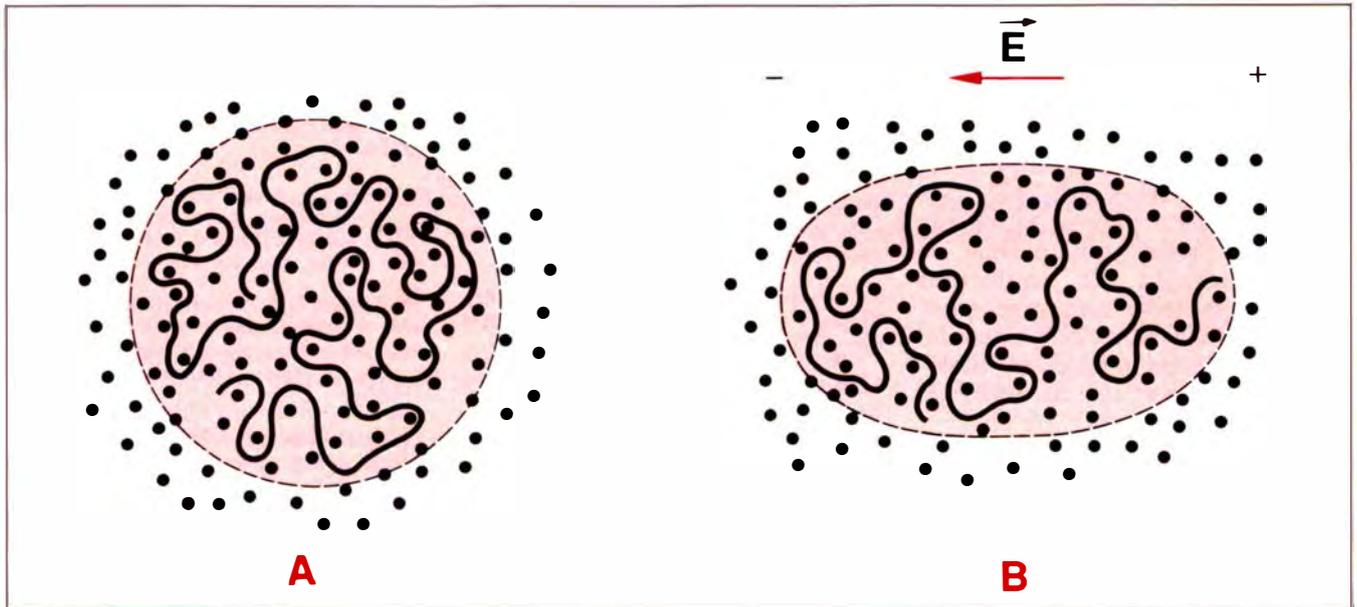


Figure 2. **Représentation schématique à deux dimensions d'une molécule d'ADN dans un gel d'agarose.** Les fibres représentées par des points sont distribuées au hasard ; l'ADN serpente entre elles, la courbure de l'ADN et la distance moyenne entre deux fibres voisines étant approximativement identiques (voir texte). **A** : En absence de champ électrique, l'enveloppe de la molécule d'ADN marquée en pointillés est isotrope (sphérique). **B** : Sous l'influence d'un champ électrique \vec{E} indiqué par une flèche rouge horizontale allant de droite à gauche, la molécule se déplace vers la droite en même temps qu'elle se déforme : son enveloppe s'allonge dans le sens du champ électrique.

pour que les molécules reviennent à leur forme de repos puis de recommencer ce cycle un grand nombre de fois. On peut également, au lieu d'arrêter le champ électrique, changer son orientation et jouer ainsi sur le temps mis par la molécule d'ADN pour s'allonger dans la nouvelle direction : en appliquant alternativement des champs électriques dans des directions orthogonales [1, 6, 9], on obtient des séparations équivalentes à celles observées avec la méthode en champ pulsé. Un exemple en est donné dans la *figure 1*.

Pour la biologie moléculaire, cette nouvelle technique confine au sensationnel. Grâce à elle, il est pratiquement possible, pour tous les organismes à petit génome, procaryotes et eucaryotes inférieurs, de localiser très rapidement un gène quelconque. Par exemple, 16 des 17 chromosomes de la levure *Saccharomyces cerevisiae* peuvent être séparés par électrophorèse en

champs pulsés (*figure 1*) et ont été identifiés par hybridation avec des gènes clonés dont on connaissait la localisation chromosomique. On peut ainsi en une journée trouver le (ou les) chromosome(s) portant un gène qu'on vient de cloner et il est même possible par une procédure simple de déterminer sa position par rapport au centromère. Les méthodes génétiques précédemment utilisées demandaient des semaines pour obtenir une telle information avec une précision bien moindre. L'ADN correspondant à un chromosome peut aussi être sorti du gel et utilisé pour faire une banque génomique restreinte. En ordonnant, par des procédés informatiques simples, les différents fragments de cette banque, on peut reconstituer la carte de l'ADN du chromosome entier.

Certains ADN chromosomiques ne peuvent pas, pour l'instant, être séparés par cette technique : il s'agit d'ADN linéaires dont la

longueur dépasse 9 000 kb (l'ADN des chromosomes d'eucaryotes supérieurs) ou des molécules d'ADN circulaires comme par exemple le génome d'*E. coli* (5 000 kb). On peut tout de même aborder leur étude en les coupant par des enzymes de restriction dont les sites de reconnaissance sont rares. Deux enzymes de restriction commercialisées, NotI et SfiI, reconnaissent des sites à 8 bases, alors que la plupart des autres enzymes reconnaissent des sites ayant, au plus, six bases. Les fragments créés par ces enzymes sont des molécules linéaires de taille inférieure, en général, à 2 000 kb. Ils peuvent alors être étudiés par électrophorèse en champ électrique pulsé. Les résultats de telles expériences permettent de situer les fragments les uns par rapport aux autres et conduisent à la construction de cartes génétiques locales extrêmement précises pouvant s'étendre sur plusieurs millions de paires de

bases. Plusieurs régions du génome humain ont fait l'objet de telles études. Citons, par exemple, celle du gène responsable de la myopathie de Duchenne [10, 11] et la région HLA [12, 13]. Grâce à l'électrophorèse en champ électrique pulsé qui permet le fractionnement de molécules d'ADN de taille comprise entre 50 et 9 000 kb, un nouveau domaine de recherches en biologie moléculaire est en train de s'ouvrir ■

François Caron

Remerciements

Je tiens à remercier plus particulièrement H. Hervet et G. Thomas pour leur lecture approfondie de ce manuscrit et pour de judicieuses corrections.

RÉFÉRENCES

1. Schwartz DC, Cantor CR. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* 1984 ; 37 : 67-75.
2. De Gennes PG. Reptation of the polymer chain in the presence of fixed obstacles. *J Chem Phys* 1971 ; 55 : 572-8.
3. Lumpkin OJ, Déjardin P, Zimm BH. Theory of gel electrophoresis of DNA. *Biopolymers* 1985 ; 24 : 1573-93.
4. Slater GW, Noolandi J. On the reptation theory of gel electrophoresis. *Biopolymers* 1986 ; 25 : 431-54.
5. Hervet H, Bean CP. Electrophoretic mobility of phage Hind III and Hae II DNA fragments in agarose gels : a detailed study. *Biopolymers* 1987 (sous presse).
6. Carle GF, Olson MV. Separation of chromosomal DNA molecules from yeast by orthogonal-field-alternation gel electrophoresis. *Nucleic Acid Res* 1984 ; 12 : 5647-64.
7. Lerman LS, Frisch HL. Why does the electrophoretic mobility of DNA in gels vary with the length of the molecule. *Biopolymers* 1982 ; 21 : 995-7.
8. Lumpkin OJ, Zimm BH. Mobility of DNA in gel electrophoresis. *Biopolymers* 1982 ; 21 : 2315-6.
9. Smith CL, Matsumoto T, Niwa O, et al. An electrophoretic karyotype for *Schizosaccharomyces pombe* by pulsed field gel electrophoresis. *Nucleic Acids Res* 1987 ; 15 : 4481-9.
10. Ommen GJB, Verkerk JMH, Hofker MH et al. A physical map of 4 million bp around the Duchenne muscular dystrophy gene on the human X-chromosome. *Cell* 1986 ; 47 : 499-504.
11. Burmeister M, Lehrach H. Long-range restriction map around the Duchenne muscular dystrophy gene. *Nature* 1986 ; 324, 582-5.
12. Hardy DA, Bell JI, Long EO, Lindsten T, McDevitt HO. Mapping of the class II region of the human major histocompatibility complex by pulsed-field gel electrophoresis. *Nature* 1986 ; 323 : 453-5.
13. Lawrence SK, Smith CL, Srivasta R, Cantor CR, Weissman SM. Megabase-scale mapping of the HLA gene complex by pulsed field gel electrophoresis. *Science* 1987 ; 235 : 1387-90.



COLLECTION PROSPECTIVES FRANCOPHONES

Co-édition
UREF/AUPELF/
John Libbey Eurotext

Francophonie Scientifique : Le Tournant

1989, broché
ISBN 0 86 196 248 6
Continent Africain : 60 FF
Autres pays : 120 FF

6, rue Blanche
92120 Montrouge
Tél. : 47.35.85.52

Hybridation moléculaire :

Les polymorphismes de taille des fragments de restriction (RFLP, Restriction Fragment Length Polymorphisms)

L'ADN diffère entre les individus, tant au niveau des gènes qui constituent la base de leur individualité, que des séquences intergéniques. La variabilité est même beaucoup plus forte pour les séquences non codantes et non régulatrices (où les mutations peuvent s'accumuler sans conséquences) qu'au niveau des exons des gènes et des zones de contrôle de leur expression. Cette variabilité de l'ADN d'un même locus entre deux individus peut modifier les sites de reconnaissance des enzymes de restriction et donc les points de clivage. La figure 1A montre ainsi que la mutation

C/G → T/A fait disparaître un site de clivage par l'enzyme de restriction Msp-1, de telle sorte que la digestion par cette enzyme donne en « Southern blot » hybridé avec la sonde 1, un fragment de 3 kilobases (kb) au lieu de 2 kb. Une autre cause importante de polymorphisme des régions non codantes est la répétition variable d'éléments d'ADN, représentée dans la figure 1B. Ici, après digestion par l'enzyme EcoR 1, la sonde 2 hybride avec un fragment de 9 kb si le motif de base est répété 9 fois, et avec un fragment de 7 kb s'il est répété 7 fois. Dans les deux

types de polymorphisme, le résultat est une modification de la taille d'un fragment d'ADN analysé par « Southern blot », d'où le nom anglais de RFLP, *Restriction Fragment Length Polymorphism*. La figure 2 montre un exemple simple de l'utilisation de tels polymorphismes pour étudier la transmission d'un gène pathologique. Une maladie récessive est liée à la transmission d'un gène situé à proximité d'un site polymorphe. L'étude familiale a permis de démontrer que le caractère pathologique était lié au polymorphisme A, donnant un fragment d'ADN de 9 kb, alors

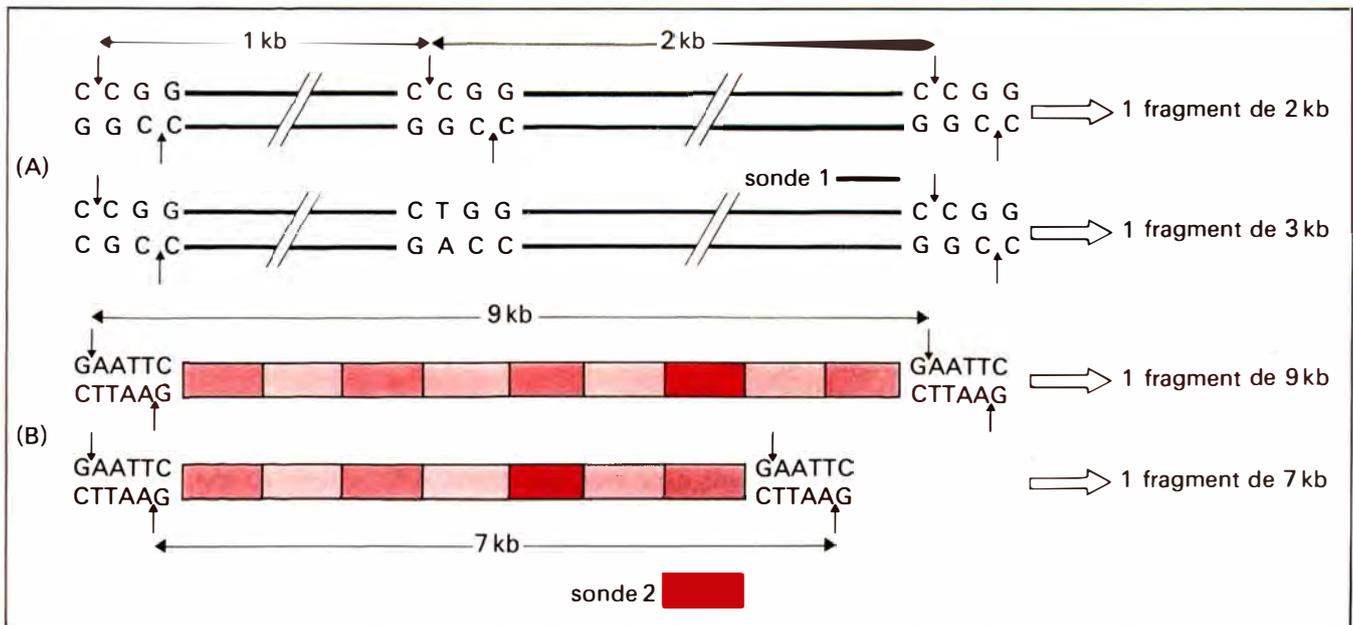


Figure 1. Mécanismes possibles du polymorphisme de taille des fragments de restriction. (A) : Au niveau d'un allèle, 3 sites de restriction clivés par l'enzyme Msp-1 sont espacés de 1 et 2 kilobases. Après coupure par l'enzyme, électrophorèse et transfert sur filtre (voir « Southern blot », Lexique médecine/sciences n° 2, vol. 2), la sonde 1 reconnaît un fragment de 2 kb. Au niveau de l'autre allèle le site Msp-1-central a été modifié, un C|G étant remplacé par T|A. L'enzyme de restriction ne clive plus ce site et la sonde 1 hybride donc avec un fragment de 3 kb. (B) : Deux allèles diffèrent par le nombre d'éléments répétés entre deux sites de clivage par l'enzyme EcoR 1. Après digestion par cette enzyme, un « Southern blot » hybridé par la sonde 2 montrera une bande de 9 kb pour le premier allèle et de 7 kb pour le second. Dans les deux exemples donnés les flèches verticales indiquent les sites de coupure.

que l'allèle normal était associé au polymorphisme B donnant un fragment d'ADN de 7 kb. L'examen de l'ADN d'un fœtus issu de deux sujets hétérozygotes permettra de prévoir si l'enfant à naître sera normal (présence du seul fragment de 7 kb), hétérozygote (présence des deux fragments) ou homozygote malade (présence du seul fragment de 9 kb). Cette méthode sera sûre si le gène étudié est très proche du site polymorphe; s'il est trop éloigné, des recombinaisons survenant à la méiose (crossing over) modifieront l'association entre eux et entraîneront de fausses prédictions. La *figure 2* montre un hétérozygote dont l'ADN se comporte comme celui d'un homozygote malade car un événement de recombinaison est survenu.

La puissance de la méthode réside en ce qu'elle constitue l'une des voies pour localiser, puis identifier des gènes inconnus responsables de maladies. Si des études génétiques familiales montrent en effet qu'un site polymorphe, éventuellement testé au hasard, est associé à la transmission d'une maladie, cela signifie que le gène modifié dans cette maladie est génétiquement proche du site étudié. Il restera alors à détecter d'autres sites polymorphes de plus en plus proches du locus morbide, puis le gène lui-même. Avant même cette étape ultime, les sondes reconnaissant les sites polymorphes seront des outils précieux de détection pré et post-natale des malades et des porteurs. C'est grâce à cette approche qu'ont pu être localisés les gènes dont les mutations sont responsables de la chorée de Huntington, de la myopathie de Duchenne, de la fibrose kystique du pancréas, etc. De nombreuses équipes tentent d'analyser selon ce principe les gènes de susceptibilité aux cancers familiaux ou à l'athérosclérose.

A. K.

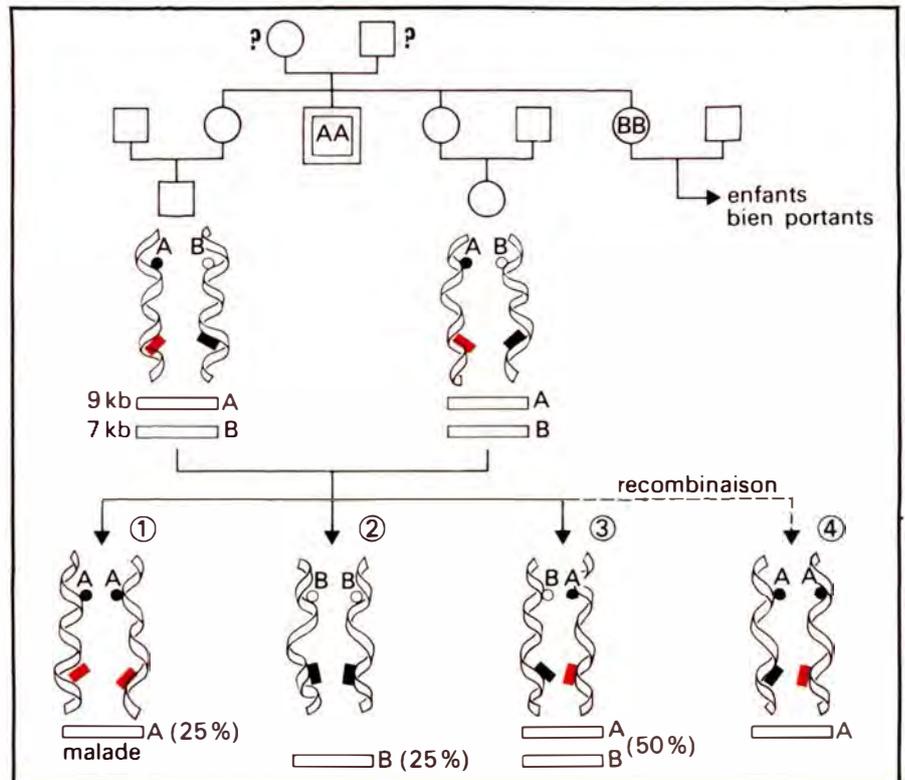


Figure 2. Exemple d'utilisation d'un polymorphisme de taille des fragments de restriction pour détecter des hétérozygotes d'une maladie autosomale récessive et pour faire un diagnostic prénatal de l'affection.

Rectangle noir : gène normal; Rectangle rouge : gène pathologique; A : allèle d'un site polymorphe donnant un fragment de 9 kb; B : allèle du même site donnant un fragment de 7 kb.

De haut en bas, des parents cousins germains ont un oncle malade (double carré), et une tante bien portante dont tous les enfants sont également bien portants. L'étude des polymorphismes de restriction indique que l'oncle n'a que l'allèle A donnant un fragment de 9 kb; il est donc homozygote AA. La tante est elle homozygote BB. L'affection est donc génétiquement liée à l'allèle A. Les deux parents sont hétérozygotes AB (l'étude de leur ADN montre deux fragments de 7 et 9 kb). Un quart des enfants risque d'être malade; le prélèvement de cellules du liquide amniotique ou d'une villosité chorionale permettra d'isoler de l'ADN du fœtus et donc de prédire s'il est homozygote pour la maladie (1 présence du seul fragment de 9 kb), tout à fait normal (2, présence du seul fragment de 7 kb), ou bien hétérozygote (3, présence des deux fragments). Si le site polymorphe est trop éloigné du gène morbide, des recombinaisons pourront survenir, schématisées en 4 : le polymorphisme A se trouve ici associé au gène normal, conduisant au faux diagnostic d'homozygotie pathologique d'un fœtus en fait hétérozygote. Ces faux diagnostics peuvent être évités en utilisant des sondes très proches du locus morbide, ou bien en combinant l'utilisation de deux sondes reconnaissant des polymorphismes situés de part et d'autre du gène morbide. A la limite, le site polymorphe se trouve dans le gène lui-même, et les recombinaisons deviennent hautement improbables.

Les différents types de mutations d'un gène et leurs conséquences

Mutations ponctuelles

Une mutation ponctuelle est le changement d'une base par une autre. Ses conséquences peuvent être multiples.

1,2 et 3. Une mutation des régions de contrôle de l'expression du gène, notamment des boîtes CAAT et ATA du promoteur, peut modifier la transcription et aboutir à une diminution de la quantité du messager, donc de la protéine correspondante.

4. Une mutation intéressant le codon d'initiation de la traduction (ATG) ou le pentanucléotide le précédant (se rapprochant du consensus CCA^ACC) peut diminuer la traductibilité du messager, donc la concentration de la protéine pour laquelle il code.

5. Une mutation d'une base dans

un exon peut être muette si elle ne change pas le sens du codon, aboutir à une protéine altérée si elle change le sens du codon (« mutation faux-sens »), ou à une absence de protéine si elle fait apparaître un codon stop très proche du début de la traduction. Une semblable « mutation non sens » située près de la fin normale de la protéine pourrait aboutir à la synthèse d'un produit raccourci. Enfin, la conséquence d'une mutation ponctuelle exonique peut être l'apparition d'un site alternatif d'épissage (voir médecine/sciences, prochain Lexique : mécanismes des anomalies d'épissage).

6. Une modification des bases GT situées au début de l'intron supprime ce site d'épissage (voir médecine/sciences, prochain Lexique).

7,8. Une mutation intronique est le

plus souvent muette; elle peut aussi faire apparaître un site alternatif d'épissage (voir prochain Lexique).

9. Une mutation du doublet AG situé à la fin des introns abolit ce site d'épissage (voir m/s, prochain Lexique).

10. Une mutation du codon stop supprime l'arrêt de la traduction à ce niveau. L'ARN est donc traduit jusqu'à ce qu'un nouveau codon stop soit rencontré, la conséquence étant la production d'une protéine allongée.

11. Une mutation du signal de polyadénylation AATAA peut diminuer la proportion des transcrits qui sont correctement clivés, puis polyadénylés. Les ARN ayant subi une maturation anormale sont instables, et la quantité du messager fonctionnel sera diminuée.

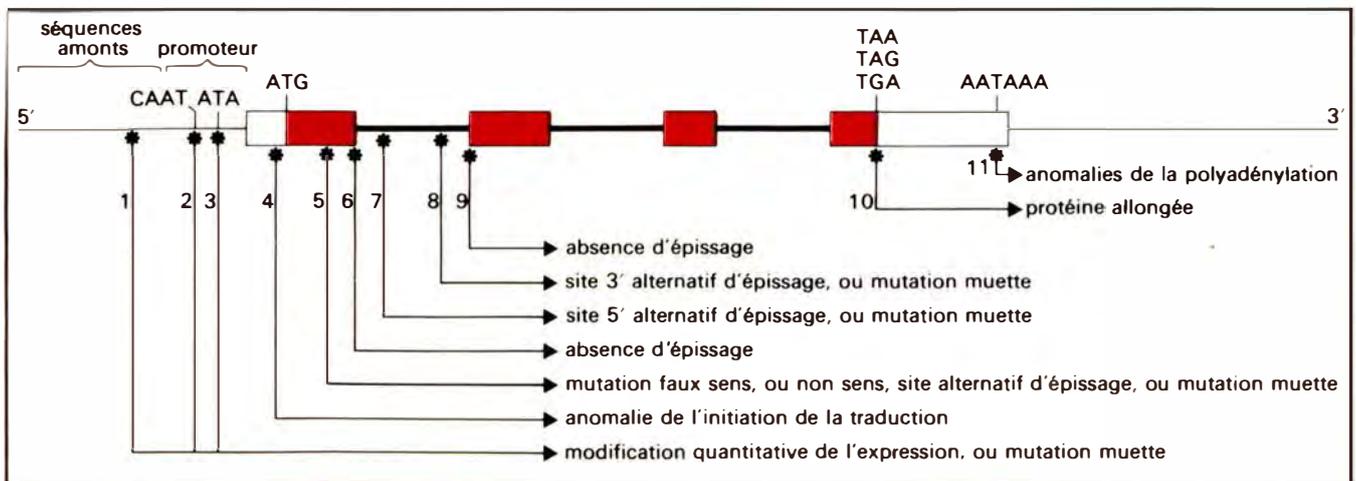


Figure 1. **Conséquences possibles des mutations ponctuelles d'un gène et de ses régions régulatrices.** Les traits fins représentent les régions intergéniques, les traits épais les introns, les rectangles blancs la partie non codante des exons et les rectangles rouges la partie codante des exons. Les « boîtes » CAAT et ATA du promoteur sont représentées, ainsi que le codon d'initiation de la traduction ATG, les codons stop TAA, TAG ou TGA, et le signal de polyadénylation AATAA.

Délétions et insertions

Les délétions de tout un gène ou d'un fragment important d'ADN sont souvent dues à un crossing-over non équationnel entre deux régions non alléliques partiellement ou totalement homologues, selon le mécanisme illustré dans le *Lexique de médecine/sciences* n° 4, vol. 1, p. 215, fig. 2. Elles aboutissent à une perte de la fonction du gène.

Des microdélétions de quelques bases peuvent, si elles emportent un multiple de 3 bases dans un exon, aboutir à la synthèse d'une protéine anormale, à laquelle manqueront quelques aminoacides. Si la délétion supprime un non-multiple de 3 bases, elle créera un décalage de phase de lecture et, en aval de la mutation, le sens des codons sera entièrement modifié; ils seront traduits en une protéine n'ayant strictement rien à voir avec celle normalement codée par le gène (*voir médecine/sciences, prochain Lexique*).

Les insertions sont des phénomènes rares qui peuvent être dus à l'intégration dans un gène d'un rétrovirus, ou d'un élément transposable. Les éléments transposables, surtout connus chez les procaryotes, les plantes et les drosophiles, sont des segments d'ADN possédant à leurs deux extrémités des séquences répétées (homologues des LTR des rétrovirus), qui ont la propriété d'être excisés du génome et de se réintégrer en d'autres sites. Certaines séquences répétitives des mammifères pourraient avoir cette propriété.

Une insertion peut également être secondaire à une duplication partielle d'un gène. Les conséquences sont parallèles à celles d'une délétion et, pour de très petites insertions, varient selon que le fragment inséré contient un multiple de trois bases (entraînant la synthèse d'une protéine altérée) ou un non-multiple de trois bases (produisant un décalage de phase de lecture).

Fréquences relatives des différents types de mutation

Elles varient évidemment selon les gènes mutés. Très approximativement, les délétions représentent de 5 à 15 % des mutations connues chez l'homme.

Parmi les mutations ponctuelles, certaines sont très rares (mutations des zones de contrôle). Les plus habituelles sont sûrement les mutations faux-sens et les anomalies d'épissage.

A. K.

Mutations des gènes : anomalies d'épissage et décalages de phase de lecture

Dans le lexique du précédent numéro de *médecine/sciences*, nous avons passé en revue les conséquences des différents types de mutation d'un gène. Nous détaillerons ici les divers mécanismes des anomalies d'épissage et des décalages de phase de lecture.

La *figure 1* rappelle que des séquences très conservées (dites séquences consensus) sont retrouvées aux deux extrémités des introns, qui commencent toujours par GT (sur l'ADN) et finissent par AG.

1) Une mutation de ces deux dinucléotides ou, à un moindre titre, d'un autre résidu des séquences consensus peut abolir l'excision des

introns. Les transcrits restent alors dans le noyau où ils sont dégradés, et le résultat net de cette mutation est un déficit quantitatif en ARN messager cytoplasmique.

2) Une mutation peut faire apparaître un site alternatif d'épissage « 1' » dans l'intron. L'excision entre les sites normaux 1 et 2 donnera l'ARN messager normal alors que l'excision entre les sites 1' et 2 donnera un ARN aberrant contenant une insertion anormale sous la forme d'un fragment d'intron persistant. Si cet ARN peut néanmoins atteindre le cytoplasme, il pourra soit entraîner la synthèse d'une protéine allongée contenant

une insertion peptidique si l'insertion contient un multiple de 3 bases et pas de codon stop, soit une protéine raccourcie si l'insertion contient un codon stop prématuré; à la limite, si le codon stop anormal est très proche du site AUG d'initiation de la traduction, il n'y aura pas de synthèse protéique. Si l'insertion ne contient pas un multiple de trois bases, cela provoquera un décalage de phase de lecture dont les conséquences sont expliquées dans la *figure 2*.

3) Une mutation exonique peut

également faire apparaître un site supplémentaire, alternatif, d'épissage « 1'' ». L'excision anormale entre les sites 1'' et 2 produit un ARN délété qui peut soit être traduit en une protéine raccourcie, soit donner un décalage de phase si la délétion ôte un non-multiple de trois bases.

La figure 2 montre le principe de ce que sont les mutations avec décalage de phase de lecture (*frameshift*, en anglais). Sur la première ligne est représenté un ARN avec le sens normal de certains de ses codons (Valine, Asparagine, Alanine), un codon stop UAA et une extension 3' d'acide polyadénylique.

1) La délétion d'une base peut entraîner la lecture UGA, qui signifie stop, au lieu de la lecture normale des codons Val et Asn. Selon que ce signal stop est proche du début ou de la fin de la séquence codante, le résultat en sera l'absence de synthèse protéique ou la synthèse d'une protéine aberrante, raccourcie.

2) La délétion des deux bases G et U du codon GUG amène à une lecture GAA (codon de l'acide glutamique)... puis à la lecture d'autres codons qui seront tous différents de ceux du messenger normal. En aval de la mutation, la séquence de la protéine mutée sera entièrement différente de celle de la protéine normale.

3) Une délétion des bases G et C du codon Ala amène à lire un codon UUA (qui signifie leucine)... et non le codon stop normal UAA du messenger. La traduction continue donc dans la région 3' non codante du messenger normal jusqu'à ce que, par hasard, un nouveau codon stop soit rencontré. Au maximum, si aucun codon stop nouveau n'apparaît, l'extension poly-A peut être traduite sous la forme de plusieurs codons AAA... qui signifie « lysine ».

De semblables « décalages de phase de lecture » sont produits par toutes insertions ou délétions de non-multiples de 3 bases. On connaît, en pathologie humaine ou animale, des exemples de tous ces types de mutation, les plus étudiées étant celles des gènes de l'hémoglobine.

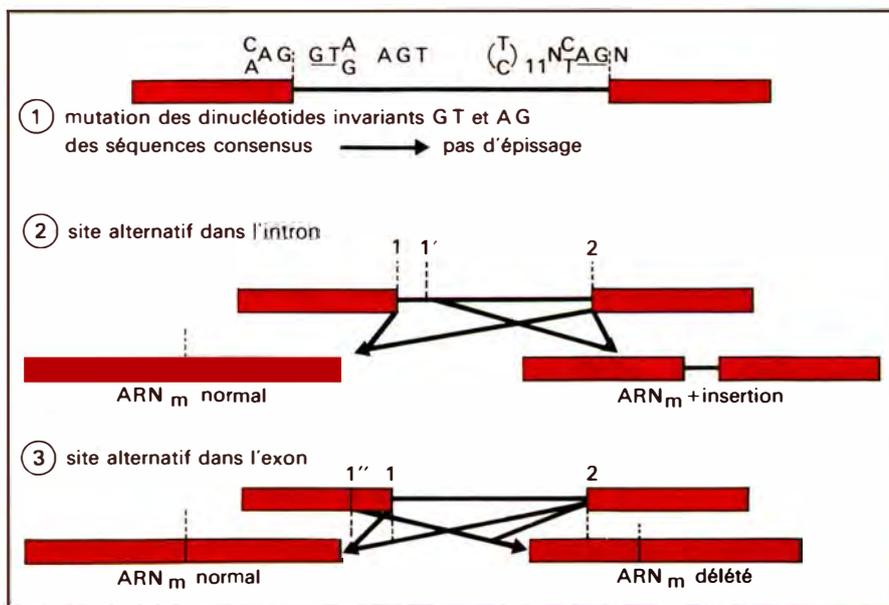


Figure 1. Différents types d'anomalies de l'épissage, et leurs conséquences.

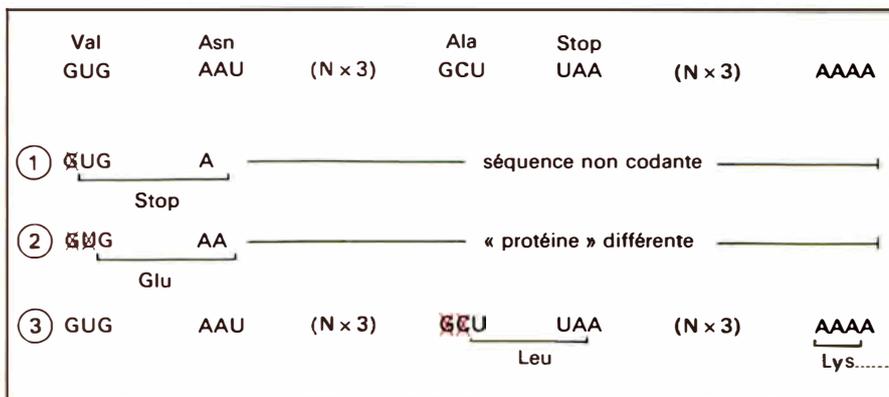


Figure 2. Différents types de « décalage » de phase de lecture, et leurs conséquences. (N x 3) indique un multiple de trois nucléotides.

Clonage de gènes

La possibilité de « cloner » un gène, c'est-à-dire d'isoler et d'amplifier des millions de fois la séquence d'ADN responsable du caractère étudié, est à la base de la génétique moléculaire; elle est notamment une méthode essentielle à la progression des connaissances sur le génome humain. Nous présentons dans la *figure 1* le principe de cette

méthode, dont il existe maintenant de nombreuses variantes. On sait ainsi, grâce à l'utilisation de vecteurs différents du phage λ (les cosmides), « cloner » des fragments beaucoup plus grands qu'avec ce dernier, de l'ordre de 35 à 45 kilobases (kb) au lieu de 15-20 kb.

A. K.

Figure 1. **Confection et criblage d'une banque d'ADN génomique dans le phage λ .**

1 : L'ADN à étudier est coupé partiellement par l'enzyme S (par exemple *Sau3A*) en fragments de 15 à 20 kilobases (kb).

2 : Ces fragments sont liés aux deux bras de l'ADN d'un phage λ . Le phage λ est constitué de deux bras, indispensables au cycle lytique, et d'une partie interne non indispensable qui peut donc, comme c'est le cas ici, être remplacée par de l'ADN étranger. La réaction, catalysée par une ligase, est appelée « ligation ».

3 : Une réaction in vitro permet d'empaqueter l'ADN des phages recombinés dans des capsides phagiques vides. Chaque phage contient de 45 à 52 kilobases d'ADN total.

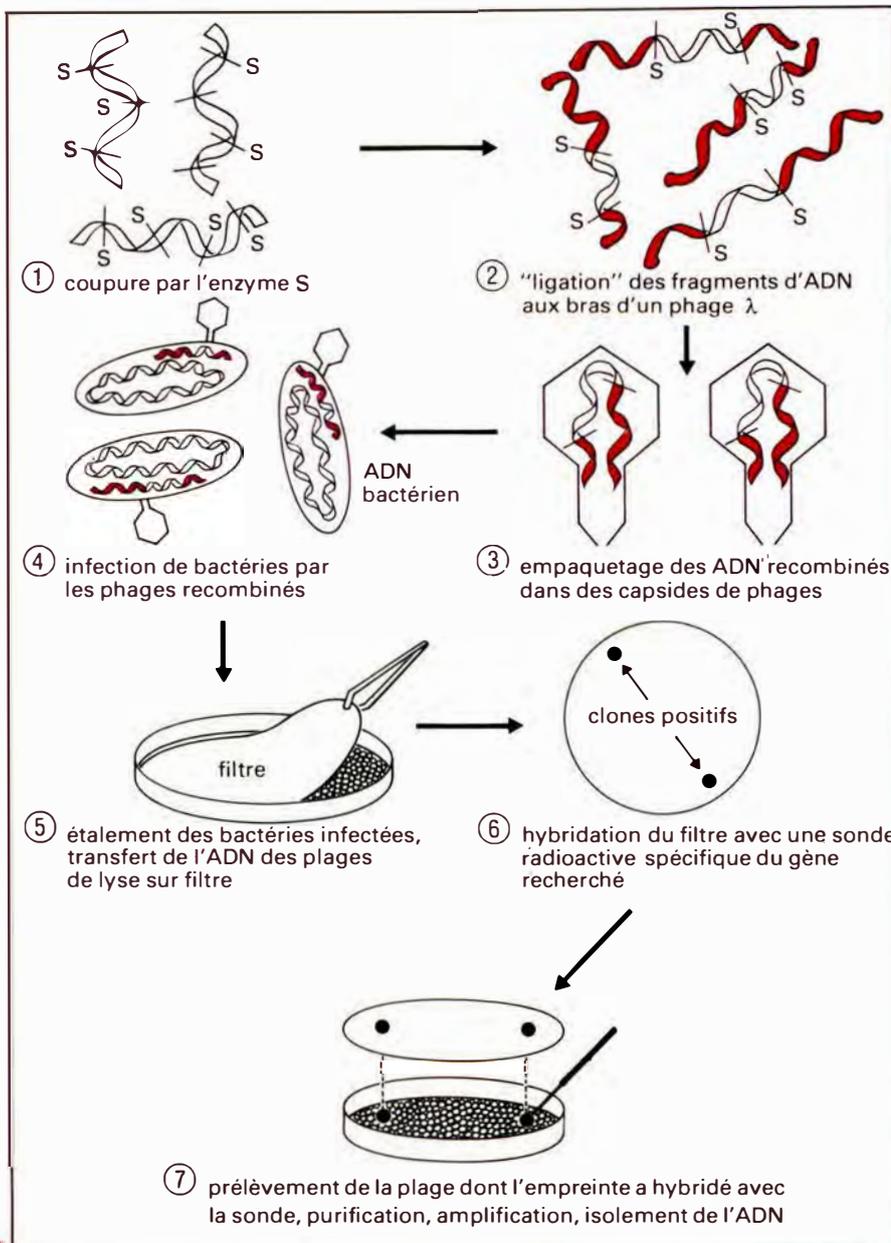
4 : La particule ainsi reconstituée est infectante, chaque bactérie n'étant infectée que par un seul phage

5 : Les bactéries infectées sont étalées sur des boîtes. Le génome humain faisant environ 3×10^6 kb, il faut au moins 200 000 phages contenant des fragments de 15 kb pour qu'il soit représenté en entier. En fait, la présence ou l'absence d'un phage contenant un fragment particulier parmi ceux analysés est aléatoire, et en général on s'efforce d'explorer au moins 1 000 000 de phages recombinés (ce qui représente 5 « équivalents génomes ») pour être plus sûr de ne pas manquer celui que l'on cherche.

Les phages sont lytiques, et aux alentours de chacune des bactéries infectées se développe une « plage de lyse » au niveau de laquelle sont libérées les particules phagiques. Celles-ci sont transférées sur filtre et l'ADN est dénaturé in situ par un traitement alcalin.

6 : Les filtres sont hybridés avec une sonde spécifique du gène recherché : seul l'ADN des plages provoquées par un phage contenant le gène ou un de ses fragments hybridera avec la sonde radioactive.

7 : A la tache du filtre correspond sur la boîte dont il est l'empreinte une plage de lyse dont les phages sont prélevés et amplifiés : ils contiennent alors une grande quantité du gène étudié !



Les récepteurs des neuromédiateurs...

ou vers la structure des protéines inconnues

La majorité des stratégies de clonage d'ADNc et de gènes fait appel à quelques caractéristiques de la protéine pour laquelle ils codent : anticorps, éléments de séquence, etc. Dans d'autres cas (illustrés par la démarche de la « génétique inverse »), le point de départ du clonage est la localisation du gène et la possession de mutants. Il est également possible de cloner l'ADNc codant pour des protéines actives dont on ne sait rien, codées par des gènes dont on ne connaît pas la localisation. Il suffit de savoir « ce que fait » cette protéine et de pouvoir le mesurer dans un système hétérologue pour réaliser cela. Le principe de cette approche est le suivant.

Étape 1. Une banque d'ADNc est construite (voir *lexique m/s, suppl. au n° 7 vol. 3, p. 14*) dans un vecteur contenant la séquence promotrice reconnue par certaines ARN-polymérases de phages, très actives (par exemple des phages T3, T7 et SP6). L'ARN utilisé a été purifié à partir du tissu ayant l'activité biologique étudiée la plus forte possible.

Étape 2. La banque est divisée en plusieurs groupes (par exemple de quelques milliers ou dizaines de milliers de clones chacun) dont les ADNc sont transcrits, grâce à l'ARN-polymérase de phage appropriée, en une population d'ARN.

Étape 3. Ces populations d'ARN sont injectées dans des ovocytes de batraciens (xénopes le plus souvent) au niveau desquels on dose l'activité biologique étudiée. Si la population d'ARN injectée contenait l'ARN codant pour la protéine responsable de cette activité, ce dernier a dû être traduit dans l'ovocyte en une protéine qui peut être fonctionnelle.

Étape 4. Le groupe de clones ayant donné un résultat positif est lui-même subdivisé en sous-groupes successifs testés de la même manière. La répétition de la procédure doit permettre d'aboutir au clone codant pour la protéine recherchée. Une telle approche, voisine de la méthode historique employée au début du génie génétique par C. Weissman pour cloner l'ADNc de l'interféron α , a été récemment mise en œuvre pour cloner les ADNc des récepteurs de deux neuromédiateurs, la substance K [1] et la sérotonine [2]. Le test biologique utilisé au niveau des ovocytes après micro-injection des ARN a été, dans les deux cas, l'ouverture d'un canal ionique induit par le neuromédiateur correspondant, et donc un courant de membrane qu'il est aisé de détecter.

Dans son principe, la méthode résumée ci-dessus n'a guère de limitations, que celle d'un système facile à doser dans l'ovocyte, et par conséquent dont il n'existe pas d'équivalent dans cette cellule. Sinon d'autres systèmes cellulaires pourraient être envisagés, dont aucun ne réunit néanmoins les avantages de l'œuf de xénope.

A. K.

**John Libbey
EUROTEXT
LONDON PARIS**

**Paris · Londres
EDITIONS MEDECINE-SCIENCES**

J—CP

à paraître

**Journal of
Cellular
Pharmacology**

n° 1, vol. 1, Mars 1990
(mensuel)

6, rue Blanche, 92120 Montrouge
Tél. : (1) 47.35.85.52

1. Masu Y, Nakayama K, Tamaki H, Harada Y, Motoy K, Nakanishi S. cDNA cloning of bovine substance-K receptor through ovocyte expression system. *Nature* 1987 ; 329 : 836-40.
2. Julius D, Mac Dermott AB, Axel R, Jessell TM. Molecular characterization of a functional cDNA encoding the serotonin 1c receptor. *Science* 1988 ; 241 : 558-64.

Analyse directe des anomalies de l'ADN chez l'homme

Une maladie en rapport avec une mutation congénitale ou acquise (cas des cancers) d'un gène peut être analysée au niveau du produit du gène muté et de son ARN, quand ils sont présents, ou bien directement au niveau du gène. C'est ce dernier point que nous analyserons ici, l'étude sur les ARN faisant l'objet du « Lexique » du prochain numéro.

1. Les délétions ou les insertions d'un nombre suffisant de bases (plus d'une centaine): elles sont aisées à détecter par *Southern blot*; elles entraînent la disparition ou le changement de taille d'un fragment révélé par hybridation moléculaire avec les sondes radioactives spécifiques du gène normal.

2. Les mutations ponctuelles, microdélétions ou microinsertions: leur identification nécessite en revanche l'utilisation de méthodes beaucoup plus complexes.

- Confection d'une banque d'ADN génomique, détection du clone muté, détermination de la séquence nucléotidique (voir *Lexique de médecine|sciences n° 6, vol. 2*). La méthode est extrêmement lourde pour des gènes qui peuvent être immenses (près de 200 kilobases pour le gène du facteur VIII!). Si l'on dispose de sondes reconnaissant tous les exons (par exemple une copie d'ADN complémentaire complète du messager), il est alors possible de n'étudier qu'eux en faisant l'impasse sur une mutation intronique qui pourrait entraîner des anomalies d'épissage.

- Digestion de l'ADN par toute une batterie d'enzymes de restriction reconnaissant des séquences différentes, et hybridation du *Southern blot* avec des sondes spécifiques. Quoique plus d'une centaine d'enzymes de restriction soient connues à l'heure actuelle, cette méthode n'explore pas toutes les séquences possibles!... et elle est de

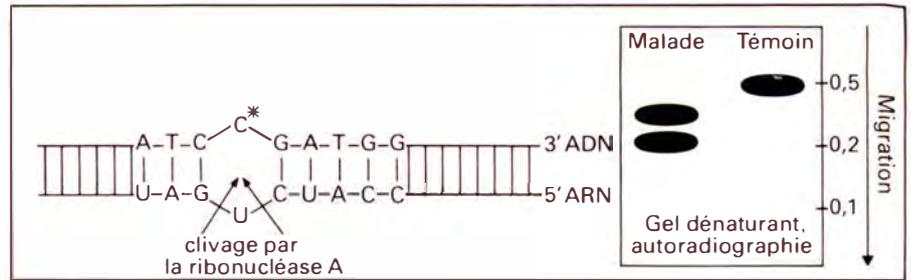


Figure 1. **Protection d'hybrides ADN-ARN contre la digestion par la ribonucléase A.** A gauche, un fragment d'ADN génomique inconnu est dénaturé (c'est-à-dire que ses brins sont dissociés), l'un des brins étant hybridé avec une sonde d'ARN radioactif connue. S'il existe un seul mauvais appariement entre les brins d'ADN et d'ARN (sur la figure, un C, marqué d'un astérisque, au lieu d'un A), la ribonucléase A pourra cliver l'ARN à ce niveau. A droite, la sonde ARN est analysée, après traitement, par migration en gel dénaturant: elle est entière si l'ADN est « normal » et clivée si l'ADN est « muté ».

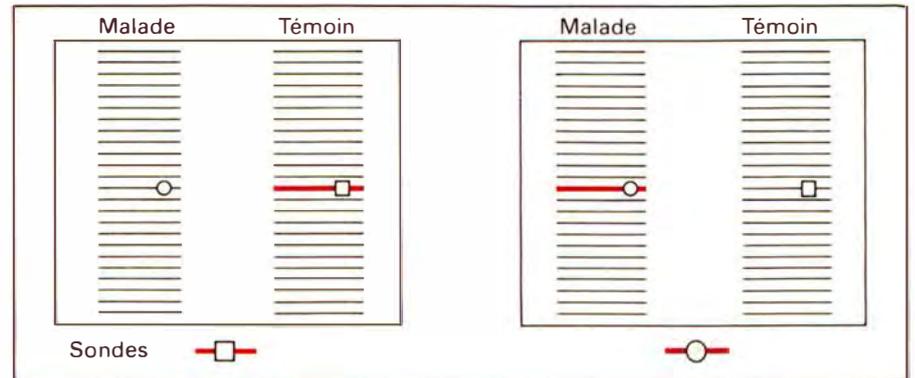


Figure 2. **Détection d'une mutation particulière par hybridation de Southern blots avec des oligonucléotides de synthèse.** Des oligonucléotides de 17 résidus complémentaires de la région à explorer sont synthétisés. Ils sont conçus de telle sorte que la base éventuellement modifiée se trouve au milieu de l'oligonucléotide. De telles sondes spécifiques du gène normal hybrident mieux avec un fragment de ce dernier qu'avec le fragment équivalent du gène muté. Par contre un oligonucléotide complémentaire du gène muté hybridera mieux avec lui qu'avec le gène normal. —□—: oligonucléotide complémentaire du gène normal (□: base éventuellement modifiée); —○—: oligonucléotide complémentaire du gène muté (○: base mutée). Le trait rouge indique le fragment d'ADN hybridant avec la sonde, les traits fins représentant les nombreux fragments d'ADN transférés sur les filtres mais non reconnus par les sondes.

plus aussi fastidieuse qu'onéreuse. Enfin, il n'est pas possible devant la découverte d'un site de clivage anormal (ou la disparition d'un site normal) de prédire si l'anomalie est vraiment la cause de la maladie ou bien ne représente qu'un polymorphisme sans retentissement fonctionnel.

- Protection d'hybrides ADN génomique ARN contre la digestion par la ribonucléase A (figure 1). Cette méthode utilise des sondes ARN faciles à obtenir depuis deux ans, grâce à la construction de plasmides contenant un site d'initiation de la transcription par une ARN polymérase très spécifique. Lorsque cet ARN rendu radioactif est hybridé avec un brin d'ADN parfaitement complémentaire, il devient insensible à la digestion par la ribonucléase A. Par contre le mauvais appariement d'une seule base crée à ce niveau un site de clivage aisé à caractériser par l'analyse de la taille de la sonde d'ARN ainsi traitée. Cette méthode, de description récente, est sûrement extrêmement prometteuse; elle devrait permettre d'étudier tout fragment d'ADN génomique non cloné pour peu que l'on dispose de sondes complémentaires. Par contre, si la méthode indique le site d'une mutation, elle n'en indique pas la nature dont la précision reste du ressort du séquençage nucléotidique du fragment muté préalablement cloné.

- Recherche d'une mutation en un site connu par hybridation à des oligonucléotides de synthèse (figure 2). Certaines mutations particulières sont fréquentes (cas de nombreuses thalassémies, ou bien des mutations des codons 12 et 61 des gènes *c-ras* au cours de cancers) et le problème est alors de les détecter chez des individus donnés, et non d'explorer la totalité du gène. Des oligonucléotides de synthèse de 17 ou 18 résidus sont alors utilisés. Leur hybridation avec une séquence totalement homologue étant significativement plus stable que celle avec une séquence ne comportant qu'une seule base modifiée, ils permettent d'explorer rapidement l'existence d'une mutation de cette séquence, et également d'en identifier fermement

la nature (voir figure 2).

- Amplification de la région mutée (figure 3). Cette méthode consiste à amplifier des dizaines de milliers de fois la région mutée en la recopiant à partir d'amorces nucléotidiques. Une telle opération facilite évidemment toute opération ultérieure de caractérisation fine de cette région. Comme dans la technique précédente, il s'agit ici de rechercher des mutations particulières préalablement caractérisées chez d'autres malades.

3. Conclusion

L'énumération qui précède n'est pas exhaustive et comporte une part de parti pris en fonction du sentiment que l'auteur a de l'avenir réservé à différentes techniques continuellement publiées. N'a pas été citée, par exemple, une méthode

électrophorétique jouant sur la migration différentielle dans un gradient d'agents dénaturants d'hybrides ADN/ADN parfaits ou partiellement mal-appariés.

Il faut retenir, en résumé, que l'identification précise de la mutation d'un gène connu est, aujourd'hui, toujours possible. Mais si certaines techniques sont relativement simples et peuvent être appliquées à un grand nombre d'échantillons (mise en évidence d'une délétion ou bien détection d'une mutation « connue » à l'aide d'oligonucléotides de synthèse), d'autres restent extrêmement lourdes et du domaine exclusif de laboratoires importants. C'est le cas chaque fois qu'il s'agit de rechercher une mutation ponctuelle en un site inconnu du gène.

A. K.

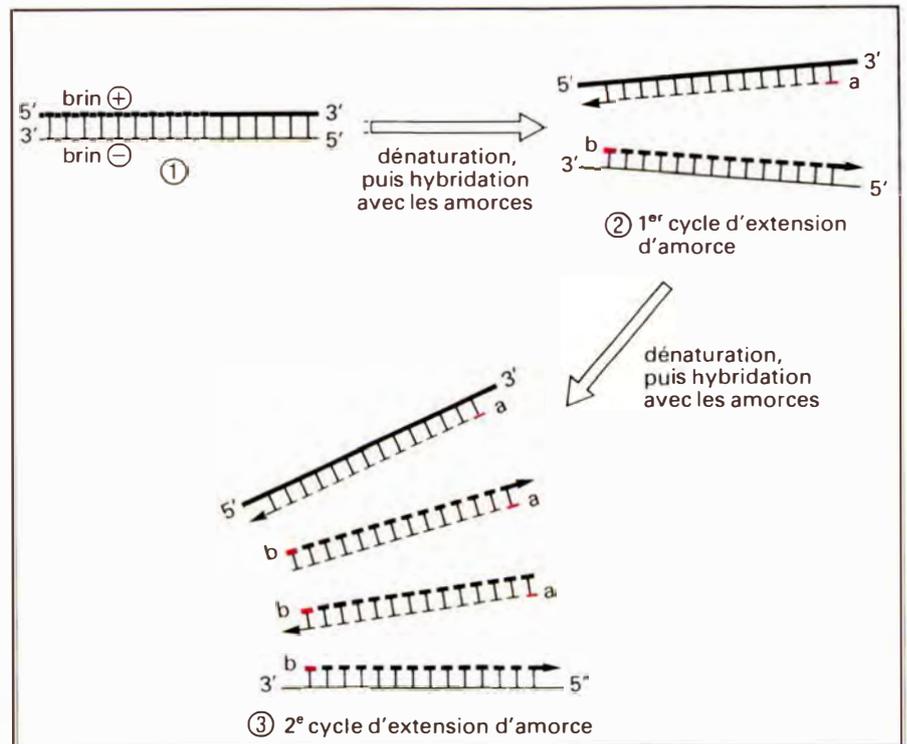


Figure 3. **Amplification d'une région particulière de l'ADN génomique.** L'ADN en 1 est dénaturé, chacun des brins étant hybridé avec un oligonucléotide de synthèse (traits rouges). En 2, les oligonucléotides « a » (complémentaires du brin +, épais) et « b » (complémentaires du brin -, fin) servent d'amorce pour la synthèse de chaînes complémentaires d'ADN sous l'action de l'ADN polymérase. Cette réaction est appelée « extension d'amorce ». En 3, l'opération est recommencée, aboutissant à une amplification des régions comprises entre les séquences reconnues par les 2 oligonucléotides de synthèse. Les brins initiaux d'ADN sont représentés en traits continus et les brins amplifiés en lignes discontinues. Les flèches indiquent le sens de synthèse des molécules d'ADN.

L'analyse de l'ARN en médecine

A l'exception notable près des gènes subissant un réarrangement au cours des processus de différenciation (ceux des immunoglobulines et du récepteur des lymphocytes T), l'ADN est le même dans toutes les cellules de l'organisme et peut être analysé au niveau des cellules et tissus les plus aisés à obtenir. Tout autre est le problème des ARN messagers qui ne sont transcrits que dans les cellules exprimant le gène étudié : une étude des ARN du fibrinogène, par exemple, exigera que l'on dispose d'une biopsie hépatique ! Une autre difficulté de ce type d'étude réside en la très grande instabilité des ARN comparée à celle des ADN.

Méthodes de base de l'étude des ARN

(a) Le « Northern blot ». Ce terme, intraduisible en français, est un jeu de mot : *Southern* ayant décrit l'hybridation moléculaire sur filtre de l'ADN, le nom de « Northern blot » fut donné à l'opération consistant à transférer sur des filtres des ARN séparés par électrophorèse, puis à les révéler par hybridation moléculaire comme nous l'avons indiqué dans le lexique de *m/s n° 1, vol. 2, p. 104-105*. Cette méthode permet d'apprécier la taille de l'ARN (dont dépend sa migration électrophorétique, d'autant plus rapide que l'ARN est petit), et sa concentration. Cette dernière peut aussi être appréciée par l'intensité de l'hybridation d'une sonde radioactive avec un échantillon d'ARN déposé sur un filtre sans séparation électrophorétique préalable (méthode du « dot blot »).

(b) La cartographie à la nucléase S1 ou à la RNA-se A (figure 1). Cette technique permet de détecter des différences de séquence entre une sonde radioactive et l'ARN étudié. La nucléase S1 clivera une sonde

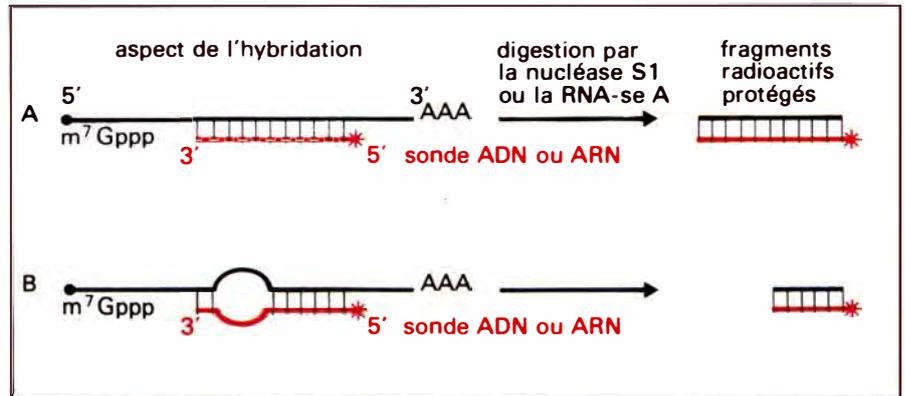


Figure 1. Principe des réactions de « cartographie » à la nucléase S1 ou à la RNAse A. Une sonde radioactive simple brin d'ADN ou d'ARN est hybridée avec le messager étudié. En « A », une hybridation parfaite protégera complètement cette sonde contre la digestion par la nucléase S1 (spécifique de l'ADN simple brin) ou par la RNAse A (spécifique de l'ARN simple brin). En « B », un clivage aura lieu au niveau d'une zone non appariée, entraînant une protection partielle de la sonde. La taille des sondes résiduelles obtenues après digestion est analysée par électrophorèse en milieu dénaturant. L'astérisque indique l'extrémité radioactive de la sonde ADN, les sondes ARN utilisées étant habituellement entièrement radioactives.

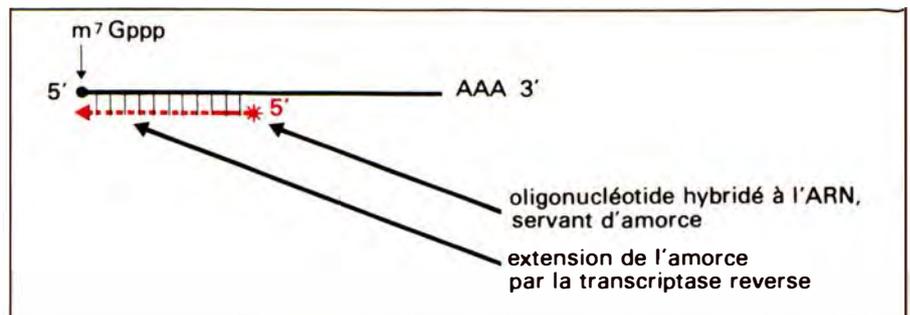


Figure 2. Principe de l'extension d'amorce. Un oligonucléotide radioactif (marqué à son extrémité 5', comme l'indique l'astérisque) est hybridé à l'ARN étudié ; il sert d'amorce pour une extension par la transcriptase inverse qui, en présence des 4 nucléotides dans le milieu, recopie l'ARN en ADN complémentaire. La longueur du fragment allongé est mesurée par électrophorèse en milieu dénaturant.

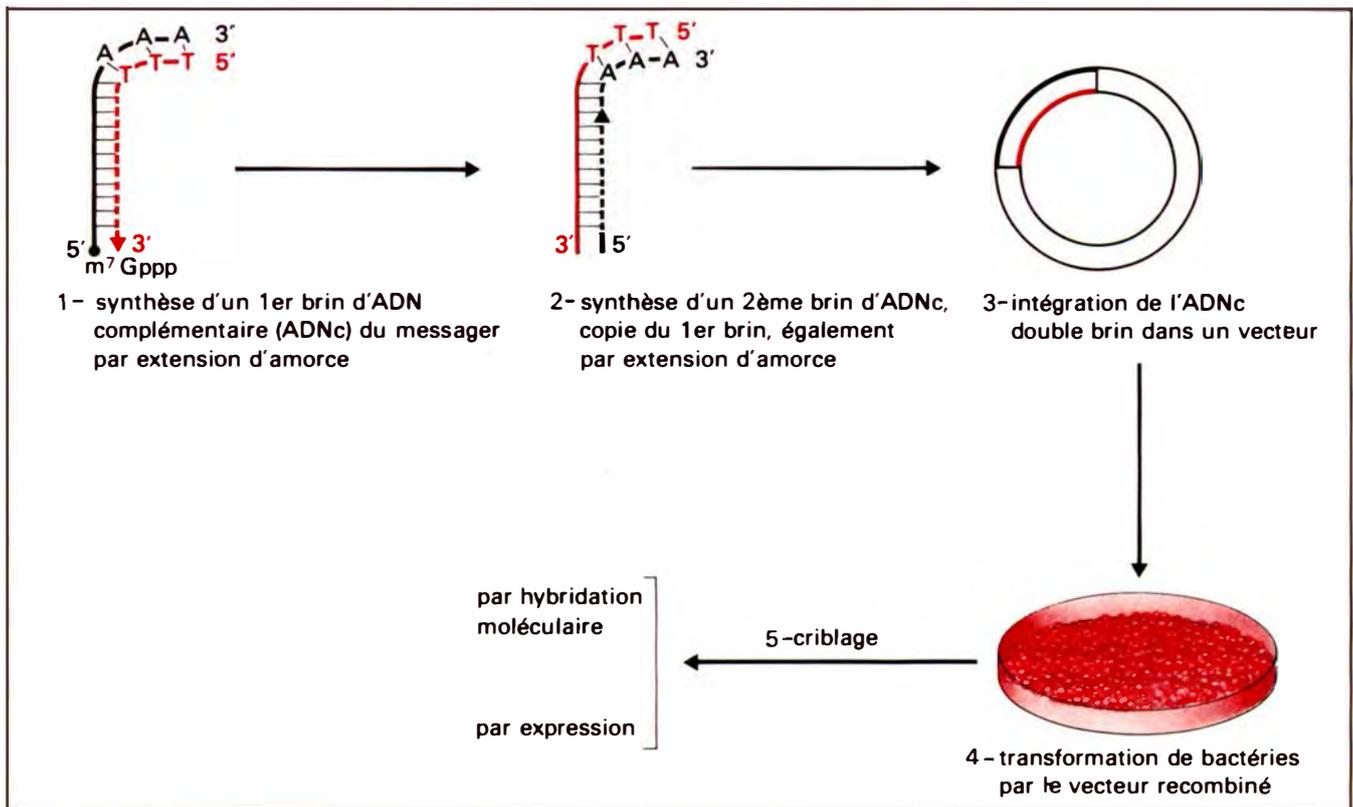


Figure 3. **Principe de la confection de banques d'ADN complémentaires.**

Étape 1 : Les ARN messagers d'un tissu sont recopiés en un brin complémentaire d'ADN par extension d'amorce à l'aide de la transcriptase reverse, comme cela est expliqué dans la figure 2. L'amorce la plus souvent utilisée est ici un homopolymère d'acide thymidilique, complémentaire de l'extension d'acide polyadénylique prolongeant généralement les ARN messagers eucaryotiques.

Étape 2 : Après destruction des ARN, ou simplement dénaturation des hybrides ARN/ADN, les premiers brins sont eux-mêmes recopiés en un deuxième brin par extension d'amorce. Les amorces sont ici diverses selon les méthodes et ne seront pas détaillées.

Étape 3 : L'ADNc double brin est intégré dans un vecteur qui est soit un plasmide (c'est-à-dire un élément génétique à réplication autonome dans la bactérie), soit, comme indiqué dans le lexique « Clonage de gènes » (m/s, n° 6, vol. 2, p. 333), un phage.

Étape 4 : Des bactéries sont « transformées » par intégration du vecteur recombiné, chacune d'entre elle, qui sera à l'origine d'une colonie, ne pouvant intégrer qu'une seule molécule de vecteur... et donc un seul type d'ADNc. Ces bactéries transformées sont diluées et étalées sur des boîtes de milieu de culture de telle sorte qu'elles soient à l'origine de colonies séparées.

Étape 5 : Le criblage, c'est-à-dire la détection parmi les colonies de celles transformées par le vecteur ayant intégré l'ADNc complémentaire du messenger étudié, se fait de différentes façons qui peuvent être ramenées à deux groupes principaux : le criblage par hybridation moléculaire, souvent à l'aide d'oligonucléotides synthétisés d'après une portion connue de la séquence de la protéine codée par le messenger étudié, se fait comme cela est indiqué dans le lexique « Clonage de gènes » (m/s, n° 6, vol. 2, p. 333); le criblage par expression consiste à reconnaître à l'aide d'anticorps, et parfois par son activité biologique, la protéine codée par l'ADNc intégré dans le vecteur transformant les colonies recherchées.

d'ADN si le mauvais appariement est de plus de 3 bases, alors que la RNase A sera capable de cliver une sonde d'ARN ne différant de l'ARN étudié que par une seule base; cette dernière est donc particulièrement précieuse pour étudier des mutations ponctuelles, héréditaires ou acquises (activation d'un oncogène). L'appréciation de l'intensité de la bande protégée est également une méthode extrêmement sensible de détection et de quantification des ARN très minoritaires.

(c) *L'extension d'amorce (figure 2)*. Moins sensible que la précédente pour détecter des messages minoritaires, cette technique a deux intérêts principaux: elle permet de préciser la localisation de l'extrémité 5' des messagers, et donc les sites d'initiation de la transcription des gènes correspondants, qui peut être anormale au cours de phénomènes de remaniement génique; elle est aussi, et surtout, une méthode irremplaçable pour établir la séquence nucléotidique d'un ARN non cloné s'il n'est pas trop minoritaire. Le fragment allongé d'ADN complémentaire peut en effet être séquencé par l'une ou l'autre des méthodes disponibles.

(d) *L'hybridation in situ sur coupes histologiques*. Toutes les méthodes résumées ci-dessus utilisent de l'ARN purifié après broyage de tissus... ce qui peut en fait entraîner le mélange entre des populations de messagers de cellules très différentes composant des tissus comme le cerveau ou le rein ou, plus encore, des embryons entiers. *L'hybridation in situ* consiste à détecter par hybridation moléculaire les ARN à l'intérieur des cellules les contenant, reconnaissables par les techniques histologiques habituelles. Elle s'applique aussi bien à la caractérisation des ARN transcrits normalement par des populations cellulaires parfois très minoritaires dans un organe donné qu'à la recherche de séquences virales.

(e) *La confection de banques d'ADN complémentaires*. Les méthodes permettant de cloner les copies ADN des messagers sont maintenant si nombreuses que la figure 3

ne fera qu'en donner le principe. Il s'agit en dernière analyse de la seule méthode permettant de déterminer la structure d'un ARN messager très minoritaire; elle est dans l'immense majorité des cas la première étape de l'étude moderne d'un gène codant pour une protéine connue.

Principales indications de l'analyse des ARN en médecine

(a) *Caractérisation d'une cellule normale ou pathologique par l'expression de gènes spécifiques*. Chaque cellule a un « phénotype » qui est fonction des gènes qu'elle transcrit en messagers. Certains d'entre eux peuvent constituer de forts marqueurs de types particuliers de différenciation normale ou pathologique. De telles sondes reconnaissant des messagers spécifiques de certaines cellules sont déjà utilisées pour l'analyse de l'oestrogène-dépendance des cancers du sein. Appliquée à des coupes de tissu, une telle caractérisation par « hybridation in situ » des populations d'ARN messagers de certains types de cellule est probablement appelée à un très grand avenir en anatomie pathologique.

(b) *Recherche de messagers anormaux*. L'étude quantitative et qualitative des messagers d'oncogènes a déjà quitté le domaine exclusif de la recherche et pourrait constituer un outil utile pour le diagnostic et le suivi de certains cancers. De même, la détection de transcrits viraux est-elle particulièrement importante pour l'étude pathogénique, voire pour la détection de certains types d'infections.

(c) *Analyse des mécanismes d'une maladie héréditaire*. L'étude des ARN au cours d'une maladie héréditaire ne comportant pas d'anomalie structurales évidentes du gène impliqué permet de déterminer si ce gène est transcrit, si les ARN subissent une maturation normale et si le messager est traductible en protéine. Le clonage moléculaire de l'ADN complémentaire du messager muté permettra, par séquençage nucléotidique, de parvenir à la reconnaissance de la mutation ponctuelle éventuellement en cause. A. K.

JL John Libbey
EUROTEXT
LONDON PARIS

Paris · Londres
EDITIONS MEDECINE-SCIENCES

Hormones et régulation cellulaire (n° 12)

J. NUNEZ, J.E. DUMONT,
E. CARAFOLI

**Co-édition, John Libbey
Eurotext/INSERM**
1988, vol. 165, 256 pages
330 FF, 28 £, 55 \$ US

Cellules hépatiques et médicaments

A. GUILLOUZO
**Co-édition, John Libbey
Eurotext/INSERM**
1988, vol. 164, 514 pages
477 FF, 40 £, 78 \$ US

Les inhibiteurs de l'hématopoïèse

A. NAJMAN, M. GUIGON,
N.C. GORIN, J.Y. MARY
**Co-édition, John Libbey
Eurotext/INSERM**
1987, vol. 162, 376 pages
330 FF, 28 £, 55 \$ US

Obstétrique

La maîtrise de la parturition

C. SUREAU, Ph. BLOT, D. CABROL,
F. CAVAILLÉ, G. GERMAIN
**Co-édition, John Libbey
Eurotext/INSERM**
1987, vol. 151, 284 pages
290 FF, 24,50 £, 48 \$ US

Pédiatrie

Physiologie cardio-vasculaire et respiratoire du fœtus et du nouveau-né

P. KARLBERG, A. MINKOWSKI,
W. OH, L. STERN,
M. MONSET-COUCHARD
**Co-édition, John Libbey
Eurotext/INSERM**
1986, vol. 133, 206 pages
210 FF, 17,50 £, 34 \$ US

6, rue Blanche, 92120 Montrouge
Tél. : (1) 47.35.85.52

L'amplification in vitro des fragments d'ADN par PCR (polymerase chain reaction) : un tournant en génétique

La biologie est une science typiquement expérimentale. De ce fait, ses progrès se sont faits au rythme de la mise au point de nouveaux outils dont l'ensemble le plus élaboré et le plus cohérent est aujourd'hui la technologie des recombinants d'ADN, ou génie génétique. Ce n'est pas, par exemple, l'absence d'imagination ou l'étroitesse intellectuelle qui empêchaient les biologistes moléculaires des années 1960 d'aller plus vite et plus loin dans l'étude des gènes mais l'impossibilité qu'il y avait alors à isoler en quantité suffisante un gène cellulaire et à déterminer sa séquence nucléotidique. Les enzymes de restriction, l'hybridation moléculaire sur filtre (le *Southern blot*), le clonage d'ADN et son séquençage ont permis le développement de la génétique moléculaire moderne.

Malgré cet arsenal technique d'une exceptionnelle puissance (à l'origine de la quasi-totalité des progrès récents en génétique, cancérologie, immunologie, virologie, etc.), des difficultés d'ordre quantitatif et économique persistaient quant à l'utilisation large du génie génétique en médecine. Analyser un gène anormal responsable d'une maladie héréditaire ou peut-être d'un cancer restait un gigantesque travail si l'anomalie en cause était très petite (mutation ponctuelle, délétion minimale créant un décalage de phase de lecture). (Voir *m/s suppl. au n° 7, vol. 3, p. 6 et 8*). Parfois, le matériel biologique disponible (par exemple le prélèvement fait sur un malade), était trop peu abondant pour autoriser un « clonage moléculaire ». D'autres fois, enfin, la représentativité d'un fragment d'ADN étudié parmi tout l'ADN de l'échantillon biologique était trop petite pour permettre aisément son « clonage ». Prenons l'exemple d'un gène humain porté par un fragment d'ADN de 15 000 bases de long. Puisqu'il y a trois milliards de bases dans un exem-

plaire d'ADN humain, il faudra, après avoir fragmenté cet ADN et l'avoir intégré dans des microorganismes, analyser au moins 200 000 clones pour détecter — en moyenne — un clone contenant le gène recherché. Compte tenu des lois de répartition aléatoire, c'est en fait un million de clones qui sont couramment analysés (*m/s suppl. au n° 7, vol. 3, p. 10*). Si maintenant le gène étudié a été modifié dans quelques cellules seulement d'un organisme... cellules à l'origine d'une tumeur cancéreuse, son « clonage » peut devenir impossible et sera de toutes façons si dif-

ficile qu'il ne pourrait jamais être appliqué à un grand nombre de malades.

• **Le principe de la méthode d'amplification par PCR.** Décrite en décembre 1985 par R.K. Saiki et ses collaborateurs de la firme Cetus Corporation (Emeryville, Californie, USA) [1], la méthode a été décrite il y a plus de deux ans dans *m/s* (n° 7, vol. 2, p. 398); nous reproduisons ici le schéma explicatif présenté à l'époque (*figure 1*). L'expérience consiste à rajouter à la préparation d'ADN contenant la séquence à amplifier deux amorces oligonucléotidiques

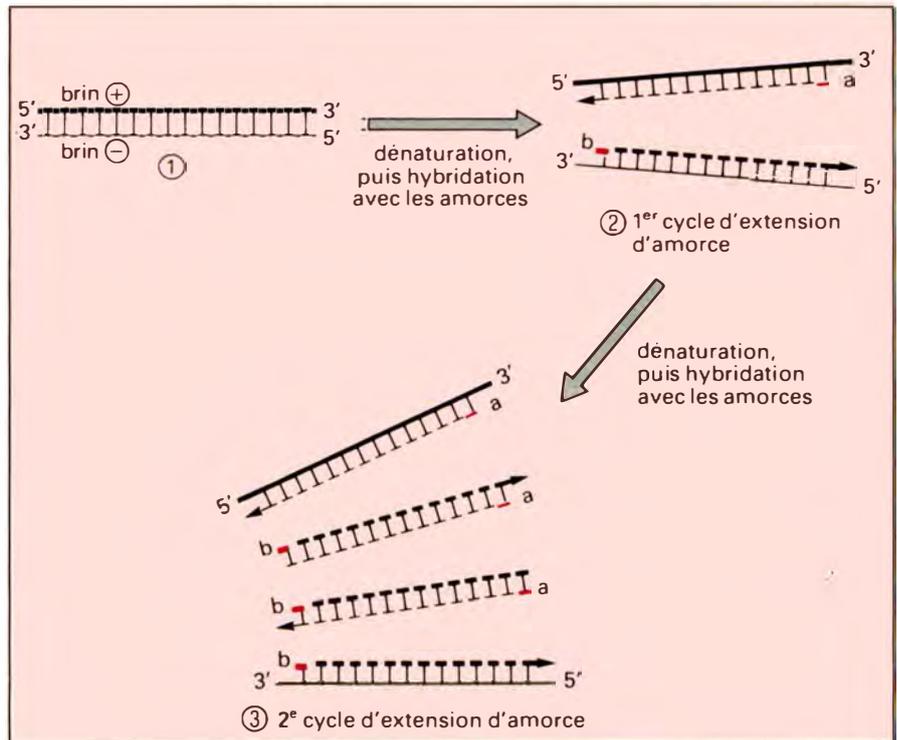


Figure 1. Amplification d'une région particulière de l'ADN génomique. L'ADN en 1 est dénaturé, chacun des brins étant hybridé avec un oligonucléotide de synthèse (traits rouges). En 2, les oligonucléotides « a » (complémentaires du brin +, épais) et « b » (complémentaires du brin -, fin) servent d'amorce pour la synthèse de chaînes complémentaires d'ADN sous l'action de l'ADN polymérase. Cette réaction est appelée « extension d'amorce ». En 3, l'opération est recommencée, aboutissant à une amplification des régions comprises entre les séquences reconnues par les 2 oligonucléotides de synthèse. Les brins initiaux d'ADN sont représentés en traits continus et les brins amplifiés en lignes discontinues. Les flèches indiquent le sens de synthèse des molécules d'ADN.

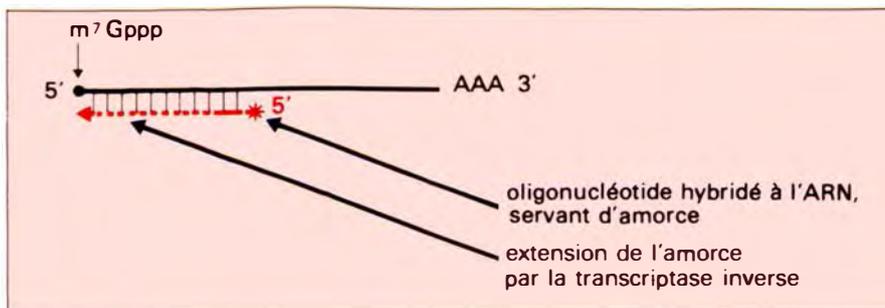


Figure 2. **Principe de l'extension d'amorce.** Un oligonucléotide radioactif (marqué à son extrémité 5', comme l'indique l'astérisque) est hybridé à l'ARN étudié ; il sert d'amorce pour une extension par la transcriptase inverse qui, en présence des 4 nucléotides dans le milieu, recopie l'ARN en ADN complémentaire. La longueur du fragment allongé est mesurée par électrophorèse en milieu dénaturant.

d'une vingtaine de bases (oligonucléotides de synthèse) complémentaires respectivement de l'un et de l'autre des deux brins d'ADN au niveau de séquences délimitant la longueur du fragment qui sera amplifié. On procède alors à des cycles successifs d'allongement des amorces par de l'ADN polymérase, puis de séparation des brins (par la chaleur). La figure 1 montre qu'en partant de deux brins, on a quatre fragments après un cycle, huit après deux cycles, seize après trois cycles, 2^{n+1} après n cycles ; c'est-à-dire, après 30 cycles, plus de deux milliards de copies... théoriquement. En fait, le rendement de la réaction n'étant pas de 100 %, la multiplication est moindre. La formule réelle figurant le nombre Y de copies obtenues à partir de A copies initiales pour n cycles, le rendement de l'amplification étant R , s'écrit $Y = A(1 + R)^n$ [2]. Cet effet du rendement influe énormément sur le niveau de l'amplification finale. Ainsi, pour une valeur $R = 0,5$ (50 % de rendement), les deux brins initiaux seront devenus environ $3,8 \times 10^9$ après 30 cycles, au lieu de 2×10^9 (deux milliards) pour un rendement de 100 %. Ce rendement, ainsi que la rapidité et le coût de l'amplification, ont été considérablement augmentés par l'utilisation d'une ADN polymérase d'une bactérie thermophile (*thermus aquaticus*) vivant dans les eaux chaudes, jusqu'à 95° C [3]. Cette enzyme n'est pas détruite par le chauffage nécessaire à chaque cycle pour dissocier les doubles brins d'ADN ; fonctionnant à très haute température (70° C), elle recopie un

ADN dont les superstructures ont toutes les chances d'être effacées, ce qui diminue les risques d'arrêt prématuré de la polymérisation.

- **Les performances de la PCR.** Il est possible grâce à cette méthode de multiplier *in vitro* un nombre considérable de fois n'importe quel fragment de séquence d'ADN connu, sans recourir au clonage moléculaire. La taille de la séquence amplifiée peut être assez grande (plusieurs kilobases), mais le rendement diminue linéairement avec l'augmentation de taille, si bien que si une sensibilité maximum est recherchée, il est préférable de choisir des amorces correspondant à des séquences distantes de seulement cent à quelques centaines de bases (J. Chelly *et al.* en préparation).

Dans ces dernières conditions, des amplifications ont été entreprises avec succès à partir de l'ADN d'une seule cellule, ou du bulbe d'un seul cheveu. Une tache de sang séché d'un test de Guthrie [4] peut également suffire. Le matériel amplifié est généralement assez abondant pour que puisse être déterminée directement, sans clonage préalable, sa séquence nucléotidique [5]. Cette non nécessité du clonage permet d'éviter les erreurs de séquences secondaires aux incorporations de bases mal appariées par la polymérase de *thermus aquaticus* : environ une tous les 9 000 résidus [6], ce qui peut faire, au terme d'une amplification de 30 cycles et compte tenu que ces erreurs sont cumulatives, une aberration toutes les 400 à 500 bases [7]. Les erreurs sont cependant réparties au hasard, chacune d'entre

elles n'intéressant qu'une partie infime de la population amplifiée. Une séquence nucléotidique déterminée sur la population dans son ensemble donnera, à chaque position, la base présente au niveau d'environ 499 fragments sur 500, c'est-à-dire la base réellement présente dans l'ADN d'origine. La polymérase utilisée pourrait également créer un décalage de phase de lecture environ toutes les 41 000 bases [5].

Il faut noter enfin que la totalité de la procédure d'amplification est très aisément automatisable, ce qui permet de conduire parallèlement un grand nombre d'expériences indépendantes. Chaque cycle dure environ 3,5 minutes, l'expérience dans son ensemble étant achevée en moins de deux heures.

- **Amplification d'ADN génomique en médecine.** La PCR permet aisément, à partir de l'ADN de cellules d'accès facile (globules blancs, villosités chorales, fibroblastes, cellules de la muqueuse buccale ou de l'urine) de vérifier une anomalie de l'ADN d'une région connue quelconque d'un gène : mutation à l'origine d'une maladie génétique répandue comme la drépanocytose ou une autre hémoglobinopathie [1, 8], polymorphisme d'ADN associé à une maladie (région du gène HLA-DQ α associée à la susceptibilité au diabète) [9, 10], ou oncogène activé par mutation ponctuelle [11, 12].

Une situation particulièrement favorable est celle de translocations chromosomiques bien définies, par exemple associées à des hémopathies malignes. Il existe ainsi, chez 90 % des malades atteints de lymphome folliculaire, une translocation t(14; 18) réarrangeant l'oncogène bcl-2 dans les cellules leucémiques. Il est possible à l'aide d'amorces reconnaissant les régions des chromosomes 14 et 18 immédiatement adjacentes à la cassure, d'amplifier très spécifiquement l'ADN du chromosome hybride 14-18 à travers le point de cassure. Aucun fragment ne peut être amplifié à partir de l'ADN des cellules non leucémiques ne comportant pas la translocation. Il serait ainsi possible de repérer une cellule leucémique résiduelle avec translocation t(14; 18) parmi 1 000 cellules saines [13].

• **Détection de séquences virales.** De par sa sensibilité, la PCR est d'ores et déjà devenue la méthode la plus fine pour détecter des infections virales. Les virus sont des agents infectieux dont le génome (ARN ou ADN) peut en effet être aisément amplifié, souvent à partir du sang. Cette approche est déjà utilisée par plusieurs laboratoires pour l'hépatite virale et le SIDA. Dans ce dernier cas, la positivité du test PCR précède de plusieurs semaines l'apparition des anticorps, c'est-à-dire la conversion à la « séropositivité » [14]. S'il était possible d'appliquer cette méthode aux donneurs de sang, on supprimerait probablement le très petit pourcentage de cas où la contamination est due à la transfusion du sang d'un donneur très récemment infecté et encore séronégatif.

• **Application à l'ARN.** Après recopiage de l'ARN en ADN complémentaire par la transcriptase inverse, la PCR peut parfaitement être appliquée. La première étape de l'expérience est donc une extension d'amorce, telle que nous la décrivions dans *m/s* (n° 9, vol. 2, p. 518) (figure 2). La seconde amorce est alors ajoutée, et la réaction d'amplification entreprise. Cette procédure peut être utilisée pour étudier, au niveau de l'ARN, le même type d'anomalies qu'au niveau de l'ADN. Par exemple, le messenger chimérique issu du remaniement chromosomique t (9; 22) (donnant le chromosome philadelphie dans la forme typique) caractéristique de la leucémie myéloïde chronique peut être détecté avec sûreté et sensibilité [15, 16]. Mieux, même: du fait de sa sensibilité, la méthode permet de détecter des ARN dont il n'existe qu'une seule copie dans une cellule sur 1 000, ce qui semble être le cas des transcrits correspondant à des gènes spécifiques d'un type particulier de différenciation tissulaire hors des cellules où ils s'expriment normalement. C'est ainsi qu'ont pu être détectés dans des fibroblastes ou des lymphoblastes des transcrits de dystrophine, β -globine, hormone antimüllérienne, facteur VIII, etc. [2, 17]. Ce phénomène, appelé « transcription illégitime » par J. Chelly *et al.* pourrait permettre d'analyser dans n'importe quelle cellule d'accès facile n'importe quel transcrit d'un

gène présomptivement pathologique [17].

• **La PCR, technique de base des laboratoires de génétique moléculaire.** Pour être complet, il faudrait ajouter aux usages possibles de la PCR toutes les manipulations de l'ADN effectuées dans les laboratoires de génétique moléculaire: constructions génétiques, séquençage d'ADN, mutagenèse dirigée. Je voudrais commenter une application particulière, très riche de promesses. Nous avons déjà vu (*m/s* n° 3, vol. 4, p. 190) que l'on pouvait maintenant, une fois sur 1 000 environ, insérer à sa place normale un gène transféré dans une cellule. Tout fragment d'ADN contenant des séquences homologues de celles d'un gène endogène peut, par recombinaison homologue, être semblablement intégré de manière « ciblée » et ainsi inactiver un gène au milieu duquel il s'insérera.

La difficulté de l'approche réside en la mise au point d'un crible destiné à reconnaître les cellules ayant subi une recombinaison homologue, en moyenne 999 fois moins nombreuses que celles ayant subi un événement de recombinaison au hasard. S'il est possible d'amplifier un fragment d'ADN à partir d'une amorce reconnaissant l'ADN exogène et d'une amorce reconnaissant l'ADN endogène cible, on pourra détecter une petite minorité de cellules contenant un tel ADN cible remanié parmi des cellules ayant intégré au hasard l'ADN exogène. Si on a ainsi à tester 1 000 clones indépendants résultant d'une expérience de transformation par de l'ADN exogène « ciblé » vers une séquence endogène, on réalisera une réaction de PCR sur 10 groupes de 100 clones. Puis, le groupe positif sera lui-même subdivisé et l'on pourra parvenir au clone recherché...

• **PCR, médecine légale et paléontologie.** Comme *m/s* l'a rapporté à de maintes reprises (*m/s* n° 6, vol. 1, p. 333; n° 3, vol. 2, p. 155), il est possible désormais de caractériser tout individu grâce à ses « empreintes génétiques », c'est-à-dire à l'organisation extrêmement variable d'un individu à l'autre de certaines régions de l'ADN. Une telle recherche peut — en théorie — être faite à partir de sperme ou de sang séchés... de fragments de peau ou de

poils. On risque cependant de buter souvent sur l'insuffisance de matériel, et la possibilité de multiplier la quantité infinitésimale d'ADN présente dans un échantillon, avant de l'étudier, étend considérablement le champ d'application de la méthode. L'ADN, contrairement à l'ARN, est chimiquement très stable. Il est de ce fait maintenant sûrement possible de caractériser grâce à des résidus de sang séché sur un vêtement, très à distance de l'événement en cause (accident, crime, catastrophe, guerre) l'empreinte génétique de l'individu d'où provient ce sang.

Puisque nous introduisons ici la notion d'étude à distance, il faut signaler que la « PCR » commence aussi à être largement utilisée par les paléontologistes. Dans des résidus d'os d'animaux fossiles, de peau de momies, il persiste des traces d'ADN très dégradé, fragmenté... mais qu'il est possible d'amplifier. La détermination de la séquence de ces fragments permet d'améliorer considérablement le positionnement d'un animal fossile dans l'arbre évolutif des espèces, ou d'apprécier la vitesse d'accumulation des changements génétiques survenus depuis les premières dynasties de l'ancienne Égypte.

En conclusion après les méthodes de base du génie génétique développées entre 1972 et 1976, un nouveau pas technique essentiel a été accompli en décembre 1985: la mise au point de l'amplification fidèle, rapide et massive de n'importe quel fragment connu d'acide nucléique. Cette méthode rend d'ores et déjà aisément accessible à l'analyse par les méthodes de l'hybridation moléculaire et du séquençage nucléotidique des fragments d'ADN humain dont on soupçonne qu'ils pourraient être responsables d'une maladie génétique héréditaire ou acquise (cancer); elle est aussi la méthode la plus sensible pour détecter précocement et fidèlement tout agresseur, viral (pathologie tumorale ou infectieuse) ou humain (médecine légale).

Nul doute que la puissance de ce nouvel outil fera rapidement progresser les connaissances dans tous les domaines nécessitant l'analyse ou la manipulation d'un acide nucléique. ►

RÉFÉRENCES

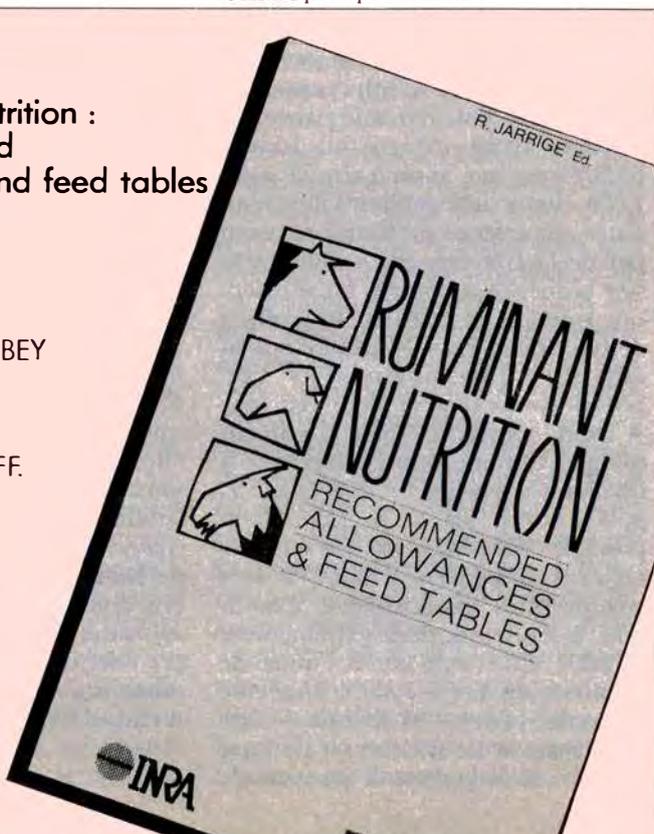
1. Saïki RK, Scharf S, Faloona F, *et al.* Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 1985 ; 230 : 1350-4.
2. Chelly K, Kaplan JC, Maire P, Gautron S, Kahn A. Transcription of the dystrophin gene in human muscle and non-muscle tissues. *Nature* 1988 : 333 : 858-60.
3. Saïki RK, Gelfand DH, Stoffel S, *et al.* Primer-directed amplification of DNA with a thermostable DNA polymerase. *Science* 1988 ; 239 : 487-91.
4. Lyonnet S, Rey F, Caillaud C, Abadie V, Munnich A, Rey J. Bases moléculaires de la phénylcétonurie en France. *médecine/sciences* 1988 (sous presse).
5. Stoffel ES, Kœberl DD, Sarkar G, Sommer SS. Genomic amplification with transcript sequencing. *Science* 1988 ; 239 : 491-4.
6. Tindall KR, Kunkel TA. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* 1988 ; 27 : 6008-13.
7. Higuchi R, Krummel B, Saïki RK. A general method of *in vitro* preparation and specific mutagenesis of DNA fragments : study of protein and DNA interactions. *Nucleic Acids Res.* 1988 ; 16 : 7351-67.
8. Amselem S, Vidaud M, Vidaud D, *et al.* Determination of the spectrum of beta-thalassemia genes in Spain by use of dot-blot analysis of amplified beta-globin DNA. *Am J Hum Genet* (sous presse).
9. Saïki RK, Bugawan TL, Horn GT, Mullis KB, Erlich HA. Analysis of enzymatically amplified β -globin and HLA-DQ α DNA with allele-specific oligonucleotide probes. *Nature* 1986 ; 324 : 163-6.
10. Gu XF, Elion J, Ouagued M, Clauser E, Assan R, Krishnamoorthy R. A simple strategy to amplify specifically the HLA-DQ β gene region with genomic DNA as template. *Febs Lett.* 1988 ; 236 : 23-6.
11. Almoguera C, Shibata D, Forrester K, Martin J, Arnheim N, Perucho M. Most human carcinomas of the exocrine pancreas contain mutant *c-K-ras* genes. *Cell* 1988 ; 53 : 549-54.
12. Farr CJ, Saïki RK, Erlich HA, McCormick F, Marshall CJ. Analysis of RAS gene mutations in acute myeloid leukemia by polymerase chain reaction and oligonucleotide probes. *Proc Natl Acad Sci USA* 1988 ; 85 : 1629-33.
13. Lee MS, Chang KS, Cabanillas F, Freireich EJ, Trujillo JM, Stass SA. Detection of minimal residual cells carrying the t(14; 18) by DNA sequence amplification. *Science* 1987 ; 237 : 175-8.
14. Loche M, Mach B. Identification of HIV-infected seronegative individual by a direct diagnostic test based on hybridisation to amplified viral DNA. *Lancet* 1988 ; 11 : 418-21.
15. Kawasaki ES, Clark SS, Coyne MY, *et al.* Diagnosis of chronic myeloid and acute lymphocytic leukemia by detection of leukemia-specific mRNA sequences amplified *in vitro*. *Proc Natl Acad Sci USA* 1988 ; 85 : 5698-702.
16. Delfau MH, Garbarz M, Chaverroche I, Grandchamp B. Detection du messenger chimérique abl-bcr dans la leucémie myéloïde chronique. *médecine/sciences* 1988 ; 4 : 402-3.
17. Chelly J, Concordet JP, Kaplan JC, Kahn A. Analyzing any transcript in any cell-type. Soumis pour publication.

Ruminant Nutrition : recommended allowances and feed tables

Co-édition
INRA/JOHN LIBBEY
EUROTEXT

R. JARRIGE
1989, relié, 300 FF.

6, rue Blanche
92120 Montrouge
Tél. : 47.35.85.52



JL John Libbey
EUROTEXT
LONDON · PARIS

Paris · Londres
EDITIONS MEDECINE-SCIENCES

Porphyries et porphyrines

Y. NORDMANN
Co-édition John Libbey
Eurotext/INSERM
1986, vol. 134, 232 pages
250 FF, 21 £, 41 \$ US

Nouvelles perspectives de la recherche sur le vieillissement

Y. COURTOIS, B. FAUCHEUX,
B. FORETTE, D.L. KNOCK,
J.A. TRÉTON
Co-édition John Libbey
Eurotext/INSERM
1986, vol. 147, 624 pages
440 FF, 37 £, 72 \$ US

Protéines de liaison des hormones stéroïdes

M.G. FOREST, M. PUGÉAT
Co-édition John Libbey
Eurotext/INSERM
1986, vol. 149, 672 pages
440 FF, 37 £, 72 \$ US

Biochimie et physiopathologie de la membrane plaquettaire

G. MARGUERIE, R.F.A. ZWAAL
Co-édition John Libbey
Eurotext/INSERM
1988, vol. 158, 200 pages
320 FF, 27 £, 53 \$ US

Les événements respiratoires du sommeil

P. LÉVI-VALENSI, B. DURON
Co-édition John Libbey
Eurotext/INSERM
1988, vol. 167, 260 pages
340 FF, 29 £, 56 \$ US

6, rue Blanche, 92120 Montrouge
Tél. : (1) 47.35.85.52

Mécanismes moléculaires de la vision des couleurs et du daltonisme

Une équipe de Stanford [1,2] vient d'élucider de façon magistrale la génétique moléculaire de la vision des couleurs ainsi que celle de ses anomalies conduisant aux diverses variétés de daltonisme. Ces résultats ont fait également l'objet d'éditoriaux dans *Science* [3] et dans *Nature* [4].

Les pigments visuels sont des molécules qui absorbent la lumière. Ils sont formés d'une apoprotéine, qui porte le nom d'opsine, et d'un cofacteur dérivé de la vitamine A, le II-cis-rétinal. Ces pigments sont localisés dans la membrane plasmique des cellules photoréceptrices. On connaît quatre pigments visuels, spécifiés par la longueur d'onde de leur maximum d'absorption. Trois d'entre eux servent à la vision colorée, absorbant à 420 nanomètres (sensible au bleu), 530 nm (sensible au vert) et 560 nm (sensible au rouge). Tous trois sont localisés dans les cônes de la rétine, un cône donné ne possédant qu'un seul pigment. Le quatrième, qui absorbe à 495 nm, est la rhodopsine, qui permet la vision crépusculaire; il siège dans les bâtonnets, et est de loin le plus abondant.

La caractérisation des gènes de la vision des couleurs repose sur l'exploitation d'une idée qui n'est plus nouvelle, mais qui s'est avérée d'une extrême fécondité: on a d'abord pensé que la reconnaissance d'un acide nucléique par un autre (au moyen de l'hybridation) exigeait leur identité presque complète, et l'on insistait sur la notion de spécificité. On s'est aperçu ensuite que la possibilité d'hybridation entre deux molécules dépend des conditions expérimentales de l'hybridation et des lavages qui la suivent: à conditions très rigoureuses (on dit strictes), par exemple à température élevée, hybridation très sélective; à conditions moins rigoureuses, hybridation moins sélective, par exemple entre deux molécules qui

ne seraient homologues qu'à 50 %. Partant de cette idée, on fait l'hypothèse qu'une famille de protéines dont le rôle est similaire dérive d'un ancêtre commun, que leurs gènes ont gardé une structure voisine, et qu'à partir d'un des gènes on doit parvenir à isoler tous les autres.

Comment appliquer cette hypothèse aux pigments visuels? Une protéine était mieux connue que les autres, la rhodopsine. Nathans *et coll.* [1] ont préparé des clones d'ADN complémentaire et génomique de rhodopsine de bœuf, puis les ont utilisés pour cribler une banque d'ADN germinal humain réalisée incidemment à partir d'un des auteurs. Dans un premier temps ils ont obtenu uniquement des clones du gène de la rhodopsine humaine. Travaillant ensuite à des températures plus basses, ils isolèrent des clones provenant d'autres gènes; ils en ont établi la séquence et déduit celle des acides aminés correspondants. Ces clones étaient de deux types: le premier avait la même structure à cinq exons et le même nombre d'acides aminés que la rhodopsine (348) avec 42 % d'identité. Par des expériences d'hybridations cellulaires somatiques homme-souris, ce gène a été localisé sur le chromosome 7 (7q 22-7qter); il n'est donc pas lié à l'X et a été identifié comme le gène du pigment «bleu». Par la même méthode, le gène de la rhodopsine a été localisé sur le chromosome 3 (3q 21-3qter).

Le deuxième type de clones et les sondes qui en dérivent peuvent être attribués par hybridation somatique à la partie distale du bras long du chromosome X (Xq 22-q 28), au voisinage du locus de la glucose 6 phosphate deshydrogénase, en accord avec les données génétiques classiques sur la transmission du daltonisme. La structure générale reste la même que celle de la rhodopsine; notamment tous ces pigments possèdent une lysine qui

fixe le rétinol par une liaison covalente. Il existe cependant une différence: la séquence des nouveaux gènes montre six exons; les exons 2 à 6 correspondent aux exons 1 à 5 de la rhodopsine et du pigment bleu, mais il y a en 5' un exon I supplémentaire. En outre, la carte de restriction permet de distinguer deux gènes différents malgré une homologie à 96 %, codant respectivement pour les pigments «vert» et «rouge». On peut extraire de la rétine des ARN messagers qui reconnaissent chacune de ces sondes.

Les auteurs ont alors fait une découverte inattendue; s'il existe toujours un, et un seul, gène rouge, le nombre de gènes verts, même chez les sujets à vision normale, varie de un à trois. Comment expliquer cette variabilité d'un seul des deux gènes? L'hypothèse est que ces deux gènes, qui ne sont distants que de cinq kilobases, sont disposés en tandem tête-queue (*figure 1*), résultat sans doute d'une duplication génique au cours de l'évolution. La très grande ressemblance des deux gènes permet une recombinaison non équationnelle (*voir médecine/sciences n° 4, vol. 1, p. 214*), donnant naissance à un chromosome à deux gènes et à un autre qui n'en porte aucun. Pour interpréter le fait que le gène rouge n'est jamais impliqué dans l'échange inégal, il faut admettre que ce gène est situé à la lisière de la zone de recombinaison.

Anomalies de la vision des couleurs. En pratique, elles portent sur la vision du vert et du rouge, en liaison avec le chromosome X; celles de la vision du bleu, transmises comme un caractère autosomique récessif, restent exceptionnelles. Seuls les sujets du sexe masculin sont donc atteints. On peut diviser les anomalies en deux grandes catégories. Dans la première figurent les individus «aveugles» pour le rouge

ou pour le vert, qui sont dichromates; on peut la subdiviser en deux sous-groupes, selon que fait défaut la vision du rouge (protanopie) ou celle du vert (deutéranopie). La deuxième catégorie comprend les sujets qui voient les trois couleurs fondamentales (trichromates) mais chez lesquels l'une d'elles est mal perçue ou déformée. Des méthodes de comparaison de couleurs permettent de bien analyser les diverses catégories.

Nathans *et coll.* [2] ont étudié par la méthode de *Southern* l'ADN de 25 sujets porteurs d'anomalies variées de la vision des couleurs. Les sondes utilisées sont spécifiques des gènes «verts» et «rouges» situés sur le chromosome X. Elles fournissent quatre fragments de restriction appelés A, B, C et D, dont les trois premiers analysent la partie initiale et le quatrième la partie terminale du gène. Comme la longueur de chacun de ces fragments diffère selon qu'il appartient au gène vert ou rouge, on peut les assigner facilement à l'un de ces gènes. Les sujets se divisent en 15 dichromates, chez lesquels manque la vision du rouge R (red) (9 cas) ou celle du vert G (green) (6 cas), et 10 trichromates. Deux types de mécanismes moléculaires principaux ont été observés. Le plus rare (6 cas) consiste en la délétion d'un gène; comme nous l'avons montré plus haut, il s'agit toujours du gène G. Le plus fréquent est la présence d'un gène hybride comportant à la fois des éléments G et R. La partie fonctionnellement importante est le segment D correspondant à l'exon 5. Quand la partie proximale (dite 5') du gène est «verte» et la partie distale (dite 3') est «rouge», c'est la vision du vert qui est supprimée ou altérée et vice versa. Il est probable que c'est la position du point de raccord des deux gènes qui détermine la disparition complète ou l'altération d'une couleur. Un modèle plus détaillé est représenté sur la figure. Enfin, un seul sujet a montré une structure apparemment normale des gènes et présente probablement une mutation ponctuelle.

Le nombre de sujets victimes d'anomalies de la vision des couleurs est

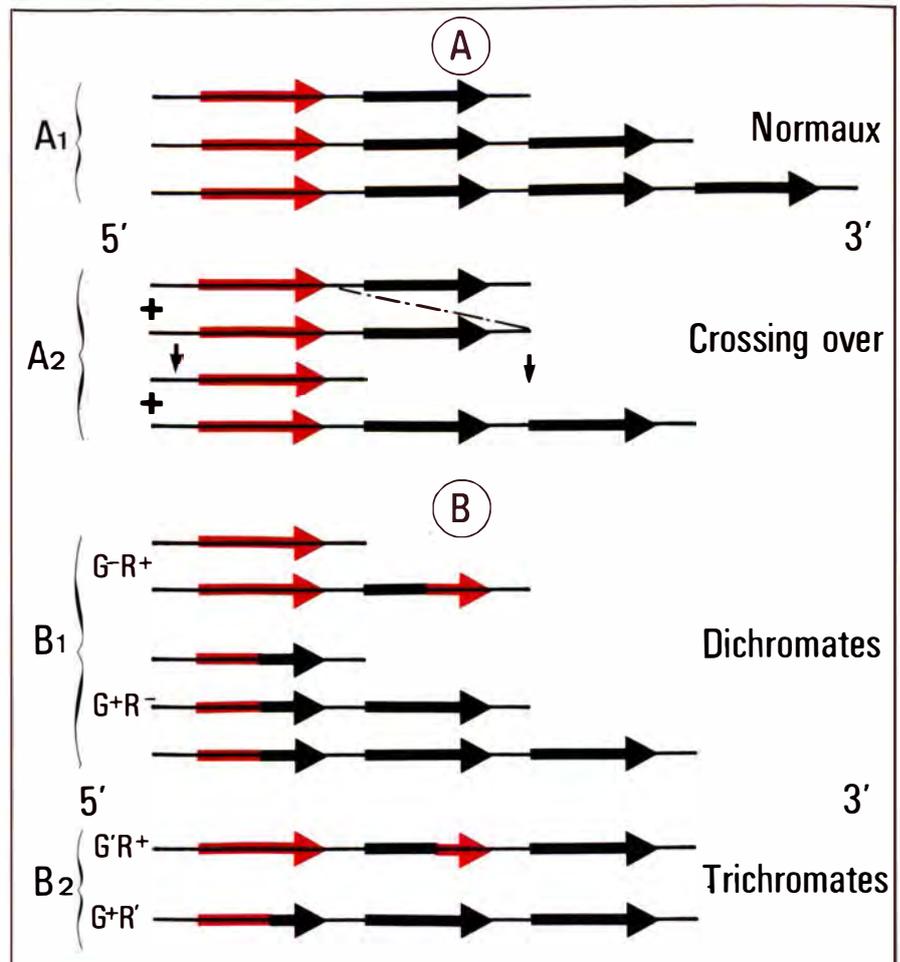


Figure 1. A. **Représentation des gènes «vert» et «rouge».** G = vert (green); R = rouge (red).

A₁, Situation normale: gènes R et G en tandem «tête-queue»; un seul gène R; un à trois gènes G.

A₂, Crossing over entre deux allèles au niveau du gène G. Il en résulte un chromosome sans gène G et un chromosome à deux gènes G.

B. Mécanismes moléculaires du daltonisme.

B₁, Dichromates

G⁻ R⁺

soit délétion de G résultant du crossing over décrit en A₂, soit gène hybride 5'G- 3' R.

G⁺ R⁻

gène hybride 5' R-3' G suivi ou non de un ou plusieurs gènes G normaux.

B₂, Trichromates. Le gène anormal est marqué par le signe ' (G' ou R')

G' R⁺

R normal suivi d'un hybride 5' R-3' G suivi lui-même ou non d'un ou deux gènes G normaux.

G⁺ R'

Hybride 5' R-3' G suivi de gènes G normaux.

On a figuré (ce n'est qu'une hypothèse) la partie 3' de l'hybride plus courte pour les tri- que pour les dichromates.

Chez les dichromates une des couleurs n'est pas perçue. Chez les trichromates elle est perçue mais la longueur d'onde d'absorption maximum est modifiée.

très élevé, puisqu'on l'estime à 8 % de la population masculine blanche. Parmi eux environ 30 % sont dichromates, avec un taux similaire d'absence de vision du vert ou du rouge. Sur les 70 % qui restent, on compte un peu plus de 55 % d'anomalies du vert pour moins de 15 % d'anomalies du rouge. On arrive donc à la conclusion surprenante que, d'après les résultats de Nathans *et coll.*, 6 % de la population seraient porteurs d'un gène hybride. Reste à comprendre le mécanisme de formation de cet hybride. Les expériences de Nathans *et coll.* semblent mettre en jeu soit des recombinaisons non équationnelles, soit des conversions géniques (voir *médecine/sciences* n° 4, vol. 1, p. 214). La possibilité d'une mutation ponctuelle, jadis considérée comme la plus probable, n'a été reconnue qu'une fois sur les 25 sujets examinés.

Les travaux de Nathans *et coll.* ouvrent enfin des perspectives intéressantes du point de vue de l'évolution. L'existence de trois variétés de cônes, porteurs chacun d'un seul type de pigment, est relativement récente. On l'observe chez l'homme, les anthropoïdes et les singes de l'Ancien Monde, mais les singes du Nouveau Monde n'en ont que deux. La séparation des deux premiers pigments remonte au moins à 500 millions d'années, alors que celle des pigments vert et rouge ne date que d'environ 30 millions d'années. Étant données les similitudes de ces deux derniers, une nouvelle duplication pourrait se produire à la suite d'un *crossing over* propice. Notre vision s'affinera peut-être encore dans l'avenir, si les hommes sont assez patients pour survivre quelques millions d'années...

J.-C. D.

1. Nathans J, Thomas D, Hogness DS. Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. *Science* 1986; 232: 193-202.
2. Nathans J, Piantanida TP, Eddy RL, Shows TB, Hogness DS. Molecular genetics of inherited variation in human color vision. *Science* 1986; 232: 303-10.
3. Botstein D. The molecular biology of color vision. *Science* 1986; 232: 142-3.
4. Mollon JD. Understanding colour vision. *Nature* 1986; 321: 12-3.

/ Identification du gène de la granulomatose chronique

La granulomatose chronique léthale de l'enfant est une maladie héréditaire dont la forme la plus fréquente est transmise selon le mode récessif lié au sexe : les garçons sont atteints, les femmes étant transmettrices. Les symptômes comportent une sensibilité aux infections microbiennes et fongiques conduisant inmanquablement à la mort. Les phagocytes du jeune malade englobent normalement les germes mais ne les détruisent pas, probablement parce qu'ils sont incapables de stimuler, lors de la phagocytose, la production d'anion superoxyde et de peroxydes. La base biochimique de cette altération fonctionnelle des leucocytes n'est pas connue, la seule anomalie identifiée étant l'absence d'une protéine héminique de spectre similaire à celui des cytochromes b. L'existence d'observations comportant des délétions entraînant à la fois une granulomatose chronique et une myopathie avait permis de déterminer que les gènes de ces deux maladies étaient liés et situés dans la bande p21 du chromosome X. Les efforts considérables consentis pour identifier le gène de la myopathie de Duchenne (DMD en anglais, pour *Duchenne Muscular Dystrophy*) eurent dès lors de rapides retombées sur l'étude du gène de la granulomatose (CGD en anglais, pour *Chronic Granulomatous Disease*) qui fut isolé par une approche uniquement génétique selon la stratégie générale suivante [1] :

Étape 1 : Des sondes génétiquement liées au locus « DMD » hybrident avec l'ADN normal mais pas avec celui d'un malade atteint de « CGD » ; elles sont utilisées pour cribler une banque d'ADN génomique normal.

Étape 2 : Plusieurs clones sont ainsi détectés ; leur ADN est immobilisé sur des filtres qui sont hybridés avec la copie ADN des ARNm de cellules supposées contenir le produit du gène CGD (lignée granulo-monocytaire). Deux clones sont positifs.

Étape 3 : L'ADN génomique identifié à l'étape 2 est utilisé pour cribler une banque d'ADNc de cellules granulo-monocytaire. Des copies ADN du messageur sont isolées et leur séquence nucléotidique est déterminée.

Étape 4 : On vérifie que l'ADNc isolé à l'étape précédente hybride avec un messageur de 5 kb présent dans les cellules phagocytaires normales mais absent dans les cellules de la plupart des observations de « CGD ». De plus, chez un malade atteint de « CGD », cette sonde d'ADNc détecte une délétion génomique de quelques centaines de bases.

La séquence en acides aminés déduite de la séquence nucléotidique de l'ADNc correspond à une protéine de 54 000 de poids moléculaire qui pourrait être transmembranaire et glycosylée et qui ne ressemble à aucune autre connue actuellement, notamment pas aux cytochromes b. Il est maintenant certain que l'obtention par génie génétique d'une grande quantité de la protéine dont le déficit est responsable de la granulomatose permettra d'en déterminer l'activité biologique, probablement enzymatique, ce qui devrait définitivement établir les bases biochimiques de la maladie. D'un autre point de vue cette découverte marque aussi une date dans l'histoire de la génétique et de la pathologie moléculaires : c'est la première fois qu'un gène responsable d'une maladie humaine est identifié, non grâce aux connaissances acquises sur son produit, mais à partir des informations recueillies sur sa localisation chromosomique et sa transmission génétique.

Un tel isolement du gène impliqué dans la maladie permet secondairement d'analyser la structure et la fonction de la protéine pour laquelle il code... et par conséquent de comprendre les mécanismes de l'affection. Nul doute que d'autres exemples remarquables ne tarderont pas à illustrer cette nouvelle démarche, concernant la myopathie de Duchenne, la fibrose kystique du pancréas et la chorée de Huntington notamment.

A. K.

La découverte du gène de la mucoviscidose

La mucoviscidose est certainement la maladie héréditaire autosomique récessive la plus fréquente (une naissance sur 2 000, un transmetteur sur 20) dans la population européenne. Les manifestations cliniques de cette affection grave, essentiellement pulmonaires et pancréatiques, résultent d'un épaissement des sécrétions exocrines qui obstruent les voies aériennes supérieures et prédisposent aux infections bronchiques répétées. Ces atteintes détruisent progressivement le poumon et conduisent à une insuffisance respiratoire létale, le plus souvent avant 30 ans. Les études électrophysiologiques les plus récentes indiquent que, chez les malades, le transport des ions à travers les membranes des cellules épithéliales est anormal. L'activité du canal à chlore (Cl⁻), qui semble sous le contrôle d'un processus complexe de régulation, est normalement stimulée par les protéines kinases C et A (c'est-à-dire dépendantes de l'AMPc); cette activation fait défaut chez les patients. Malgré ces progrès, la lésion biochimique primaire n'est toujours pas connue, et une approche alternative, celle de la génétique inverse [1], a été mise en œuvre par plusieurs laboratoires dans le but d'isoler le gène responsable en partant de sa localisation chromosomique.

Il y a un peu moins de 18 mois, en avril 1988, la revue scientifique nord-américaine *Science*, dans un article intitulé « La course pour le gène de la mucoviscidose touche à sa fin » [2], commentait longuement les enjeux de cette intense compétition et les rivalités des quelques laboratoires encore en piste. Le groupe dirigé par R. Williamson (Londres, GB), avait en effet laissé entendre un an plus tôt, qu'il avait isolé le gène CF (de *cystic fibrosis*, pour fibrose kystique du pancréas, ou mucoviscidose), une nouvelle qui avait découragé certains des protagonistes. Il s'agissait en fait d'une séquence codant pour une protéine (*int related protein*, IRP) sans aucun rapport avec la maladie, mais située, on le sait aujourd'hui, à quel-

ques 150 kilopaires de base (kpb) du gène tant convoité. Trois ans auparavant, la découverte de la localisation du gène CF sur le chromosome 7 avait brutalement accéléré la course que viennent de gagner L.C. Tsui, R. Riordan (Université de Toronto, Canada), F.S. Collins (Université du Michigan, USA) et leurs collaborateurs, qui ont, en 1987, regroupé efforts et compétences sur le sujet. Dans sa parution du 8 septembre 1989, le même journal, *Science*, publie trois articles impressionnants décrivant l'identification du gène, la nature de son produit, et la caractérisation de la principale mutation responsable de la maladie [3-5]. Il s'agit d'un événement scientifique majeur. Pour les malades et leurs familles, cette découverte donne immédiatement accès à de nouveaux moyens de diagnostic et permet d'espérer des jours meilleurs en matière de prévention et de traitement. Pour ceux qui essaient, à l'aide de méthodes semblables, de caractériser d'autres maladies génétiques, c'est un encouragement qui vient à point tant sont ardues ces approches et peuvent sembler aléatoires leurs succès.

La marche vers le gène. Depuis 1985, à la suite des travaux des groupes de L.C. Tsui, de R. White et R. Williamson, on savait que le gène CF est situé entre deux marqueurs liés, MET et D7S8 (J3.11), séparés par approximativement un à deux millions de paires de bases [6]. Pour pouvoir, avec quelques chances de succès, commencer à marcher sur la région, c'est-à-dire loner toutes les séquences qu'elle contient, il fallait réduire cette distance en trouvant des marqueurs plus proches. De la bouche même de Lap-Chee Tsui, dont la détermination et la ténacité tout au long de l'entreprise sont à saluer, le succès de la stratégie choisie (*a brute force approach*) dépendait avant tout des moyens massifs mis en jeu. Pour se rapprocher au maximum du gène convoité, son groupe entreprit de saturer la région cible en

marqueurs ADN qui furent localisés et ordonnés à l'aide d'analyses physiques et génétiques. Pas moins de 258 séquences spécifiques du chromosome 7 furent isolées [7], dont 53 localisées en 7q31-32. Deux de ces marqueurs (D7S122 et D7S340), étroitement liés au gène CF, se révélèrent contenus dans la région délimitée par MET et D7S8. L'ordre des marqueurs déterminé (MET-D7S340-DS7122-J3.11), la carte physique de la région établie à l'aide des techniques d'électrophorèse en champ pulsé (*m/s n° 1, vol. 4, p. 46*), il restait à passer au peigne fin cette région du génome (environ 1 500 kpb).

Pour accélérer la recherche et restreindre l'étendue de la zone à cibler, les méthodes classiques de marche sur le chromosome furent doublées de tentatives de saut à partir de chaque extrémité de la région. Cette seconde approche, outre son intérêt pour rejoindre plus vite le but, permet de s'affranchir des embûches dressées par les régions « inclonables » qui, on le sait maintenant, peuvent sérieusement handicaper, sinon arrêter la progression des marcheurs. A cet effet, F. Collins avait construit en 1987 une banque de saut* constituée d'une collection de séquences d'ADN génomique distantes d'une centaine de kpb en moyenne. Des séquences contiguës d'ADN s'étendant sur une région de 280 kpb furent clonées à l'aide de ces deux approches qui permirent d'isoler quelques 49 clones de marche (cosmides ou phages) et neuf clones de saut, chacun permettant de fran-

* Banque de saut : technique qui consiste à couper de grands fragments d'ADN (autour de 100 kpb), à les recirculariser en liant leurs deux extrémités à un marqueur de sélection, puis à cloner, après une nouvelle coupure, l'ADN des deux extrémités maintenant réunies au marqueur. Si un tel clone est détecté par hybridation avec une sonde de l'une des extrémités, l'ADN de l'autre extrémité fournira lui-même une sonde correspondant à une région du chromosome située, dans l'exemple pris ici, 100 kpb plus loin.

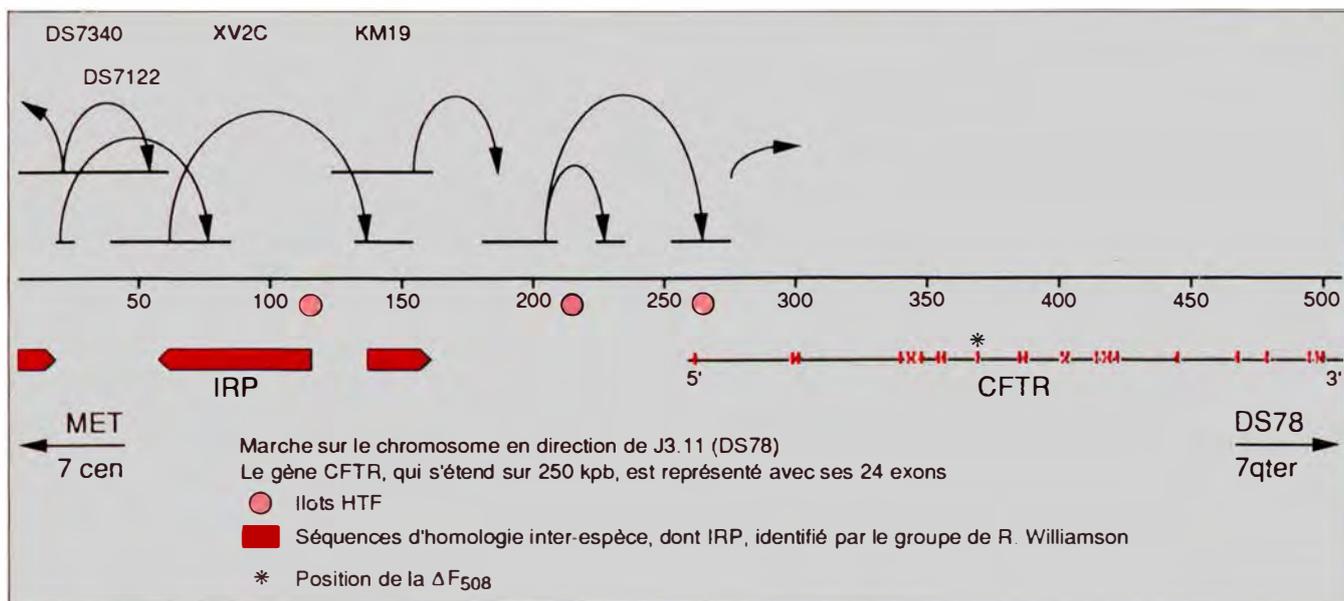


Figure 1. **Marche vers le gène de la mucoviscidose.** Le point de départ est représenté à gauche sur le schéma et correspond à la localisation de la sonde caractérisée par le groupe de L.C. Tsui comme la plus proche du gène CF. Les 280 kpb séparant ce point du début du gène CFTR ont été franchies en marchant et en sautant (flèches courbes) sur le génome en direction de DS 78. Le gène CFTR est représenté avec ses 24 exons et, sur la gauche du schéma, sous la ligne d'échelle, sont positionnées les trois séquences codantes conservées entre les espèces, dont IRP, caractérisée par le groupe de R. Williamson. Les cercles roses représentent les îlots HTF retrouvés dans la région qui, souvent, balise le début des gènes. 7 cen = vers le centromère ; 7 qter = vers le télomère du bras long.

chir entre 30 et 70 kpb (figure 1). Pour épingler les séquences codantes CF, la stratégie des *zoo-blots*** fut choisie en raison de sa relative facilité et de ses succès antérieurs, notamment pour la caractérisation du gène de la myopathie de Duchenne. Quatre régions contenant des séquences nucléotidiques homologues à celles d'autres espèces (en l'occurrence le bœuf, la souris et le poulet) furent retenues comme de possibles candidats, et les séquences conservées utilisées comme sondes pour rechercher leurs copies dans sept banques d'ADNc construites à partir d'ARN messagers de plusieurs tissus : lignée tumorale T84, cultures de cellules de glandes sudoripares normales ou de malades atteints de mucoviscidose, pancréatiques ou pulmonaires. Les trois premières zones d'homologie inter-espèces révélèrent la présence d'une séquence répétée transcrite et de deux gènes, dont IRP, déjà épinglé par le groupe de Londres grâce à la présence d'îlots HTF*** et, un temps candidat, CF. C'est la quatrième région conservée qui autorisa la pêche, difficile mais réussie, d'un seul clone contenant seulement 113 paires de bases (pb) du premier

exon du gène CF. Cette séquence permet de reconstituer, en agencant plus de 20 recombinants isolés grâce à elle, le *puzzle* de l'ADNc complet (6 129 pb) et, partant, de positionner les 24 exons du gène sur 250 kpb de génome. La distribution tissulaire des ARN messagers s'avère compatible avec les données physiopathologiques, de même que le produit du gène, une protéine de transport membranaire, que ses découvreurs proposent d'appeler CFTR (*cystic fibrosis transductor regulator*).

Le CFTR. Le gène ainsi caractérisé code pour une protéine de 1 480 acides aminés dont la séquence indique à l'évidence son appartenance à une super-famille de protéines membranaires. Le CFTR présente une homologie structurale considérable avec un ensemble de transporteurs dépendants de l'ATP décrits chez les procaryotes, et plus récemment chez les eucaryotes [8], telles les protéines MDR (*multidrug resistance*), CQR (*chloroquine resistance*) ou STE6 (de levure). Il associe deux motifs répétés formés d'un domaine hydrophobe transmembranaire, et une importante région hydrophile exposant des séquences consensus NBF

(*nucleotide (ATP)-binding fold*) pouvant lier l'ATP. Ces deux parties de la protéine sont reliées par un domaine cytoplasmique (domaine R), codé par un grand exon unique, présentant de nombreux résidus chargés, contenant la majorité des sites potentiels de phosphorylation par les protéines kinase A ou C, d'où sa fonction régulatrice probable (figure 2).

Le rôle exact joué par cette protéine reste à déterminer. Elle est très probablement impliquée dans le transport transmembranaire actif des ions au niveau du pôle apical des cellules épithéliales. Il reste à établir s'il

** *Zoo blot*: technique qui se base sur le contraste entre la conservation au cours de l'évolution des séquences codantes et la non-conservation des séquences introniques et intergéniques pour identifier des fragments exoniques ; seuls de tels exons peuvent hybrider en Southern blot avec des fragments d'ADN d'espèces variées.

*** *îlots HTF*: *HpaII* tiny fragments, îlots contenant de nombreux sites de coupure par l'enzyme de restriction, *HpaII* qui reconnaît la séquence 5'CCGG3' lorsque les C sont méthylés. Des régions riches en CG méthylées sont caractéristiques des séquences bordantes 5' des gènes actifs, principalement ceux à expression ubiquitaire.

s'agit d'un canal ionique, ou plutôt d'un régulateur de l'activité d'un tel canal, et si elle a par ailleurs d'autres fonctions cellulaires sans rapport avec le transport actif des ions. Quoiqu'il en soit, comme le suggèrent les données structurales, l'hydrolyse de l'ATP et la phosphorylation par les protéines kinases semblent nécessaires à cette activité de transport actif, et à sa régulation.

La comparaison de la structure des clones d'ADN complémentaire obtenus à partir de cellules de glandes sudoripares de sujets sains ou de malades confirme ce que les études génétiques antérieures avaient permis de postuler, à savoir qu'une majorité des patients sont porteurs de la même mutation. Il s'agit d'une délétion de trois nucléotides de l'exon 10 qui ampute la protéine d'une phénylalanine en position 508 (ΔF_{508}), c'est-à-dire au niveau d'une des séquences NBF. Cette anomalie pourrait empêcher la liaison correcte de l'ATP, ou perturber la dynamique conformationnelle indispensable à l'activité du CFTR, expliquant par là les observations faites sur l'épithélium bronchique des malades grâce aux expériences de *patch-clamp* [9].

Génétique de la mucoviscidose : prévention et traitement. Les données de l'analyse génétique apportent les arguments les plus solides en faveur du fait que le gène codant pour le CFTR est bien le gène CF. Environ 70 % des défauts moléculaires observés chez les patients ou les transmetteurs sont constitués par la délétion ΔF_{508} . Une étude exhaustive des marqueurs ADN de la région a permis d'établir une série d'haplotypes utiles à la caractérisation des divers chromosomes sains ou CF. La délétion ΔF_{508} est presque toujours associée, sur les chromosomes CF, à un haplotype majoritaire ou à ses dérivés ; en revanche, elle n'est jamais retrouvée sur les chromosomes sains portant cet haplotype. Cette observation rend extrêmement probable l'hypothèse qu'il s'agit d'une mutation délétère, et non d'un polymorphisme. En une occasion, elle est retrouvée associée à un haplotype différent, ce qui suggère la survenue d'un événement mutationnel récurrent, phénomène déjà observé pour le gène β -globine. Les autres lésions moléculaires constituent un groupe hétérogène et

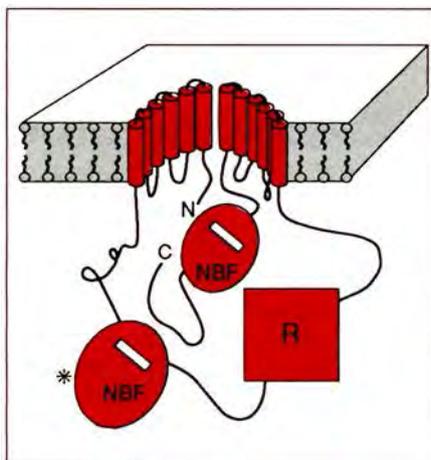


Figure 2. **Schéma de la protéine CFTR, produit du gène CF.** Les deux régions transmembranaires sont séparées par une région de liaison potentielle de l'ATP (NBF, nucleotide binding fold) et par la grande sous-unité R. Il existe un autre domaine NBF à proximité de l'extrémité carboxyterminale de la protéine. L'astérisque indique la position de la ΔF_{508} . (Modifié d'après [4].)

restent à caractériser, mais l'analyse génétique donne déjà quelques clefs importantes. Deux formes peuvent être distinguées cliniquement, PI (insuffisance pancréatique), associée à ΔF_{508} , et PS (sans insuffisance pancréatique). L'analyse génétique fait apparaître que les mutations responsables de ces deux formes sont de gravité différente, et que le phénotype PS est dominant sur PI. L'iléus méconial****, qui survient chez 5 à 10 % des nouveaux-nés, surviendrait chez des sujets porteurs de la délétion ΔF_{508} , dans un contexte génétique ou d'environnement qui reste à élucider. Ces découvertes sont d'une très grande importance, à terme, pour la compréhension de la biologie et de la physiologie du système. Elles concernent déjà les patients et leurs familles en bouleversant les moyens de prévention. Près de 70 % des transmetteurs de la mucoviscidose peuvent d'ores et déjà être identifiés facilement, à l'aide de la sonde spécifique ΔF_{508} , par des techniques maintenant bien standardisées. En dépit des problèmes posés par la non-caractérisation de 30 % des muta-

**** Iléus méconial : occlusion intestinale néonatale due à l'épaississement du méconium qui ne peut s'évacuer.

tions, rien ne s'oppose à ce que le dépistage des transmetteurs soit mis en place dès maintenant, en commençant avec des échantillons ciblés, judicieusement définis, de la population. Bien sûr, pour les mutations encore inconnues du gène, il n'est pas encore possible de s'affranchir des méthodes d'analyse de liaison génétique, et le diagnostic prénatal, notamment, requiert pour sa mise en œuvre la maîtrise des deux approches. Cependant, la mucoviscidose est désormais accessible au dépistage à grande échelle.

Cette percée laisse-t-elle envisager, à moyen terme, la mise au point d'un traitement curatif ? Certes, il est d'autres maladies pour lesquelles le défaut moléculaire est connu de longue date et que l'on ne sait toujours pas guérir. Mais la connaissance de la structure et du fonctionnement du CFTR sont maintenant une question de temps, et l'on peut raisonnablement espérer que l'on pourra à terme agir avec efficacité contre les effets délétères du gène anormal en modulant l'activité de son produit.

RÉFÉRENCES

1. Jordan B. Grandeur et servitudes de la génétique inverse. *médecine/sciences* 1988 ; 4 : 138-40.
2. Roberts L. The race for the cystic fibrosis gene. *Science* 1988 ; 240 : 141-4 ; 282-5.
3. Rommens JM, Tannuzzi MC, Kerem BS, et al. Identification of the cystic fibrosis gene : chromosome walking and jumping. *Science* 1989 ; 245 : 1059-64.
4. Riordan JR, Rommens JM, Kerem BS, et al. Identification of the cystic fibrosis gene : cloning and characterization of complementary DNA. *Science* 1989 ; 245 : 1066-72.
5. Kerem BS, Rommens JM, Buchanan JA, et al. Identification of the cystic fibrosis gene : genetic analysis. *Science* 1989 ; 245 : 1073-80.
6. Kitzis A, Warren P, Kaplan JC. *médecine/sciences* 1988 ; 4 : 151-6.
7. Rommens JM, Zengerling S, Burns J, et al. Identification and regional localization of DNA markers on chromosome 7 for the cloning of the cystic fibrosis gene. *Am J Hum Genet* 1988 ; 43 : 645-63.
8. Higgins CF. Export-import family expands. *Nature* 1989 ; 340 : 342.
9. Sauvé R. Le *patch clamp*, une nouvelle façon de voir le canaux ioniques. *médecine/sciences* 1987 ; 3 : 538-45.

Michel Goossens

Professeur à la faculté de médecine, Université Paris XII.

Laboratoire de biochimie et Inserm U. 91, hôpital Henri-Mondor, 51, avenue du Maréchal-de-Lattre-de-Tassigny, 94010 Créteil, France.

Génie génétique, production de médicaments et thérapie génique

Production par génie génétique de protéines d'intérêt biologique et thérapeutique.

La purification en grande quantité d'hormones peptidiques, de facteurs de coagulation, de protéines vaccinales... ou de tout autre polypeptide d'intérêt biologique (interférons, interleukines, facteurs de croissance, etc.), est le domaine privilégié des biotechnologies utilisant les méthodes des recombinants d'ADN, ou « génie génétique ». Le principe de ces méthodes est de transférer un gène naturel « cloné » et plus ou moins modifié *in vitro*, voire même un gène artificiel synthétisé chimiquement, dans une cellule vivante qui servira à l'« exprimer » en la protéine biologiquement active désirée.

Les trois exigences requises pour entreprendre une telle étude sont celles de l'activité du produit obtenu, de son innocuité et de sa production avec un rendement et à des coûts en faisant une méthode concurrentielle de celles, plus classiques, consistant à l'extraire de tissus humains ou animaux.

L'aspect quantitatif du problème exige la construction de gènes hybrides auxquels ont été rajoutées différentes séquences exogènes destinées à augmenter la transcription du gène, la stabilité du messenger (et parfois de la protéine), sa traductibilité et, chaque fois que possible, à permettre l'exportation de la protéine dans le milieu extra-cellulaire (*figure 1, page suivante*). L'innocuité du produit requiert, notamment s'il

s'agit d'un facteur de substitution à administrer à des malades, que soient éliminées de la protéine à utiliser toutes séquences protéiques potentiellement immunogènes.

L'activité biologique exige enfin dans un certain nombre de cas que la protéine synthétisée subisse toute une série de modifications post-traductionnelles transformant un précurseur inactif en principe biologiquement actif : glycosylation, carboxylation, coupure protéolytique, par exemple. Or, seules les cellules eucaryotiques, et parfois même des cellules animales différenciées, sont capables d'effectuer ces opérations. En l'absence de modifications post-synthétiques indispensables, les bactéries (et surtout *E. coli*) sont les meilleures cellules productrices, peut-être concurrencées par les levures qui sont elles, en plus, capables d'effectuer certaines opérations de maturation post-traductionnelle (comme certains types de glycosylation).

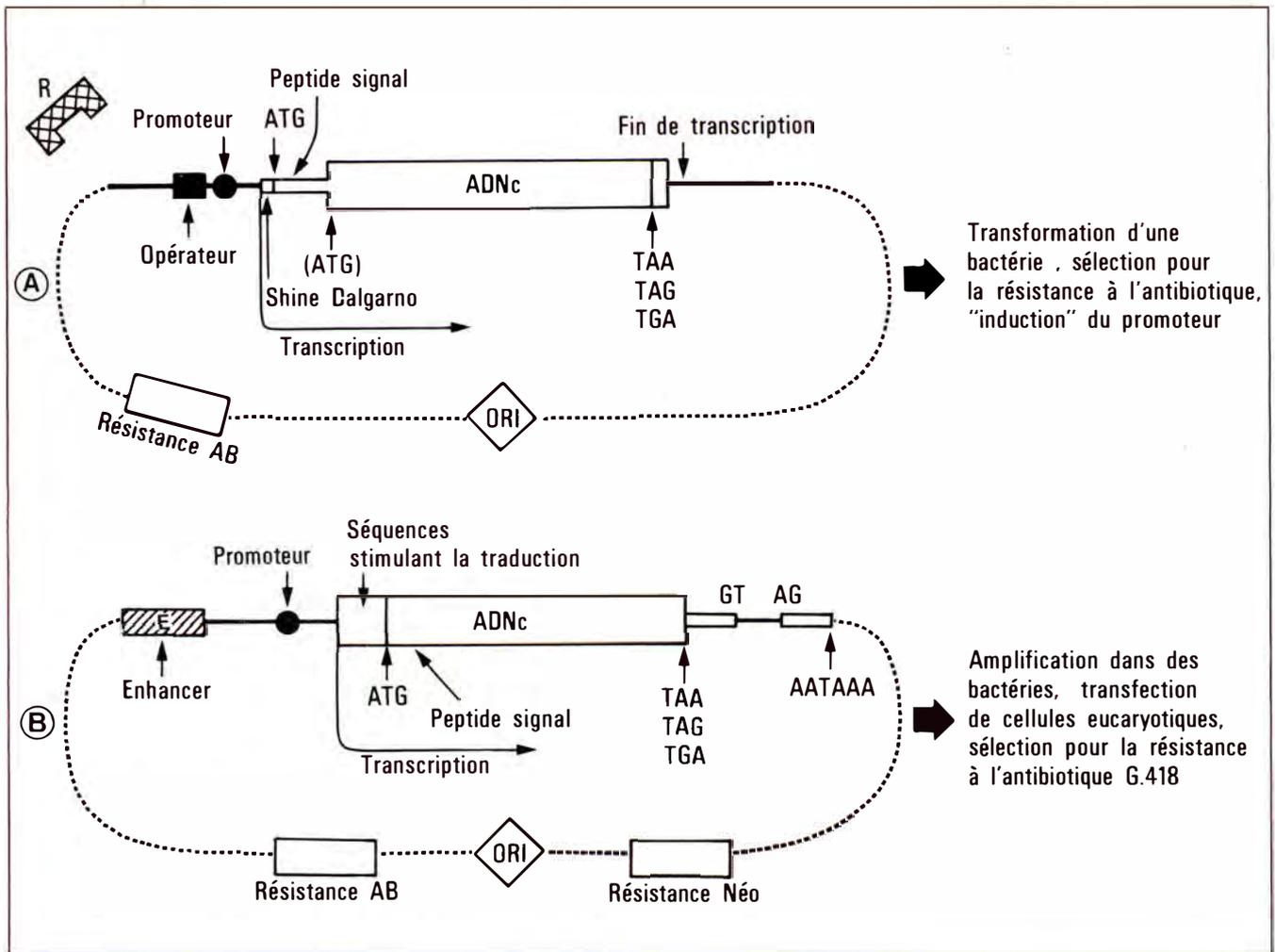
Les cellules animales sont plus difficiles à cultiver en très grande quantité et produisent moins de la protéine désirée. Elles sont cependant le seul système utilisable dans des cas comme la production de facteurs de coagulation.

Thérapie génique : tendances pour demain.

Quoique la « chirurgie » et la « greffe » des gènes n'aient encore jamais été réalisées chez l'homme, il nous a semblé difficile de clore un « lexique » consacré à la génétique moléculaire et à la médecine sans au moins évoquer cet aspect.

La manipulation génétique d'embryons dans le but de remplacer un gène absent ou défectueux est techniquement concevable (quoique probablement très difficile) mais se heurte à de redoutables problèmes éthiques. Plus que le transfert direct de gènes dans l'œuf fécondé *in vitro*, ce qui pourrait être envisagé est le traitement *in utero* du fœtus par l'injection intra-amniotique de particules virales contenant un gène exogène destiné à remplacer ou à compléter celui du fœtus (*voir plus loin*). L'avantage d'une telle approche serait de pouvoir être utilisée après un diagnostic prénatal... son désavantage étant de ne s'appliquer qu'à des maladies dont les effets ne seraient pas irréversiblement délétères pour l'embryon et le fœtus jeune.

Plus proches sont probablement les premières tentatives de correction post-natale de maladies génétiques monogéniques par autogreffe de cellules génétiquement recombinées. Les affections types qui pourraient bénéficier d'un tel traitement sont les rares déficits enzymatiques associés à un déficit immunitaire sévère, tel le déficit en adénosine désaminase. Une restitution partielle de l'activité enzymatique serait suffisante pour restaurer l'immunité. Il devrait donc être possible de prélever des cellules souches médullaires des malades, de leur transférer un gène (ou un ADN complémentaire, comme dans le schéma de la *figure 1 B, page suivante*) codant pour l'adénosine désaminase, et



de greffer aux malades ces cellules ainsi modifiées. Selon le même principe de restitution partielle, on pourrait envisager de transférer dans des cellules variées (fibroblastes ou cellules médullaires) un ADN complémentaire codant pour un facteur antihémophilique, la sécrétion dans le sang circulant d'une faible quantité d'un tel facteur pouvant mettre les hémophiles à l'abri des accidents hémorragiques les plus graves. Une telle manipulation dans le cas des maladies moléculaires de l'hémoglobine, autrement plus fréquentes que les déficits immunitaires secondaires à une enzymopathie, est beaucoup plus difficile. Il faudrait, en effet,

modifier ici toutes les cellules souches de telle sorte qu'elles produisent une quantité normale d'hémoglobine normale... ce qui semble aujourd'hui difficile. Les vecteurs les plus étudiés actuellement pour introduire un gène exogène dans des cellules, médullaires ou autres, sont des rétrovirus dont tous les gènes potentiellement pathogènes ont été délévés et remplacés par le gène à insérer dans le génome des cellules à modifier. Enfin, une dernière voie théorique de transfert de gène dans l'organisme serait d'injecter au receveur des particules virales infectieuses contenant un génome recombiné et capables de l'introduire dans

différentes cellules. Ces particules pourraient être, là encore, rétrovirales, ou bien dérivées de virus à ADN provoquant normalement des infections bénignes ou inapparentes (certains adénovirus, ou des virus associés aux adénovirus). Les obstacles sont nombreux et divers, comportant le risque qu'un rétrovirus, même non pathogène par lui-même, s'insère à proximité d'un oncogène et l'active, et les incertitudes persistent sur l'avenir d'une infection par un virus à ADN dont le cycle normal ne comporte pas d'intégration dans le génome cellulaire et contre lequel l'organisme doit développer une réaction immunitaire.

A. K.

Figure 1. **Construction de gènes artificiels à visée biotechnologique.** A) Transformation de cellules bactériennes. L'ADN complémentaire du messager que l'on veut exprimer à un haut niveau dans les cellules bactériennes (en général *E. coli*) est placé sous le contrôle d'un promoteur bactérien inductible et fort. Il comprend, au-delà du site d'initiation de la transcription : la séquence de Shine Dalgarno, indispensable à la fixation des ribosomes bactériens au messager, et donc à sa traduction ; un codon d'initiation de la traduction (ATG) ; une séquence codant pour un peptide signal d'une protéine bactérienne, permettant la sécrétion du produit final du gène étudié dans le milieu de culture bactérienne ; éventuellement, un ATG codant pour une méthionine et non utilisé comme initiateur de la traduction, situé juste entre les séquences bactériennes à gauche (rectangle étroit) et eucaryotiques à droite (rectangle large). Au-delà du codon de terminaison de la traduction (TAA, TAG ou TGA), un signal spécifique de fin de transcription est indispensable dans un système procaryotique. Le plasmide dans lequel est intégré ce gène comprend au moins une origine de replication (ORI) et un gène de résistance aux antibiotiques. Les bactéries sont « transformées » par un tel plasmide recombinant qui leur confère la résistance à un antibiotique, ceci permettant d'éliminer les bactéries non transformées. Le gène hybride n'est alors pas exprimé, le promoteur étant inhibé par la fixation d'un répresseur (R) sur une région régulatrice appelée « opérateur ». Lorsque cette répression cesse (par exemple, parce que le répresseur est thermo-instable et que le milieu de culture est placé à une température le détruisant, ou bien par addition d'une substance bloquant la fixation du répresseur sur l'opérateur), le gène est transcrit très activement et l'ARN est traduit. Le produit protéique, grâce à la présence d'un peptide signal de protéine bactérienne (m/s n° 6, vol. 2, p. 341) est sécrété dans le milieu à partir duquel il peut être purifié. Les restes de protéine bactérienne persistant après l'excision du peptide signal lors du passage transmembranaire du produit peuvent être éliminés par clivage de la méthionine rajoutée entre les séquences procaryotiques et eucaryotiques... à condition qu'il n'existe pas d'autres méthionines dans le gène.

B) Transfection de cellules eucaryotiques. Le gène (ou plus souvent l'ADN complémentaire dépourvu d'introns) dont on désire faire synthétiser le produit par des cellules eucaryotiques est placé sous le contrôle d'un promoteur et d'une séquence stimulatrice (enhancer) forts et actifs chez les eucaryotes, tels par exemple des promoteurs et des enhancers viraux (LTR du virus de Rous, séquences régulatrices de gènes d'adénovirus ou de virus SV-40). Au-delà du site d'initiation de la transcription, on peut introduire dans cette construction plusieurs types de signaux : une séquence localisée dans la partie 5' non codante du messager et destinée à en augmenter la traductibilité, comme cela a été démontré dans le virus LAV-HTLVIII (ou HIV) (séquence TAR, m/s n° 5, vol. 2, p. 285). Au-delà de l'ATG initiateur de la traduction il est, là aussi, possible d'introduire une séquence de peptide signal permettant la sécrétion du produit dans le milieu ; après le signal de fin de traduction, l'introduction d'une séquence comportant un intron, avec ses séquences consensus d'excision-épissage (m/s n° 3, vol. 1, p. 158) est de nature à augmenter l'expression du gène hybride ; enfin, la présence d'un signal de polyadénylation AATAAA est absolument indispensable à la stabilité du messager. Le plasmide dans lequel est insérée cette construction peut être similaire au précédent (ce qui permet son « amplification » dans des cellules bactériennes) et comporter en plus le gène « Néo » qui code pour la résistance des cellules eucaryotiques à l'antibiotique G 418. Ainsi, après amplification du plasmide dans des cellules bactériennes, celui-ci est transféré par divers procédés (on parle alors de « transfection ») dans des cellules eucaryotiques en culture. Le plasmide s'intègre alors au bout de quelques jours dans l'ADN chromosomique des cellules transfectées qui sont sélectionnées grâce à leur résistance au G 418. Les « clones » ainsi isolés sont testés pour la production de la protéine codée par le gène introduit.

STRATÉGIE THÉRAPEUTIQUE DANS LA MALADIE DE HODGKIN

Symposium International, Paris, 28, 29 et 30 juin 1989

Michel Henri-Amar, Hélène Sancho-Garnier (Institut Gustave-Roussy, Villejuif)

- Résultats de l'analyse des facteurs pronostiques
Stades I-II/Stades III-IV
- Morbidité à long terme, incidence des seconds cancers
- Maladie de Hodgkin chez l'enfant
- Clinique et histologie, nouvelles techniques diagnostiques
- Place de l'autogreffe de moelle osseuse aux stades III-IV
- Place de l'aide à la décision dans le choix d'une thérapeutique adaptée

Co-édition John Libbey Eurotext/INSERM

Nouvelles orientations pour la thérapie génique

Axel Kahn,
Pascale Briand

Etant acquis qu'il n'y a ni indication ni légitimité à la modification génétique héréditaire de l'embryon précoce (*Tableau 1*), au moins en l'état actuel de nos connaissances, les efforts des « thérapeutes géniques » se concentrent déjà depuis plusieurs années sur la thérapie génique somatique, c'est-à-dire sur l'apport d'un gène fonctionnel à des cellules n'intervenant pas dans la gamétogenèse et possédant un équivalent anormal du gène transféré.

Les greffes

En un sens, les greffes d'organes pratiquées pour traiter des maladies héréditaires peuvent être considérées comme une thérapie génique, le but en étant bien, au moins en partie, l'apport d'un gène fonctionnel. A dire vrai, ce sont même pour l'instant les seules thérapies géniques à avoir fait la preuve de leur efficacité : greffes de moelle osseuse pour des hémoglobinopathies graves ou des déficits immunitaires congénitaux, greffes de foie pour des affections en rapport avec des troubles héréditaires du métabolisme hépatique, voire greffes de poumons ou de blocs cœur-poumons chez des mucoviscidosiques, etc. Les progrès extrêmement rapides réalisés dans le domaine des greffes, l'apparition d'immunosuppresseurs de plus en plus efficaces et de moins en moins toxiques font des greffes une réelle perspective dans nombre de maladies héréditaires. Les exigences de compatibilité immunologique, le

manque de disponibilité des organes à greffer, les risques de contamination virale des greffons, la nécessité d'un traitement immunosuppresseur et d'une surveillance prolongée des malades greffés ne permettent cependant pas de considérer cette stratégie autrement que comme un pis-aller justifié par l'absence actuelle d'autres perspectives.

Autogreffes de cellules génétiquement modifiées

Dans son principe, cette approche est extrêmement attrayante : il s'agit en effet d'utiliser les propres cellules du malade (ce qui supprime tout problème immunologique), de les modifier par introduction du gène sensé compenser l'anomalie génétique, puis de les réintroduire dans l'organisme (*figure 1*). Avant même de passer à l'expérimentation, des questions surgissent cependant. La première est celle de l'inocuité de l'introduction, au hasard dans le génome, d'un fragment d'ADN, le plus souvent apporté par un vecteur rétroviral (*figure 2*). Les risques potentiels d'activation d'un oncogène situé à proximité du site d'intégration de l'ADN exogène sont bien sûr diminués autant qu'il est possible par utilisation de vecteurs dont les *enhancers* (situés dans les LTR, *long terminal repeats*) sont inactivés. De plus, l'absence d'apparition de cancers spontanés chez les souris transgéniques ayant intégré,

m/s

vient de paraître

Mars 90

n° 3, vol. 6

NUMÉRO
ANNIVERSAIRE
1985-1990

Thérapie génique germinale chez l'homme ?

- Fécondation in vitro
- Embryon → morula ou blastocyste
- Prélèvement d'un blastomère, PCR
 - atteint (homozygote, maladie récessive hétérozygote, maladie dominante)
 - sain

TRI, et non THÉRAPIE !

Tableau I. *La thérapie germinale chez l'homme ne serait envisageable qu'après un diagnostic de la maladie sur des embryons issus de FIV et maintenus in vitro jusqu'à un stade morula ou blastocyste. Le diagnostic se ferait par utilisation de la PCR. Puisque dans tous les cas, on se retrouverait devant un mélange d'embryons sains et malades, c'est le tri d'embryon qui s'imposerait, pas la thérapie génique.*

Les cellules utilisables pour la thérapie génique somatique

• Cellules souches médullaires	• Fragiles
• Fibroblastes	• Divisions limitées
• Kératinocytes	• Sécrétion externe
• Cellules endothéliales	
• Hépatocytes	• Difficiles à transformer, à établir
• Lymphocytes	• Durée de vie limitée (?)
• Cellules nerveuses embryonnaires	• (?)

Tableau II.

TYPE DE MALADIE HÉRÉDITAIRE	THÉRAPIE GÉNÉRIQUE		ALLOGREFFE
	Autogreffe	Vecteur in vivo	
1 organe responsable non atteint (maladies de la coagulation, hypercholestérolémie, oxalose, déficit immunitaires)	+		+, Fait
1 organe responsable et atteint, greffable (maladies de l'hémoglobine, déficit en a1 AT, tyrosinémie, glycoséose)	+ (Moelle)		+, Fait
Organe atteint non greffable, ou tous les tissus (Tay Sachs, myopathie, mucoviscidose)		+	

Tableau III. *Types de thérapie génique ou de greffe d'organe adaptées à diverses situations de maladies héréditaires.*

Thérapie génique somatique

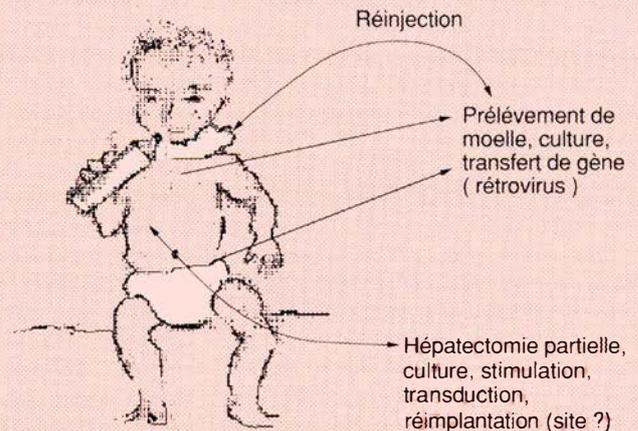


Figure 1. *La thérapie génique somatique peut consister en un prélèvement chez l'enfant atteint de cellules malades (moelle osseuse ou hépatocytes, par exemple), en l'introduction in vitro du gène curateur, puis en la réimplantation des cellules génétiquement modifiées.*

au hasard dans leur génome, des séquences d'ADN non oncogéniques semble indiquer que le processus d'activation insertionnelle n'est pas fréquent. Le risque théorique en persiste néanmoins et, quand bien même la méthode aurait par ailleurs démontré son efficacité, resterait probablement un obstacle durable aux autorisations larges d'essais chez l'homme ; les décisions seraient alors prises au cas par cas, tenant compte, comme dans tout essai thérapeutique, de l'efficacité supposée, de la gravité de l'affection à traiter et de l'évaluation du risque potentiel.

Les incertitudes persistantes à propos des conséquences éventuelles d'une insertion au hasard d'ADN exogène dans le génome cellulaire expliquent que beaucoup d'espoirs soient désormais placés dans la « recombinaison homologue » qui permet d'intégrer un fragment d'ADN au sein de la séquence homologue de l'ADN chromosomique, voire de réellement « réparer » un gène modifié par mutation. L'approche consisterait ici à sélectionner par culture *in vitro* les cellules des malades sièges d'un événement de recombinaison homologue, comme cela est décrit en détail dans l'article de Valérie Lemarchandel et Xavier Montagutelli (*m/s n° 1, vol. 6, p. 18*), puis à réimplanter ces cellules. Malheureusement, les récents résultats obtenus par R.L. Brinster *et al.*, en collaboration avec R.D. Palmiter (Philadelphie, PA, et Seattle, WA, USA) indiquent que le processus de recombinaison homologue peut s'accompagner de l'introduction de maintes mutations risquant d'inactiver le gène que l'on désirait réparer [1].

Une autre difficulté, évidente *a priori* et confirmée *a posteriori*, réside en l'exigence que les cellules génétiquement modifiées soient, notamment, des cellules souches douées d'autorenouvellement, et non des cellules déjà engagées dans une voie donnée de différenciation terminale. Dans ce dernier cas, en effet, on peut s'attendre à la disparition rapide des cellules modifiées, à la vitesse de leur différenciation irréversible en des cellules à durée de vie limitée.

Enfin, exigence supplémentaire, il faut s'assurer que le gène transféré s'exprime quantitativement à des

niveaux suffisants pour avoir une réelle action curative. A la lumière de tous ces prérequis, il semblait probable, jusqu'à il y a une année, que les maladies curables par l'intermédiaire de cellules médullaires génétiquement modifiées seraient les premières à faire l'objet de tentatives thérapeutiques. Il ne s'agissait pas tant des maladies moléculaires de l'hémoglobine exigeant que les cellules modifiées synthétisent la chaîne de globine codée par le gène transféré en quantité énorme, ajustée à la synthèse endogène de la chaîne non intéressée par la maladie, que de carences immunitaires rares dues à des déficits enzymatiques intéressant le métabolisme des nucléotides puriques (adénosine désaminase et purine nucléosides phosphorylase). Les conditions semblaient particulièrement favorables : la moelle est facile à prélever, le succès des greffes de moelle indique que des cellules souches sont bien présentes dans ces prélèvements, et beaucoup d'éléments indiquent qu'une correction partielle des déficits enzymatiques devrait être suffisante pour supprimer la carence immunitaire. En réalité les difficultés expérimentales rencontrées chez l'animal furent considérables, liées soit à l'altération des cellules souches médullaires au cours de la culture *in vitro* nécessaire à l'infection par les rétrovirus recombinés, soit à la moindre infectabilité des cellules souches comparées aux cellules déjà engagées dans les différentes voies hématopoïétiques. De plus, des réarrangements associés à l'intégration de l'ADN exogène sont apparus fréquents. Jusqu'à très récemment, par conséquent, on n'était pas parvenu à maintenir durant une longue période un haut niveau d'expression de l'ADN exogène transféré à l'animal d'expérience via l'infection rétrovirale de cellules médullaires [2]. Une série de progrès récents permettent néanmoins d'être plus optimiste. Il s'agit de la simplification maximale des vecteurs, de l'utilisation de promoteurs internes de préférence au promoteur rétroviral des LTR et de la limitation au strict minimum des manipulations *in vitro* des cellules médullaires.

De plus, l'infectabilité des cellules souches par les rétrovirus peut être

considérablement stimulée en provoquant leur entrée dans le cycle cellulaire grâce à l'action de cytokines, notamment les interleukines 3 et 6 [3, 4, 5]. Grâce à ces perfectionnements, des expressions très prolongées de gènes exogènes introduits dans des cellules médullaires greffées ont été obtenues chez la souris.

Les candidats « thérapeutes géniques » se sont aussi tournés vers d'autres cellules receveuses éventuelles qui accepteraient de se laisser modifier et d'engendrer *in vivo* des lignées permanentes (*Tableau III*). Les fibroblastes ne semblent pas de bons candidats, car leur capacité de division est limitée (*m/s n° 7, vol. 3, p. 425*). Les kératinocytes ont quelques propriétés intéressantes... mais auraient plus tendance à sécréter les produits des gènes transférés à l'extérieur, vers la peau, que dans la circulation sanguine. Les cellules endothéliales ont des propriétés très attrayantes : elles sont au contact direct du flot sanguin, semblent douées d'autorenouvellement et peuvent être apportées par l'intermédiaire de greffons vasculaires (*m/s n° 7, vol. 5, p. 511*). Fibroblastes, kératinocytes et cellules endothéliales génétiquement recombinés peuvent non seulement apporter à l'organisme des protéines dont elles assurent habituellement la synthèse, mais aussi des molécules normalement produites par d'autres cellules. Ainsi la production de facteur IX normalement synthétisé par les hépatocytes peut être assurée par des kératinocytes ou des cellules endothéliales qui vont le déverser en quantité appréciable dans la circulation [6].

Les lymphocytes entrent aussi dans cette catégorie de cellules, potentielles usines de production d'un facteur manquant. Ils doivent être mis à part, principalement parce qu'ils sont le premier type de cellules à avoir été utilisé pour une expérience de transfert de gène chez l'homme et que nous possédons par conséquent des premières indications sur leur comportement après infection par un rétrovirus transportant un gène étranger, en l'occurrence un marqueur de résistance à un antibiotique actif sur les cellules de mammifères (*m/s n° 9, vol. 5, p. 698*). Il s'agissait

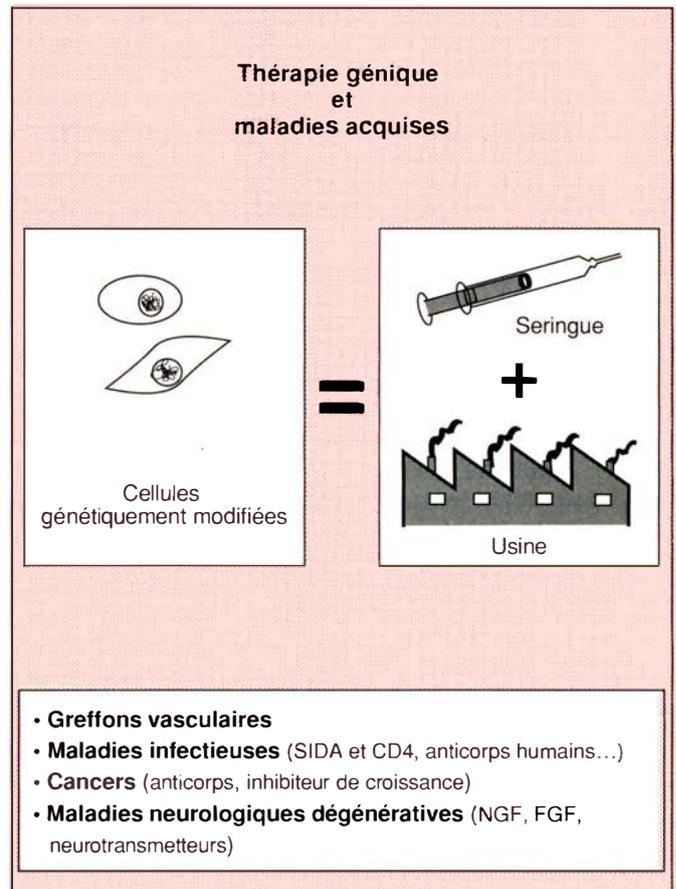
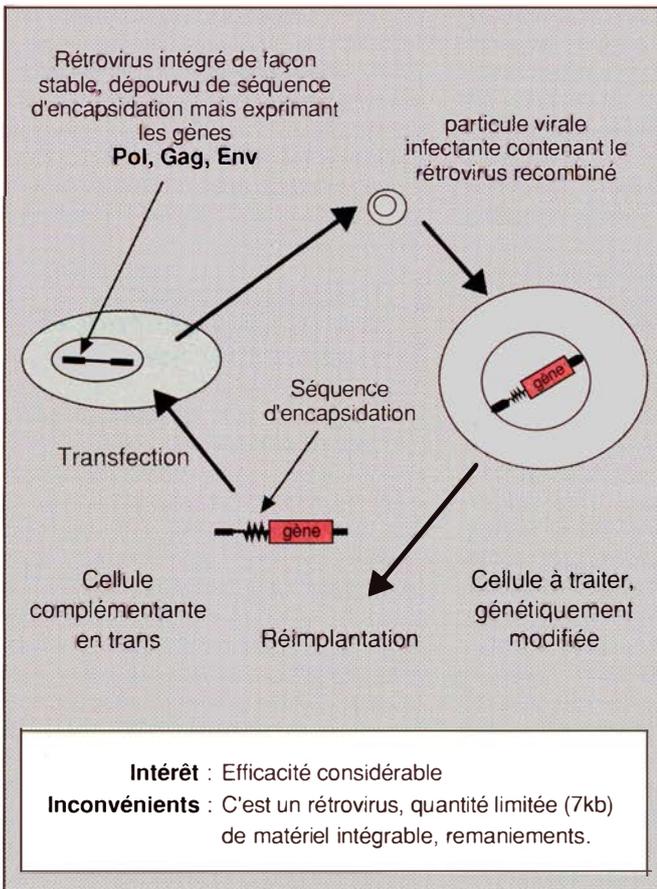


Figure 2. **Utilisation de rétrovirus à des fins de thérapie génique somatique** Figure 4.

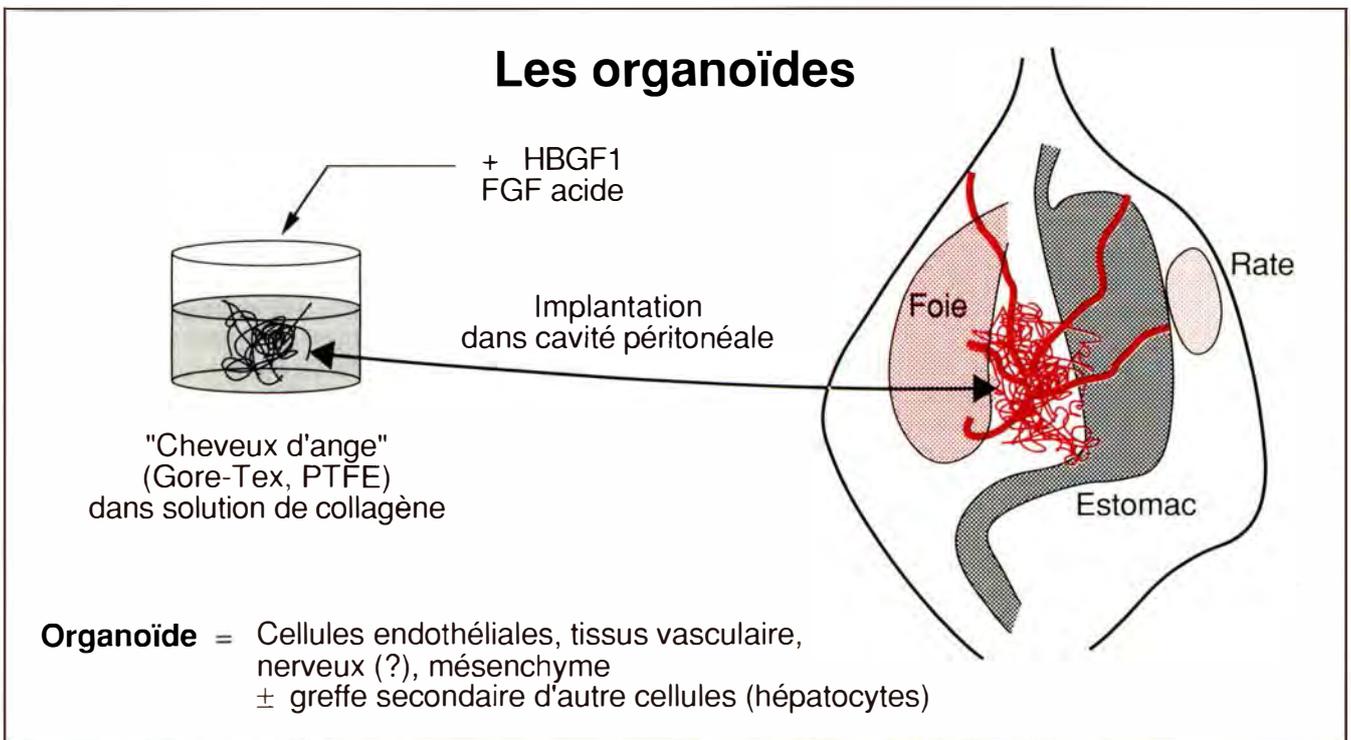


Figure 3.

de suivre le devenir de TIL (*tumor infiltrating lymphocytes*) chez des malades souffrant de mélanomes métastasés. Ces lymphocytes peuvent en effet, après activation par l'interleukine-2, avoir un effet antitumoral, les échecs pouvant être dus à l'incapacité des TIL réinjectés de se concentrer au niveau des tumeurs. Quoiqu'il soit très improbable avec ces cellules de parvenir à une colonisation définitive *in vivo*, les expériences menées chez les malades atteints de mélanome ont montré qu'elles pouvaient survivre, dans de bonnes conditions, plusieurs mois. Faciles à prélever, faciles à injecter, les lymphocytes sont donc peut-être une solution intéressante, soit que l'on envisage de traiter une maladie génétique (par exemple un déficit immunitaire) en répétant les injections de cellules génétiquement modifiées, soit que la thérapie génique soit utilisée, comme nous le verrons plus loin, pour traiter une maladie non génétique, éventuellement transitoire.

Les hépatocytes, malgré la difficulté de leur manipulation *in vitro* et leur faible degré d'infectabilité par les rétrovirus, ont fait l'objet de nombreux essais de transfert parce que le traitement de maladies héréditaires variées du métabolisme nécessite une correction de l'anomalie dans ce type cellulaire. La faisabilité des réimplantations d'hépatocytes, recombinaison génétiquement ou non, a été démontrée. De telles mini-greffes ont par exemple permis de faire apparaître de la bilirubine conjuguée dans la bile de rats *Gunn* déficients en UDP-glucuronyl transférase, ainsi que de l'albumine dans le sang de rats albuminémiques (*m/s n° 5, vol. 3, p. 303*).

En revanche, des hépatocytes infectés par des rétrovirus contenant le gène du récepteur des LDL et réinjectés dans la cavité péritonéale de lapins *Watanabe*, déficients en ce récepteur, n'ont pas permis d'obtenir un résultat appréciable et durable sur l'hypercholestérolémie de ces animaux [7].

Par conséquent, si l'on peut considérer que la réintroduction des hépatocytes ne pose pas de problèmes particuliers, les niveaux d'expressions et la persistance d'une telle expression après transfert restent de faible

importance. De plus, une stratégie classique de thérapie génique effectuée sur les hépatocytes nécessiterait leur prélèvement sur le malade par hépatectomie partielle. Sans être impossible, la difficulté de chaque étape rend probablement lointaine l'application d'une telle approche.

Les organoïdes

Très récemment, plusieurs équipes de la côte Est des États-Unis ont décrit une technique remarquable qui pourrait faciliter considérablement la thérapie génique somatique, notamment celle utilisant les hépatocytes : il s'agit de la création d'organoïdes [8]. *médecine/science* a longuement traité de l'angiogénèse, préalable indispensable à la croissance des tissus, normaux et tumoraux (*m/s n° 5, vol. 4, p. 318 et n° 7, vol. 5, p. 516*). Les facteurs angiogéniques sont nombreux ; beaucoup d'entre eux ont la propriété de se fixer à l'héparine ou à des glycosaminoglycane de la matrice extracellulaire. Ils forment ainsi la famille des HBGF (*heparin-binding growth factors*). HBGF-1 est identique au FGF acide (*fibroblast growth factor*) alors que HBGF-2 est le FGF basique.

La formation d'organoïdes est simple (*figure 3*). Un support non résorbable et bien toléré par l'organisme, le polytétrafluoréthylène (PTFE, Gore-Tex) est utilisé sous la forme de fins filaments non tissés (« cheveux d'anges ») ; il est incubé dans une solution de collagène de manière à recouvrir les fibres textiles de cette protéine de la matrice extracellulaire, puis dans une solution d'HBGF-1/FGF acide. Ce dernier facteur de croissance est si puissant que quelques centaines de nanogrammes suffisent. La mixture est alors implantée dans la cavité péritonéale de rats. En quelques semaines, de complexes structures vasculaires (artériolaires, capillaires, lymphatiques et veineuses) se développent au sein d'un mésenchyme abondant qui semble aussi contenir, fait plus surprenant, des formations nerveuses. Selon le site d'implantation, l'organoïde, qui semble parfaitement viable aussi longtemps que dure l'expérience (plus de six mois) échange de nombreuses connections vasculaires avec les tissus environnants, foie, rate,

péritoine, anses intestinales, etc. Des hépatocytes en suspensions injectés dans un organoïde constitué s'y établissent et restent fonctionnels durant toute la période d'examen, c'est-à-dire six mois. Ils sont capables de provoquer une diminution de plus de moitié de la bilirubine sérique de rats *Gunn*.

Thérapie génique de maladies génétiques et non génétiques

Les indications de la thérapie génique somatique par autogreffe de cellules génétiquement modifiées incluent des maladies génétiques et des affections acquises. Les maladies génétiques les plus accessibles à ce traitement (*Tableau III*) sont celles qui sont secondaires au déficit d'une protéine circulante (maladies de la coagulation, qu'ils s'agissent des hémophilies ou des maladies thrombotiques congénitales [9], déficits hormonaux congénitaux, etc.) et celles dans lesquelles un organe, le foie par exemple est la cause de la maladie, mais non sa victime : oxalose ; maladie de Crigler-Najjar, équivalent humain du déficit du rat *Gunn* ; hypercholestérolémie familiale ; anomalies du cycle de l'urée ou de la néoglucogénèse, etc. [10]. Un déficit hétérozygote en α -1 antitrypsine pourrait être une indication de thérapie génique selon ces méthodes, afin d'éviter la destruction du tissu pulmonaire et l'emphysème. En revanche, parce qu'il s'accompagne d'un risque important de cirrhose et d'hépatocarcinome par surcharge en la protéine mutée, un déficit homozygote de phénotype Z resterait, pour l'instant, une indication de la greffe hépatique (*m/s n° 3, vol. 3, p. 181 et n° 3, vol. 5, p. 182*).

Les maladies non génétiques que l'on peut envisager de traiter par thérapie génique sont, en théorie, nombreuses. Les cellules génétiquement modifiées sont ici considérées comme des mini-usines de production et de distribution d'une substance d'intérêt biologique : facteurs thrombolytiques libérés par des cellules endothéliales tapissant un grefon vasculaire ; anticorps monoclonal « humanisé » (c'est-à-dire possédant des régions constantes humaines) dans une immunothéra-

pie passive anti-infectieuse ou anti-tumorale ; dérivés solubles de CD4 chez les malades atteints de SIDA, facteurs neurotrophiques libérés par des cellules neuronales embryonnaires implantées dans le cerveau de sujets atteints de maladies neurodégénératives [11], etc.

Injection de particules virales

L'introduction directe d'un gène correcteur dans l'organisme pourrait permettre d'ajouter à la liste des candidats à une thérapie génique les malades atteints d'anomalies héréditaires nécessitant d'être corrigées dans une très large proportion de leurs cellules. C'est le cas, par exemple, des mucoviscidosiques ou des myopathes. Cette approche permettrait en outre un traitement très précoce parfois rendu nécessaire par l'apparition de symptômes dès la période néonatale. La thérapie génique directe se heurte malheureusement à deux difficultés majeures. La première est celle déjà soulevée des risques d'activation d'oncogènes par l'insertion des séquences introduites. La recombinaison homologue, qui à ce jour n'est pas contrôlée mais seulement repérable *ex vivo*, ne peut ici être utilisée pour écarter ce danger. La recherche de vecteurs viraux performants mais présentant le minimum de risques pathogènes est donc particulièrement importante pour rendre applicable cette stratégie. A côté des rétrovirus délétés de leurs séquences régulatrices, les adénovirus peuvent être considérés comme de bons candidats [12]. Il n'est cependant pas exclu qu'ils puissent occasionner des symptomatologies plus sévères que la classique rhinopharyngite. La seconde difficulté majeure de cette approche est que les cellules germinales ne sont pas à l'abri du traitement. Même l'absence de tropisme d'un virus donné vis-à-vis des spermatozoïdes et des ovules n'écarterait pas tout danger car l'on sait bien que le tropisme est plus lié à la possibilité d'expression que de pénétration du virus dans une cellule. Or le danger est bien dans la pénétration et l'intégration dans le *germen* de cet ADN viral recombinant qui pourrait modifier l'espèce. Par conséquent, s'il s'avérait nécessaire, dans quelques cas

particuliers, d'opter pour cette stratégie, il faudrait peut-être envisager d'adjoindre un traitement contraceptif définitif. Il faut mentionner qu'une telle contraception est en fait déjà associée à des traitements de maladies graves (chimiothérapies anticancéreuses en particulier, cancers du sein hormono-dépendants, certains cancers du testicule).

A côté des vecteurs viraux, d'autres véhicules ont été envisagés pour introduire de l'ADN *in vivo*, mais les résultats sont restés décevants. Citons pour mémoire les liposomes, initialement développés pour tenter des thérapies enzymatiques. Ces véhicules sont potentiellement intéressants car ils devraient permettre de cibler l'ADN vers un type cellulaire donné et par conséquent d'éliminer tout risque d'atteinte des cellules germinales. Par exemple, les liposomes portant à leur surface des glycoprotéines dont les acides sialiques ont été hydrolysés peuvent être reconnus par les récepteurs hépatiques des asialoglycoprotéines et permettre un ciblage dans les hépatocytes [13]. Il n'existe cependant pas assez de résultats expérimentaux pour affirmer qu'ils constituent un outil adapté, efficace et sans conséquences délétères propres.

En conclusion, en partie à cause des difficultés rencontrées dans la mise au point des thérapies géniques de déficits héréditaires initialement supposés être de bons candidats à une telle approche thérapeutique, les chercheurs se sont tournés vers la recherche de transporteurs de gènes, de cellules cibles, et de modes de réintroduction dans l'organisme des cellules corrigées, entièrement nouveaux, tels les cellules endothéliales et les organoïdes.

Cette évolution des méthodes envisagées s'est accompagnée d'une dérive progressive des indications de la thérapie génique : alors que le traitement par ces moyens des maladies héréditaires les plus fréquentes et les plus graves (mucoviscidose, maladies de l'hémoglobine, myopathies, maladies lysosomales avec encéphalopathie, etc.) reste probablement encore très lointain, il est récemment apparu que nombre de maladies acquises pourraient bénéficier dans un avenir proche de l'introduction de cellules génétiquement modifiées ■

RÉFÉRENCES

1. Brinster RL, Braun RE, Lo D, Avarbock MR, Oram F, Palmiter R. Targeted correction of a major histocompatibility class II E α gene by DNA microinjected into mouse eggs. *Proc Natl Acad Sci USA* 1989 ; 86 : 7087-91.
2. Culliton BJ. Designing cells to deliver drugs. *Science* 1989 ; 246 : 746.
3. Bodine DM, Karlsson S, Nienhuis AW. Combination of interleukins 3 and 6 preserves stem cell function in culture and enhances retrovirus-mediated gene transfer into hematopoietic stem cells. *Proc Natl Acad Sci USA* 1989 ; 86 : 8897-901.
4. Correll PM, Fink JK, Brady RO, Perry LK, Karlsson S. Production of human glucocorticoid synthase in mice after retroviral gene transfer into multipotential hematopoietic progenitor cells. *Proc Natl Acad Sci USA* 1989 ; 86 : 8912-6.
5. Lim B, Apperley JF, Orkin SH, Williams DA. Long term expression of human adenosine deaminase in mice transplanted with retrovirus-infected hematopoietic stem cells. *Proc Natl Acad Sci USA* 1989 ; 86 : 8892-6.
6. Garver RI, Chiytil A, Courtney M, Crystal RG. Clonal gene therapy: transplanted mouse fibroblast clones express human α -1 antitrypsin gene *in vivo*. *Science*, 1989 ; 237 : 762-4.
7. Wilson JM, Johnston DE, Jefferson DM, Mulligan RC. Correction of the genetic defect in hepatocytes from the Watanabe heritable hyperlipemic rabbit. *Proc Natl Acad Sci USA* 1988 ; 85 : 4421-5.
8. Thompson JA, Haudenschild CC, Anderson KD, Dipietro JM, Anderson F, Maciag T. Heparin binding growth factor 1 induces the formation of organoid neovascular structures *in vivo*. *Proc Natl Acad Sci USA* 1989 ; 86 : 7928-32.
9. Boyer-Neumann C, Wolf M, Larrieu MJ. Les maladies thrombo-emboliques constitutionnelles. *médecine/sciences* 1987 ; 4 : 216-22.
10. Bernard O, Fischer A, Grünfeld JP. La transplantation d'organe comme traitement des maladies métaboliques héréditaires. *médecine/sciences* 1988 ; 7 : 406-12.
11. Wolff JA, Fisher LJ, Xu L *et al.* Grafting fibroblasts genetically modified to produce L-dopa in a rat model of Parkinson disease. *Proc Natl Acad Sci USA* 1989 ; 86 : 9011-4.
12. Chasse JF, Levrero M, Kamoun P, *et al.* L'adénovirus : vecteur de thérapie génique ? *médecine/sciences* 1989 ; 5 : 331-7.
13. Wu CH, Wilson JM, Wu GY. Targeting genes : delivery and persistent expression of a foreign gene driven by mammalian regulatory elements *in vivo*. *J Biol Chem* 1989 ; 264 : 16985-7.

Le locus des gènes β -globine et les perspectives de transfert de gènes dans les hémoglobinopathies

Certains types de thalassémie β et de syndrome de persistance de l'hémoglobine fœtale sont dus à des délétions emportant de larges régions d'ADN, dont une partie du « locus β », c'est-à-dire de la portion d'environ 60 kilobases (kpb) comportant tous les gènes de type β [1]: le gène embryonnaire ϵ , exprimé au niveau du sac vitellin; les gènes fœtaux γ ($A\gamma$ et $G\gamma$), exprimés dans le foie fœtal; les gènes adultes δ et β , exprimés dans les érythroblastes adultes; le pseudo-gène $\psi\beta$, enfin, non exprimé.

Dans deux types de thalassémie $\gamma\beta$ avec délétion, le type anglais et le type hollandais (figure 1) [2, 3], le gène β et ses régions bordantes sont intacts; dans la forme anglaise la délétion s'arrête même à environ 25 kpb en 5' du gène β . Quoique non modifié, le gène β ne s'exprime cependant pas chez ces malades et semble dans une conformation chromatinienne « inactive » (*m/s, suppl. au n° 7, vol. 2, p. 11*). En pratique, la conformation chromatinienne d'un gène est appréciée par son accessibilité à la

DNase I: lorsque la chromatine entourant un gène est dense, l'ADN est peu accessible à la DNase I, une nucléase, et est donc peu digéré par cette enzyme. En revanche, un gène actif est entouré d'une chromatine probablement moins dense qui le protège mal contre la digestion par la DNase I. La formation de complexes nucléoprotéiques dans le voisinage des gènes actifs (*m/s, n° 7, vol. 3, p. 428 et n° 8, vol. 3, p. 487*) crée, de plus, des contraintes dans le double brin, éventuellement la formation de boucles

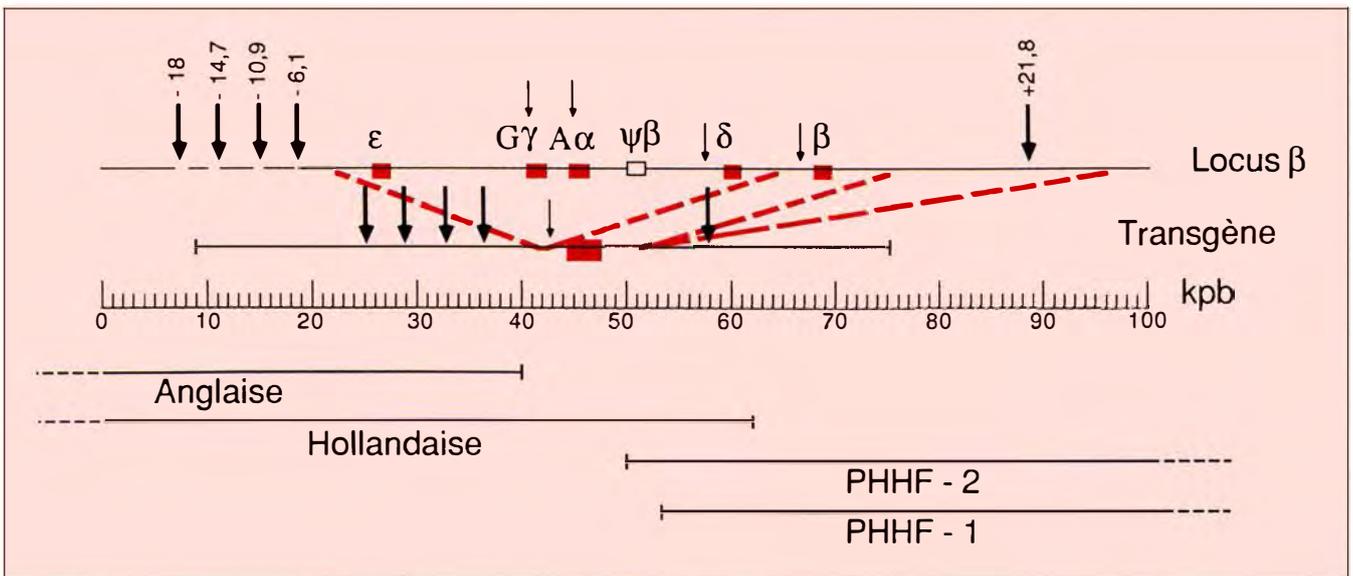


Figure 1. **Locus β et mini-locus utilisé comme transgène.** Rectangles rouges = gènes actifs; rectangle blanc = pseudogène; zones ombrées = délétions effectuées pour construire le transgène; flèches épaisses = sites d'hypersensibilité à la DNase I spécifiques des cellules érythroïdes mais ne changeant pas au cours du développement; flèches minces = sites d'hypersensibilité à la DNase I spécifiques des cellules érythroïdes et apparaissant parallèlement à l'expression des différents gènes. L'importance de différentes délétions rencontrées dans des syndromes thalassémiques (formes anglaise et hollandaise) et dans des syndromes de persistance héréditaire de l'hémoglobine fœtale (type 1 et 2) est matérialisée par les barres sous l'échelle des tailles en kilobases. Les limites 5' des délétions 5' et 3' des délétions 3' sont hors du locus β . Dans les thalassémies de type anglais et hollandais, les gènes non délétés ne sont pas exprimés et sont dans une conformation chromatinienne « inactive ». Au contraire, dans les persistances héréditaires de l'hémoglobine fœtale, les gènes intacts sont parfaitement actifs.

d'ADN exposant particulièrement une région à la digestion nucléasique ; ceci crée un site hypersensible, dont des exemples sont montrés sur la *figure 1*. Il existe, dans le locus β , différentes classes de sites hypersensibles à la DNase I, spécifiques de la différenciation érythroïde. Certains sont présents dès les premiers stades de l'érythropoïèse embryonnaire et persistent tout au long du développement et de la vie adulte ; ils sont situés entre 6 et 18 kpb en amont du gène ϵ (soit à plus de 45 kpb en amont du gène β), et à près de 22 kpb en aval du gène β [4, 6]. D'autres sites hypersensibles sont spécifiques du stade développement, apparaissant parallèlement à l'activation des gènes en amont desquels ils sont situés (*figure 1*). Il faut remarquer que dans les thalassémies de type anglais et hollandais, la région contenant les sites situés en 5' du locus β est déléetée, les autres sites n'apparaissant plus au cours du développement. Les délétions entraînant les syndromes de persistance héréditaire de l'hémoglobine foétale (PHHF) de type 1 et 2 intéressent les gènes δ et β [7] ; les gènes intacts, particulièrement ceux codant pour les chaînes γ (A γ et G γ) sont en revanche parfaitement actifs, les sites d'hypersensibilité à la DNase I des régions non déléetées étant présents et apparaissant selon la chronologie normale. La *figure 1* montre que le site hypersensible à la DNase I situé en aval du gène β , constant au cours du développement, est intéressé par les délétions entraînant les PHHF. Ces résultats suggèrent qu'existaient, loin en amont des gènes β , des régions dont la présence était indispensable à l'expression de tous les gènes du locus au cours de la différenciation érythroïde. Les sites hypersensibles indépendants du développement, décrits tout d'abord par D. Tuan [4], semblaient de bons candidats pour indiquer la localisation précise de ces éléments de contrôle régional.

Par ailleurs, en quelque sorte surajouté aux éléments discutés plus haut, chaque gène actif du locus possède ses propres éléments de contrôle. Le gène β , par exemple, a un promoteur fonctionnant mieux dans les érythroblastes adultes que dans d'autres cellules et un *enhancer* absolument spécifique de la différenciation érythroïde adulte, situé entre 600 et 800 pb en aval du gène. De plus, un motif du troisième exon, encore mal défini, coopérerait avec les éléments précédents pour stimuler l'expression du gène dans les cellules érythroïdes [6, 8]. Quoique ces régions bordantes situées de part et d'autre du gène β permettent, à elles seules, de commander une expression spécifique des cellules érythroïdes adultes, in vitro dans des cellules en culture [9-11] ou in vivo dans des souris transgéniques [6, 8, 11, 12], le niveau d'expression atteint dans les souris transgéniques est souvent très bas par comparaison à l'expression du gène endogène et varie considérablement d'une souris à l'autre, probablement en fonction du site d'intégration. Des concentrations appréciables d'hémoglobine issue du transgène ne sont atteintes que lorsque ce dernier est intégré en de multiples copies (*m/s n°1, vol. 3, p. 49*). Cette faiblesse d'expression, observée quelque soit le mode de création des animaux transgéniques [13, 14] constitue un obstacle qui semblait insurmontable à la tentative de traiter les malades souffrant d'hémoglobinopathies sévères (principalement de thalassémie) par transfert de gène dans les cellules médullaires. L'idée est venue à plusieurs laboratoires, associant ces données éparses, que la faiblesse de l'expression des transgènes β globine pouvait être la conséquence de l'absence des éléments de contrôles localisés loin en amont du gène β , au niveau des sites d'hypersensibilité indépendants du développement. Grosveld *et al.* [6] ont donc construit, dans un vecteur acceptant plus de 30 kpb

d'ADN exogène (un cosmide), un « mini locus β » représenté sur la *figure 1* ; les sites hypersensibles amont se trouvent rapprochés, grâce à une délétion, entre 6 et 20 kpb du gène β . Une délétion rapproche aussi le site hypersensible aval. Des souris transgéniques pour cette construction se comportent de façon remarquable : elles expriment toutes le transgène, proportionnellement au nombre de copies intégrées, chaque copie contribuant pour un niveau d'expression identique à celui du gène endogène ; l'expression est indépendante du site d'intégration. Ces résultats s'interprètent en supposant que, comme le suggéraient les thalassémies $\gamma\beta$ de type anglais et hollandais, la région des sites hypersensibles amont est indispensable à l'expression normale des gènes du locus β . Il pourrait s'agir d'un *enhancer* stimulé par des protéines communes à toutes les étapes de la différenciation érythroïde et indispensable à la modification de la chromatine de la région du locus β , modification elle-même requise pour que s'expriment, sous l'effet de protéines transactivatrices caractéristiques des différentes étapes de la différenciation érythroïde, les différents gènes du locus. Chez la drosophile, Gasser et Laemmli [15] ont étudié en détail les relations entre les gènes transcrits et l'association de l'ADN à la matrice nucléaire, c'est-à-dire au « squelette » du noyau. L'ADN semble être attaché à la matrice en des sites disposés environ toutes les 50 kpb ou plus [16], formant ainsi des boucles qui pourraient peut-être correspondre aux unités transcriptionnelles. Les sites d'attachement peuvent coïncider avec la localisation des *enhancers* et d'autres séquences de contrôle de la transcription. On peut donc faire l'hypothèse que les régions de contrôle en amont du locus β globine pourraient fonctionner comme un *enhancer* et (ou) un site de fixation à la matrice nucléaire, ces fonctions étant nécessaires

à l'acquisition par la chromatine d'une conformation permettant l'expression des gènes.

Quoiqu'il en soit des mécanismes d'action des séquences activatrices amont du locus β globine, leur découverte ouvre d'importantes perspectives pour la thérapeutique génique des hémoglobinopathies. L'équipe de R.C. Mulligan [14] a montré qu'il était possible de prélever la moelle d'une souris, de l'infecter par un rétrovirus recombinant contenant un gène de β globine humaine, puis de greffer cette moelle à l'animal rendu aplasique par chimiothérapie ou irradiation. Les cellules souches infectées par le rétrovirus recombiné permettent de repeupler la moelle, notamment en cellules de la lignée érythroïde synthétisant de la globine β humaine. Le niveau d'expression obtenu par Mulligan reste faible, mais sa construction ne contenait pas la région des sites hypersensibles amonts. S'il était possible de mieux délimiter les éléments actifs de cette région, et ainsi de les intégrer à proximité d'un gène β dans une construction rétrovirale qui ne peut guère admettre plus de 7 kpb d'ADN exogène..., nous ne serions peut-être pas éloigné des premiers essais thérapeutiques.

Axel Kahn

RÉFÉRENCES

- Maniatis T, Fritsch EF, Lauer J, Lawn RM. Molecular genetics of human hemoglobins. *Ann Rev Genet* 1981 ; 14 : 145-78.
- Curtin P, Pirastu M, Kan YW, Gobert-Jones JA, Stephens AD, Lehman H. A distant gene deletion affects β -globin gene function in an atypical $\gamma\delta\beta$ thalassemia. *J Clin Invest* 1985 ; 76 : 1554.
- Kioussis D, Vanin E, de Lange T, Flavell RA, Grosveld F. β -globin gene inactivation by DNA translocation in $\gamma\beta$ thalassemia. *Nature* 1983 ; 306 : 662-6.

- Tuan D, Solomon W, Quiliang LS, Irving ML. The « β -like-globin» gene domain in human erythroid cells. *Proc Natl Acad Sci USA* 1985 ; 32 : 6384-8.

- Forrester WC, Tagekawa S, Papayannopoulou T, Stamatoyannopoulos G, Groudine M. Evidence for a locus activation region : the formation of developmentally stable hypersensitive sites in globin-expressing hybrids. *Nucleic Acids Res* 1987 ; 15 : 10159-77.

- Grosveld F, Van Assendelft GB, Greaves DR, Kollias G. Position independent, high level expression of the human β -globin gene in transgenic mice. *Cell* 1987 ; 51 : 975-85.

- Vanin EF, Henthorn PS, Kioussis D, Grosveld F, Smithies O. Unexpected relationships between four large deletions in the human β -globin gene cluster. *Cell* 1983 ; 35 : 701-9.

- Behringer RD, Hammer RE, Brinster R, Palmiter RD, Townes TM. Two 3' sequences direct adult erythroid specific expression of human β -globin genes in transgenic mice. *Proc Natl Acad Sci USA* 1987 ; 84 : 7056-60.

- Charnay P, Treisman R, Mellon P, Chao M, Axel R, Maniatis T. Differences in human α and β -globin gene expression in mouse erythroleukemia cells : the role of intragenic sequences. *Cell* 1984 ; 38 : 251-63.

- Cone RD, Benarous-Weber A, Baorto D, Mulligan RC. Regulated expression of a complete human β -globin gene encoded by a transmissible retrovirus vector. *Mol Cell Biol* 1987 ; 7 : 887-97.

- Wright S, Rosenthal A, Flavell RA, Grosveld F. DNA sequences required for regulated expression of β -globin genes in murine erythroleukemia cells. *Cell* 1984 ; 38 : 265-73.

- Costantini F, Lacy E. Introduction of a rabbit β -globin gene into the mouse germ line. *Nature* 1981 ; 294 : 92-4.

- Costantini F, Radice G, Magram J, Stamatoyannopoulos G, Papayannopoulou T, Chada K. Developmental regulation of human globin genes in transgenic mice. Cold Spring Harbor Symposium. *Quant Biol* 1985 ; 50 : 361-70.

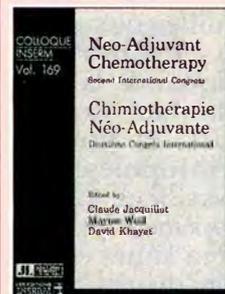
- Dzierzak EA, Papayannopoulou T, Mulligan RC. Lineage-specific expression of a human β -globin gene in murine bone marrow transplant recipients reconstituted with retrovirus transduced stem cells. *Nature* 1988 ; 331 : 35-41.

- Gasser SM, Laemmli UK. Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D melanogaster*. *Cell* 1986 ; 46 : 521-30.

- Collins FS, Cole JL, Lockwood WK, Ianuzzi MC. The deletion in both common types of hereditary persistence of fetal hemoglobin is approximately 105 kilobases. *Blood* 1987 ; 70 : 1797-803.

John Libbey
EUROTEXT
LONDON · PARIS

Paris · Londres
EDITIONS MEDECINE-SCIENCES



**Chimio-
thérapie
Néo-
Adjuvante**

C. JACQUILLAT
M. WEIL
D. KHAYAT

**Co-édition John Libbey
Eurotext/INSERM**

- Associations Radiothérapie/Chimiothérapie
- Cancer du sein
- Voies aérodigestives
- Cancers du poumon
- Urologie
- Tumeurs de l'enfant sarcomes des os et des tissus mous
- Gynécologie
- Biologie des métastases, chimiorésistance
- Traitements locorégionaux
- Nouveaux développements dans le mélanome
- THP-adriamycine (pirarubicine)
- Carboplatine (paraplatine)
- Lymphokines et autres outils immunologiques
- Nausées, vomissements et chimiothérapie
- Mitomycine C

Novembre 1988, broché
vol. 169, 1 100 pages
780 FF, 66 £, 130 \$ US

6, rue Blanche, 92120 Montrouge
Tél. : (1) 47.35.85.52

Souris transgéniques et modèles du diabète

Le diabète insulino-dépendant de type I est un désordre métabolique résultant de l'absence de sécrétion endogène d'insuline [1]. Des mécanismes auto-immuns semblent impliqués dans l'étiopathogénie de ce type de diabète aussi bien chez l'homme [2] que dans les modèles animaux que constituent les souris NOD (*non-obese diabetic*) et les rats BB. Les arguments en faveur de ces mécanismes sont l'apparition d'une auto-immunité cellulaire et humorale caractérisée par une infiltration leucocytaire des îlots du pancréas et une production d'auto-anticorps dirigés contre les cellules β . La façon

dont procède le système immunitaire pour distinguer le soi du non-soi, et pour maintenir un état de tolérance à l'égard des antigènes du soi, n'est pas complètement comprise mais l'on sait qu'il fait intervenir les molécules de classe II du complexe majeur d'histocompatibilité (CMH: H2 chez la souris, HLA chez l'homme) (voir lexique *m/s* n° 4, vol. 4, p. 118). L'intervention du locus H2-I (figure 1) dans la prédisposition au diabète caractéristique de la lignée de souris NOD vient, grâce à la technologie des souris transgéniques* (*note et figure 2, p. 445*), d'être démontrée [3]. Des études

génétiques avaient permis d'affirmer que la susceptibilité au diabète est sous le contrôle d'au moins trois gènes dont l'un, localisé sur le chromosome 17, est étroitement lié au locus H2-K [4]. L'on savait par ailleurs que les souris NOD n'expriment pas l'antigène I-E. Des croisements effectués entre des souris NOD et des souris transgéniques C57BL/6 exprimant $E\alpha$ permettent d'obtenir des souris NOD $E\alpha+$ chez lesquelles aucune autre modification du chromosome 17 n'a été introduite. Ainsi le fait que ces souris ne développent pas de diabète peut être expliqué par l'expression des molécules I-E dont

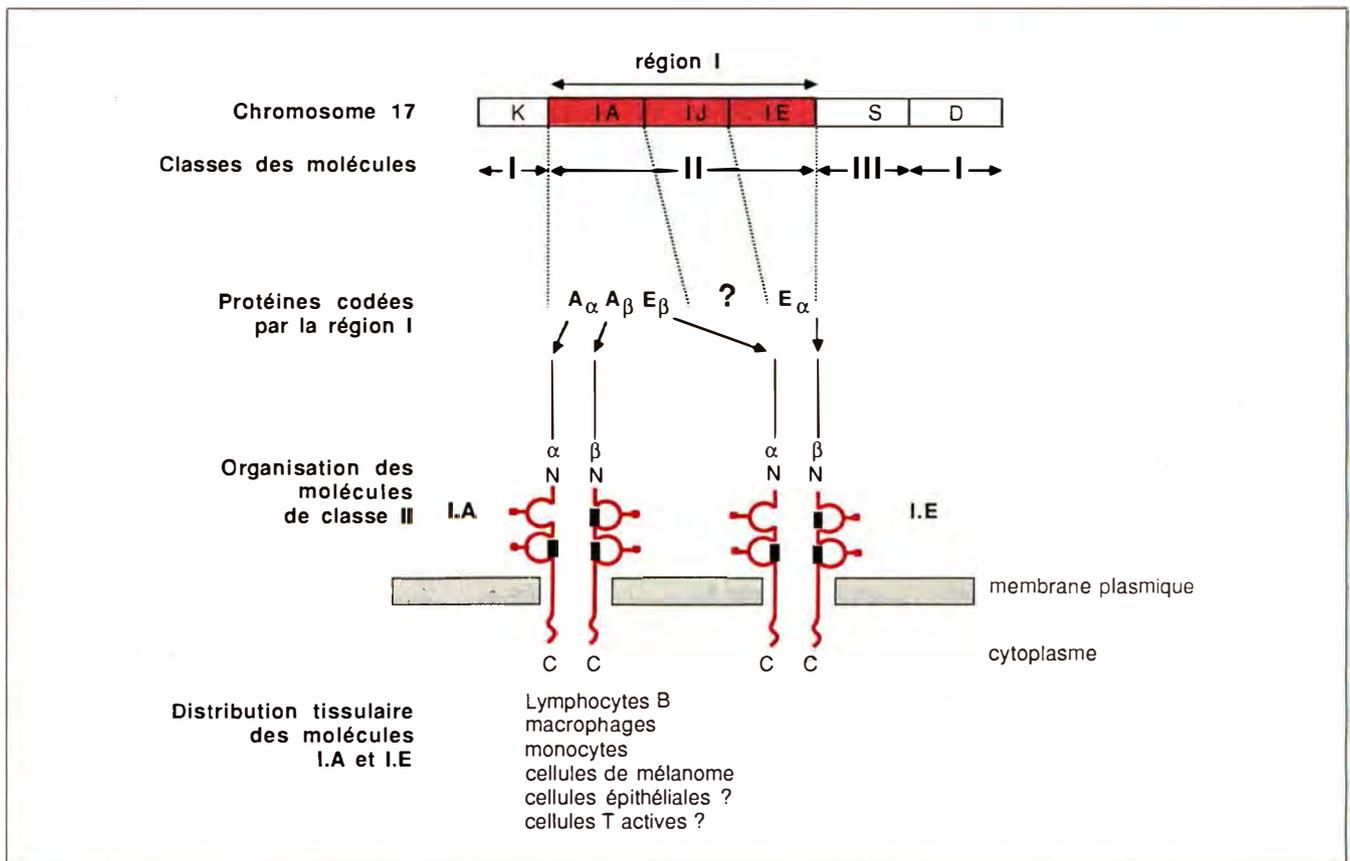


Figure 1. Le complexe H-2 de la souris. - : pont disulfure; • : sucre (CHO).

dans la réponse immunitaire aux polymères synthétiques vient par ailleurs d'être clairement démontrée [5-7]. Les conséquences de l'expression des molécules I-E chez les souris NOD permettent de proposer divers mécanismes pour expliquer leur rôle éventuel dans la mise en place d'une tolérance immunitaire: (1) elles pourraient induire la production de cellules suppressives qui bloqueraient le développement d'une réaction auto-immune vis-à-vis des cellules β ; (2) elles pourraient présenter des épitopes identiques à des antigènes présents à la surface de la cellule β . La tolérance vis-à-vis de ces antigènes serait acquise via celle développée précocément vis-à-vis des molécules I-E. L'absence de molécule I-E chez les souris NOD et une expression de l'antigène spécifique des cellules β , postérieure aux possibilités d'acquisition d'une tolérance immunitaire, serait à l'origine du processus auto-immun. Il faut noter que seules les cellules β semblent touchées par cette réaction. Ce serait donc la conjonction d'un antigène spécifique du pancréas et d'une absence de molécules I-E qui provoquerait la non-tolérance vis-à-vis de cet antigène du soi. A l'inverse, dans le modèle des rats diabétiques BB comme dans les diabètes auto-immuns humains, il n'y a pas défaut mais expression anormalement localisée dans le pancréas, de molécules de classe I et de classe II. Existe-t-il une relation de cause à effet entre l'expression des molécules du CMH

dans les cellules β du pancréas et l'apparition du diabète auto-immun caractérisé par l'infiltration lymphocytaire? En d'autres termes est-ce l'expression aberrante de molécules du CMH et la présentation d'antigènes de la cellule β qui, stimulant des cellules T spécifiques, entraîne l'in-

filtration lymphocytaire ou bien l'infiltration précède-t-elle l'apparition des molécules du CMH? Trois groupes ont très récemment choisi la technologie des souris transgéniques pour résoudre ce problème. Différentes molécules de classe I (H-2K^b) ou de classe II (I-E^b et I-A^d) [8-11]

* La technique des souris transgéniques fut présentée dans un numéro précédent de médecine/science par C. Babinet [14]. Elle permet d'obtenir des animaux qui possèdent, en plus du matériel héréditaire qui leur est propre, un gène supplémentaire ou « transgène » qui a été, le plus souvent par micro-injection dans l'œuf fécondé, ajouté à leur patrimoine héréditaire. Comme il est possible de construire des séquences chimériques composées d'un promoteur A, capable de diriger l'expression d'un gène dans le tissu A, couplées à une séquence codant pour une protéine B, l'on peut obtenir des animaux transgéniques exprimant spécifiquement la protéine B dans le tissu A. Un nouveau caractère héréditaire est ainsi conféré à la descendance de l'animal transgénique.

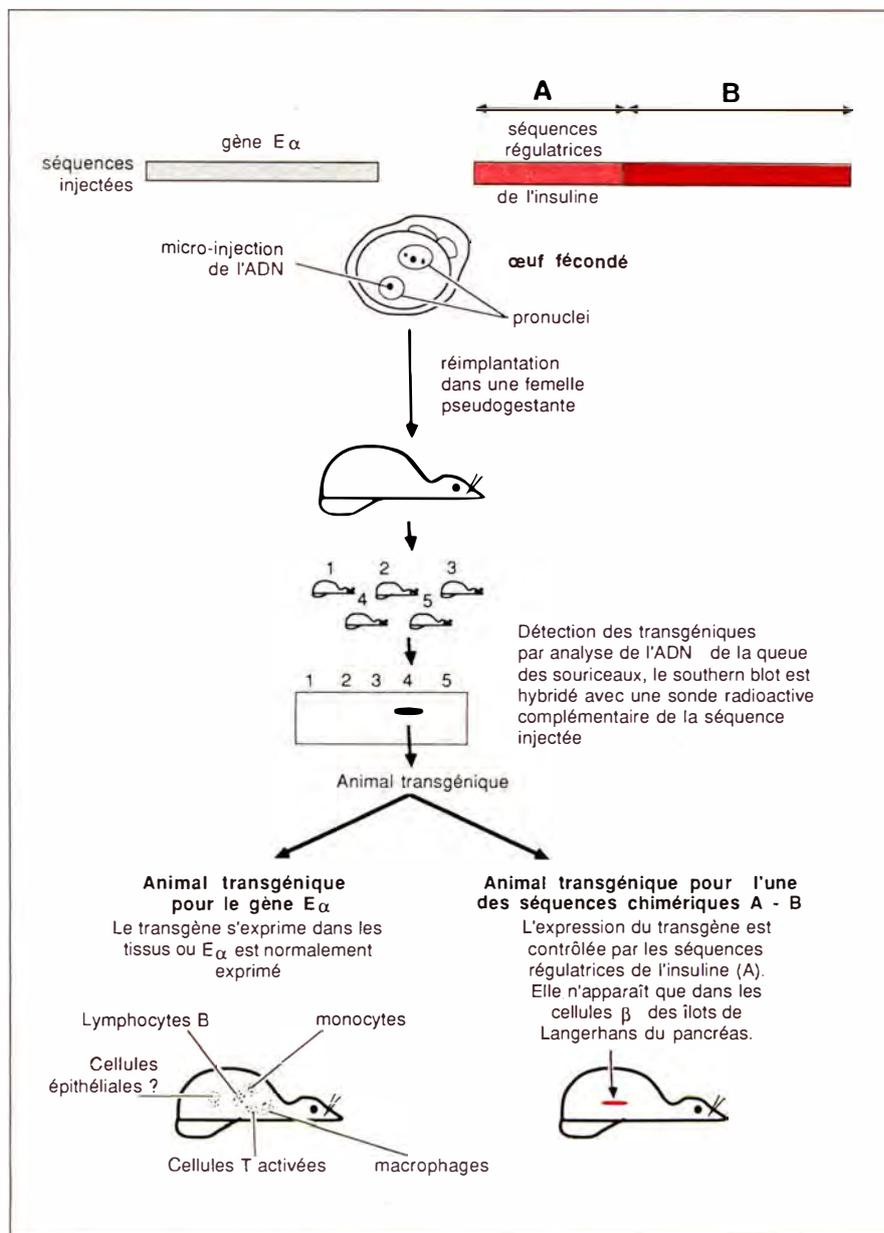


Figure 2. La technologie des souris transgéniques. Micro-injection de gènes (E α) ou de séquences chimériques (séquences régulatrices d'un gène A — l'insuline dans les cas décrits — couplées au gène B-T de SV40, I-A, I-E ou H-2k.

furent placées sous le contrôle du promoteur de l'insuline et micro-injectées dans des embryons de souris. Une très forte expression de ces molécules fut obtenue dans les cellules β du pancréas et comme l'attendaient les auteurs un diabète se développa chez les animaux transgéniques. Ces modèles n'ont cependant pas les caractéristiques escomptées. En effet, si à des degrés divers les cellules β dégèrent, aucun élément permettant de rattacher cette dégénérescence à un mécanisme auto-immun n'est observé. En particulier, il n'y a pas d'infiltration lymphocytaire du pancréas. De plus une tolérance immunitaire vis-à-vis des molécules H2 et I-E est rapportée. Par conséquent, l'expression aberrante des molécules du CMH dans les cellules β semble à elle seule suffisante pour induire un diabète dont le mécanisme ne met pas en jeu les cellules T. La tolérance vis-à-vis des molécules codées par les transgènes pourrait résulter d'une expression précoce au cours du développement qui les fait considérer comme des protéines du soi. A l'appui de cette hypothèse il faut citer le résultat récent obtenu par Adams *et al.* [12]. Des lignées de souris transgéniques exprimant l'antigène T de SV 40 dans les cellules β des îlots de Langerhans ont été obtenues par micro-injection du gène T mis sous contrôle des séquences régulatrices du gène de l'insuline. L'antigène T est naturellement considéré, dans la plupart de ces lignées, comme une protéine du soi et n'induit par conséquent aucune réaction immunitaire. De façon tout à fait inattendue, les auteurs virent apparaître dans l'une de leurs lignées transgéniques, des anticorps anti-T circulant. Ils démontrèrent que, chez ces animaux, pour une raison non élucidée, l'expression de T était retardée par rapport à celle observée chez les animaux transgéniques sans anticorps anti-T. On peut donc suspecter que ce simple retard dans l'apparition d'un antigène à la surface d'une cellule entraîne l'absence d'établissement d'une tolérance immunologique à son égard. Dans le diabète juvénile insulino-dépendant, des auto-anticorps reconnaissent un

antigène présent à la surface des cellules β . Il est possible, conformément à l'expérience précédemment décrite, qu'une expression retardée de cette protéine, au-delà de la période permettant le développement de la tolérance immunologique, soit à l'origine de l'atteinte auto-immune. Cependant, dans la plupart des cas, les molécules codées par les transgènes n'induisent pas de réaction auto-immune et l'apparition du diabète chez les animaux portant les transgènes I-A^d, I-E^b et H-2K^b doit avoir une autre origine. Il faut à ce stade noter que les productions aberrantes des molécules I-E^b, I-A^d et H-2K^b dans le pancréas, si elles entraînent toutes un diabète, n'ont pas les mêmes conséquences histologiques. Dans le premier cas, le pancréas est intact alors que dans les deux derniers le nombre d'îlots est à divers degrés réduit. Dans tous les cas la production d'insuline est diminuée. L'hypothèse la plus probable pour expliquer cette diminution est qu'il existe une interaction entre les molécules du CMH et les mécanismes de synthèse, maturation ou sécrétion de l'insuline. Cela peut impliquer une compétition quant à l'utilisation des protéines et autres composants nécessaires à ces opérations ou une interaction directe entre l'insuline et les molécules du CMH. On peut faire par exemple l'hypothèse que la pré-pro-insuline se fixe au niveau du sillon des molécules du CMH dont le rôle est de présenter les fragments peptidiques de l'antigène [13]. Cette interaction pourrait perturber le repliement de la molécule dans le réticulum endoplasmique ou l'appareil de Golgi ou bien encore inhiber la protéase qui libère le peptide C. La destruction progressive des cellules β pourrait résulter d'une destruction due à l'accumulation de peptides dérivés de l'insuline non excrétée. Un autre modèle de souris diabétique a été obtenu par micro-injection d'une séquence codant pour l'interféron γ (IFN), placée sous le contrôle du promoteur de l'insuline [10], l'objectif étant d'induire une expression des molécules de classe II équivalente à celle que l'on observe après certaines infections virales. Dans ce cas

un diabète se développe avec une réaction inflammatoire importante dont l'origine peut être rapprochée de la réaction auto-immune décrite par Adams *et al.* [12]. Les molécules du CMH induites par l'IFN pourraient dans ce cas être produites plus tardivement que dans les modèles précédents (peut être du fait de l'expression tardive dans l'ontogénèse des récepteurs de l'IFN), ou bien l'interféron pourrait induire la synthèse d'un peptide qui serait présenté à la surface des cellules β par les molécules de classe II, elles-mêmes induites par l'IFN. Ainsi les souris transgéniques IFN constituent, outre un modèle de diabète, un modèle de l'inflammation créée par l'interféron permettant de comprendre la cascade des réactions induites par la production locale de cette substance. En conclusion, les modèles obtenus ne sont pas d'exactes copies phénotypiques des diabètes humains insulino-dépendants. Ils permettent d'éliminer l'hypothèse selon laquelle l'expression dans les cellules β des molécules de classe II puisse, quel que soit le stade ontogénique où elle se produit, déclencher une réaction auto-immune. En revanche, ils mettent en évidence leur intervention directe dans la perturbation de la production d'insuline et probablement leur rôle dans la réaction auto-immune lorsque leur expression est tardive ou lorsque leur présence autorise la présentation de peptides produits tardivement au cours de l'ontogénèse.

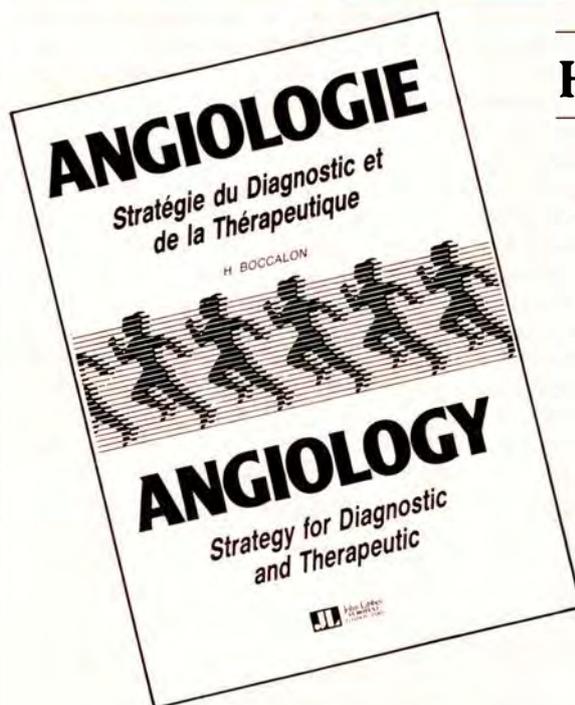
Pascale Briand

RÉFÉRENCES

1. Eisenbarth GS. Type I diabetes mellitus : a chronic autoimmune disease. *N Engl J Med* 1986 ; 314 : 1360-8.
2. Gepts W, Lecompte PM. The pathology of type I (juvenile) diabetes. In : Volk BW, Arguilla ER, *The Diabetic Pancreas*. New York : Plenum, 1985.
3. Nishimoto H, Kikutani H, Yamasura K, Kishimoto T. Prevention of autoimmune insulinitis by expression of I-E molecules in NOD mice. *Nature* 1987 ; 432-4.

RÉFÉRENCES

4. Prochazka M, Leiter EH, Serreze DV, Coleman DL. Three recessive loci required for insulin-dependent diabetes in non obese diabetic mice. *Science* 1987 ; 237 : 286-8.
5. Le Meur M, Gerlinger P, Benoist C, Mathis D. Correcting an immune response deficiency by creating E_α gene transgenic mice. *Nature* 1985 ; 316 : 38-42.
6. Yamamura KI, Kikutani H, Folsom V, et al. Functional expression of a microinjected E_α gene in C57BL/6 transgenic mice. *Nature* 1985 ; 316 : 67-9.
7. Pinkert CA, Widera G, Cowing C, et al. Tissue-specific, inducible and functional expression of the E_α MHC class II gene in transgenic mice. *EMBO J* 1985 ; 4 : 2225-30.
8. Allison J, Campbell IL, Morahan C, Mandel TE, Harrison CL, Miller JF. Diabetes in transgenic mice resulting from over-expression of class I histocompatibility molecules in pancreatic B cells. *Nature* 1988 ; 333 : 529-33.
9. Lo D, Burkly LC, Widera G, et al. Diabetes and tolerance in transgenic mice expressing class II MHC molecules in pancreatic B cells. *Cell* 1988 ; 53 : 159-68.
10. Sarvetnick N, Liggitt D, Pitts SL, Hansen SE, Stewart TA. Insulin-dependent diabetes mellitus induced in transgenic mice by ectopic expression of class II MHC and interferon-gamma. *Cell* 1988 ; 52 : 773-82.
11. Parham P. Intolerable secretion in tolerant transgenic mice. *Nature* 1988 ; 333 : 500-3.
12. Adams TE, Alpert S, Hanahan D. Non-tolerance and autoantibodies to a transgenic self antigen expressed in pancreatic β cells. *Nature* 1987 ; 325 : 223-8.
13. Kahn A. La troisième dimension pour les molécules de classe I du complexe majeur d'histocompatibilité : un cristal de HLA-A2. *médecine/sciences* 1988 ; 4 : 52-4.
14. Babinet C, Morello D. Animaux transgéniques : une voie nouvelle pour l'étude du développement. *médecine/sciences* 1986 ; 2 : 253-9.



H. BOCCALON

1988, broché, 696 pages
Complété par un supplément
de 196 pages

450 FF, 38 £, 74 \$ US

Un ouvrage qui s'adresse aux spécialistes
en Angiologie médico-chirurgicale

**Le point sur la stratégie du diagnostic
et de la thérapeutique en Angiologie**

- Ultrasonologie (Doppler transcrânien)
- Échographie veineuse
- Angiologie interventionnelle
- Stimulation épidurale
- Système expert en Angiologie

Les données les plus récentes en physiologie

- Comportement cardio-vasculaire dans l'espace

BON DE COMMANDE

NOM _____ Prénom _____

Adresse _____

Désire recevoir l'ouvrage de H. BOCCALON, au prix de 450 FF + 30 F de frais de port, soit 480 FF.

Ci-joint mon règlement à l'ordre de **John Libbey Eurotext**
6, rue Blanche - 92120 MONTRouGE - France - Tél. : (1) 47.35.85.52

Génétique moléculaire et sauce tomate

Les perspectives nouvelles des ARN antisens en génétique moléculaire végétale

Les industriels de l'agro-alimentaire se trouvent devant un redoutable dilemme : soit vendre des tomates mûres et savoureuses à souhait... dont plus de la moitié sera écrasée au cours des manipulations et du transport, soit proposer des fruits verts, durs, totalement insipides... mais faciles à conditionner. Il semble que ce soit cette dernière solution qui ait été adoptée généralement aux USA [1]. Le problème vient de ce que la maturation de la tomate s'accompagne d'une synthèse importante de l'enzyme polygalacturonase, qui digère les parois des cellules végétales, effondre les travées assurant la rigidité de l'architecture du fruit et aboutit à cet objet mou ayant une tendance naturelle à s'écraser, et pas seulement sur le visage des acteurs, des chanteurs... ou des premiers ministres contestés. Jusqu'à présent, la solution à un seul de ces problèmes avait été trouvée : le fruit vert traité à l'éthylène rougit... ce qui ne modifie en rien ses piètres qualités gustatives.

La solution génétique consisterait à dissocier la maturation du fruit de l'hyperactivité de l'enzyme responsable de son ramollissement. Cela est maintenant possible — et en partie testé par deux équipes anglaises — grâce à l'utilisation d'ARN antisens complémentaire du messenger de la polygalacturonase [2]. La tomate étant un dicotylédone, le système classique de transformation par le plasmide Ti de la bactérie du sol *Agrobacterium tumefaciens* fonctionne parfaitement. Des plantes transgéniques ont par conséquent été créées par introduction d'un transgène constitué de l'ADN codant pour l'ARN antisens, contrôlé par le promoteur fort du gène de l'ARN 35S du virus de la mosaïque du chou-fleur et terminé par l'extrémité 3' du gène de la nopaline synthase. Il s'agit là d'une construction parfaitement classique en génétique végétale... hormis le fait qu'elle comporte un

ADN transcrit en un ARN antisens. Les tomates transgéniques ont une réduction de 90 % de l'activité de la polygalacturonase, associée à une diminution du même ordre de la concentration du messenger. Le mécanisme de cette baisse de la quantité de messenger peut être une interférence de l'ARN antisens avec la maturation nucléaire des transcrits primaires, ou une instabilité du messenger cytoplasmique partiellement double brin.

L'intérêt de cette expérience dépasse le douloureux problème de la saveur des tomates ; il est en fait de confirmer l'efficacité dans les plantes des ARN antisens, déjà démontrée de façon spectaculaire dans les pétunias par une équipe néerlandaise. L'expression, dans des pétunias transgéniques, issus d'une fleur parentale rouge, d'un ARN antisens complémentaire du messenger de la chalcone synthase (enzyme principale de la synthèse des flavonoïdes, responsables de la couleur de la fleur) aboutit à des fleurs allant du rose au blanc [3, 4]. Des succès peuvent aussi être obtenus avec le tabac.

Les raisons de ces succès répétés, alors qu'une même approche s'est révélée peu efficace chez les souris, restent énigmatiques. On peut supposer que le promoteur végétal utilisé est très fort. Peut-être existe-t-il également dans les plantes un système nucléasique particulièrement efficace pour dégrader l'ARN double brin. Quoiqu'il en soit, nul doute que cette stratégie, utilisant les ARN antisens ou les ribozymes naturellement actifs dans les plantes (*m/s n° 8, vol. 4, p. 522*), est appelée à connaître un impérieux développement en recherche végétale fondamentale aussi bien qu'en biotechnologie lorsqu'il s'agira d'inhiber ou de freiner l'expression d'un gène aux effets défavorables. Ainsi peut-on imaginer des plantes pratiquement imputrescibles, du café sans caféine et du tabac sans nicotine ! Des ARN

antisens antiviraux pourraient aussi protéger les plantes contre une grande variété d'infections.

A. K

1. Roberts L. Genetic engineers build a better tomato. *Science* 1988 ; 241 : 1290.
2. Smith CJS, Watson CF, Ray J, *et al.* Anti-sense RNA inhibition of polygalacturonase gene expression in transgenic tomatoes. *Nature* 1988 ; 334 : 724-6.
3. Van der Krol AR, Lenting PE, Veenstra J, *et al.* An anti-sense chalcone synthase gene in transgenic plants inhibits flower pigmentation. *Nature* 1988 ; 333 : 866-9.
4. Lichenstein C. Anti-sense RNA as a tool to study plant gene expression. *Nature* 1988 ; 333 : 801-2.

ms

à paraître

Avril 90

n° 4, vol. 6

ONCOGÈNES

Éditorial de
Pierre Tambourin

Mécanismes moléculaires de la recombinaison homologue

La recombinaison homologue est un phénomène d'échange de matériel entre deux molécules d'ADN de séquence très voisine au cours duquel aucun nucléotide n'est perdu ni ajouté.

Dans un premier temps, les molécules mises en jeu s'apparient base à base dans la région d'homologie.

La molécule intermédiaire, qui réunit deux double brins d'ADN recombinants, contient quatre brins d'ADN échangeant leurs appariements au point de contact. Cette structure en brins croisés est appelée « structure de Holliday » (figure IA). L'existence de cette espèce moléculaire a été suggérée par la mise en évidence de régions de double brins hétérologues, formées de deux monobrins qui ne sont pas exactement complémentaires, dans les deux molécules d'ADN issues d'une recombinaison.

Un modèle moléculaire a permis de mettre en évidence deux propriétés physiques de la structure de Holliday : la migration des branches et l'isomérisation.

La rotation du squelette de la molécule permet aux monobrins croisés de passer librement d'une hélice d'ADN à l'autre. Le point d'intersection entre les deux hélices peut donc se déplacer aisément grâce à un mécanisme de fermeture à glissière au cours duquel des bases équivalentes des deux molécules originelles sont échangées : c'est le phénomène de « migration des branches » ou de « diffusion rotatoire » (figure IB). Une région d'une molécule d'ADN donnée peut donc se déplacer d'une hélice à l'autre en produisant des double brins hétérologues plus ou moins étendus. Des brins peuvent ainsi être échangés entre deux double brins homologues d'ADN à une certaine distance de la cassure originelle.

La seconde propriété physique de la structure de Holliday est son aptitude à subir une isomérisation (figure IIA). Un simple réarrangement stérique (deux rotations successives de 180° d'une moitié du complexe moléculaire autour du point de croisement) place les brins croisés « à l'extérieur » de la molécule et croise les brins non recombinés (figure IIB).

L'ultime étape de la recombinaison homologue est la résolution de la structure de Holliday en deux molécules double brins. Ce phénomène met en jeu des nucléases et engendre deux types de paires de double brins : des double brins non recombinants qui n'ont échangé qu'une région monobrin (sans isomérisation de la structure de Holliday) et des double brins recombinants qui résultent d'un échange des quatre brins (après isomérisation de la structure de Holliday).

Des études menées chez la levure ont montré que la recombinaison n'est initiée que si une des deux hélices d'ADN présente une discontinuité : cassure ou région de séparation des deux brins. Deux modèles ont été proposés, toujours chez la levure, afin d'expliquer la for-

mation de la structure de Holliday. Le premier modèle a été établi par Meselson et Radding en 1975 et suggère que la recombinaison est initiée par une cassure simple brin de l'un ou l'autre des double brins impliqués (figure III). L'extrémité 3' du brin interrompu est utilisée comme amorce de réplication par l'ADN polymérase. Cette néosynthèse implique un déplacement de l'extrémité 5' du brin, qui envahit alors l'autre double brin dans la région d'homologie et donne naissance à une structure tricaténaire (A). Le brin envahisseur du double brin donneur déplace le brin homologue du double brin receveur, et la région double brin hétérologue est étendue grâce à la synthèse du brin donneur, concomitante de la dégradation du brin homologue déplacé du double brin receveur (B). La structure de Holliday est engendrée par ligation des extrémités 3' du brin néosynthétisé et 5' du brin dégradé (C, D, E).

Le second modèle a été proposé par Szostak et al. en 1983 (figure IV). Il implique la double cassure d'un des deux double brins et la création, par des exonucléases, d'une délétion bordée d'extrémités 3' simple brin. Une des extrémités 3' envahit l'autre double brin (A), est utilisée comme amorce de réplication pour l'ADN polymérase et déplace le brin homologue de la molécule acceptrice (B). La synthèse d'ADN à partir des deux extrémités 3' de la molécule donneuse permet de combler la délétion (C). Le phénomène de migration des branches engendre deux jonctions de Holliday qui sont ensuite résolues par l'action d'exonucléases (D).

V.L.
X.M

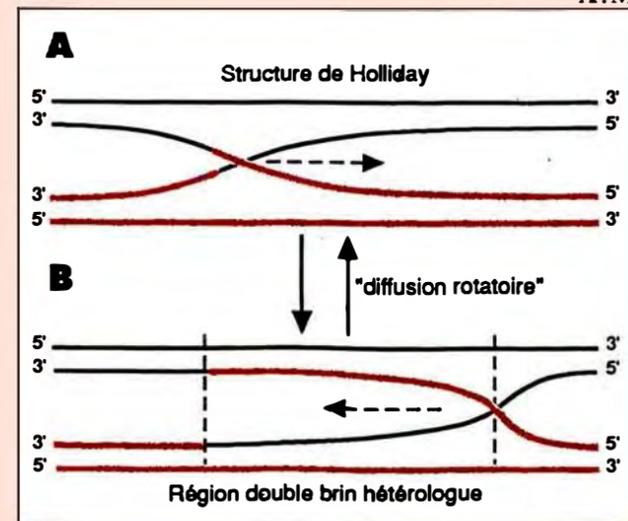


Figure I. Diffusion libre du point de jonction entre deux hélices d'ADN recombinantes.

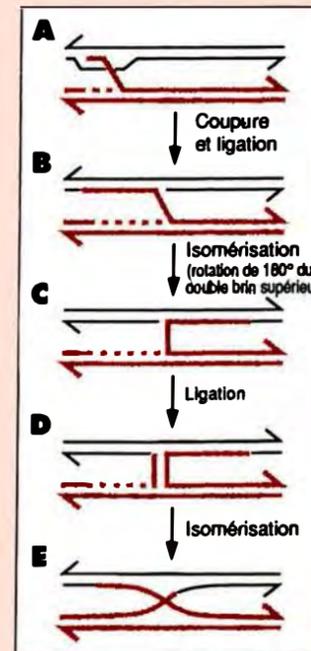
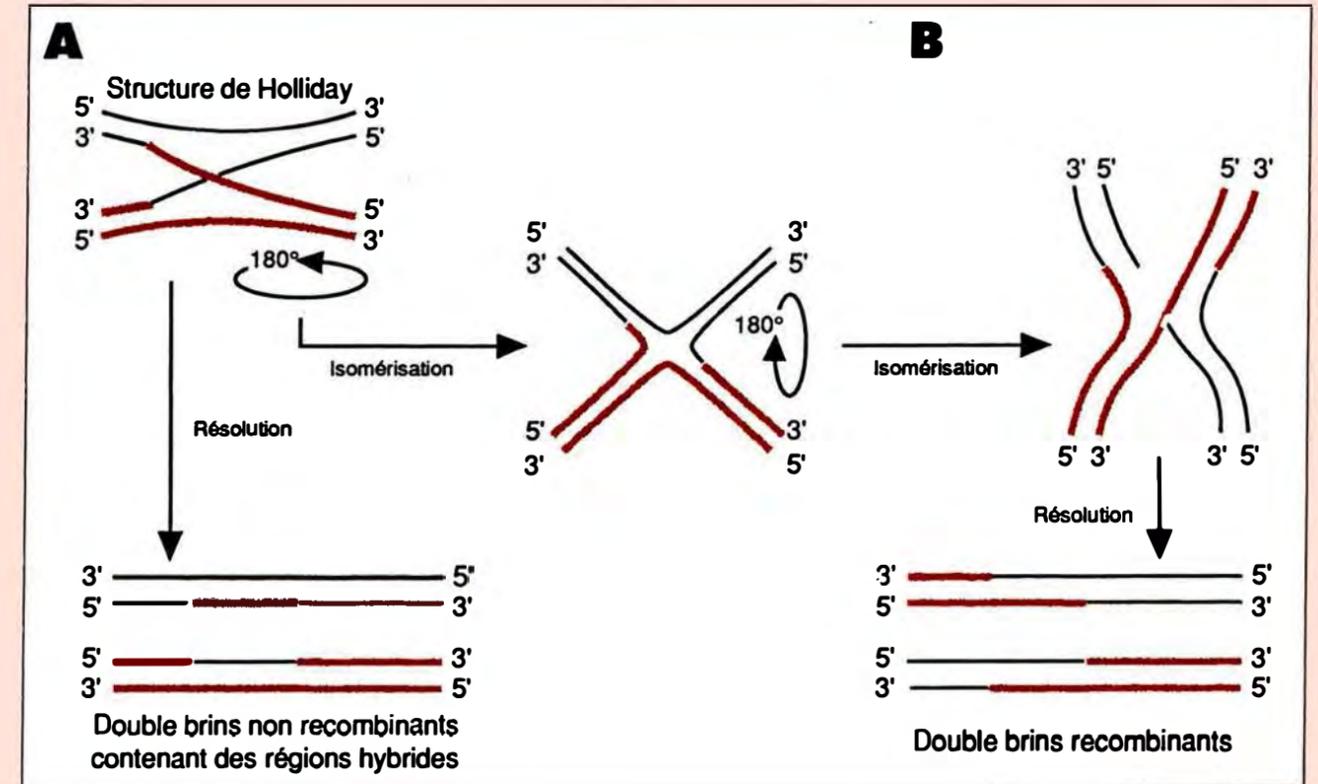


Figure II (en haut). Isomérisation et résolution de la structure de Holliday.

Figure III (à gauche). Le modèle de Meselson-Radding : formation de la structure de Holliday par cassure simple brin.

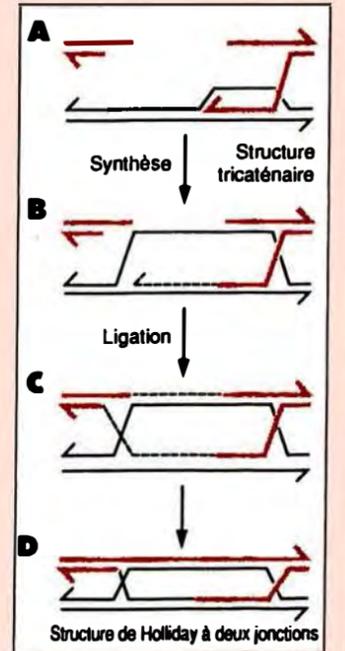


Figure IV (à droite). Le modèle de Szostak et Orr-Weaver : formation de la structure de Holliday par cassure double brin.